Computing Surface

CS-2 Hardware Reference Manuals

83-MS050

meiko

Acceptance

All Meiko software and associated manuals ("the Software") is provided by the Meiko Group of Companies ("Meiko") either directly or via a Meiko distributor and is licensed by Meiko only upon the following terms and conditions which the Licensee will be deemed to have accepted by using the Software. Such terms apply in place of any inconsistent provisions contained in Meiko's standard Terms and Conditions of Sale and shall prevail over any other terms and conditions whatsoever.

Copyright

All copyright and other intellectual property rights in the software are and shall remain the property of Meiko or its Licensor absolutely and no title to the same shall pass to Licensee.

Use

Commencing upon first use of the Software and continuing until any breach of these terms, Meiko hereby grants a non-exclusive licence for Licensee to use the Software.

Copying

Copying the Software is not permitted except to the extent necessary to provide Licensee with back-up. Any copy made by Licensee must include all copyright, trade mark and proprietary information notices appearing on the copy provided by Meiko or its distributor.

Assignment

Licensee shall not transfer or assign all or any part of the licence granted herein nor shall Licensee grant any sub-licence thereunder without prior written consent of Meiko.

Rights

Meiko warrants that it has the right to grant the licence contained under "Use" above.

Warranty

Meiko warrants that its software products, when properly installed on a hardware product, will not fail to execute their programming instructions due to defects in materials and workmanship. If Meiko receives notice of such defects within ninety (90) days from the date of purchase, Meiko will replace the software. Meiko does not warrant that the operation of the software shall be uninterrupted or error free.

Unless expressly stated in writing, Meiko gives no other warranty or guarantee on products. All warranties, express or implied, whether statutory or otherwise [except the warranty hereinbefore referred to], including warranties of merchantability or fitness for a particular purpose, are hereby excluded and under no circumstances will Meiko be liable for any consequential or contingent loss or damage other than aforesaid except liability arising from the due course of law.

Notification of Changes

Meiko's policy is one of continuous product development. This manual and associated products may change without notice. The information supplied in this manual is believed to be true but no liability is assumed for its use or for the infringements of the rights of others resulting from its use. No licence or other rights are granted in respect of any rights owned by any of the organisations mentioned herein.

Nuclear and Avionic Applications

Meiko's products are not to be used in the planning, construction, maintenance, operation or use of any nuclear facility nor for the flight, navigation or communication of aircraft or ground support equipment. Meiko shall not be liable, in whole or in part, for any claims or damages arising from such use.

Termination

Upon termination of this licence for whatever reason, Licensee shall immediately return the Software and all copies in his or her possession to Meiko or its distributor.

Important Notice

FEDERAL COMMUNICATIONS COMMISSION (FCC) NOTICE

Meiko hardware products ("the Hardware") generate, use and can radiate radio frequency energy and, if not installed and used in accordance with the product manuals, may cause interference to radio communications. The Hardware has been tested and found to comply with the limits for a Class A computing device pursuant to Subpart J of Part 15 of FCC Rules which are designed to provide reasonable protection against such interference when operated in a commercial environment. Operation of the Hardware in a residential area is likely to cause interference in which case the user at his or her own expense will be required to take whatever measures may be required to correct the interference.

Computing Surface

CS-2 Hardware Overview

S1002-10M141.00

meiko

Meiko's address in the US is:

Meiko
130 Baker Avenue
Concord MA01742

508 371 0088
Fax: 508 371 7516

Meiko's address in the UK is:

Meiko Limited
650 Aztec West
Bristol
BS12 4SD

Tel: 01454 616171
Fax: 01454 618188

Issue Status:		
	Draft	
	Preliminary	
	Release	x
	Obsolete	

Circulation Control: *External*

Contents

Hardware Overview 1

The CS-2 consists of the following key hardware components:

- The Bay.
- The Modules.
- Processor Boards.
- Network Switch Boards.
- I/O Devices
- The CS-2 data network.
- The CS-2 control and diagnostic network.

This document provides a brief overview of these components and includes pointers to more detailed information in your CS-2 documentation set.

CS-2 Bays

The Bay provides the physical infrastructure for the modules; it provides the power distribution, cabling for the data and control networks, ducting to maximise the flow of cooling air through the modules, and a firm and level securing point for the modules.

The Bay is a scalable tubular structure which can be adapted to suit a number of machine configurations, from the 4 module Half Bay system to the 24 module 3-Bay system. Larger systems are configured as a number of interconnected bays.

Figure 1-1 A Half Bay System

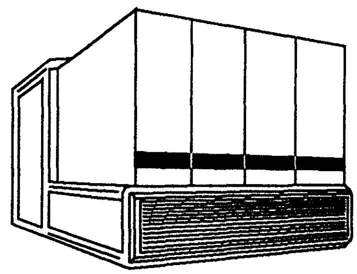


Figure 1-2 A Single Bay System

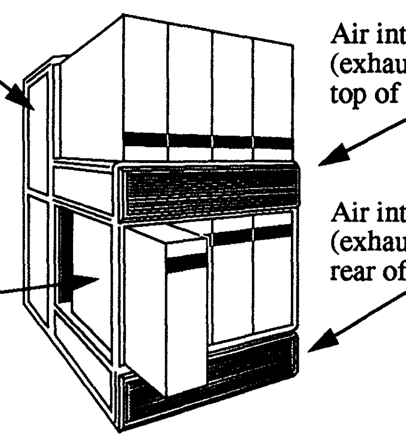
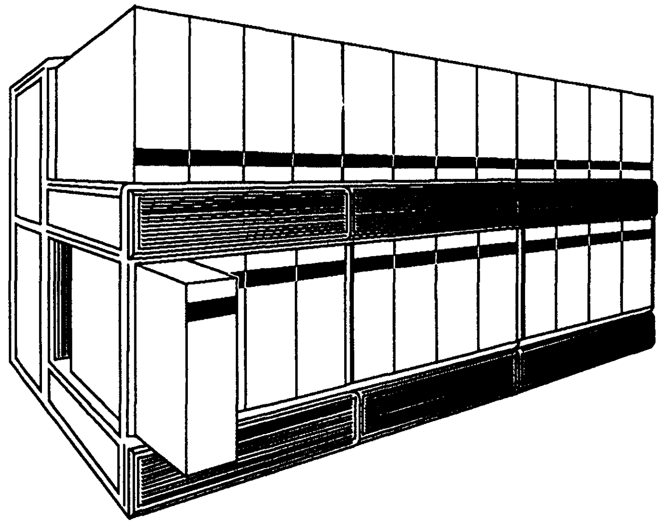


Figure 1-3 A Three Bay System (also referred to as a Cluster)



CS-2 Modules

There are currently three types of module that can be fitted into a Bay: the Processor Module, containing processor boards and a number of SCSI disk devices, the Switch Module, containing switching components for the CS-2 data network, and the Peripheral Module, which contains an array of SCSI disk devices.

All modules include cooling fans, power supplies, and a dedicated module controller processor which is the module's interface to the CS-2 control and diagnostics network. A printed circuit board within each module — the module's backplane board — distributes the power and control signals to the module's contents, and also carries the data buses to the module's rear connectors. Most connections from the module to the Bay are via Beta Flex connections which are fixed to the Bay at each module's mounting position; these are zero insertion force connectors that open and close (thus gripping the module's connections) under the control of an electric current.

More information about the modules can be found in:

- *The CS-2 Processor Module*, Meiko document S1002-10M128.
- *The CS-2 Switch Module*, Meiko document S1002-10M132.

- *The CS-2 Peripheral Module*. Meiko document S1002-10M142.

Processor Boards

Three types of processor board are currently available:

- The SPARC/IO board.
- The multi-SPARC/Compute board.
- The VPU board.

Each board includes one or more SPARC processors, each running a full port of the Solaris operating system. All SPARC processors have an associated Elan Communications Processor (ECP) sharing a cache coherent M-Bus memory interface; the ECP is the SPARCs interface to the CS-2 data network. In addition all boards include a board control processor which is the boards interface to the CS-2 control and diagnostics network.

The SPARC/IO board includes all the functionality of a high performance Unix workstation, with up to 2 SPARC M-Bus processor modules, memory, a range of standard I/O interfaces (including SCSI, Ethernet, and serial), and 3 standard S-Bus peripheral interfaces.

The multi-SPARC/compute board consists of up to 4 SPARC processors each with a dedicated cache coherent M-Bus interface to the memory system and an Elan Communications Processor. The SPARC processors on this board have no additional I/O capabilities and must be configured as diskless Unix clients serving via the CS-2 data network from one or more of the single SPARC boards.

The VPU board has a single SPARC processor, Elan Communications Processor, memory system, and some I/O capabilities (including standard SBus slots). In addition this board includes 2 Fujitsu micro vector co-processors which cooperate with the SPARC processors via a cache coherent M-Bus interface.

More information about the processor boards can be found in:

- *SPARC/IO Processing Element User's Guide*, document S1002-10M137.
- *Quad SPARC/Compute Processing Elements User's Guide*, S1002-10M138.
- *Vector Processing Element User's Guide*, document S1002-10M139.

Network Switch Boards

There are two types of network switch card, both built from one or more Elite Network Switches (ENS).

- Backplane switch cards.
- Module switch cards.

The backplane switch cards are fitted into the rear of the Processor Module directly to the module's backplane; they offer the first level of switching between the processors in the module. Up to 8 backplane switch cards may be fitted to offer full interconnection, on two independent network layers, of 4 multi-SPARC boards (up to 16 SPARC processors).

The module switch cards must be installed in a Switch Module. They provide network switching for higher levels of the network. Three variants are available which may be used in combination to build a switch network for any size of system.

More information about the Meiko Switch boards can be found in:

- *The CS-2 Processor Module*, Meiko document S1002-10M128
- *The CS-2 Switch Module*, Meiko document S1002-10M132.

I/O Devices

The use of standard interfaces on the processor boards (such as SCSI and S-Bus) allows a large range of I/O devices to be connected to your CS-2. These devices will fall into one of the following categories:

- System Console.
- SCSI disk devices.
- Other I/O devices.

The System Console must be a colour graphics workstation connected to one of the processor boards via an S-Bus graphics cards.

SCSI disk devices will typically be fitted into the Processor Module alongside their host processor, or in a dedicated Peripheral Module which may be configured either as a RAID device or as an array of disks.

Other I/O devices will be mounted externally to the CS-2. Cables taken from a processor board's front panel are routed to the rear of the Bay and then via underfloor cable ducts to a dedicated Peripheral Box. In most systems a Peripheral Box containing CD-ROM, QITC, and Exabyte will be located near the system console.

The CS-2 Data Network

The data network is built upon the Elan Communications Processors (ECP), of which there is one for each processing element in your system, and Elite Network Switches (ENS). Interconnection of the Elan and Elite devices is currently by 50-way ribbon cables which are housed in the cable ducts at the rear of the Bay.

The Elan Communications Processor shares a cache coherent M-Bus interface with its host SPARC processor. It offers two independent byte-wide data links onto the network, each capable of bidirectional operation at 50Mbytes/s (100Mbytes/s total). The ECP supports remote read, write, and synchronisation operations, with virtual processor addressing and virtual address translation all handled in hardware. All communications between parallel applications are routed via this device, as well as NFS mounting of filesystems.

The data network is built upon the Elite Network Switches. These are 4x4 cross point switches with byte-wide bidirectional data paths. Switching is byte steered, with the first byte of an incoming message identifying the link used for the outgoing message; route bytes are pre-pended to network packets by the ECP. Broadcasts over a contiguous range of links are supported, with hardware support for the recombination of acknowledge or not-acknowledge tokens from the broadcast destinations. The number of network switches increases logarithmically with the number of processing elements to provide full point to point connectivity and increased network bandwidth for any size of machine with minimal increases in component counts and cabling.

The ECP and ENS are described in more detail in the following Meiko manuals:

- *The Communications Processor Overview*, document S1002–10M100.

- *The Communications Network Overview*, document S1002-10M105.

The Control and Diagnostics Network

The control and diagnostic network is a hierarchical, low bandwidth serial network that runs throughout the system. All modules, boards, and processors have an interface to this network.

The control network carries heartbeat signals from all components (to signal continued operation) and is also used by system software, such as the machine manager and Pandora, to query the operating status of system components, and to reset or reconfigure those components. Remote console connections to the SPARC processors can also be carried over the control network.

The control network is independent of the data network and does not impact on its performance.

The control network is described in the following Meiko document:

- *An Overview of the Control Area Network (CAN)*, document S1002-10M140.

Computing
Surface

The CS-2 Bay

S1002-10M143.01

meiko

Meiko's address in the US is:

Meiko
130 Baker Avenue
Concord MA01742

508 371 0088
Fax: 508 371 7516

Meiko's address in the UK is:

Meiko Limited
650 Aztec West
Bristol
BS12 4SD

Tel: 01454 616171
Fax: 01454 618188

Issue Status:		
	Draft	
	Preliminary	
	Release	x
	Obsolete	

Circulation Control: *External*

Contents

Overview 1

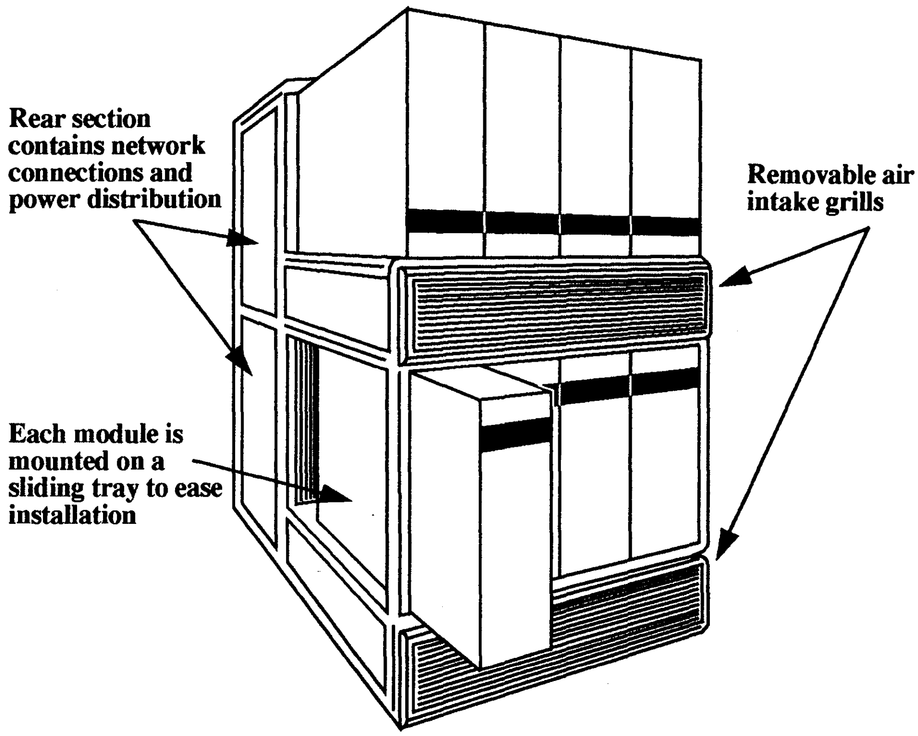
The CS-2 Bay is a tubular structure that forms a platform for the CS-2 modules. It includes the interconnect for the modules' control and data networks, and the power distribution circuits.

The Bay includes the following features:

- Modules are mounted on sliding trays to enable easy installation and removal.
- Power distribution box.
- Enclosed cable ducts for the data network cables.
- Service access panels to front and rear.
- Air channels to maximise flow of cooling air through the modules.

The Bay is a scalable structure which can be adapted to suit a number of machine configurations, from the 4 module Half Bay configuration, to the 24 module 3 Bay system. Larger systems are configured as a number of interconnected Bays. Figure 1-1 shows a Single Bay system.

Figure 1-1 A Single Bay



Service Access 2

The Bay includes service access panels at its front and rear.

You will need to remove the access panels when removing or installing modules. You may also need to remove the access panels when routing peripheral cables from a module; typically these cables will be routed from the front of a module to the rear of the Bay via the ducts beneath the module.

Front Panel (Air Grill) Removal

The lower (charcoal colored) grill covers the air intake and cable guide for the lower modules; the upper grill (purple colored) covers the air intake and cable guide for the upper row of modules, and also hides the baffles which route the warm air exhaust from the lower modules through to the rear of the Bay.

At each end of the grill, and partly hidden from view, is a single captive screw. Use a screwdriver to rotate this screw a quarter turn and remove the grill by pulling away from the Bay. When installing the grill ensure that the locating pins at its ends mate with the holes on the Bay; this will ensure the correct location of the captive screws. Secure the grill in position by a quarter turn of the captive screws.

Rear Panel Removal

Two rear panels are fitted to the rear of each Bay. To remove a rear panel locate the two captive screws along its top edge and rotate each by a quarter turn. Pull the top edge of the panel away from the Bay and lift the bottom edge from its guide. Installation is the reverse of removal; ensure that the small holes along the bottom edge of the panel mate with the pins on the Bay.

Figure 4-1 on page 10 shows the location of the rear access panels and of the retaining screws.

Power Connections 3

The Bay has no power requirements itself but does include a power distribution circuit to provide power to the modules.

Each CS-2 Bay requires the following power connections:

- EUR: 415V, 30Amp, 3 Phase (5 wire).
- US: 208V, 45Amp, 3 Phase (5 wire).

Each CS-2 Bay is supplied with a 3-phase distribution panel and 3m detachable flying lead that is terminated by a 3 phase plug (pin style to IEC 309-1 and IEC 309-2):

- EUR: series 1, 240–415V, 30Amp, 3 pole+Neutral+Earth
- US: series II, 120–208V, 60Amp, 4 pole-5 wire.

Matching female sleeve connectors are required at the customer site.

Module Installation 4

Introduction

To ease installation and removal each module is located on a sliding tray.

Connection of the module to the CS-2 data network is via Beta Flex connectors fitted to the bay. Connection to the control network and power distribution is via cable connection to the module's power supply. Additional connections from the module's boards to external peripherals may be made from the front of the module to the rear of the Bay via cable ducts beneath the module.

Removal and Installation Instructions

The following steps should be used to remove a module from the Bay. Reverse the procedure when installing a module.

Your attention is drawn to the illustrations at the end of this chapter.

1. **Switch-off the module.**
   Use the power switch located at the rear of the module's power supply. In the case of processor modules you must first shutdown each processor via its system console. You should also use Pandora to Configure-Out the processors in the module and to change the link state of it's network switches.

2. **Open the Beta Flex Connectors.**
Modules connect with the data network via Beta Flex connectors that are mounted on the bay. These connectors are electro-mechanical devices that are opened and closed by the application (or removal) of a controlling electric current.

Warning - Before installing or removing a module you must ensure that the Beta Flex connectors are open; failure to do so may result in permanent damage.

Use the Beta Phase controller box to open and close the connectors. From the rear of the Bay connect the controller to a suitable mains outlet and to the 37-way connector located directly behind the module (refer to Figure 4-3). Use the controller's Activate button to open in a single operation all the Beta Flex connectors used by that module — pressing the same button a second time will close those connectors. Status lights on the box show the condition of each connector; red indicates open, amber indicates power is being applied, green indicates no power (connector closed). When using the Beta Phase Controller it is safe to hold the connectors open for prolonged periods.

3. **Disconnect control bus and power cables.**
From the rear of the Bay the module's control bus and power cables are readily accessible. Remove the power supply and the external CAN connections (if any).

4. **Disconnect peripheral cables from front of module.**
Some switch and processor boards may have cable connections to their front panels. Remove the front of the module and its LED board (by simply pulling forward) and remove any cable connections. Unscrew the cable retainer on the front face of the sliding tray and allow the cables to hang clear of the module and the tray (refer to Figure 4-2).

5. **Slide-out module tray.**
The slide tray is held closed by two captive retaining screws (refer to Figure 4-2). Release these screws and carefully pull the tray and the module forward.

Warning - Never extend more than one module at any one time.

---

6. **Release the module from the slide tray.**
   The module is held in position by two captive screws on the underside of the sliding tray (refer to Figure 4-2). These screws must be released before the module is lifted clear of the tray.

   When re-installing a module you must ensure that the four feet on its base mate with the holes provided in the sliding tray. Correct alignment ensures that the captive screws are easily secured.

7. **Lift module clear of tray.**
   The module is heavy and must be lifted by two people or by using the Genie lifting equipment that is supplied by Meiko (large multi-module systems only).

   When using the Genie fork lift you must first attach the two lift clamp assemblies (black steel boxes) to either side of the module. Undo the screw that holds the telescopic section of the lift-clamp closed. Insert the rear locating pins into the holes at the rear edge of the module (the open end of the box section should face the front). Align the front locating pins with the holes at the front of the module and tighten the holding screw. Locate the forks of the genie lift into the two box sections in the lift-clamps, and raise the module by turning the lift's handle anti-clockwise.

Warning - Your attention is drawn to the operating instructions that are supplied with the Genie lift.

Illustrations

Figure 4-1 shows a Single Bay with its front, side, and top panels removed. This illustration shows the Bay's tubular construction, and the position of the sliding trays, the rear access panels and their retaining screws, and the air and cable ducts below the modules.

Figure 4-1 Skeletal Illustration of a Single Bay

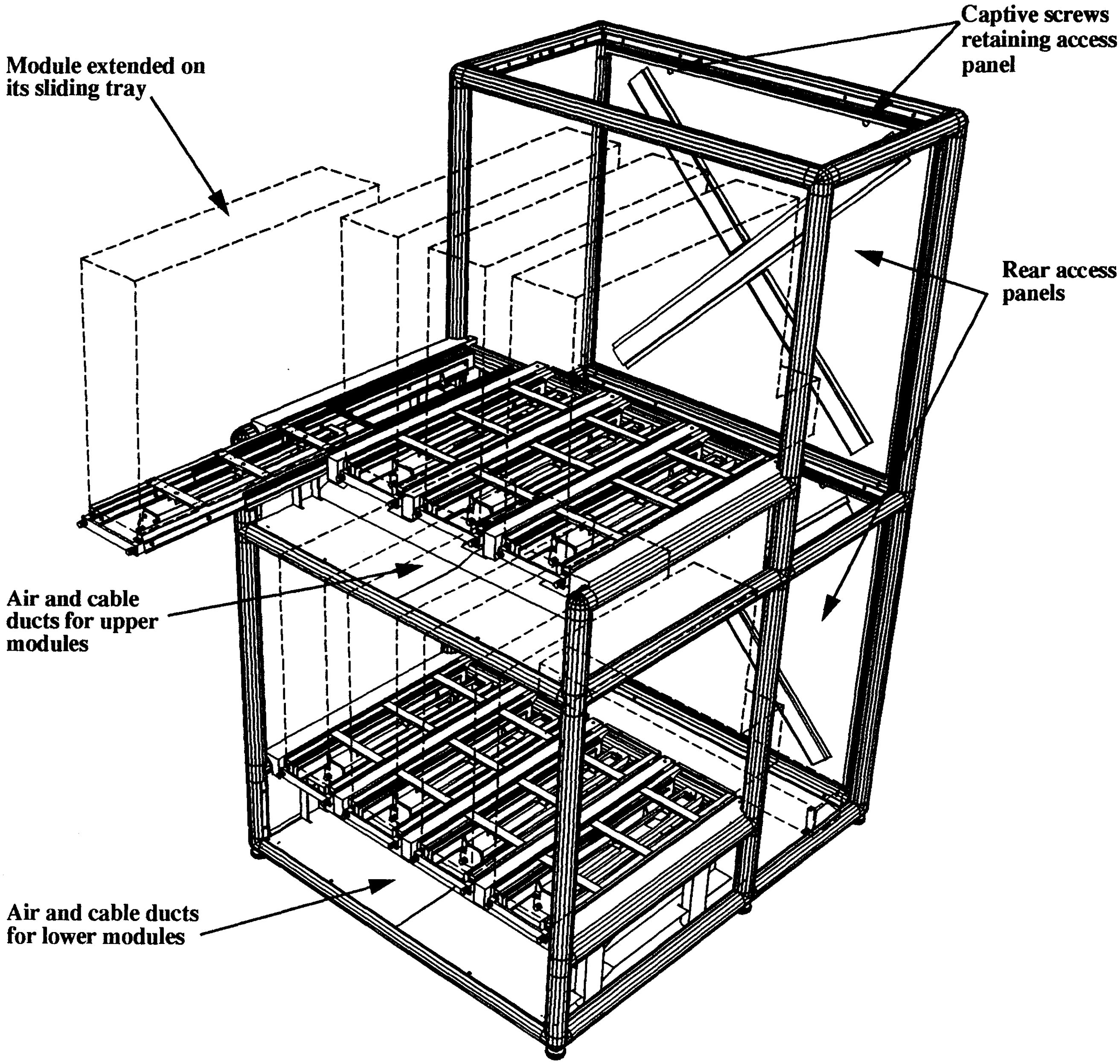


Figure 4-2 shows a module's sliding tray when extended and viewed from its underside. It shows the two captive screws that are used to hold the module in place, and the two screws that hold the tray closed. Cables that are connected to the front of the module's boards are routed behind the retaining bar and via cable ducts to the rear of the Bay.

Figure 4-2 A Module Sliding Tray Viewed from Below

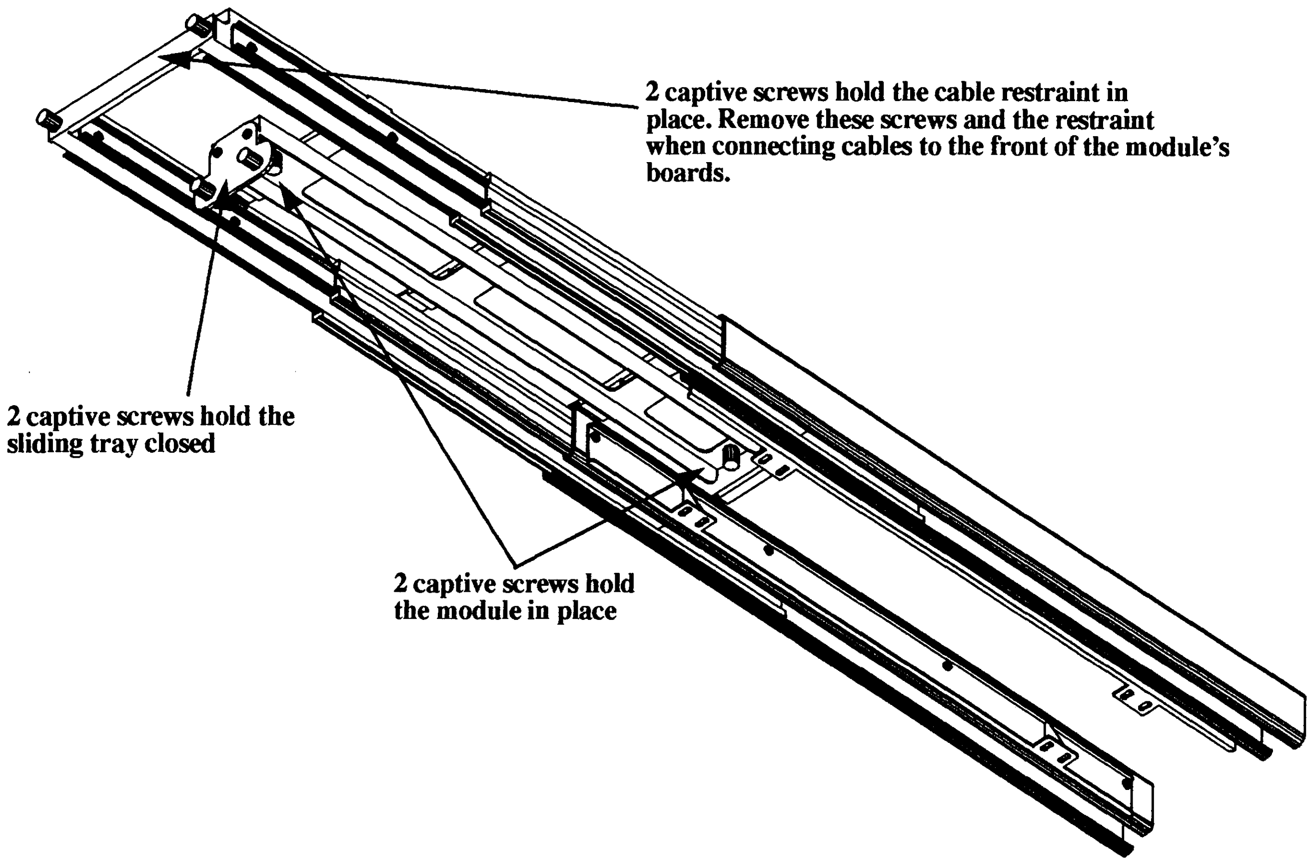
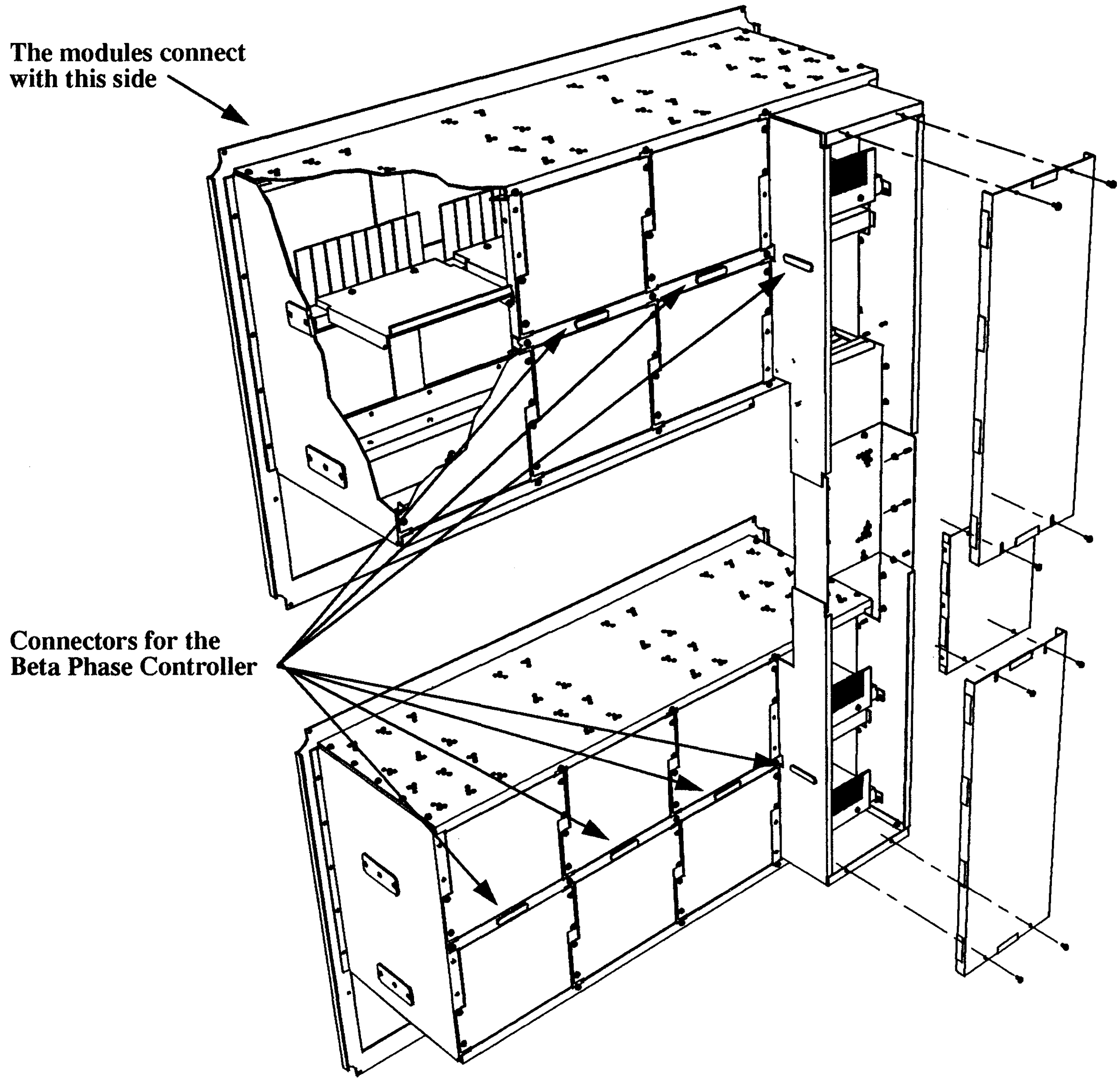


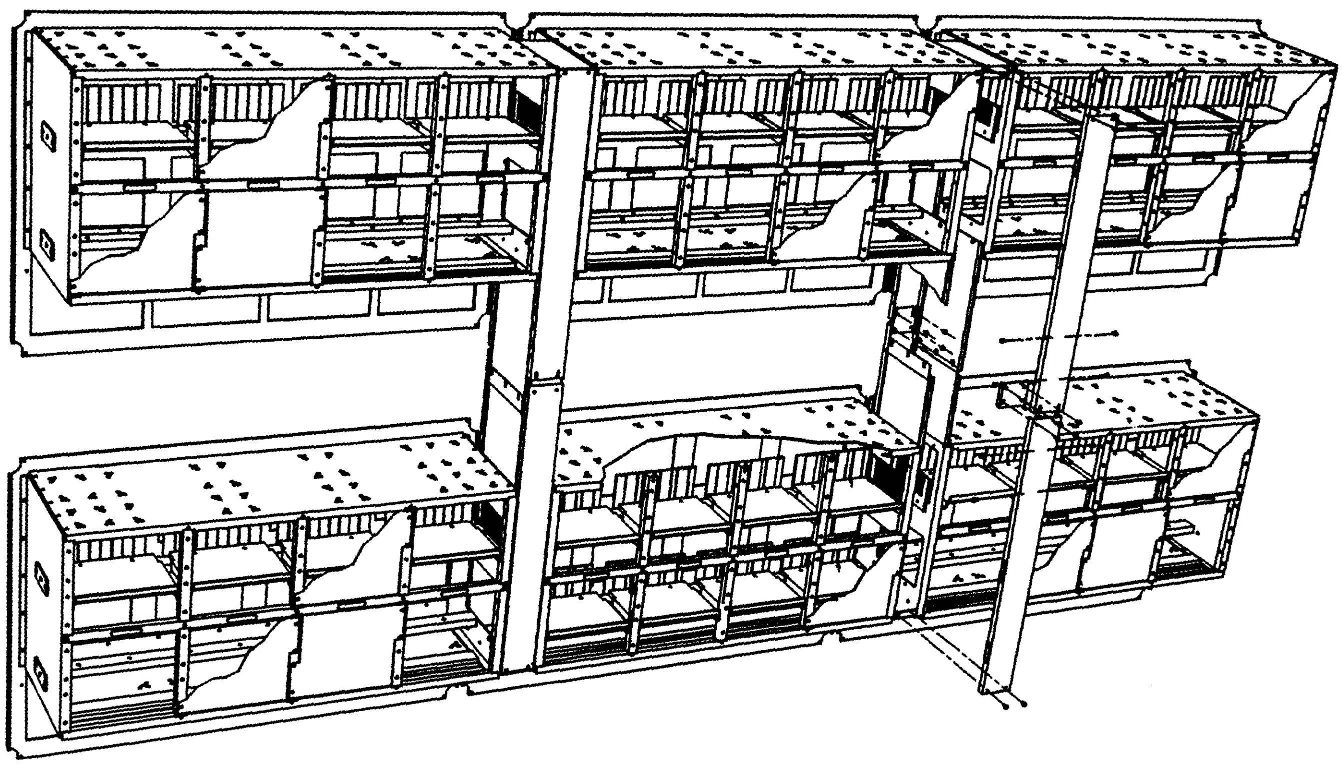
Figure 4-3 shows the backplane box for a Single Bay. This box encloses the cabling for the data network. It is located at the rear of the Bay, directly behind the rear service panels. Note the 8 connectors that are used by the Beta Phase Controller Box to open and close each module's Beta Phase connectors.

Figure 4-3 Backplane Box for a Single Bay System



The backplane box for 1, 2, and 3 Bay systems are slightly different in configuration. Figure 4-4 shows the backplane box used for a 3 Bay system.

Figure 4-4 Backplane Box for a 3 Bay System



Computing Surface

The CS-2 Processor Module

S1002-10M128.02

meiko

Meiko's address in the US is:

Meiko
130 Baker Avenue
Concord MA01742

508 371 0088
Fax: 508 371 7516

Meiko's address in the UK is:

Meiko Limited
650 Aztec West
Bristol
BS12 4SD

01454 616171
Fax: 01454 618188

Issue Status:		
	Draft	
	Preliminary	
	Release	X
	Obsolete	

Circulation Control: *External*

Contents

Federal Communications Commission (FCC) Notice

Warning: Changes or modifications to this unit not expressly approved by the party responsible for compliance could void the user's authority to operate the equipment.

NOTE: This equipment has been tested and found to comply with the limits for a Class A digital device, pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. The equipment generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the instruction manual, may cause harmful interference to radio communications. Operation of this equipment in a residential area is likely to cause harmful interference in which case the user will be required to correct the interference at his own expense.

Shielded cables must be used with this unit to ensure compliance with the FCC Class A limits.

General Precautions

- Adhere strictly to the instructions and warnings provided in this documentation.
- Do not push objects into the interior of the equipment through any openings. Hazardous voltages and moving parts are present. Conductive objects could cause electric shock, fire, or damage to the equipment.
- It is not permitted to make mechanical or electrical modifications to the equipment. The manufacturer is not responsible for regulatory compliance of equipment that has been modified. You may also invalidate your warranties if you make unauthorised changes to your equipment.
- When connecting a visual display unit to your module your attention is drawn to the instructions that are issued by the device manufacturer.
- Your module may contain disk devices. Your attention is drawn to the instructions that are issued by the device manufacturers.
- A lithium battery is installed on many of the boards in the module — it is an integral component of the non-volatile RAM (or NVRAM). Lithium batteries are not customer replaceable. If they are mishandled there is a danger that they may explode. Always consult Meiko if you suspect that the battery needs replacing. Never dispose of lithium batteries in a fire or attempt to dismantle.
- Many of the electronic components fitted in the module are fragile or static sensitive. They should only be handled by trained engineers. You must observe anti-static precautions when handling boards or electronic devices.
- Operation of this equipment in a residential area may cause unacceptable interference to radio and TV reception, requiring the operator to take whatever steps are necessary to correct the interference.

Lifting

The module is heavy and must be lifted by two people or by using the Genie lifting equipment that is supplied by Meiko (large multi-module systems only).

When using the Genie fork lift you must first attach the two lift-clamp assemblies to either side of the module. Unscrew the securing screw that holds the telescopic section of the lift-clamp closed. Insert the rear locating pins into the holes at the rear edge of the module (the open end of the box section should face the front of the module). Align the front locating pins with the holes at the front of the module and tighten the holding screw.

Locate the forks of the Genie lift into the two box sections in the lift clamps. Raise the module by turning clockwise the handle on the Genie lift. Lower the module by turning the handle anti-clockwise.

Warning - Your attention is drawn to the operating instructions for the Genie lift that are supplied by the lift's manufacturers.

- Genie fork lift, part number 73-FORK-GENIE.
- Module lifting gear, part number 65-MODULELIFT

Overview 1

The Processor Module is a physical enclosure for up to 4 processor boards. Supporting infrastructure within the module consists of cooling fans, power supply, and connections to the CS-2 data network, the CS-2 control network, and external SCSI peripheral buses. Up to 4 SCSI disk devices may also be fitted within the module and connected to the processor boards in one of three standard configurations (all disks connected to one board, two disks to each of two boards, or one disk per board). An LED panel at the front of the module offers a 4x4 array of LEDs per board which are software controlled. The module is fully enclosed and FCC compliant.

The CS-2 processor module is intended for installation on a CS-2 bay which provides a firm level support, network interconnect, and power distribution.

Module Packaging 2

CS-2 modules are shipped in units of 1 to 4. The packaging consists of a wood base, packing foam, antistatic bag, and enclosing triple wall card outer.

Packing Dimensions

All dimensions are approximate.

Packing dimensions (single module):

Height	74cm (29")
Width	34cm (13.5")
Length	112cm (44")

Packing weight (without module):

4 module skid	20Kg (44lbs)
Module base	4.5Kg (10lbs)
Filler	1.8Kg (4lbs)
Outer card cover	4.5Kg (10lbs)
Packing foam	0.2Kg (0.4lbs)
Total (1 module)	**11Kg (24.4lbs)**
Total (4 modules)	**31Kg (68.4lbs)**

Module weight (without processor boards or peripherals):

Module weight	43.5Kg (96lbs)

Processor board weights (all processor options fitted):

MK401 (Dino)	2Kg (4.4lbs)
MK405 (Quatro)	3Kg (6.6lbs)
MK403 (VPU)	2.7Kg (6.0lbs)

Backplane boards:

MK515 Controller	0.15Kg (0.33lbs)
MK516 SCSI card	0.15Kg (0.33lbs)
MK511/2 Switch	0.15Kg (0.33lbs)

Peripherals:

Disk carrier	0.6Kg (1.3lbs)
Disk device	0.8Kg (1.8lbs)
Total per disk	**1.4Kg (3.1lbs)**

Unpacking

Warning - You are reminded of the safety precautions listed in *General Precautions* **on page ii.**

Modules are shipped in groups of 4 or 1.

For 4 module shipments each module is packaged individually and secured with the others onto a large skid. To unpack a four module skid cut the outer banding and remove one module. Space the remaining three module packs uniformly over the skid.

- To unpack each module first cut the banding and remove the outer card carton by lifting it clear of the module.
- Remove the protective foam from the top and front of the module.
- Remove the antistatic bag by pulling upwards.
- Lift the module from the packing base. Note: the module is heavy and must be lifted by at least two people or by using a Genie lift.
- Check the module for damage and advise the transportation company immediately if any is found.

Retain all packaging and use it when shipping the module.

Installing the Processor Module 3

You must read this chapter thoroughly and completely before using your CS-2 module.

Location

The processor module must be located in a CS-2 Bay. This provides a firm level support for the module and has been designed to maximise the flow of cooling air through the module.

Warning - Overheating can damage this product. Do not block or cover any openings that are built into the module, and do not place the module near any sources of heat.

Modules sit on module trays that are fixed to the bay by telescopic rails. The tray must be extended from the bay before a module may be loaded onto it or removed from it.

Warning - Never extend more than one loaded module tray.

To extend the tray first unscrew the two retaining screws and then pull the tray forward. Lift the module onto the tray ensuring that the feet on the base of the module mate with the holes in the tray. Fix the module into position using the captive screws on the underside of the tray.

Warning - Your attention is drawn to the lifting instructions in *Safety Precautions* on page i.

The bay may carry a number of Beta Phase connections to the module. These connectors must be opened using the Meiko Beta Phase Controller before the module tray is pushed into position.

Warning - Failure to open the Beta Flex connectors before pushing the module trays into position will damage your equipment.

With the Beta Phase connectors held open by the Beta Phase Controller push the module tray into position. On each of the telescopic arms is a spring slip that must be pushed-in to allow the arm to slide. Push the tray into position carefully, ensuring that the Beta Phase connectors are correctly situated. Lock the module tray into position using the two captive screws at the front of the tray.

After pushing the module tray into position use the Beta Phase Controller to close the connectors.

Using the Meiko Beta Phase Controller

Connect the Beta Phase controller to the 37-way connector located behind the module mounting position. Connect the Beta Phase Controller to a mains outlet using the supplied cable. Use the Activate button on the box to open all the Beta Phase connectors used by that module — pressing the same button a second time will close the connectors. Status lights on the box show the condition of each connector; red indicates open, amber indicates power is being applied, green indicates no power. When using the Controller it is safe to hold the Beta Phase connectors open for prolonged periods.

Power Supply

Check that the voltage and frequency of your power supply is suitable for your equipment. The power supply unit is auto-ranging, 110–230 V, 50/60/400 Hz. Maximum current is 15 A.

Power Connections

Power is supplied to the processor module by a 3-way cable that is supplied by Meiko.

The following cable types are available:

- Harvey Hubble to Harvey Hubble.
- Harvey Hubble to UK 13 A outlet.
- Harvey Hubble to US 15 A outlet.

Where several processor modules are used Meiko may also supply a mains distribution board. This connects eight modules to one power supply outlet.

- Power distribution panel, European. Part number 65-DIST-EUR-01
- Power distribution panel, USA. Part number 65-DIST-US-01

Warning – You must only use power cords and distributions panels that are supplied by Meiko. These have the correct power rating for your system.

Warning – You should check that you have been supplied with the correct type of power cord and distribution boards.

Fuses

Five fuses are located at the rear of the processor module as shown in Figure 3-1 on page 15. One fuse protects the power supply unit from the incoming mains supply. Two fuses protect the fans and 2 protect the disk drives.

- Main fuse: standard 250 V, 15 A. Part number 22-F0100-04E150.
- Fan fuse: anti-surge 250 V, 10 A. Part number 22-F0100-04E100.
- Fan fuse: anti-surge 250 V, 5 A. Part number 22-F0100-04E500.
- Disk fuse: anti-surge 250 V, 10 A. Part number 22-F0100-04E100.
- Disk fuse: anti-surge 250 V, 5A. Part number 22-F0100-04E500.

There are 4 fuses located on the module backplane alongside the 4 SCSI card connectors (see Figure 3-2 on page 16). These fuses protect the ±12 V and the ±5 V circuits on both the backplane and all boards that draw their power supply from it (including the backplane cards and the LED panel). Access to these fuses is gained from the rear of the module and requires the SCSI backplane cards to be removed. All 4 fuses are the same type:

- 1.25×0.52", 4 A anti-surge, HCR ceramic. Part number 22-FU100-03E400.

Warning - The backplane fuses may only be changed by trained engineers.

Operation

The power supply is operated by the main on/off switch located above the power connector. 0 represents off, 1 represents on.

When first switched on five LEDs are visible through the rear of the power supply unit. Initially three of these should be green, two are red. After a short delay all LEDs should be green.

Warning - A fault is present if any light remains red for more than 10 seconds after first switching on, or changes from green to red during operation. Switch the power supply off, disconnect from the main supply, and contact Meiko for advice.

Removable Panels

Warning - Ensure the power supply is switched off and disconnected from the main supply before removing the module's front or side panels.

The front and side panels may be removed from the module. Removing the front panel gives access to the LED board, the processor boards, and the disk devices. The side panels may be removed but there is no requirement to do so, except to protect them during transit or when lifting the module into a bay.

Warning - Processor boards and disk devices should only be fitted or removed by trained engineers.

To remove the front panel pull it forward. The front panel is retained by four clips, one in each corner of the panel. When fitting the front panel ensure that the tinted window is aligned with the module's LED display.

The side panels are held in place by 5 key slots on the side of the module. To remove the a side panel slide it forward and pull it off.

External Connections

The processor module includes a number of connectors to the front and rear of the module.

At the Rear of the Power Supply Unit

1×25-way D-type connector:
RS232 connector for module control card diagnostics. This connection is for use by Meiko's trained engineers only.

2 ×9-way D-type connectors:
X-CAN and G-CAN connectors. The X-CAN is the middle connector — this is used to connect the modules within a cluster (up to 3 bays/24 module systems). The G-CAN is the right-most connector — this is used to interconnect clusters. When using more than one cluster at least two modules in each cluster must be connected to the G-CAN; all unused G-CAN connectors must have terminators fitted. Within a cluster the X-CAN connections are daisy-chained using Meiko supplied cables.

- G-CAN terminators, part number 60-CS2MA009-1T.
- X-CAN interconnect, part number 60-CA0246-1T

At the Rear of the Module

Upper 8 cards slots:
Up to eight module switch cards (MK511 or MK512) are fitted in the upper card slots.

Lower 5 card slots:
Up to 4 module SCSI cards (MK516) and 1 module control card (MK515) may be fitted to the lower card slots.

Warning - The module switch cards, module controller card, and the module SCSI cards are not user serviceable. They should only be fitted or removed by trained engineers.

At the Front of the Module

Behind the front panel are four polarised 50-way connectors. Two of these are used to connect to the MK525 module LED board. Fit the board to the upper connectors when the module is mounted on the lower shelf of a bay, and the lower connectors when the module is mounted on the upper shelf.

The connectors for the 4 processor boards and the 4 disk carriers are also located behind the front panel.

Warning - Processor cards and disk devices are not user serviceable. They should only be fitted or removed by trained engineers.

External Indicators

A number of LEDs are visible through the tinted window of the module's front panel.

Four 4×4 matrices of red LEDs are driven by the processor boards. The SPARC processor on single SPARC boards controls all 16 LEDs, whereas each processor on a 4-SPARC board controls one row or four lights. The default patterns shown on the LEDs are generated by the Boot PROM and the Operating System — the pattern may be changed by the user under software control. When power is first applied to the boards all red LEDs should illuminate, fade, and then display a random pattern.

Below each matrix of red LEDs are 3 more LEDs. The green and amber lights are driven by the processor board's CAN interface — the green light is the board's heart-beat, the amber light indicates a transmit of data. The red light is a status light for an optional hard disk that is installed below the processor board.

Two LEDs in the top right hand corner of the display are driven by the module controller. The green light is a heart-beat signal, the amber light flashes each time the module controller is sending information over the CAN bus.

Module Identification Number

At the rear of the power supply unit are three dials — each represents one nibble from a 12 bit network address, or *netid*. The netid for each module must be unique and modules must be numbered sequentially starting from 0.

Enter each nibble of the address onto the dials using a suitable screwdriver.

Note: for two module systems the module id's must differ in the least significant bit only. Module id's of 0 and 1, or 2 and 3 are therefore permitted, but 1 and 2 are not.

Warning - You must ensure that the power is disconnected before setting the module address.

Additional Note for Revision A MK526:

Having identified a module's network address each nibble of the 12 bit address is bit-flipped (e.g. 0111 becomes 1110). The resulting 3 nibbles are then entered onto the dials using a suitable screw driver.

The following table summarises the module id's for systems of up to 32 modules. The dials for module 5, for example, must be set to 00A.

Module	Dials			
0-3	0 0 0	0 0 8	0 0 4	0 0 C
4-7	0 0 2	0 0 A	0 0 6	0 0 E
8-11	0 0 1	0 0 9	0 0 5	0 0 D
12-15	0 0 3	0 0 B	0 0 7	0 0 F
16-19	0 8 0	0 8 8	0 8 4	0 8 C
20-23	0 8 2	0 8 A	0 8 6	0 8 E
24-27	0 8 1	0 8 9	0 8 5	0 8 D
28-31	0 8 3	0 8 B	0 8 7	0 8 F

Operating Conditions

Recommended operating temperature	10–25°C (50–80°F)
Peak operating temperature	32°C (90°F)
Temperature gradient	10°C (18°F) per hour
Storage temperature	-18–60°C (0–140°F)
ANSI media data integrity (max.)	32°C (89.6°F)
Relative humidity	20–80% non-condensing
Storage humidity	10–90% non-condensing
Altitude	3000m (10000 feet)
Operating Shock	1G

Figure 3-1 The CS-2 Processor Module

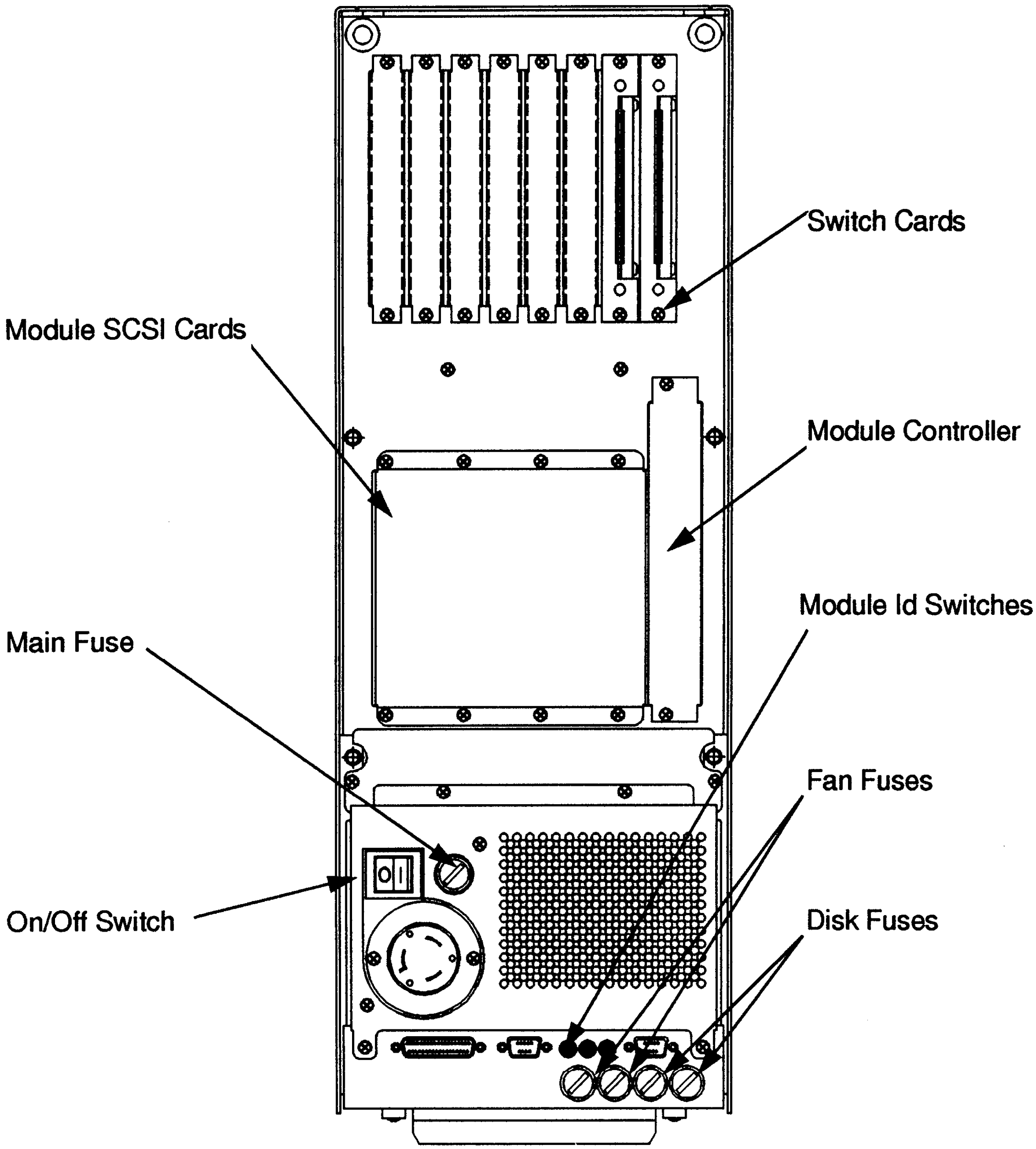
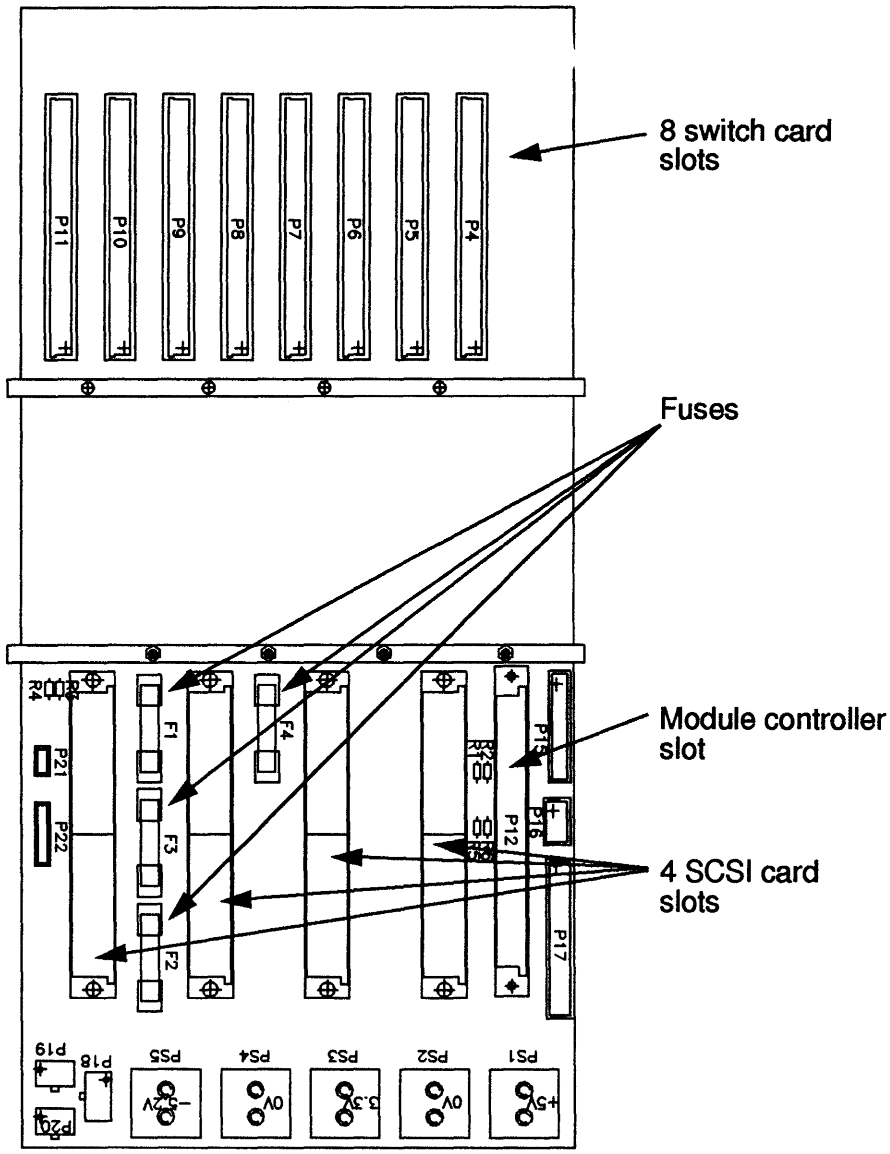


Figure 3-2 The MK510 Processor Module Backplane (not user serviceable)



Installing Processor Boards 4

Warning - The procedures in this chapter must only be undertaken by trained engineers.

Warning - You must disconnect the power supply before installing processor boards into the processor module.

Three types of processor board are available:

- MK401 (single-SPARC with I/O), also known as Dino. This is a general purpose compute processor and operating system server.
- MK405 (quad-SPARC board), also known as Quatro. This is a high performance compute server.
- MK403 (single-SPARC, dual Fujitsu VPU), also known as VPU. This is a high performance vector processor.

Blanking plates must be installed in unused board slots to ensure correct cooling and compliance with RFI regulations.

Installation

The same installation procedure is used to install all three board types.

Insert the board so that it fits into the guide rails at the top and bottom of the module, ensuring that the component side is to the left (viewed facing the module). Gently push the board squarely on the front panel. Before pushing the board fully into position fold back the levers at each end of the front panel so that they are at 90° to the board. Push the board in further until the base of the two levers is touching the card cage; push the board into position by pushing firmly on both levers until they are flat. Secure using the captive screw at each end of the board's front panel.

To remove the board pull both levers out — this will lever the board away from the module's connections. Pull the board clear of the module.

You should take care not to damage the connectors within the module.

You should also take care not to damage the RFI (copper) seals along the edge of the board's front panel.

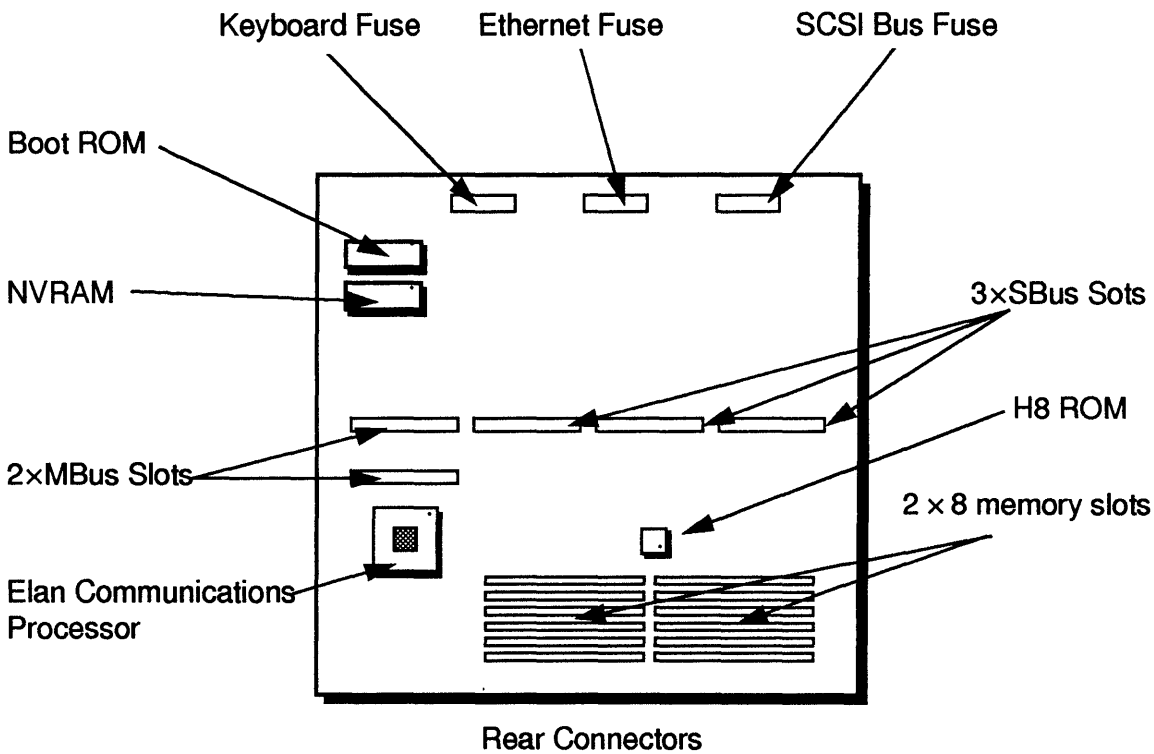
MK401 Single-SPARC Board

Field Upgradeable Components

The MK401 has the following field upgradeable components (see Figure 4-1):

- Superscalar SPARC processor module fitted to MBus slot.
- Three SBus slots.
- 16 memory slots.
- Boot ROM.
- H8 ROM.
- Non-volatile RAM (NVRAM).
- Fuses.

Figure 4-1 MK401 Components



Take care not to damage the component or its neighbours when installing or removing these devices.

MBus Processor Modules

Superscalar SPARC modules may be installed in the two MBus slots — when using just one module it must be installed in the slot nearest the communications processor (the one nearest the back of the board). Installation is simple — push the processor into the connector on the MK401 and secure into position with a screw at each corner.

As newer higher performance SPARC modules become available, Meiko will offer these as field upgrade options. Simply unplug the old unit and replace with the new.

SBus Modules

Three SBus slots are provided and these may be fitted with standard SBus modules. As with the processor modules SBus modules are simply plugged into the SBus connectors and secured with 2 screws. When using SBus cards that have external connections, for example a graphics card, remove the appropriate panel from the front of the MK401 — the panel is held in place by two small screws.

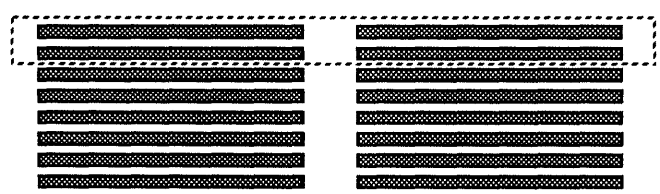
SBus devices are numbered from 0 to 2, device 0 being next to the processor slots. Device 4 is the on-board Ethernet and SCSI bus. Device 5 is the second onboard SCSI bus.

Memory

Up to 16 single in-line memory modules (SIMMs) may be fitted to the MK401 board (JEDEC 36 bit SIMMs). The array is constructed of 4 groups of 4 SIMMs, arranged as shown in Figure 4-2. Within each group the SIMMs must be identical, but there is no requirement for the groups to be the same.

SIMMs can be either 4Mbit DRAM or 16Mbit DRAM technology, either single or double sided. This gives a minimum memory configuration of 16Mbytes and a maximum of 512Mbytes.

Figure 4-2 Memory slots on an MK401; the dotted region shows one of the 4 groups of 4 memory slots.



Boot ROM

The boot ROM may be upgraded from time to time. It is held in a DIL socket and is easily replaced.

H8 ROM

The H8 ROM may be upgraded from time to time. It is held in a socket and is easily replaced using the appropriate tool.

NVRAM

The non-volatile RAM holds system configuration information, environment variables, and the clock. It is held in a DIL socket and is easily replaced. Note however that the information within the NVRAM can only be restored by Meiko's engineers.

The NVRAM contains lithium batteries which have special handling and disposal requirements — your attention is drawn to the *General Precautions* on page ii.

Fuses

There are 3 fuses on the MK401.

Keyboard fuse: 20 × 5mm 1 A quick blow ceramic.

Ethernet fuse: 20 × 5mm 1 A quick blow ceramic.

SCSI bus fuse: 20 × 5mm 500mA quick blow ceramic.

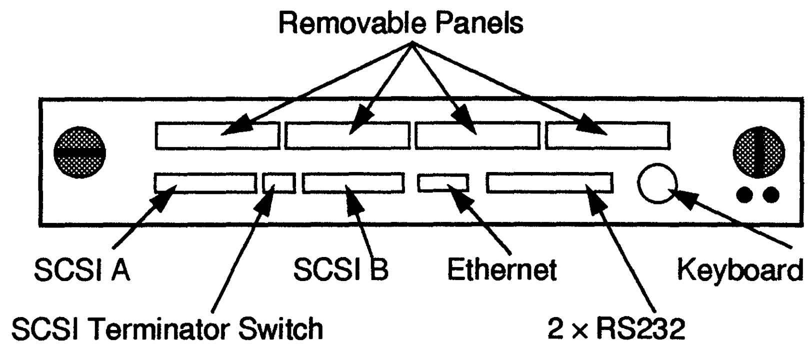
External Connections

External connections are provided for a keyboard/mouse (8 pin circular socket), RS232 interfaces (2 channels provided by one 25-way D-type connector), Ethernet (15-way D-type connector), and two independent SCSI buses (each via a 50-way miniature connector).

Front Panel Connections

Removable panels provide access to connectors on the optional SBus boards.

Figure 4-3 MK401 Front Panel Connections



RS232 Connections

The two RS232 channels are connected as follows. Signal ground is pin 7, chassis ground in pin 1.

Table 4-1 RS232 Channel A Pinout.

Signal	Input/Output	Pin number
TXD	Out	2
RXD	In	3
RTS	Out	4
CTS	In	5
CSR	In	6
DCD	In	8
DB	In	15

Table 4-1 RS232 Channel A Pinout.

Signal	Input/Output	Pin number
DD	In	17
DA	On	24
DTR	On	20

Table 4-2 RS232 Channel B Pinout.

Signal	Input/Output	Pin number
TXD	Out	14
RXD	In	16
RTS	Out	19
CTS	In	13
DCDB	In	12

Adding SCSI Peripherals

The MK401 includes two SCSI-2 controllers — both SCSI buses are available via connections on the front panel or via the processor module backplane. Up to 3 additional SCSI buses can be added using standard SBus cards, and these are available via connections on the SBus cards.

The 2 on-board SCSI buses can be taken from the processor module backplane by using an MK516 module SCSI card. This transfers the buses onto polarised ribbon cable connectors which can be wired to SCSI disks within the processor module.

The MK516 can support up to 5 SCSI buses, only 2 are currently used.

SCSI Termination

Two switches on the MK401 front panel allow the SCSI bus termination to be switched; with the switches down termination is on.

If *either* the front panel connections *or* the backplane connections are used, or no devices are attached at all, then the on-board termination must be turned on. Only when both front-panel and backplane connectors are used together should the on-board termination switches be turned off, and in this case the SCSI bus must be terminated at both ends.

External Indicators

Two LEDs (one green, one amber) are included on the front panel. The green LED is the heart beat from the board's CAN controller. The amber light illuminates each time the CAN controller transmits on the CAN bus. Both should flash steadily. These indicators are also displayed on the module's LED display.

The green LED flashes at a slow steady rate (once per second) when operating normally. A quicker flash rate (2×normal) indicates that the SPARC processor is not responding; a very quick flash rate (3×normal) indicates that the H8 processor on the module controller is not responding.

Each processor board within a processor module controls a 4×4 matrix of red LEDs on the module's front panel. The MK401 displays a random pattern on these when running the Boot ROM. When Solaris has been booted a circulating pattern is displayed. The pattern can be changed by user programs.

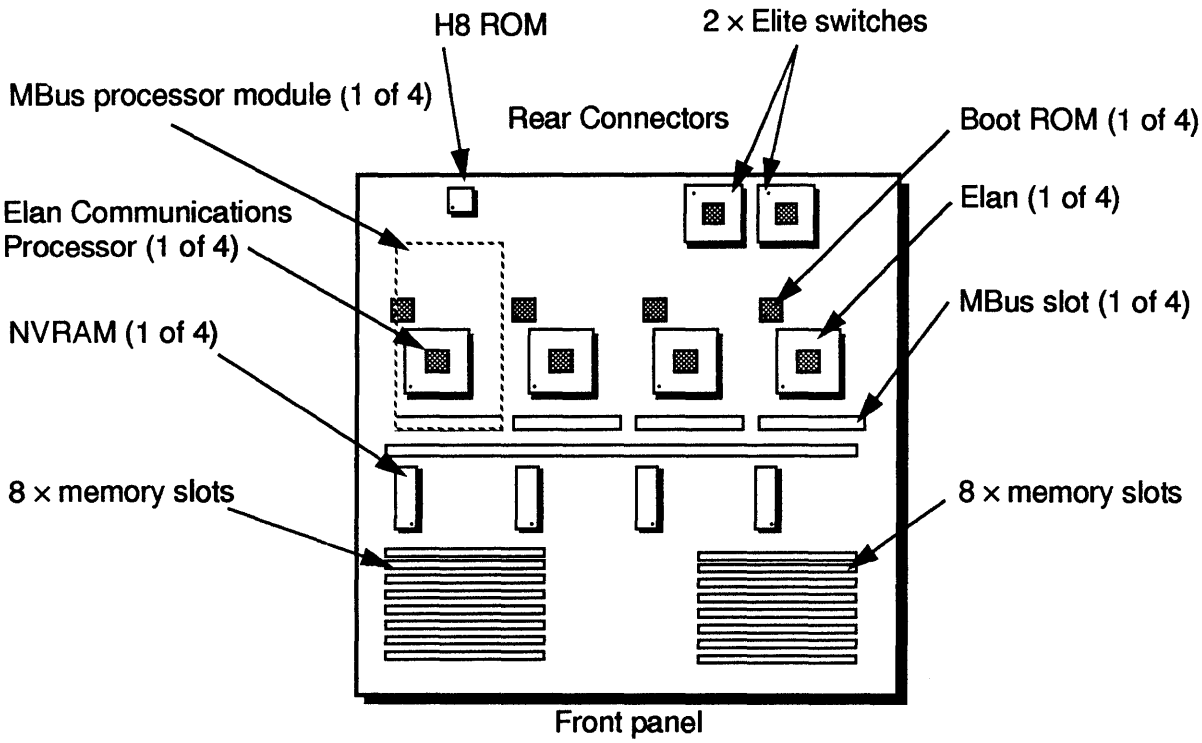
MK405 Quad-SPARC Board

Field Upgradeable Components

The MK405 has the following field upgradeable components (see Figure 4-4):

- Superscalar SPARC processor modules fitted to MBus slots.
- 16 memory slots.
- Boot ROMs.
- H8 ROMs.
- Non-volatile RAM (NVRAM).

Figure 4-4 MK405 Components



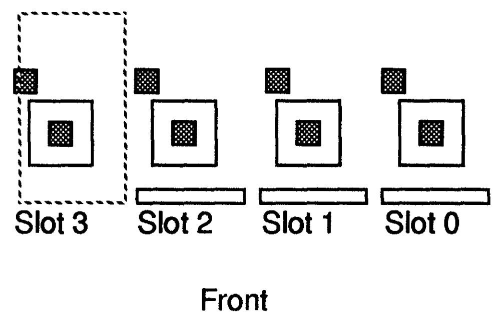
MBus Processor Modules

Up to 4 Superscalar SPARC modules may be installed in the MBus slots. Installation is simple — push the processor into the connector on the MK405 and secure into position with a screw at each corner.

As newer higher performance SPARC modules become available, Meiko will offer these as field upgrade options. Simply unplug the old unit and replace with the new.

The processor slots are numbered from 0 to 3, from right to left (when viewed from the front) as shown in Figure 4-5.

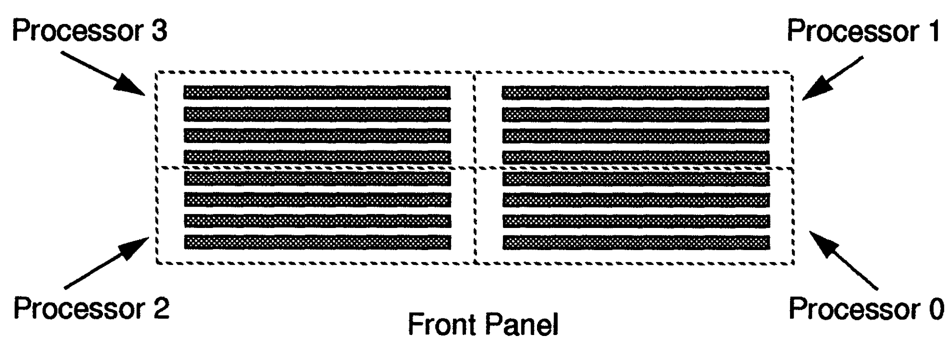
Figure 4-5 MK405 Processor Numbering



Memory

Up to 16 single in-line memory modules (SIMMs) may be fitted to the MK405 board (JEDEC 36bit SIMMs). The array is constructed of 4 groups of 4 SIMMs, one group for each processor — see Figure 4-6. Within each group the SIMMs must be identical but there is no requirement for the groups to be the same.

Figure 4-6 Memory slots on an MK405 showing the allocation of slots to each of the 4 processors.



SIMMs can be either 4Mbit DRAM or 16Mbit DRAM technology, either single or double sided. This gives a total minimum memory configuration of 64Mbytes and a maximum of 512Mbytes (16 - 128Mbytes per processor).

Boot ROM

The boot ROMs may be upgraded from time to time.They are held in sockets and are easily replaced using the appropriate tool.

H8 ROM

The H8 ROMs may be upgraded from time to time. They are held in sockets and are easily replaced using the appropriate tool.

NVRAM

The non-volatile RAMs hold system configuration information, environment variables, and the clock. They are held in a DIL sockets and are easily replaced. Note however that the information within the NVRAM can only be restored by Meiko's engineers.

The NVRAM contains lithium batteries which have special handling and disposal requirements — your attention is drawn to the precautions listed in the *General Precautions* on page ii.

External Indicators

The MK405 front panel includes two pairs of LEDs. In the middle are two amber LEDs — illumination of these indicates an error with the Elite switches. The green/amber LED pair are the board's CAN bus indicators — green is the board's heartbeat, and amber indicates a transmission onto the bus.

The green LED flashes at a slow steady rate (once per second) when operating normally. A quicker flash rate (2×normal) indicates that the SPARC processor is not responding; a very quick flash rate (3×normal) indicates that the H8 processor on the module controller is not responding.

Each processor board within a processor module controls a 4×4 matrix of red LEDs on the module's front panel. Each of the four processors on an MK405 controls one row of 4 lights. Processor 0 controls the top row, processor 3 the bottom.

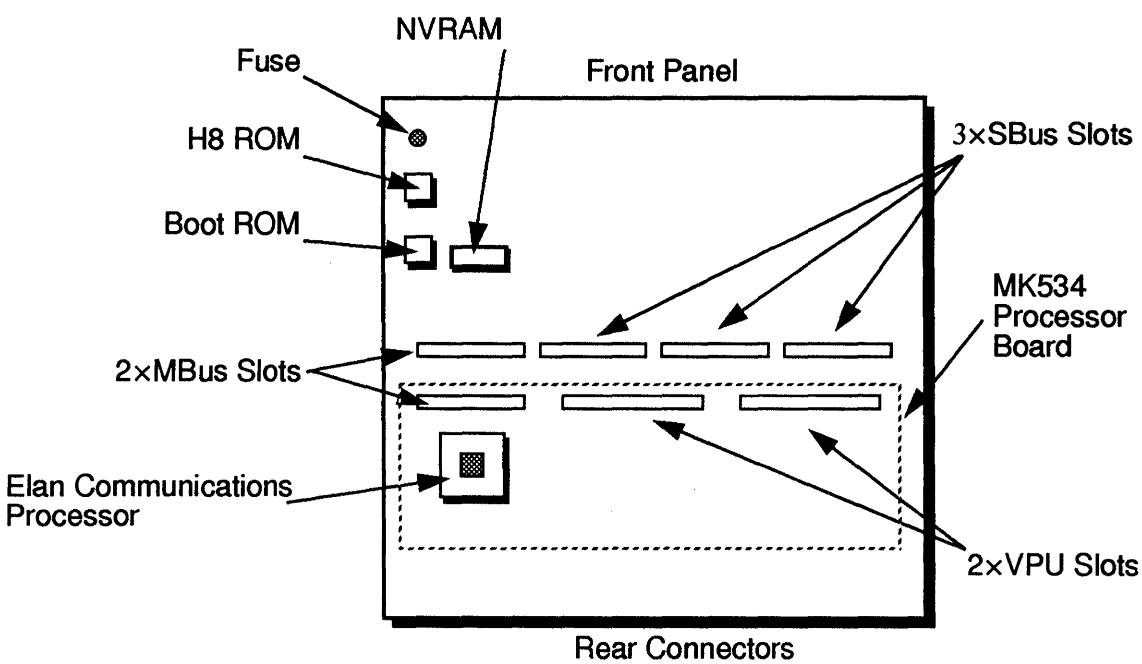
MK403 Vector Processor Board

Field Upgradeable Components

The MK403 has the following field upgradeable components (see Figure 4-7):

- Superscalar SPARC processor module fitted to MBus slot.
- Three SBus slots.
- Two VPU slots.
- Boot ROM.
- H8 ROM.
- Non-volatile RAM.
- Fuses.

Figure 4-7 MK403 Components



MBus/VPU Processor Modules

Early VPU boards are fitted with standard MBus SPARC modules and two VPU modules — these plug into the MBus/VPU slots on the MK403 mother board.

Later shipments include a single cache coherent processor board that includes both the dual VPU processors on a single plug-in board; this board plugs into the MBus and 2 VPU sockets on the MK403 motherboard. A standard Superscalar SPARC module plugs into the second MBus slot.

SBus Modules

Three SBus slots are provided and these may be fitted with standard SBus modules. As with the processor modules SBus modules are simply plugged into the SBus connectors and secured with 2 screws.

When using SBus cards that have external connections, for example a graphics card, remove the appropriate panel from the front panel of the MK403 — the panel is held in place by two small screws. When using SBus SCSI cards the external connections are made via the SCSI connectors on the MK403 front panel.

SBus devices are numbered from 0 to 2, device 0 being next to the processor slots.

Boot ROM

The boot ROM may be upgraded. It is held in a socket and is easily replaced using the appropriate tool.

H8 ROM

The H8 ROM may be upgraded from time to time. It is held in a socket and is easily replaced using the appropriate tool.

NVRAM

The non-volatile RAM holds system configuration information, environment variables, and the clock. It is held in a DIL socket and is easily replaced. Note however that the information within the NVRAM can only be restored by Meiko's engineers.

The NVRAM contains lithium batteries which have special handling and disposal requirements — your attention is drawn to the precautions listed in the *General Precautions* on page ii.

Fuses

One fuse is used to protect the external keyboard/mouse. This fuse is 250mA quick blow, Meiko part number 22-fu400-02e250.

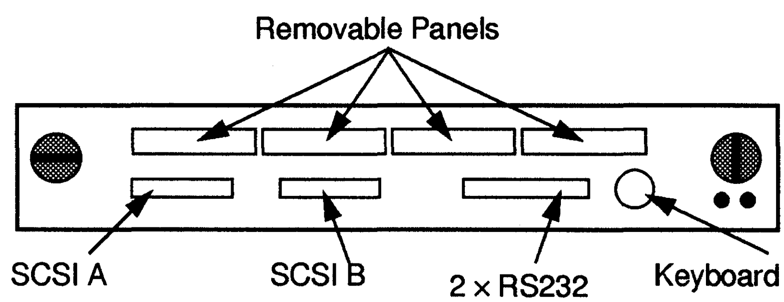
External Connections

External connections are provided for a keyboard/mouse (8 pin circular socket), RS232 interfaces (2 channels provided by one 25-way D-type connector), and two independent SCSI buses (each via a 50-way connector). Note that the MK403 does not include SCSI controllers — the SCSI connectors must be connected to SCSI SBus cards.

Front Panel Connections

Removable panels provide access to connectors on the optional SBus boards.

Figure 4-8 MK403 Front Panel Connections



RS232 Connections

The two RS232 channels are connected as follows. Signal ground is pin 7, chassis ground in pin 1.

Table 4-3 RS232 Channel A Pinout.

Signal	Input/Output	Pin number
TXD	Out	2
RXD	In	3
RTS	Out	4
CTS	In	5
CSR	In	6
DCD	In	8
DB	In	15
DD	In	17
DA	On	24
DTR	On	20

Table 4-4 RS232 Channel B Pinout.

Signal	Input/Output	Pin number
TXD	Out	14
RXD	In	16
RTS	Out	19
CTS	In	13
DCDB	In	12

External Indicators

Two LEDs (one green, one amber) are included on the front panel. The green LED is the heart beat from the board's CAN controller. The amber light illuminates each time the CAN controller transmits on the CAN bus. Both should flash steadily. These indicators are also displayed on the module's LED display.

The green LED flashes at a slow steady rate (once per second) when operating normally. A quicker flash rate (2×normal) indicates that the SPARC processor is not responding; a very quick flash rate (3×normal) indicates that the H8 processor on the module controller is not responding.

Each processor board within a processor module controls a 4×4 matrix of red LEDs on the module's front panel. The MK403 displays a random pattern on these when running the Boot ROM. When Solaris has been booted a circulating pattern is displayed. The pattern can be changed by user programs.

Installing Backplane Boards 5

Warning - The procedures in this chapter must only be undertaken by trained engineers.

Warning - You must disconnect the power supply before installing boards into the processor module.

A number of boards can be connected to the rear of the processor module. Along the top edge of the module backplane are connectors for up to 8 switch cards. Below these are another 5 slots; 4 are used for the module SCSI cards, and one (the larger one) for the module control card.

Blanking plates must be fitted over unused backplane slots to ensure correct cooling and compliance with RFI regulations.

Installation

The installation procedure for all backplane cards is similar. Note however that SCSI cards are normally installed during the module's manufacture and are not readily removed.

Remove the blanking plate from the rear of the module by removing the screws at each end. In some case, particularly when removing boards, it is useful to remove the covers from the neighbouring card slots.

Insert the card into the card guides ensuring that the component side of the card faces the right. Firmly push the card into the position ensuring that the card mates correctly with the backplane connection.

After installing a module controller card you must reinstall the blanking plate.

MK511/2 Module Switch Cards

Up to 8 module switch cards can be connected to the module backplane to provide a first layer of switching. Even numbered backplane connectors carry plane 0 of the switch network, the odd numbered connectors carry plane 1. A fully connected module containing single SPARC boards requires just two cards in slots 0 and 1, whereas a module containing quad-SPARC boards requires 8 switch cards.

The MK511 contains a single Elite that provides the first level of network switching. The MK512 is a buffer card that has no switching and is used solely when directly connecting two modules — in this case one module contains MK511 switch cards, the second contains MK512 buffer cards.

Neither card has special installation requirements, and neither card includes field serviceable components.

MK515 Module Control Card

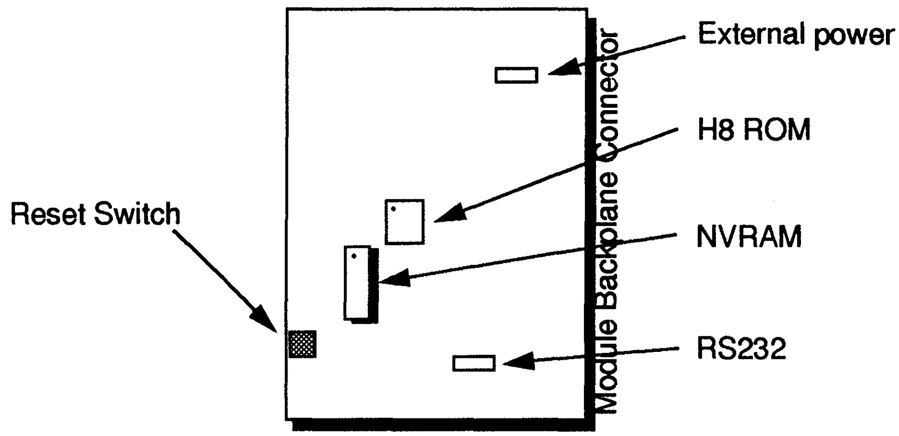
Monitors the status and configuration of the module, propagating error conditions over the CAN bus (for interception by the System Software) and initiating module shutdown under critical conditions (such as major component failure).

Field Upgradeable Components

The MK515 has the following field upgradeable components:

- H8 ROM.
- Non-volatile RAM (NVRAM).

Figure 5-1 MK515 Components



Meiko may upgrade the H8 ROM from time to time.

The non-volatile RAM (NVRAM) holds system configuration information, such as the number of module switches. The NVRAM contains Lithium batteries which have special handling and disposal requirements — your attention is drawn to the *General Precautions* on page ii.

External Indicators

Two LEDs (one green, one amber) are visible from the rear of the module. The green LED is the module controller's CAN heart beat, the amber light illuminates each time the controller writes to the CAN bus. These indicators are also displayed in the top right hand corner of the module's LED display.

Reset Switch

The reset switch is not accessible when the board is mounted in the module. It is used by Meiko engineers when bench testing the board.

External Connections

Two 3 pin connectors are provided for Meiko engineering use. They are used for bench testing of the board, and allow power and diagnostic RS232 connections to be made.

MK516 SCSI Cards

Transfers up to 5 SCSI buses from the processor module backplane to external connectors. Five ribbon cable connectors transfer the buses to disk devices mounted within the module, and one Beta Phase connector connects with externally mounted devices.

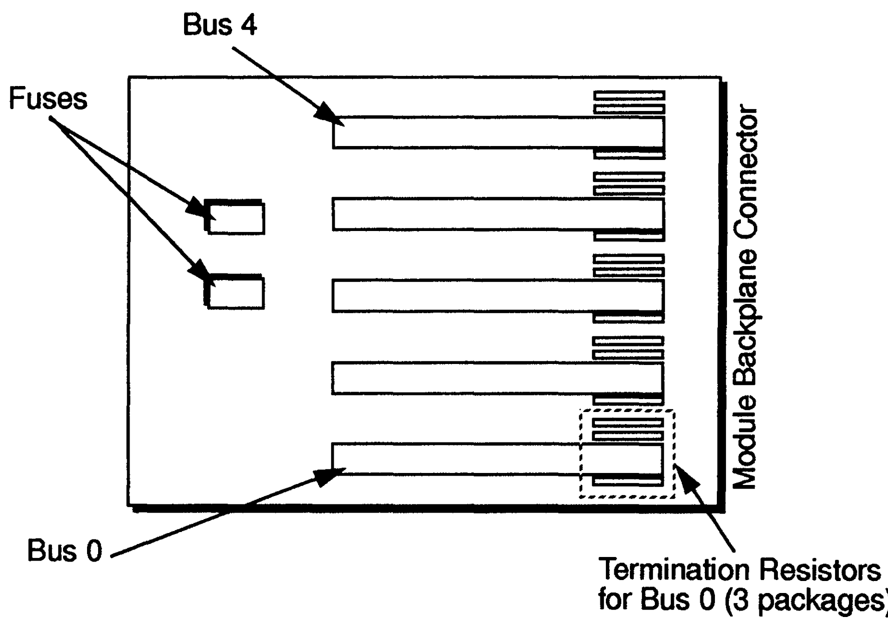
Field Upgradeable Components

The MK516 has the following field upgradeable components:

- Fuses
- Termination resistors

The fuses should only be fitted if on-board SCSI bus termination is used (rarely). Fuses are 125V, 2A micro 273 series. The termination resistors must be soldered in place — there are 3 SIL packages for each bus, each package is an 8 pin, 6 terminal 220/330R device.

Figure 5-2 MK516 Components



Installing SCSI Devices 6

Warning - The procedures in this chapter must only be undertaken by trained engineers.

The MK401 processor board is equipped with two independent SCSI-2 controllers and includes 3 SBus slots that may be used to add additional SCSI controllers. The MK403 Vector Processing board has no on-board SCSI controllers, but does have 3 SBus slots that can be used to add SCSI devices.

The CS-2 Processor Module includes space for up to 4 SCSI disk devices — these are connected to the SCSI controllers on the processor boards.

Additional SCSI disk devices or SCSI devices requiring frequent access (CD, tape, etc.) may be mounted externally to the processor module.

Module Disk Devices

The SCSI bus connection between the processor boards and the disk devices is via the Module SCSI Cards installed at the rear of the module. The Module SCSI cards take up to five[1] independent SCSI buses from each processor board and transfer these to ribbon cable connections and a Beta Phase connector.

1. Only 2 are currently used; the remainder are for future expansion.

Disk devices mounted within the processor module are wired directly to the ribbon cable connectors. This wiring is added during manufacture of the module; two connection options are available:

- One disk per processor board — each of the four disks is connected to SCSI bus 0 on the processor board.
- A pair of disks connected to processor boards 0 and 1, each using a separate SCSI bus.

Installing the Disk Devices

The four disk devices within the processor module are each mounted in a removable disk carrier. The disk carrier is connected to the module backplane by a 50-way connector that carries power, indicator signals, and the SCSI bus.

Warning - You must disconnect the module from the power supply before installing a disk device.

Fixing the Disk Device into the Carrier

To install a 3.5" SCSI disk into a disk carrier:

- Remove the base plate from the carrier by removing the 6 screws.
- Connect the disk to the power supply and SCSI bus connectors.
- At the front of the disk connect the LED activity output to the disk carrier.
- Mount the disk drive into the carrier using four screws that are inserted from two sides of the carrier. Note that the screw must be long enough to accommodate the thickness of the carrier wall.

Fixing the Carrier into the Processor Module

Insert the carrier into the processor module so that the indicator lights on the front of the carrier are at the top. Push into position and secure with the two captive screws.

External Indicators

Three LEDs show from the front of the disk carrier. The green LED is the power light, the red LED is the activity light, and the yellow is the fail indicator. The activity signal is also displayed on the module's LED panel using the small circular red LED.

External SCSI Devices

External SCSI devices may be connected in two ways:

Front Panel Connections:
Some processor board types have SCSI bus connections on their front panels. When using these the cable should be routed behind the module's front panel and into the cable guide below the module.

Module SCSI Card Connections:
The module SCSI cards (MK516) fitted to the rear of the module has an external Beta Phase connection. This may be used to connect the processor boards to a Peripheral Module.

Computing Surface

The CS-2 Switch Module

S1002-10M132.03

meiko

Meiko's address in the US is:

Meiko
130 Baker Avenue
Concord MA01742

508 371 0088
Fax: 508 371 7516

Meiko's address in the UK is:

Meiko Limited
650 Aztec West
Bristol
BS12 4SD

01454 616171
Fax: 01454 618188

Issue Status:		
	Draft	
	Preliminary	
	Release	X
	Obsolete	

Circulation Control: *External*

Contents

Safety Precautions

Federal Communications Commission (FCC) Notice

Warning: Changes or modifications to this unit not expressly approved by the party responsible for compliance could void the user's authority to operate the equipment.

NOTE: This equipment has been tested and found to comply with the limits for a Class A digital device, pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. The equipment generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the instruction manual, may cause harmful interference to radio communications. Operation of this equipment in a residential area is likely to cause harmful interference in which case the user will be required to correct the interference at his own expense.

Shielded cables must be used with this unit to ensure compliance with the FCC Class A limits.

General Precautions

- Adhere strictly to the instructions and warnings provided in this documentation.
- Do not push objects into the interior of the equipment through any openings. Hazardous voltages and moving parts are present. Conductive objects could cause electric shock, fire, or damage to the equipment.
- It is not permitted to make mechanical or electrical modifications to the equipment. The manufacturer is not responsible for regulatory compliance of equipment that has been modified. You may also invalidate your warranties if you make unauthorised changes to your equipment.
- A lithium battery is installed on some of the boards in the module — it is an integral component of the non-volatile RAM (or NVRAM). Lithium batteries are not customer replaceable. If they are mishandled there is a danger that they may explode. Always consult Meiko if you suspect that the battery needs replacing. Never dispose of lithium batteries in a fire or attempt to dismantle.
- Many of the electronic components fitted in the module are fragile or static sensitive. They should only be handled by trained engineers. You must observe anti-static precautions when handling boards or electronic devices.
- Operation of this equipment in a residential area may cause unacceptable interference to radio and TV reception, requiring the operator to take whatever steps are necessary to correct the interference.

Lifting

The module is heavy and must be lifted by two people or by using the Genie lifting equipment that is supplied by Meiko (large multi-module systems only).

When using the Genie fork lift you must first attach the two lift-clamp assemblies to either side of the module. Unscrew the securing screw that holds the telescopic section of the lift-clamp closed. Insert the rear locating pins into the holes at the rear edge of the module (the open end of the box section should face the front of the module). Align the front locating pins with the holes at the front of the module and tighten the holding screw.

Locate the forks of the Genie lift into the two box sections in the lift clamps. Raise the module by turning clockwise the handle on the Genie lift. Lower the module by turning the handle anti-clockwise.

Warning - Your attention is drawn to the operating instructions for the Genie lift that are supplied by the lift's manufacturers.

- Genie fork lift, part number 73-FORK-GENIE.
- Module lifting gear, part number 65-MODULELIFT

Overview 1

The Switch Module is a physical enclosure for up to 4 switch boards — the components used to build the CS-2 data network. Supporting infrastructure within the module consists of cooling fans, power supply, and connections to the processor modules (often via other switch modules) and control the network. The module is fully enclosed and similar in appearance to the Processor Module, and is FCC compliant.

The Switch Module is intended for installation in a CS-2 bay which provides a firm level support, network interconnect, and power distribution panels.

Module Packaging 2

CS-2 modules are shipped in units of 1 to 4. The packaging consists of a wood base, packing foam, antistatic bag, and enclosing triple wall card outer.

Packing Dimensions

All dimensions are approximate.

Packing dimensions (single module):

Height	74 cm (29")
Width	34 cm (13.5")
Length	112 cm (44")

Packing weight (without module):

4 module skid	20Kg (44lbs)
Module base	4.5Kg (10lbs)
Filler	1.8Kg (4lbs)
Outer card cover	4.5Kg (10lbs)
Packing foam	0.2Kg (0.4lbs)
Total (1 module)	**11 Kg (24.4lbs)**
Total (4 modules)	**31 Kg (68.4lbs)**

Module weight (without switch boards):

Module weight	41.8Kg (92lbs)

Switch board weights:

MK522 (8 Elites)	2.0 Kg
MK523 (2 Elites)	1.45Kg (3.19lbs)
MK529 (4 Elites)	2.0 Kg

Unpacking

Warning - You are reminded of the safety precautions listed in *General Precautions* **on page ii.**

Modules are shipped in groups of 4 or 1.

For 4 module shipments each module is packaged individually and secured with the others onto a large skid. To unpack a four module skid cut the outer banding and remove one module. Space the remaining three module packs uniformly over the skid.

- To unpack each module first cut the banding and remove the outer card carton by lifting it clear of the module.
- Remove the protective foam from the top and front of the module.
- Remove the antistatic bag by pulling upwards.
- Lift the module from the packing base. Note: the module is heavy and must be lifted by at least two people or by using a Genie lift.
- Check the module for damage and advise the transportation company immediately if any is found.

Retain all packaging and use it when shipping the module.

Installing the Switch Module 3

You must read this chapter thoroughly and completely before using your CS-2 module.

Location

The switch module must be located in a CS-2 Bay. This provides a firm level support for the module and has been designed to maximise the flow of cooling air through the module.

Warning - Overheating can damage this product. Do not block or cover any openings that are built into the module, and do not place the module near any sources of heat.

Modules sit on module trays that are fixed to the bay by telescopic rails. The tray must be extended from the bay before a module may be loaded onto it or removed from it.

Warning - Never extend more than one loaded module tray.

To extend the tray first unscrew the two retaining screws and then pull the tray forward. Lift the module onto the tray ensuring that the feet on the base of the module mate with the holes in the tray. Fix the module into position using the captive screws on the underside of the tray.

Warning - Your attention is drawn to the lifting instructions in *Safety Precautions* on page i.

The module backplane carries a number of Beta Phase connections to the bay. These connectors must be opened using the Meiko Beta Phase Controller before the module tray is pushed into position.

Warning - Failure to open the Beta Flex connectors before pushing the module trays into position will damage your equipment.

With the Beta Phase connectors held open by the Beta Phase Controller push the module tray into position. On each of the telescopic arms is a spring slip that must be pushed-in to allow the arm to slide. Push the tray into position carefully, ensuring that the Beta Phase connectors are correctly situated. Lock the module tray into position using the two captive screws at the front of the tray.

After pushing the module tray into position use the Beta Phase Controller to close the connectors.

Using the Meiko Beta Phase Controller

Connect the Beta Phase controller to the 37-way connector located at the rear of the module's power supply unit. Connect the Beta Phase Controller to a mains outlet using the supplied cable. Use the Activate button on the box to open all the Beta Phase connectors used by that module — pressing the same button a second time will close the connectors. Status lights on the box show the condition of each connector; red indicates open, amber indicates power is being applied, green indicates no power. When using the Controller it is safe to hold the Beta Phase connectors open for prolonged periods.

Power Supply

Check that the voltage and frequency of your power supply is suitable for your equipment. The power supply unit is auto-ranging, 110–230 V, 50/60/400 Hz. Maximum current is 15 A.

Power Connections

Power is supplied to the switch module by a 3-way cable that is supplied by Meiko.

The following cable types are available:

- Harvey Hubble to Harvey Hubble.
- Harvey Hubble to UK 13 A outlet.
- Harvey Hubble to US 15 A outlet.

Where several switch modules are used Meiko may also supply a mains distribution board. This connects eight modules to one power supply outlet.

- Power distribution panel, European. Part number 65-DIST-EUR-01
- Power distribution panel, USA. Part number 65-DIST-US-01

Warning - You must only use power cords and distributions panels that are supplied by Meiko. These have the correct power rating for your system.

Warning - You should check that you have been supplied with the correct type of power cord and distribution boards.

Fuses

Three fuses are located at the rear of the switch module. One fuse protects the power supply unit from the incoming mains supply. Two fuses protect the fans.

- Main fuse: standard 250 V, 15 A. Part number 22-F0100-04E150.
- Fan fuse: anti-surge 250 V, 10 A. Part number 22-F0100-04E100.
- Fan fuse: anti-surge 250 V, 5 A. Part number 22-F0100-04E500.

There are 3 fuses located on the module backplane along the bottom edge. These fuses protect the ±12 V and the +5 V circuits on both the backplane and all boards that draw their power supply from it (including the LED panel and fan trays). Access to these fuses is gained by removing the module side panels. Fuse types are:

- Fuse F1 is 5 A, Meiko part number 22-FU100-03E500.
- Fuse F2 is 1 A, Meiko part number 22-FU100-03E100.
- Fuse F3 is 5 A, Meiko part number 22-FU100-03E500.

Warning - The backplane fuses may only be changed by trained engineers.

Operation

The power supply is operated by the main on/off switch located above the power connector. 0 represents off, 1 represents on.

When first switched on five LEDs are visible through the rear of the power supply unit. Initially three of these should be green, two are red. After a short delay all LEDs should be green.

Warning - A fault is present if any light remains red for more than 10 seconds after first switching on, or changes from green to red during operation. Switch the power supply off, disconnect from the main supply, and contact Meiko for advice.

Removable Panels

Warning - Ensure the power supply is switched off and disconnected from the main supply before removing the module's front or side panels.

The front and side panels may be removed from the module. Removing the front panel gives access to the LED board, and the switch boards. The side panels may be removed but there is no requirement to do so, except to protect them during transit or when lifting the module into a bay.

Warning - Switch boards should only be fitted or removed by trained engineers.

To remove the front panel pull it forward. The front panel is retained by four clips, one in each corner of the panel. When fitting the front panel ensure that the tinted window is aligned with the module's LED display.

The side panels are held in place by 5 key slots on the side of the module. To remove the a side panel slide it forward and pull it off.

External Connections

The switch module includes a number of connectors to the front and rear of the module.

At the Rear of the Power Supply Unit

1×25-way D-type connector:
RS232 connector for module control card diagnostics. This connection is for use by Meiko's trained engineers only.

2×9-way D-type connectors:
X-CAN and G-CAN connectors. The X-CAN is the middle connector — this is used to connect the modules within a cluster (up to 3 bays/24 module systems). The G-CAN is the right-most connector — this is used to interconnect clusters. When using more than one cluster at least two modules in each cluster must be connected to the G-CAN; all unused G-CAN connectors must have terminators fitted. Within a cluster the X-CAN connections are daisy-chained using Meiko supplied cables.

- G-CAN terminators, part number 60-CS2MA009-1T.
- X-CAN interconnect, part number 60-CA0246-1T

1×37-way D-type connector:
Connection for the Meiko Beta Phase Controller; attach the controller box here to operate the Beta Phase connectors on the module backplane.

At the Rear of the Module

Beta Phase Connectors:

Up to 16 Beta Phase Connectors, each carrying four links of the switch network.

At the Front of the Module

Behind the front panel are four polarised 50-way connectors. Two of these are used to connect to the MK525 module LED board. Fit the board to the upper connectors when the module is mounted on the lower shelf of a bay, and the lower connectors when the module is mounted on the upper shelf.

The connectors for the 4 switch boards are also located behind the front panel.

Warning - Switch cards are not user serviceable. They should only be fitted or removed by trained engineers.

External Indicators

A number of LEDs are visible through the tinted window of the module's front panel.

Four 4×4 matrices of red LEDs are currently unused by Switch boards.

Below each matrix of red LEDs are 3 small circular LEDs. The green and amber lights are driven by the switch board's CAN interface — the green light is the board's heart-beat, the amber light indicates a transmit of data. The red light is not used.

Two LEDs in the top right hand corner of the display are driven by the module controller. The green light is a heart beat signal, the amber light flashes each time the module controller is sending information over the CAN bus.

Module Identification Number

At the rear of the power supply unit are three dials — each represents one nibble from a 12 bit network address, or *netid*. The netid for each module must be unique and modules must be numbered sequentially starting from 0.

Enter each nibble of the address onto the dials using a suitable screwdriver.

Warning - You must ensure that the power is disconnected before setting the module address.

Additional Note for Revision A MK526:

Having identified a module's network address each nibble of the 12 bit address is bit-flipped (e.g. 0111 becomes 1110). The resulting 3 nibbles are then entered onto the dials using a suitable screw driver.

The following table summarises the module id's for systems of up to 32 modules. The dials for module 5, for example, must be set to 00A.

Module	Dials			
0-3	0 0 0	0 0 8	0 0 4	0 0 C
4-7	0 0 2	0 0 A	0 0 6	0 0 E
8-11	0 0 1	0 0 9	0 0 5	0 0 D
12-15	0 0 3	0 0 B	0 0 7	0 0 F
16-19	0 8 0	0 8 8	0 8 4	0 8 C
20-23	0 8 2	0 8 A	0 8 6	0 8 E
24-27	0 8 1	0 8 9	0 8 5	0 8 D
28-31	0 8 3	0 8 B	0 8 7	0 8 F

MK515 Module Control Card

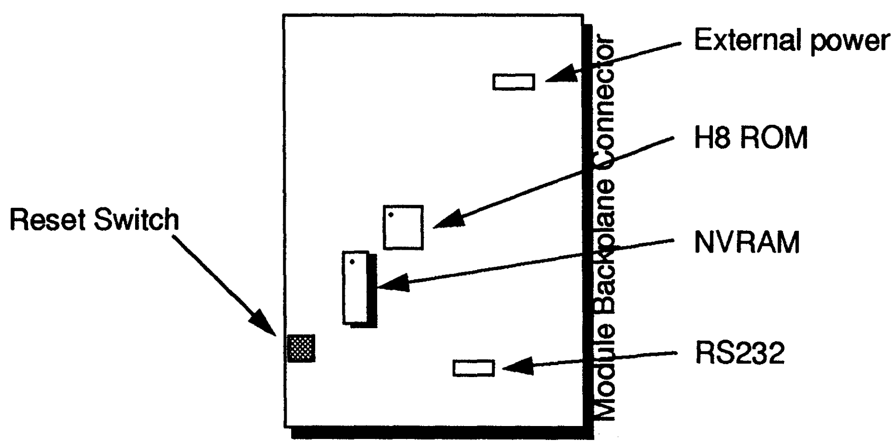
The module control card is fitted into the switch module alongside the switch cards. It monitors the status and configuration of the module, propagating error conditions over the CAN bus (for interception by the System Software) and initiating module shutdown under critical conditions (such as major component failure).

Field Upgradeable Components

The MK515 has the following field upgradeable components:

- H8 ROM.
- Non-volatile RAM (NVRAM).

Figure 3-1 MK515 Components



Meiko may upgrade the H8 ROM from time to time.

The non-volatile RAM (NVRAM) holds system configuration information, such as the number of module switches. The NVRAM contains Lithium batteries which have special handling and disposal requirements — your attention is drawn to the *General Precautions* on page ii.

External Indicators

Two LEDs (one green, one amber) are visible from the rear of the module. The green LED is the module controller's CAN heart beat, the amber light illuminates each time the controller writes to the CAN bus. These indicators are also displayed in the top right hand corner of the module's LED display.

Reset Switch

The reset switch is not accessible when the board is mounted in the module. It is used by Meiko engineers when bench testing the board.

External Connections

Two 3 pin connectors are provided for Meiko engineering use. They are used for bench testing of the board, and allow power and diagnostic RS232 connections to be made.

Operating Conditions

Recommended operating temperature	10–25°C (50–80°F)
Peak operating temperature	32°C (90°F)
Temperature gradient	10°C (18°F) per hour
Storage temperature	-18–60°C (0–140°F)
ANSI media data integrity (max.)	32°C (89.6°F)
Relative humidity	20–80% non-condensing
Storage humidity	10–90% non-condensing
Altitude	3000m (10000 feet)
Operating Shock	1G

Figure 3-2 The CS-2 Switch Module

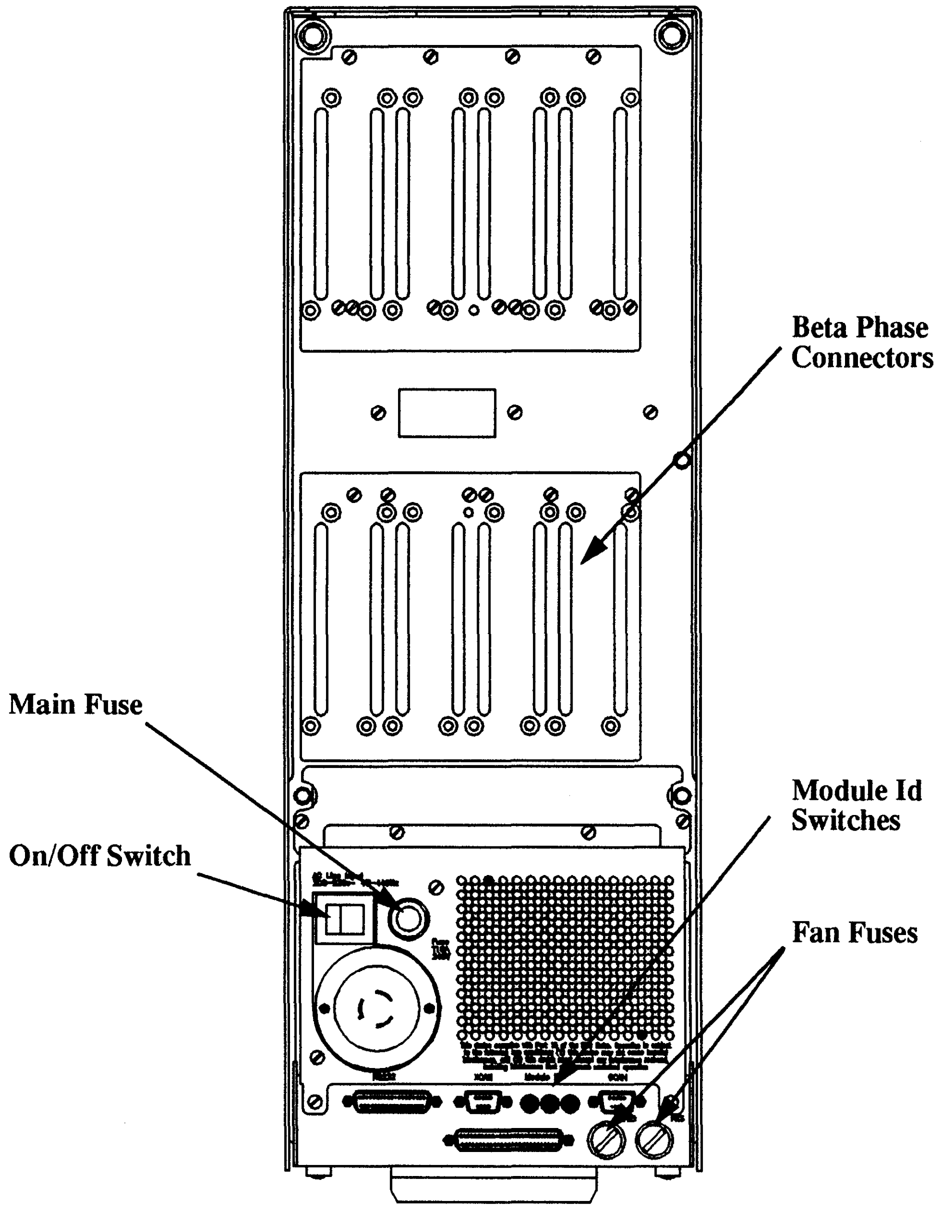
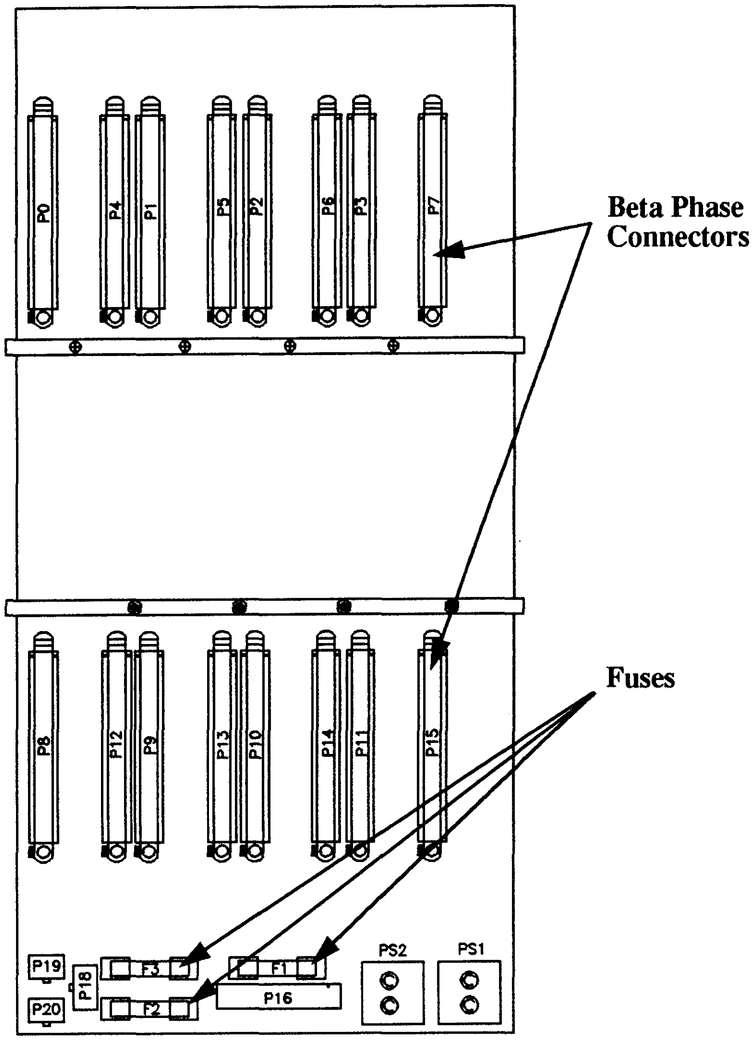


Figure 3-3 The MK520 Switch Module Backplane (not user serviceable)



Installing Switch Boards 4

Warning - The procedures in this chapter must only be undertaken by trained engineers.

Warning - You must disconnect the power supply before installing switch boards into the processor module.

Three types of switch board are available:

- MK522 — 8 interconnected Elite switches.
- MK523 — 2 unconnected Elite switches.
- MK529 — 4 unconnected Elite switches.

Blanking plates must be installed in unused board slots to ensure correct cooling and compliance with RFI regulations.

Installation

The same installation procedure is used to install all three board types.

Insert the board so that it fits into the guide rails at the top and bottom of the module, ensuring that the component side is to the left (viewed facing the module). Gently push the board squarely on the front panel. Before pushing the board fully into position fold back the levers at each end of the front panel so that they are at 90^{o} to the board. Push the board in further until the base of the two levers is touching the card cage; push the board into position by pushing firmly on both levers until they are flat. Secure using the captive screw at each end of the board's front panel.

To remove the board pull both levers out — this will lever the board away from the module's connections. Pull the board clear of the module.

You should take care not to damage the connectors within the module.

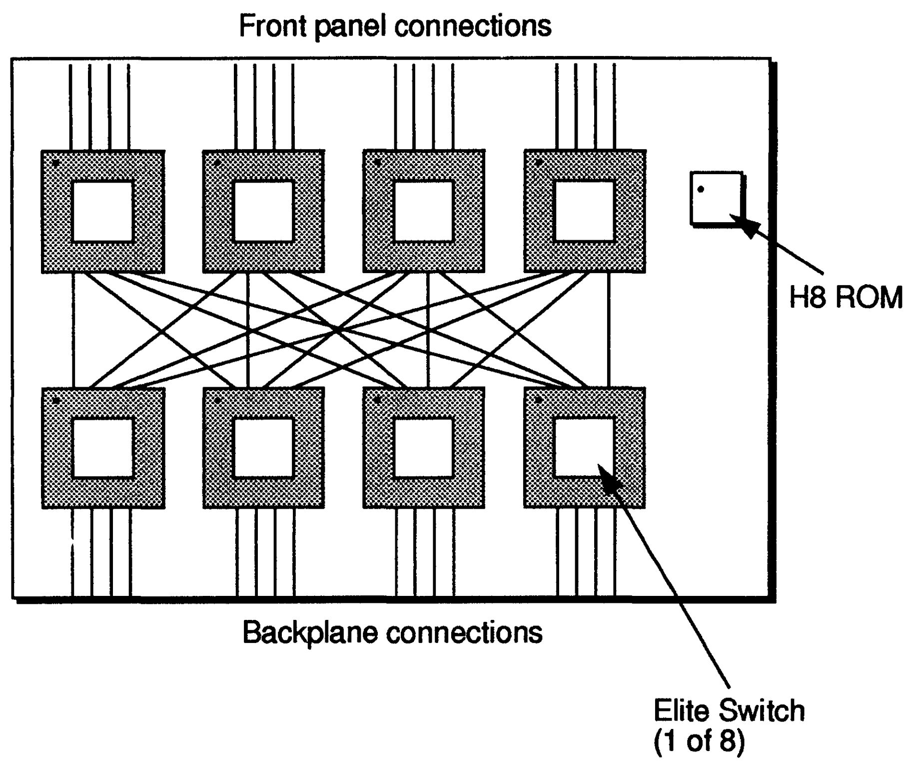
You should also take care not to damage the RFI (copper) seals along the edge of the board's front panel.

MK522 — 8 Elite Switch Card

The MK522 has the following field serviceable components:

- 8 Meiko Elite network switches.
- H8 ROM.

Figure 4-1 MK522 Components



Both the Elite switches and the H8 ROM are held in sockets and can be removed using the appropriate tool. Take care not to damage the component or its neighbours when installing or removing these devices.

External Connections

Sockets are provided on the board's front panel for external link connections. Sixteen 50-way connectors are used, one connector for each link. The connectors are numbered P10 to P25, as shown in Figure 4-2 and Figure 4-3.

Figure 4-2 External Link Numbering on the MK522

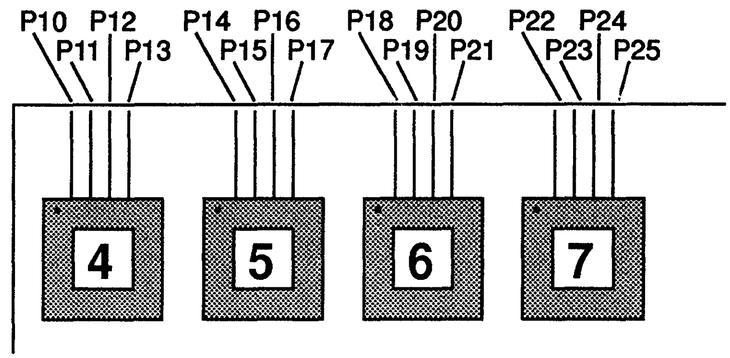
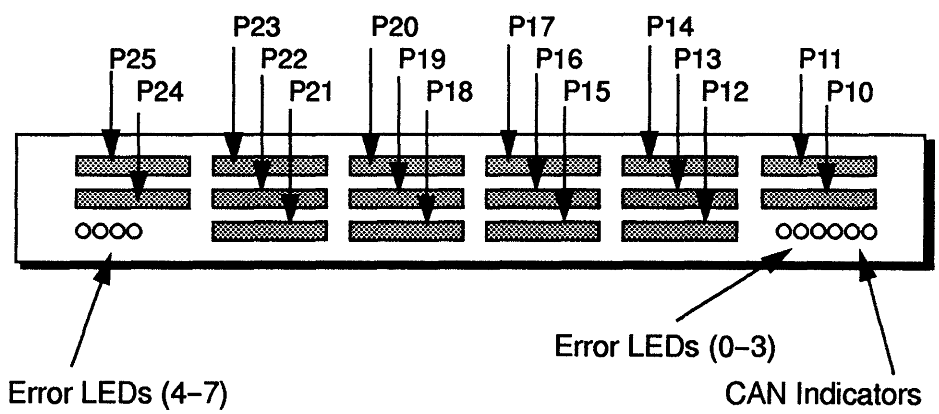


Figure 4-3 The MK522 Front Panel



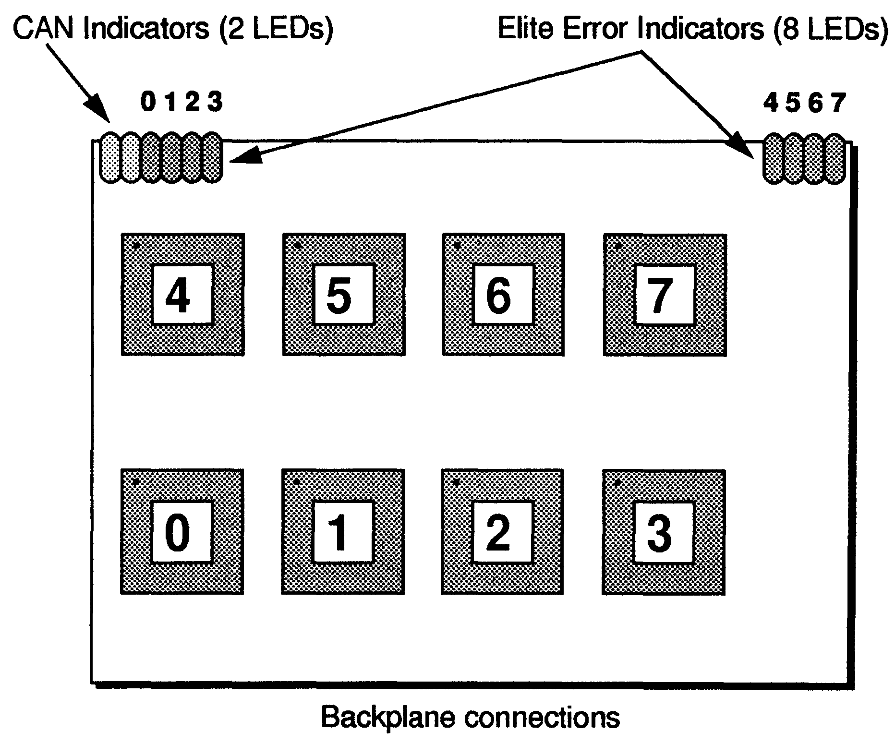
External Indicators

10 LEDs are included on the front panel. At the one end of the front panel there are 6 LEDs — 4 red, 1 green, and 1 amber. At the other end are a second group of 4 red LEDs.

The green and amber LEDs are the CAN bus indicators. The green LED is the heart beat from the board's CAN controller; this flashes at a slow steady rate (once per second) when the board is operating normally. The amber light illuminates each time the CAN controller transmits on the CAN bus.

The red LEDs are error lights, one for each of the 8 Elite switches. Figure 4-4 shows the numbering of the switches and their error lights:

Figure 4-4 Switch and Error Light Numbering

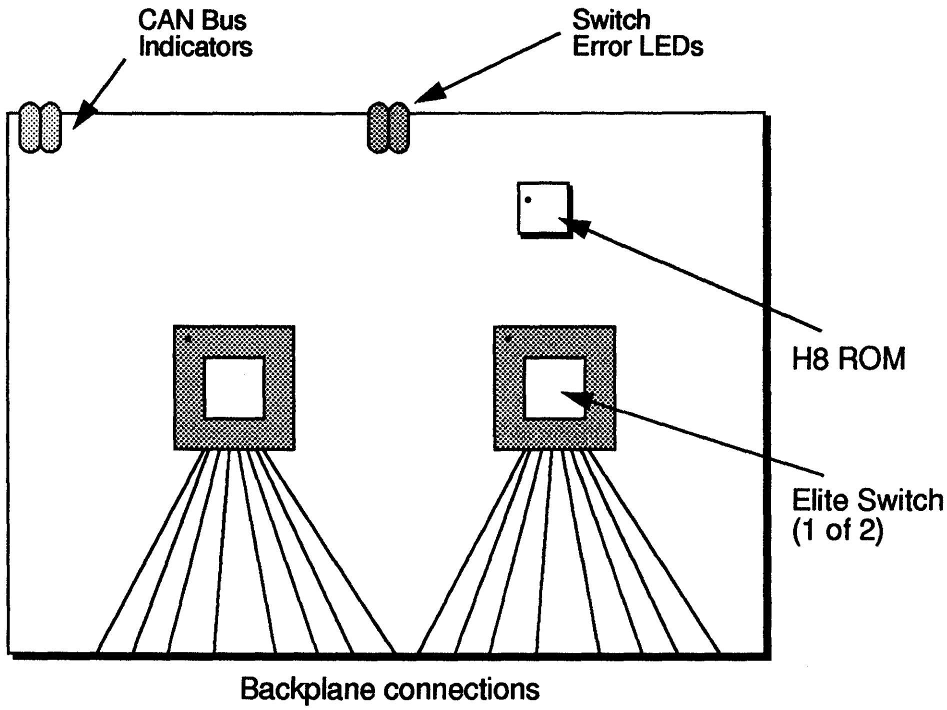


MK523 — 2 Elite Switch Card

The MK523 has the following field serviceable components:

- 2 Meiko Elite network switches.
- H8 ROM.

Figure 4-5 MK523 Components



Both the Elite switches and the H8 ROM are held in sockets and can be removed using the appropriate tool. Take care not to damage the component or its neighbours when installing or removing these devices.

External Indicators

4 LEDs are included on the front panel. At one end of the front panel there are 2 LEDs — 1 green, and 1 amber. In the middle are 2 red LEDs.

The green and amber LEDs are the CAN bus indicators. The green LED is the heart beat from the board's CAN controller; this flashes at a slow steady rate (once per second) when the board is operating normally. The amber light illuminates each time the CAN controller transmits on the CAN bus.

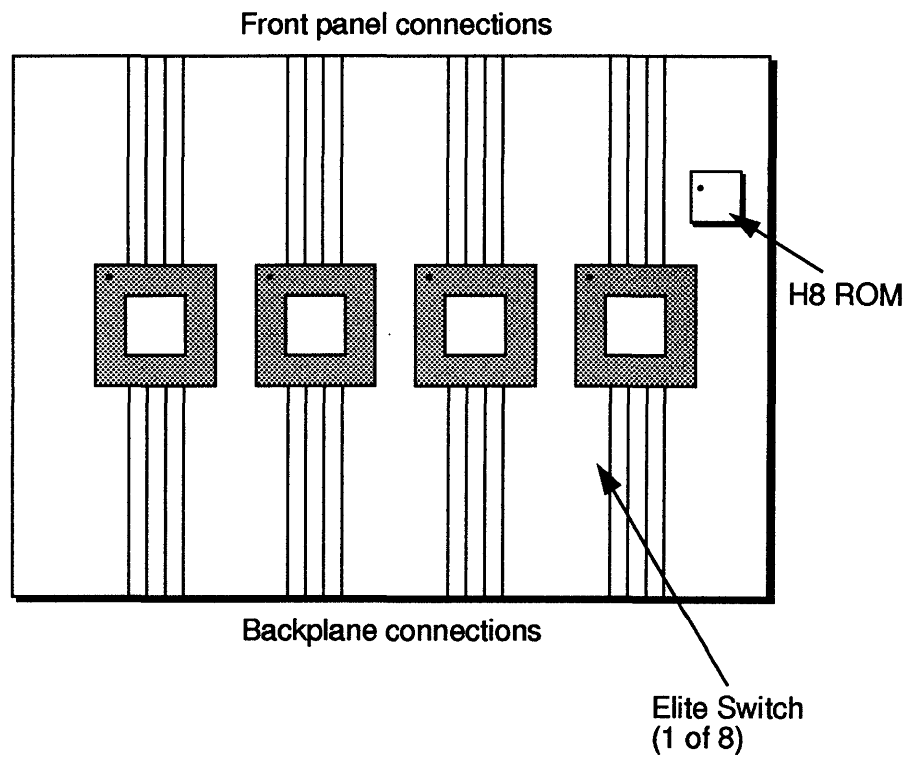
The red LEDs are error lights, one for each of the Elite switches.

MK529 — 4 Elite Switch Card

The MK529 has the following field serviceable components:

- 4 Meiko Elite network switches.
- H8 ROM.

Figure 4-6 MK522 Components



Both the Elite switches and the H8 ROM are held in sockets and can be removed using the appropriate tool. Take care not to damage the component or its neighbours when installing or removing these devices.

External Connections

Sockets are provided on the board's front panel for external link connections. Sixteen 50-way connectors are used, one connector for each link. The connectors are numbered P10 to P23.

Figure 4-7 External Link Numbering on the MK522

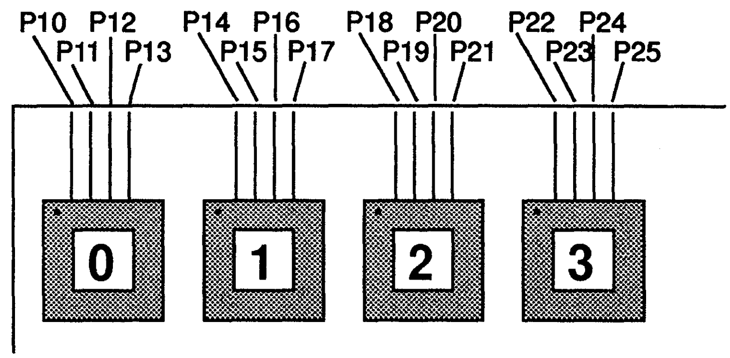
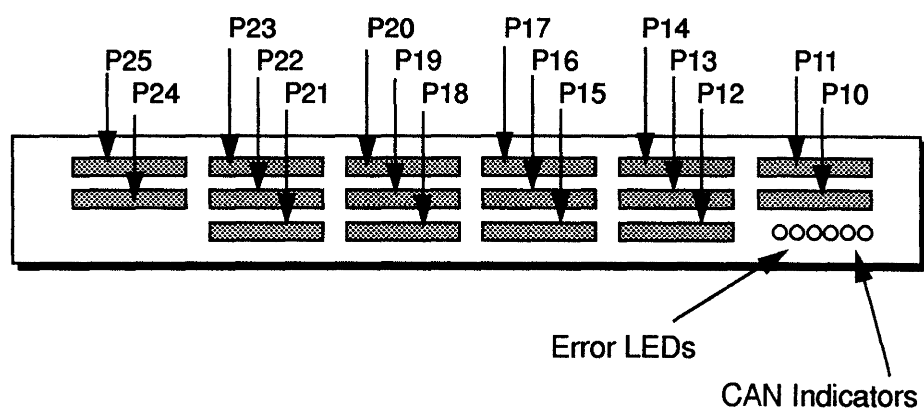


Figure 4-8 The MK522 Front Panel



External Indicators

At the one end of the front panel there are 6 LEDs — 4 red, 1 green, and 1 amber.

The green and amber LEDs are the CAN bus indicators. The green LED is the heart beat from the board's CAN controller; this flashes at a slow steady rate (once per second) when the board is operating normally. The amber light illuminates each time the CAN controller transmits on the CAN bus.

The red LEDs are error lights, one for each of the 4 Elite switches.

Computing
Surface

The CS-2 Peripheral Module

S1002-10M142.00

meiko

The specifications listed in this document are subject to change without notice.

Meiko's address in the US is:

Meiko
130 Baker Avenue
Concord MA01742

508 371 0088
Fax: 508 371 7516

Meiko's address in the UK is:

Meiko Limited
650 Aztec West
Bristol
BS12 4SD

01454 616171
Fax: 01454 618188

Issue Status:		
	Draft	
	Preliminary	
	Release	X
	Obsolete	

Circulation Control: *External*

Contents

Safety Precautions i

Federal Communications Commission (FCC) Notice

Warning: Changes or modifications to this unit not expressly approved by the party responsible for compliance could void the user's authority to operate the equipment.

NOTE: This equipment has been tested and found to comply with the limits for a Class A digital device, pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. The equipment generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the instruction manual, may cause harmful interference to radio communications. Operation of this equipment in a residential area is likely to cause harmful interference in which case the user will be required to correct the interference at his own expense.

Shielded cables must be used with this unit to ensure compliance with the FCC Class A limits.

General Precautions

- Adhere strictly to the instructions and warnings provided in this documentation.
- Do not push objects into the interior of the equipment through any openings. Hazardous voltages and moving parts are present. Conductive objects could cause electric shock, fire, or damage to the equipment.
- It is not permitted to make mechanical or electrical modifications to the equipment. The manufacturer is not responsible for regulatory compliance of equipment that has been modified. You may also invalidate your warranties if you make unauthorised changes to your equipment.
- When connecting a visual display unit to your module your attention is drawn to the instructions that are issued by the device manufacturer.
- Your attention is drawn to the instructions that are issued by the manufacturers of your disk devices.
- A lithium battery is installed on some of the boards in the module. Lithium batteries are not customer replaceable. If they are mishandled there is a danger that they may explode. Always consult Meiko if you suspect that the battery needs replacing. Never dispose of lithium batteries in a fire or attempt to dismantle.
- Many of the electronic components fitted in the module are fragile or static sensitive. They should only be handled by trained engineers. You must observe anti-static precautions when handling boards or electronic devices.
- Operation of this equipment in a residential area may cause unacceptable interference to radio and TV reception, requiring the operator to take whatever steps are necessary to correct the interference.

Lifting

The module is heavy and must be lifted by two people or by using the Genie lifting equipment that is supplied by Meiko (large multi-module systems only).

When using the Genie fork lift you must first attach the two lift-clamp assemblies to either side of the module. Undo the securing screw that holds the telescopic section of the lift-clamp closed. Insert the rear locating pins into the holes at the rear edge of the module (the open end of the box section should face the front of the module). Align the front locating pins with the holes at the front of the module and tighten the holding screw.

Locate the forks of the Genie lift into the two box sections in the lift clamps. Raise the module by turning clockwise the handle on the Genie lift. Lower the module by turning the handle anti-clockwise.

Warning - Your attention is drawn to the operating instructions for the Genie lift that are supplied by the lift's manufacturers.

- Genie fork lift, part number 73-FORK-GENIE.
- Module lifting gear, part number 65-MODULELIFT

Overview 1

The Peripheral Module is a physical enclosure for up to 16 disk devices. Supporting infrastructure within the module consists of cooling fans, 3 redunded power supplies, connection to the CS-2 control network, and SCSI bus connections to each disk. The disks may be configured for direct connection, or chained connection, to a processor's SCSI bus, or may be configured as a RAID array with an optional RAID controller installed in the module. Where the Peripheral Module is located some distance from the hosting processor a differential SCSI link may be employed, in which case differential SCSI conversion hardware may also be included in the module. An LED panel at the front of the module provides visual confirmation of each disks availability and activity. The module is fully enclosed and FCC compliant.

The CS-2 Peripheral Module is intended for installation on a CS-2 Bay which provides a firm level support, network interconnect, and power distribution.

Module Packaging 2

CS-2 modules are shipped in units of 1 to 4. The packaging consists of a wood base, packing foam, antistatic bag, and enclosing triple wall card outer.

Packing Dimensions

All dimensions are approximate.

Packing dimensions (single module):

Height	74cm (29")
Width	34cm (13.5")
Length	112cm (44")

Packing weight (without module):

4 module skid	20Kg (44lbs)
Module base	4.5Kg (10lbs)
Filler	1.8Kg (4lbs)
Outer card cover	4.5Kg (10lbs)
Packing foam	0.2Kg (0.4lbs)
Total (1 module)	**11Kg (24.4lbs)**
Total (4 modules)	**31Kg (68.4lbs)**

Module weight (without disk devices or RAID controller):

Module weight	50Kg (110lbs)

RAID Controller:

RAID Controller	9.9Kg (4.51bs)

Disk devices:

Disk carrier	0.6Kg (1.3lbs)
Disk device	0.8Kg (1.8lbs)
Total per disk	**1.4Kg (3.1lbs)**

Unpacking

Warning - You are reminded of the safety precautions listed in *General Precautions* **on page ii.**

Modules are shipped in groups of 4 or 1. For 4 module shipments each module is packaged individually and secured with the others onto a large skid. To unpack a four module skid cut the outer banding and remove one module. Space the remaining three module packs uniformly over the skid.

- To unpack each module first cut the banding and remove the outer card carton by lifting it clear of the module.
- Remove the protective foam from the top and front of the module.
- Remove the antistatic bag by pulling upwards.
- Lift the module from the packing base. Note: the module is heavy and must be lifted by at least two people or by using a Genie lift.
- Check the module for damage and advise the transportation company immediately if any is found.

Retain all packaging and use it when shipping the module.

Installing the Peripheral Module 3

You must read this chapter thoroughly and completely before using your CS-2 module.

Location

The peripheral module must be located in a CS-2 Bay. This provides a firm level support for the module and has been designed to maximise the flow of cooling air through the module.

Warning - Overheating can damage this product. Do not block or cover any openings that are built into the module, and do not place the module near any sources of heat.

Modules sit on module trays that are fixed to the bay by telescopic rails. The tray must be extended from the bay before a module may be loaded onto it or removed from it.

Warning - Never extend more than one loaded module tray.

To extend the tray first unscrew the two retaining screws and then pull the tray forward. Lift the module onto the tray ensuring that the feet on the base of the module mate with the holes in the tray. Fix the module into position using the captive screws on the underside of the tray.

Warning - Your attention is drawn to the lifting instructions in *Safety Precautions* on page i.

Power Supply

Check that the voltage and frequency of your local power supply is suitable for your equipment. The power supply units are auto-ranging, 110–230 V, 47–63 Hz. Current is 10 A.

Power Connections

Power is supplied to the peripheral module by a 3-way cable that is supplied by Meiko.

Where several processor modules are used Meiko may also supply a mains distribution board. This connects eight modules to one power supply outlet.

- Power distribution panel, European. Part number 65-DIST-EUR-01.
- Power distribution panel, USA. Part number 65-DIST-US-01.

Warning - You must only use power cords and distributions panels that are supplied by Meiko. These have the correct power rating for your system.

Fuses

Eight fuses are located at the rear of the Peripheral Module as shown in Figure 3-3 and Figure 3-3 on page 13.

One fuse, situated alongside the mains inlet cable, protects the power supplies from the incoming current.

- Main fuse: 250 V, 13 A. 1.25"x0.25" HRC Ceramic. Part number 22-F0100-04E130.

Each power supply also includes its own fuse, located directly beneath it's on-off switch.

- Power supply fuse: anti-surge 250V, 5A. 1.25"x0.25" HRC Ceramic. Part number 22-FU100-03E500.

An array of four fuses are located below the raid controller's mounting position, near the top of the module. When viewed from the rear of the module, and from right to left, these fuses are:

- LED display (5V circuit): anti-surge 4A. 1.25"x0.25" HRC ceramic. Part number 22-FU100-03E400.
- RAID and differential SCSI (5V circuit): anti-surge 10A. 1.25"x0.25" HRC ceramic. Part number 22-FU100-04E100.
- RAID and differential SCSI (12V circuit): anti-surge 4A. 1.25"x0.25" HRC ceramic. Part number 22-FU100-03E400.
- Fans (12V circuit): anti-surge 4A. 1.25"x0.25" HRC ceramic. Part number 22-FU100-03E400.

Operation

The power supplies are collectively operated by the main on/off switch located alongside the power inlet. In addition each power supply includes its own switch. In all cases 0 represents off, 1 represents on.

During normal operation the green LED on the rear of each power supply must be illuminated. The red LED illuminates if the power supply fails.

Warning – A fault is present a power supply's red LED illuminates. Switch the power supply off and contact Meiko for advice.

Replacement

The Peripheral Module requires two functioning power supply units and the failure of any one power supply will not immediately impact on the module's functionality. The failed unit should be replaced as soon as possible.

You can safely remove a power supply while the rest of the system remains operational. To do this:

1. **Switch-off the failed unit.**
   Use the power supply's own power switch.
2. **Withdraw the failed unit.**
   Undo the captive screw which is positioned next to the unit's handle. Slide the power supply out of the module.

Installation is the reverse of removal.

Removable Panels

Warning - Ensure the power supply is switched off and disconnected from the main supply before removing the module's front or side panels.

The front and side panels may be removed from the module. Removing the front panel gives access to the LED board and the disk devices. The side panels may be removed but there is no requirement to do so, except to protect them during transit or when lifting the module into a bay.

Warning - Circuit boards and disk devices should only be fitted or removed by trained engineers.

To remove the front panel pull it forward. The front panel is retained by four clips, one in each corner of the panel. When fitting the front panel ensure that the tinted window is aligned with the module's LED display.

The side panels are held in place by 5 key slots on the side of the module. To remove the a side panel slide it forward and pull it off.

External Connections

The Peripheral Module includes a number of connectors to the front and rear of the module.

At the Rear of the Power Supply Unit

1×25-way D-type connector:
RS232 connector for module control card diagnostics. This connection is for use by Meiko's engineers only.

1×25-way D-type connector:
RS232 connector to the RAID controller, for use in accordance with the RAID controller's operating instructions.

2 ×9-way D-type connectors:
X-CAN and G-CAN connectors. The X-CAN is the middle connector — this is used to connect the modules within a cluster (up to 3 bays/24 module systems). The G-CAN is the right-most connector — this is used to interconnect clusters. When using more than one cluster at least two modules in each cluster must be connected to the G-CAN; all unused G-CAN connectors must have terminators fitted. Within a cluster the X-CAN connections are daisy-chained using Meiko supplied cables.

- G-CAN terminators, part number 60-CS2MA009-1T.
- X-CAN interconnect, part number 60-CA0246-1T

At the Rear of the Module

16 miniature 50-way connectors:
Optionally provides SCSI connections to the disk devices. The wiring of these connectors to the disks, RAID controller, and differential SCSI converters is site specific. Refer to *External Connection* on page 16 for more information.

At the Front of the Module

Behind the front panel are four polarised 50-way connectors. Two of these are used to connect to the module LED board. Fit the board to the upper connectors when the module is mounted on the lower shelf of a bay, and the lower connectors when the module is mounted on the upper shelf.

The mounting positions for the 16 disk carriers are also located behind the front panel.

Warning - Disk devices are not user serviceable. They should only be fitted or removed by trained engineers.

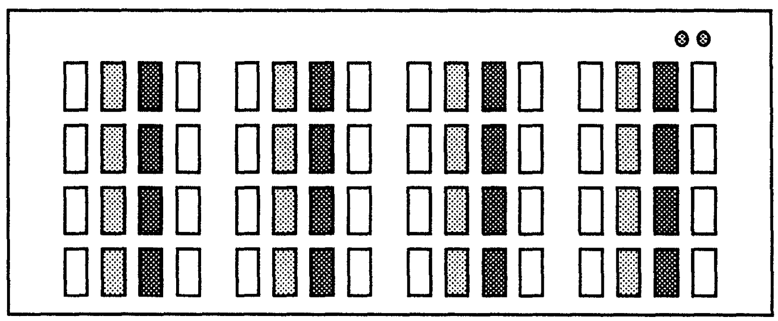
External Indicators

A number of LEDs are visible through the tinted window of the module's front panel.

Four 4x4 arrays of LEDs are used to display the power and activity status of the disks. Within each array 8 lights are used, 1 green and 1 red per disk. The ordering of the light-pairs matches the physical ordering of the disks.

Two LEDs in the top right hand corner of the display are driven by the module controller. The green light is a heart-beat signal, the amber light flashes each time the module controller is sending information over the CAN bus.

Figure 3-1 The Peripheral Module's LED Display



Green LEDs show disk power status

Red LEDs show disk activity

Module Identification Number

At the rear of the power supply unit are three dials — each represents one nibble from a 12 bit network address, or *netid*. The netid for each module must be unique and modules must be numbered sequentially starting from 0.

Enter each nibble of the address onto the dials using a suitable screwdriver.

Note: for two module systems the module id's must differ in the least significant bit only. Module id's of 0 and 1, or 2 and 3 are therefore permitted, but 1 and 2 are not.

Warning - You must ensure that the power is disconnected before setting the module address.

Additional Note for Revision A MK526:

Having identified a module's network address each nibble of the 12 bit address is bit-flipped (e.g. 0111 becomes 1110). The resulting 3 nibbles are then entered onto the dials using a suitable screw driver.

The following table summarises the module id's for systems of up to 32 modules. The dials for module 5, for example, must be set to 00A.

Module	Dials			
0-3	0 0 0	0 0 8	0 0 4	0 0 C
4-7	0 0 2	0 0 A	0 0 6	0 0 E
8-11	0 0 1	0 0 9	0 0 5	0 0 D
12-15	0 0 3	0 0 B	0 0 7	0 0 F
16-19	0 8 0	0 8 8	0 8 4	0 8 C
20-23	0 8 2	0 8 A	0 8 6	0 8 E
24-27	0 8 1	0 8 9	0 8 5	0 8 D
28-31	0 8 3	0 8 B	0 8 7	0 8 F

Operating Conditions

Recommended operating temperature	10–25°C (50–80°F)
Peak operating temperature	32°C (90°F)
Temperature gradient	10°C (18°F) per hour
Storage temperature	-18–60°C (0–140°F)
ANSI media data integrity (max.)	32°C (89.6°F)
Relative humidity	20–80% non-condensing
Storage humidity	10–90% non-condensing
Altitude	3000m (10000 feet)
Operating Shock	1G

Module Schematics

The following figures show the components and connections fitted to the rear of the Peripheral Module.

Figure 3-2 Power and Module Control Connections (Expanded View)

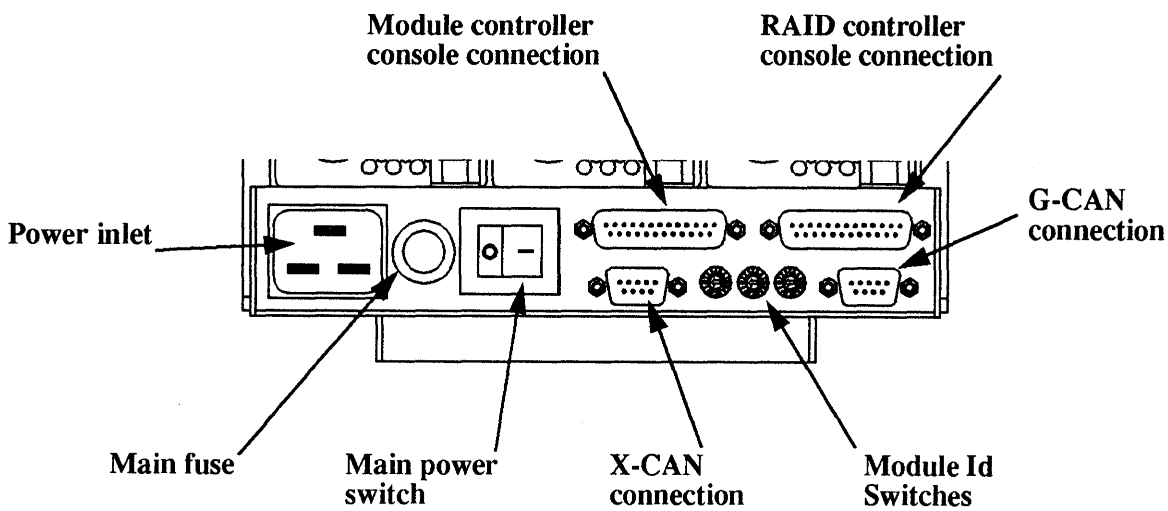
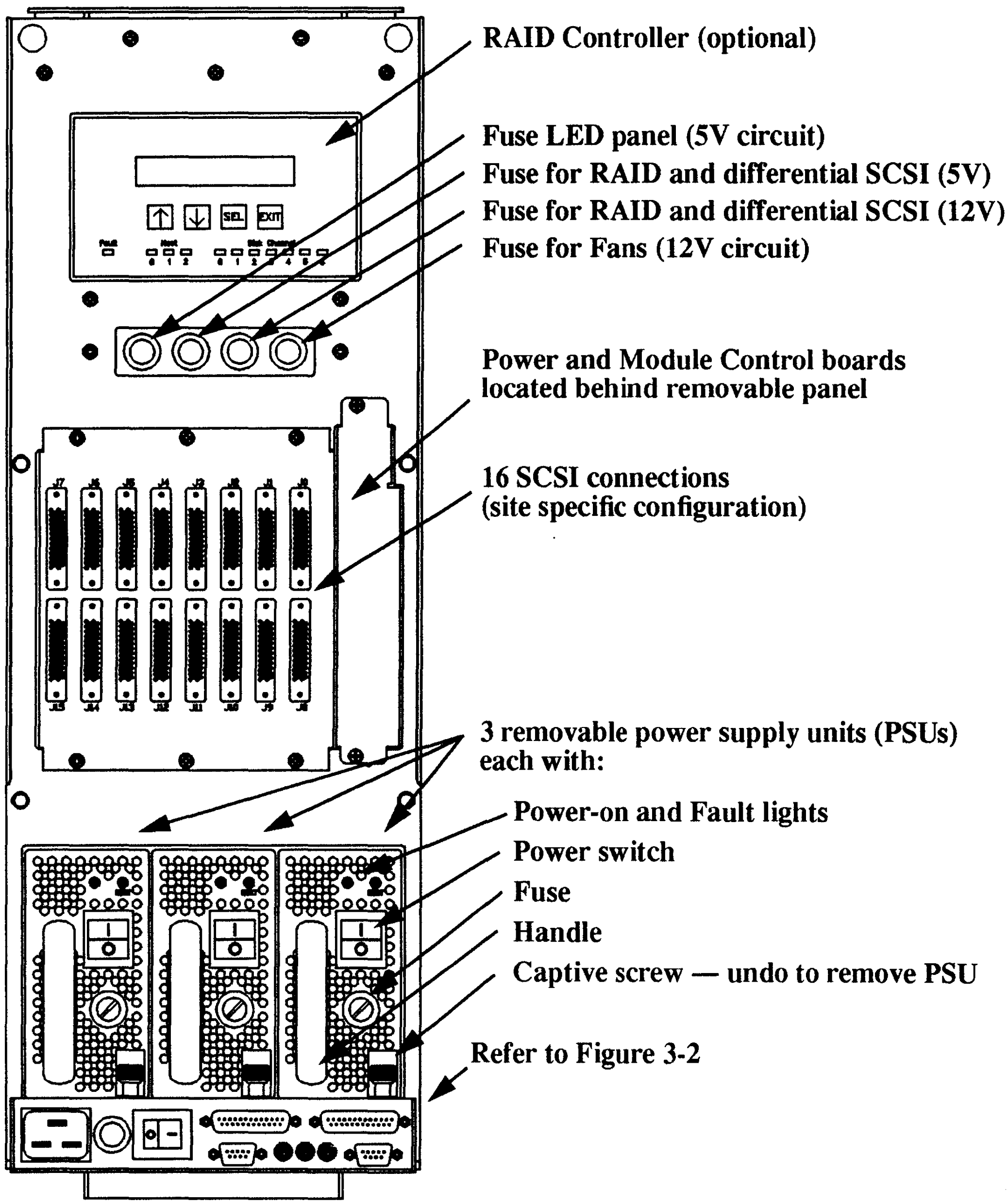


Figure 3-3 The CS-2 Peripheral Module



Installing SCSI Disks 4

The Peripheral Module may be fitted with up to sixteen 3.5" SCSI disk devices.

Warning - The procedures described in this chapter must only be undertaken by trained engineers.

Installing the Disk Devices

Each disk device must be fitted into a Disk Carrier which includes voltage step-down circuits and makes the disk easier to handle.

Fixing the Disk Device into the Disk Carrier

To install a 3.5" SCSI disk into a disk carrier:

- Remove the base plate from the carrier by removing the 6 screws.
- Connect the disk to the power and SCSI connectors.
- At the front of the disk connect the LED activity output to the disk carrier.
- Mount the disk drive into the carrier using four screws that are inserted from two sides of the carrier. Note that the screw must be long enough to accommodate the thickness of the carrier wall.
- Reinstall the carrier's base plate.

Fixing the Carrier into the Processor Module

Insert the carrier into the module so that the red and green indicator lights on the front of the carrier are at the top. Push the carrier firmly into position and secure with the two captive screws.

External Indicators

Three LEDs show from the front of the disk carrier. The green LED is the power light, the red LED is the activity light, and the yellow is the fail indicator. The activity and power signals are also displayed on the module's LED panel.

External Connection

At the time the Peripheral Module is manufactured the sixteen 50-way connectors at the rear of the module have no internal connections; they are transferred directly to 50-way receptacles within the module. Similarly the 16 disk carriers each connect with sixteen 50-way receptacles on the rear of the module's backplane.

The interconnection of the disk devices with the external connectors is via 50-way ribbon cable, and is arranged to suit the specific requirements of each system. Common configurations are:

- **RAID array.** A RAID controller is fitted into the Peripheral Module. The *host* interfaces to this controller (usually two) are each transferred via differential SCSI converters to the external connectors at the rear of the module. The *disk* interfaces on the RAID controller connect with the disk devices.
- **Disk array.** Disks may be daisy chained within the module, and one or both ends of each chain may be taken to the external connectors at the rear of the module.

You must contact Meiko to determine the configuration of your system.

Installing Backplane Boards 5

Two boards are plugged into the module's backplane via a single slot at the rear of the module.

Warning - The procedures in this chapter must only be undertaken by trained engineers.

Warning - You must disconnect the power supplies before installing boards into the Peripheral Module.

Installation

Insert the Power Supply Board (MK536) first and then the Module Controller (MK515) into the rear edge of the MK536.

Insert each card into the guides ensuring that the component side of the card faces the right. Firmly push the card into the position ensuring that the card mates correctly with its connection.

After installing the cards you must reinstall the blanking plate to ensure correct cooling and compliance with RFI regulations.

MK536 Power Supply Card

The MK536 converts the 24V produced by the power supplies into 5V and ±12V for use by the cooling fans, the front panel LED display, the module controller, the RAID controller, and the differential SCSI converter.

This board has no field serviceable components.

MK515 Module Control Card

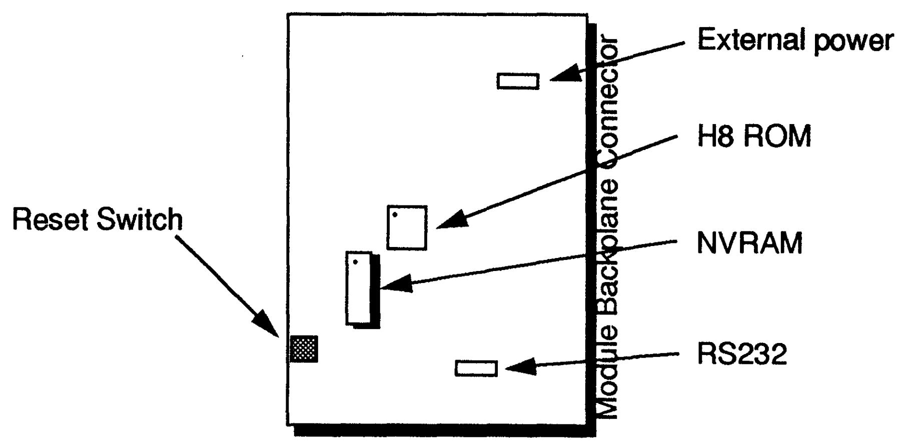
The MK515 monitors the status and configuration of the module, propagating error conditions over the CAN bus (for interception by the system software) and initiating module shutdown under critical conditions (such as major component failure).

Field Upgradeable Components

The MK515 has the following field upgradeable components:

- H8 ROM.
- Non-volatile RAM (NVRAM).

Figure 5-1 MK515 Components



Meiko may upgrade the H8 ROM from time to time.

The non-volatile RAM (NVRAM) holds system configuration information. The NVRAM contains Lithium batteries which have special handling and disposal requirements — your attention is drawn to the *General Precautions* on page ii.

External Indicators

Two LEDs (one green, one amber) are visible from the rear of the module. The green LED is the module controller's CAN heart beat, the amber light illuminates each time the controller writes to the CAN bus. These indicators are also displayed in the top right hand corner of the module's LED display.

Reset Switch

The reset switch is not accessible when the board is mounted in the module. It is used by Meiko engineers when bench testing the board.

External Connections

Two 3 pin connectors are provided for Meiko engineering use. They are used for bench testing of the board, and allow power and diagnostic RS232 connections to be made.

Using the RAID Controller A

Peripheral Modules that are configured as RAID arrays currently include the CRD-5000 RAID Controller.

Detailed usage instructions for this device are provided in the *CRD-5000 SCSI RAID Controller User's Guide* (Revision 1.1, MAN-005000-000), published by CMD Technology Inc. A copy of this document is included in your hardware documentation set.

Components

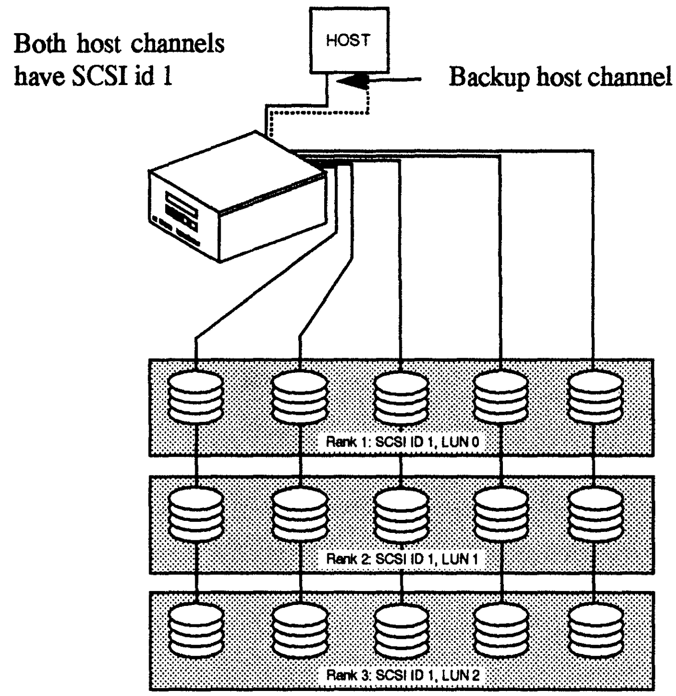
The typical RAID configuration supplied by Meiko includes two *host modules* and 5 *disk modules*[1]. Each of the disk modules is connected to 3 disks. The cache is configured with 32Mbytes of memory.

For systems in which the Peripheral Module is situated some distance away from its host processor a differential SCSI link may be employed; a differential SCSI interface will be fitted into the Peripheral Module and an S-Bus slot of the hosting processor.

A typical RAID configuration is shown in Figure A-1.

1. Host and disk modules are describe on page 2-3 of the RAID Controller User's Guide.

Figure A-1 Typical RAID Configuration



Setup and Performance Monitoring

Meiko recommends that you use a terminal connection, via the RAID console socket at the rear of the module, to setup and monitor the activity of the RAID controller.

You may also setup and monitor activity using the two-line 20 character LCD display provided on the controller's front panel. However to view this display you must either remove the module from the Bay, or remove service access covers within the Bay.

Computing
Surface

SPARC/IO Processing Element (MK401) Users Guide

S1002-10M137.00

meiko

Meiko's address in the US is:

Meiko
130 Baker Avenue
Concord MA01742

508 371 0088
Fax: 508 371 7516

Meiko's address in the UK is:

Meiko Limited
650 Aztec West
Bristol
BS12 4SD

Tel: 01454 616171
Fax: 01454 618188

Issue Status:		
	Draft	
	Preliminary	
	Release	x
	Obsolete	

Circulation Control: *External*

Contents

Overview 1

The MK401 SuperSPARC/IO board offers high performance scalar computing power and flexible I/O options. It provides a general purpose compute server capability, typically offering to the user community shared Solaris processing resource and access to disk, networking, and other I/O resources.

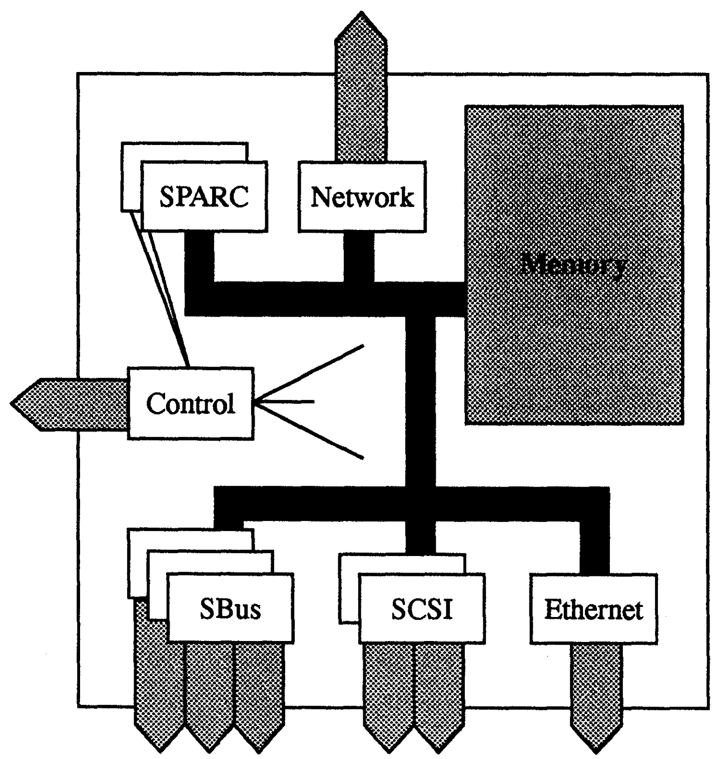
The board design encompasses at the lowest level the principle design objectives for the CS-2, offering a scalable modular construction with easy upgrade options, a reliance on state of the art commodity components, leading edge proprietary network components, and support for system-wide fault tolerance.

In outline the SuperSPARC/IO board offers:

- Dual Superscalar SPARC MBus modules. Easy upgrade to next generation SPARC technology protects investment and extends system life.
- Meiko Elan Communications Processor offering a high bandwidth, low latency interface to the CS-2 data network.
- Up to 512Mbytes of field configurable memory with error correction, detection and logging.
- Dual SCSI-2 controllers for external storage devices.
- Three full size SBus slots for more network connectivity, storage capacity, or other third party options.
- On-board Ethernet controller for use with thin or thick wire Ethernet.

- Interface to the machine-wide control area network (CAN) offering remote diagnostic control and error logging facilities.
- Keyboard, mouse, and dual serial connections.
- Fully compliant with the SPARC Compliance Definition (SCD); applications can be ported between SCD environments without change.

Figure 1-1 Board Overview



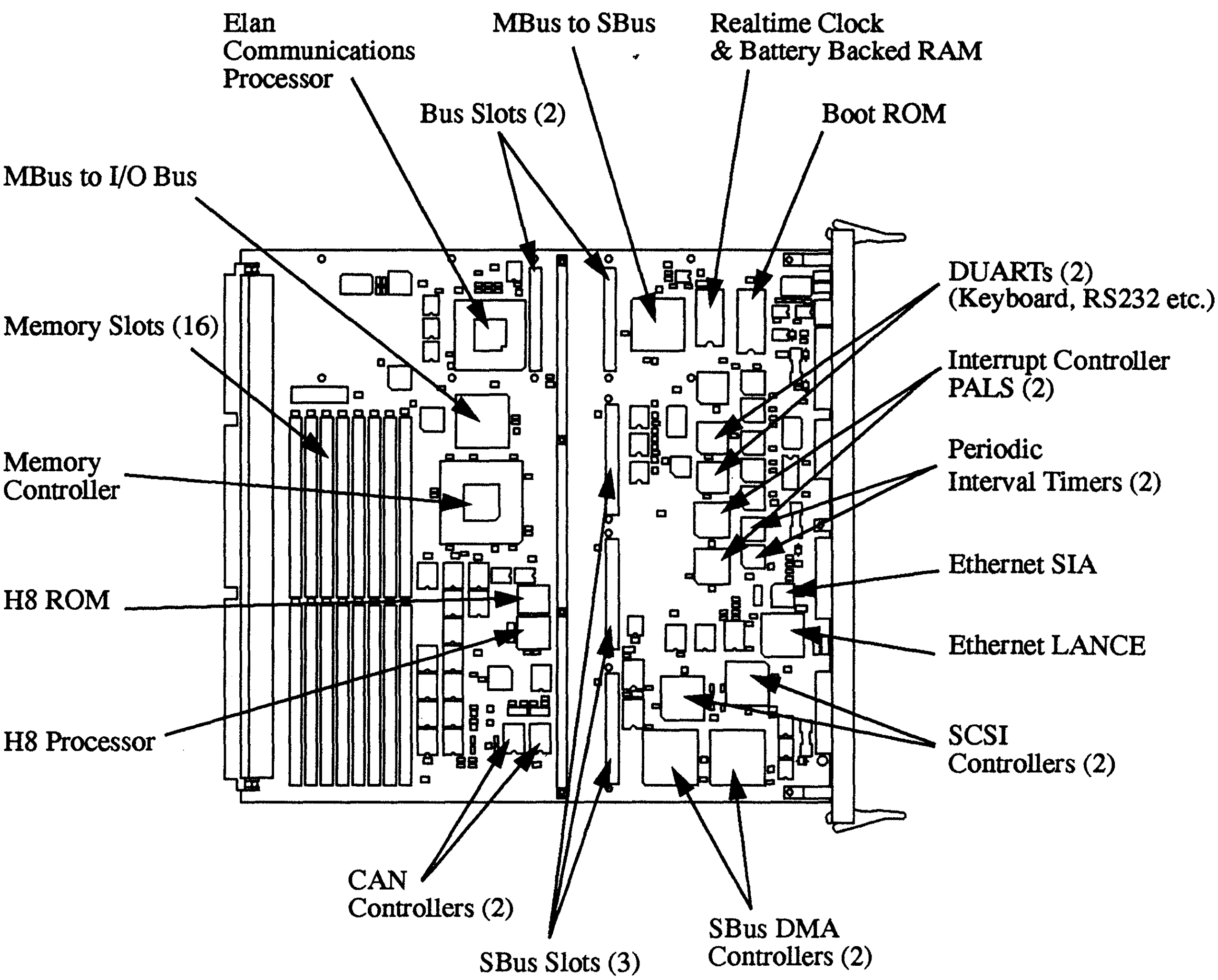
MK401 Board Description 2

Interconnection of the processing devices and the memory system is via the industry standard SPARC MBus architecture using a 40MHz, level 2 cache coherent implementation. Interfacing the MBus are two additional buses; the SBus connects the major I/O devices (such as the SCSI controllers), and the I/O bus connects minor I/O devices (such as the real time clock).

Running throughout the whole CS-2 system is a control network (CAN) used to distribute status and configuration information, to provide remote control and diagnostics of all processors, and to create remote console connections to the processors. The MK401 has two interfaces to this network; one connected to the I/O bus (and thus providing a direct interface to the SPARC processors) and one via a dedicated micro-controller which provides board control.

The major components and their placement on the MK401 motherboard are shown in Figure 2-1.

Figure 2-1 Board Schematic Showing Major Components



MBus and SBus plug-in modules not shown.

MBus

Two full size MBus sites are provided allowing two SPARC CPUs to be supported either on a single dual processor daughter board, or two single processor boards. The MBus boards simply plug into the slots provided.

The MBus is fully level 2 compliant and runs at 40MHz. The two SPARC processors share the MBus with the Elan Communications Processor, the MBus-to-SBus interface, and the I/O bus controller; the allocation of MBus Id's is:

- MBus id 0 is the I/O bus controller.
- MBus id 4 is MBus to SBus controller.
- MBus id 6 is Elan Communications processor.
- MBus id 8 and 9 are MBus slot 0.
- MBus id 10 and 11 are MBus slot 1.

The use of MBus sockets on the motherboard allows SPARC modules to be easily upgraded as SPARC technology develops. The TI Viking and ROSS Pinnacle SPARC technologies are current options.

Note that MBus modules from different manufacturers may not be compatible. If two MBus cards are fitted they must use the same technology and originate from the same manufacturer.

ROSS Pinnacle Module

Two variants of the Pinnacle MBus module are available; you may use either two single processor modules or one dual processor module. Both variants include external second level cache.

The Pinnacle MBus modules are built upon a tightly coupled set of three ROSS devices: the RT620 HyperSPARC CPU, the RT625 cache controller, memory management, and tagging unit (CMTU), and the RT627 cache data units (CDUs).

Features of the RT620 CPU are:

- SPARC version 8 conformance.
- 90MHz clock rate.

- 4 execution units offering parallel execution of major instruction types: Load/Store, Branch/Call, integer and floating point units.
- Dual instruction fetch per clock cycle.
- 8Kbyte 2-way set-associative on-chip instruction cache.
- Instruction pipelining including a cache stage to accommodate the latency for second level cache accesses on data. Simultaneous accesses to on-chip and second level cache for each instruction fetch.
- High bandwidth 64bit Intra Module Bus (IMB) provides the interface between the CPU and the second level cache. Use of second level cache decouples the processor clock rate from the lower MBus clock rate.

Key features of the RT625 (CMTU) and RT627 (CDU) devices are:

- Full level 2 cache-coherent MBus compatibility.
- Each CDU has integral 16Kbytes x 32bit SRAM. MBus modules use either 2 or 4 CDUs for 128Kbyte or 256Kbyte second level direct-mapped cache.
- Physical cache tagging with virtual indexing allow the cache coherency logic to determine snoop hits and misses without stalling the CPU's access to the cache.
- Both copy-back and write-through cache modes supported.
- 32 byte read buffer and 64 byte write buffer for buffering the 32 byte cache lines in and out of the second level cache.
- SPARC reference MMU offering 64 entry, fully set-associative TLB with 4096 contexts.

Texas Instruments Viking Module

Two variants of the TI Viking MBus module are available. One contains a Viking SPARC processor with direct connection to the MBus. The second includes a Viking processor with additional Cache Controller and 1Mbyte of second level external cache (E-cache).

Key features of the TMS390Z50 SuperSPARC are:

- SPARC Version 8 conformance.
- 3 instructions per cycle, instruction pipelining, 150MIPs peak performance.
- SPARC Integer Unit.
- SPARC Reference MMU. Cached translation lookaside buffers (TLBs). 32bit virtual addresses, tagged with a 16bit context (65,536 contexts), map to 36bit physical addresses.
- Single and Double precision FPU. Tightly coupled to the integer execution pipeline and allowing one floating-point operation and one memory reference to be issued in each clock cycle. The FPU maintains a 4 entry FIFO queue for FP operations.
- 20Kbyte instruction cache, 16Kbyte data cache. The instruction cache is 5-way set associative, physically addressed, and non-writable. The data cache is 4-way associative and physically addressed. Both cache's are coherent with each other and with optional E-cache or MBus. Without E-cache the instruction and data caches operate in write-through mode; otherwise they are copy-back mode.
- Store buffer. A FIFO queue of 8 entries, each 64bits, decouples the instruction execution pipelines from the E-cache or MBus.
- Multiprocessor cache coherent support; highly pipelined and non-multiplexed VBus interface to optional external cache and the TMS390Z55 Cache Controller, or direct connection to the MBus.
- Prefetch buffer.
- Support for system and software debugging including hardware breakpoint.

The external cache is managed by the TMS390Z55 Cache Controller. Key features of this device are:

- Built in support for cache coherent multiprocessing; multiple Viking modules can share a single MBus and remain fully cache coherent.

- High performance VBus interface with the SPARC processor; decouples the SPARC processor from the MBus clock speed allowing higher processor performance. Reduced MBus traffic reduces contention when multiple processors share a single MBus.
- E-cache is direct mapped, copy-back, and unified: there is a single cache location where a particular byte of the physical address space can reside in the cache (direct mapping); writes by the main processor into the E-cache do not propogate to main memory until the cache is flushed or replaced (copy-back); both instructions and data are supplied to the processor from the same cache (unified).
- 1Mbyte SRAM cache.

SBus Interfaces

The MBus to SBus interface supports up to five SBus devices. Device allocation on the MK401 is as follows:

- Devices 0, 1, and 2 are the SBus slots. Slot 0 is nearest the processor modules.
- Device 4 is an SBus DVMA device serving both the Ethernet controller and one of the SCSI-2 controllers.
- Device 5 is an SBus DVMA device serving the second SCSI-2 bus controller.

The Ethernet controller connects to an external Ethernet transceiver via a front panel connection and standard Ethernet drop cable. Connection to either thin or thick wire Ethernet is supported.

The SCSI-2 controllers support an 8bit single ended SCSI bus. External connections are via the backplane connectors (allowing connection to disks within the Processor Module) or via standard high density connectors on the front panel (allowing connection to external SCSI devices). Switchable SCSI terminators are accessible from the board's front panel and are used to terminated the bus when not in use or when SCSI devices are connected to just one end (either the front panel connectors or the backplane connectors); termination of the external buses is also required.

The SBus runs at a clock speed of 20Mhz.

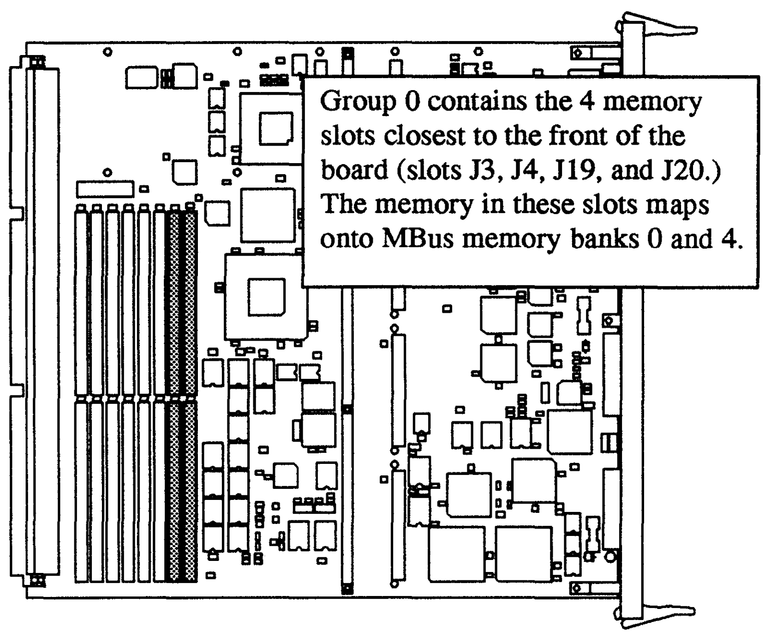
Memory Configurations

The memory controller used by the MK401 offers a 128bit wide data bus for fast memory access, with single bit error correction, 2 bit error detection, and multiple bits within nibble detection. Memory errors are logged by kernel software and propagated via the CAN bus.

Up to 16 memory modules may be fitted to the MK401 offering a maximum of 512Mbytes. Both single and double sided, 4Mbit and 16Mbit SIMM may be used. The 16 sockets on the board are configured as 4 groups of 4 sockets; within each group the SIMMs must be identical, but there is no requirement for the groups to be the same.

Memory is mapped into the MBus address space in 8 banks of 64Mbytes. Bank 0 maps onto one side of the SIMMS in physical group 0. Bank 4 maps onto the second side of the SIMMS in physical bank 0. If less than 64Mbytes of memory is present in a bank it is echoed throughout the 64Mbytes. For each 128bit MBus data access 4 bytes are read from each SIMM.

Figure 2-2 MBus Memory Mapping and Physical Memory Groups



IO Bus

The IO bus is a slave-only bus used for the connection of minor peripherals to the MBus. The following devices are connected to this bus:

- A 512Kbyte EPROM holding bootstrap and diagnostic programs. The bootstrap code initialises the hardware devices, initial page table construction, and the booting of the Unix kernel. Normally only processor 0 in a dual-processor configuration handles the hardware initialisation; processor 1 sleeps until it is woken by processor 0 with a software interrupt.
- A realtime clock module with battery backed SRAM. This holds configuration information and node fault logs, including the log of uncorrectable memory errors. The realtime clock provides year, month, day, hour, minute, and second times.
- Dual DUART devices; one for connection of keyboard and mouse, the other for two general purpose serial ports. In the absence of a keyboard the bootstrap code in the EPROM will usually direct console I/O via serial port A. The serial ports are clocked at 4.9152MHz, with a working capability of 38.4KBaud. Both serial ports share the same 25-way front panel connection; port A has full synchronous/asynchronous operation and a full complement of modem control lines; port B has a limited set of control lines and is asynchronous only.
- Interrupt controllers, one per SPARC processor. These use registers to mask out certain types of interrupt to relieve the SPARC processor from unnecessary interrupt loading, and to share the handling between the two processors. In addition some levels of interrupt are also presented direct to the Elan allowing them to be handled there and thus masked out of both the SPARC processors. The programming of the interrupt controllers is handled by the kernel device drivers.
- Periodic interrupt timers; used for maintaining the kernel clock and for kernel profiling. Two clock devices are used, one per SPARC processor. The lower level (kernel clock) interrupts from one of the clock devices are shared by both SPARC processors to remove the need for the two processor clocks to synchronise.

- A single CAN device provides both SPARCs with an interface to the machine-wide control area network (CAN). The SPARC processors write diagnostic information to this bus, and can also act as X-CAN or G-CAN routers. CAN routers transfer data between two levels of the CAN network; an X-CAN router handles transfers between the modules in a Cluster, and a G-CAN router handles transfers between Clusters. The configuration of a SPARC as a router will cause it's CAN device to generate numerous level 2 interrupts which will impact on processor performance.

Board Control Processor

The MK401 board uses an Hitachi H8/534 micro-controller (commonly referred to by Meiko as the H8) to perform basic node control functions. This controller is a single-chip 16bit RISC microcomputer with integral 2KBytes RAM, 32Kbytes EPROM, 16bit RISC CPU, and a number of I/O ports and timers.

The H8 processor runs independently of the other processors on the board and is used solely for control and diagnostic purposes. It has its own interface to the CAN bus via the second of the board's CAN interface devices. This processor receives diagnostic messages, via the CAN bus, from the local SPARC processors, and interprets incoming control messages, such as board reset, console connections, and network configuration.

Using the MK401 3

This chapter describes the usage of the MK401 in terms of its installation, hardware interfaces, and field serviceable components.

Installation

The MK401 is designed for use solely in a CS-2 Processor Module. The Processor Module supplies the board's power, cooling, and connection to the CS-2 data and control networks. The MK401 is fitted into one of the four vertical board slots behind the Processor Module's removable front panel.

Warning - You must disconnect the power from the Processor Module before removing or installing processor boards.

Warning - The board may be fitted with fragile or static sensitive devices. You must handle with care and observe anti-static precautions.

Removing the Module's Front Panel

The module's front panel is held in position by four clips, one in each corner. To remove the panel pull firmly away from the module.

The module's LED display is fitted to the module by two 50-way connectors. To remove the LED display pull firmly away from the module.

Use the reverse procedure to install the LEDs and front panel.

Installing the Processor Board

Insert the board so that it fits into the guide rails at the top and bottom of the module's board rack, ensuring that the component side is to the left (viewed facing the module). Gently push the board squarely on its front panel. Before pushing the board fully into position fold back the levers at each end of the front panel so that they are at 90° to the board; now push the board (while holding the levers) until the base of the two levers is touching the card cage. To lever the board into its final position push both levers until they lie flat on the board's front panel. Secure the board by tightening the two captive screws.

Use the reverse procedure to remove the board.

Warning - You should take care not to damage the connectors at the rear of the board and on the module's backplane. Ensure that the board mates squarely with the module's backplane.

Warning - When removing or installing a board you should take care not to damage the RFI (copper) seals along the edge of the board's front panel.

Warning - To maintain proper circulation of cooling air and to conform to RFI regulations all board slots must be fitted with a processor board or blanking plate.

Field Serviceable Components

The MK401 has the following field serviceable components.

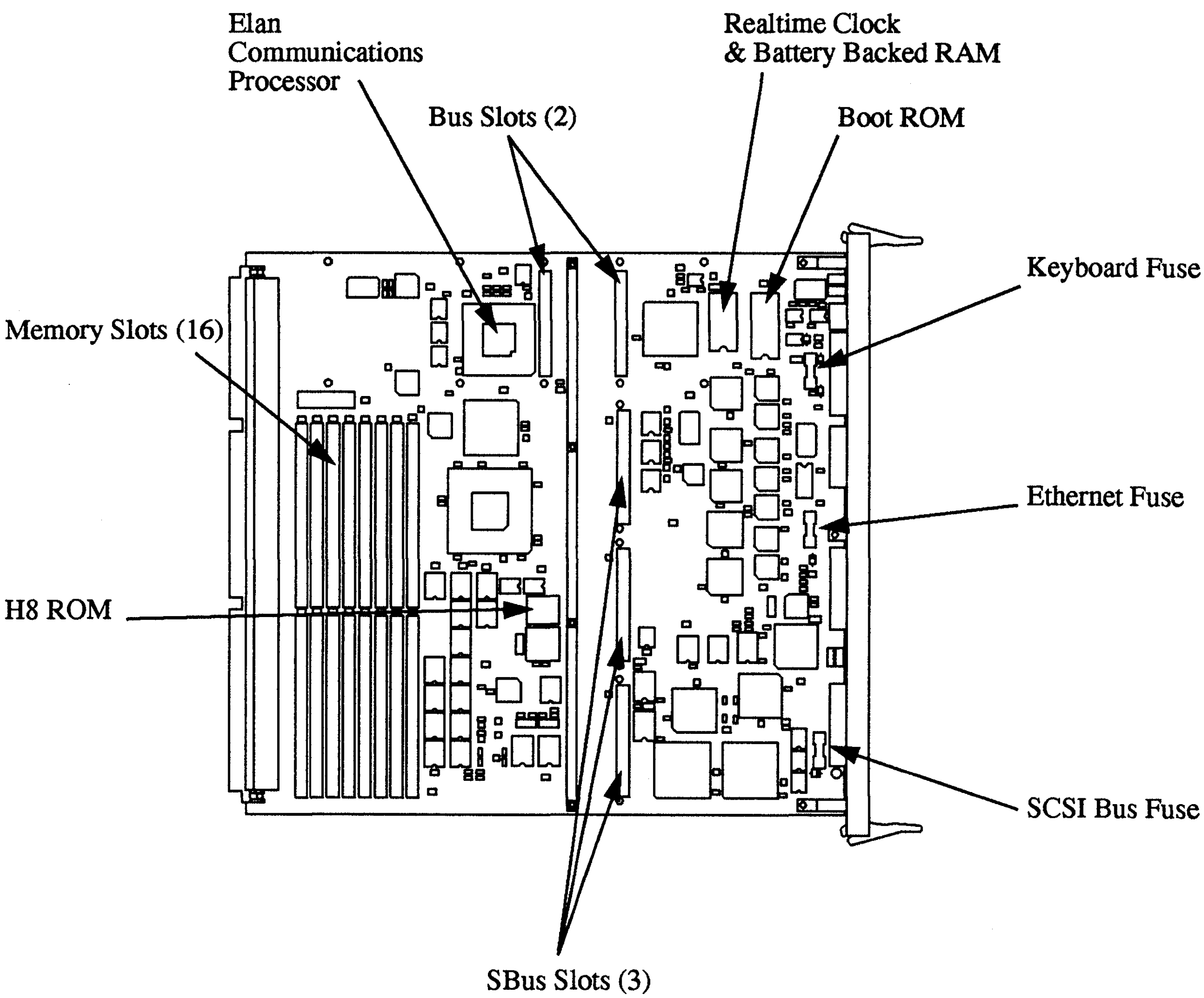
- Superscalar SPARC processor modules fitted to MBus slots.
- Three SBus slots.
- 16 memory slots.
- Boot ROM.
- H8 ROM.
- Realtime clock and non-volatile RAM (NVRAM).
- Fuses.

The location of these components on the MK401 is shown in Figure 3-1.

Warning – User's are not permitted to make mechanical or electrical modifications to the equipment. Meiko is not responsible for regulatory compliance of equipment that has been modified. You will invalidate your warranty if you make unauthorised changes.

Warning – These components are fragile or static sensitive. Handle with care and observe anti-static precautions.

Figure 3-1 Field Serviceable Components



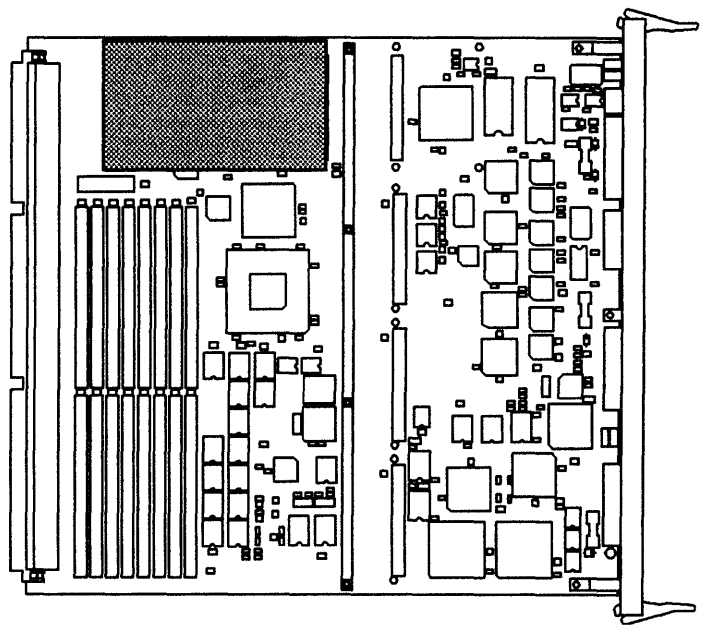
MBus (processor) and SBus (peripheral) plug-in modules are not shown.

Installing MBus Processor Modules

Superscalar SPARC modules may be installed in the two MBus sockets — when using just one module it must be installed in position 0 (the slot nearest the communications processor and the rear of the board). The sockets are polarised to prevent incorrect installation.

Installation is simple — push the processor module into the connector on the MK401 and secure into position with an M3 screw at each corner.

Figure 3-2 Position of MBus Module 0

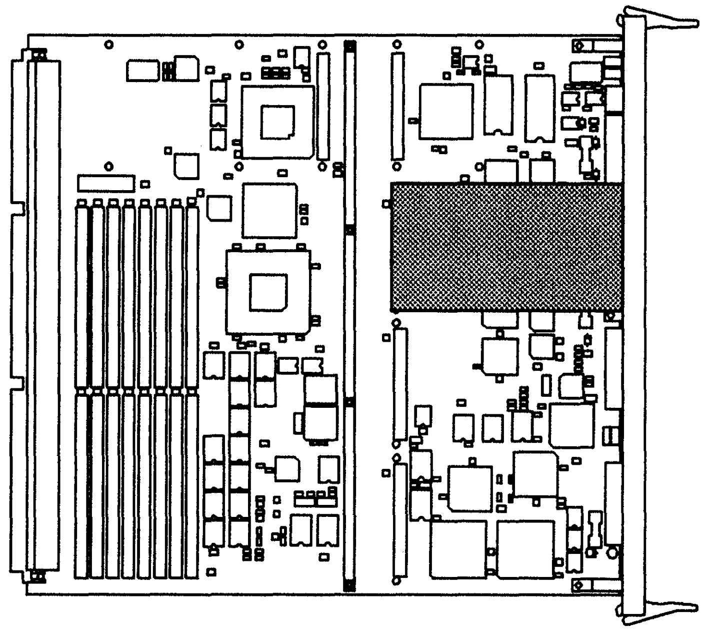


Installing SBus Modules

Three SBus slots are provided and these may be fitted with standard SBus modules; these are plugged into the SBus connectors and secured with two M3 screws. When using SBus cards that have external connections, for example a graphics card, remove the appropriate panel from the front of the MK401 — the panel is held in place by two small screws.

SBus devices are numbered from 0 to 2, device 0 being next to the processor slots. Device 4 is the on-board Ethernet and SCSI bus. Device 5 is the second on-board SCSI bus.

Figure 3-3 Position of SBus Module 0

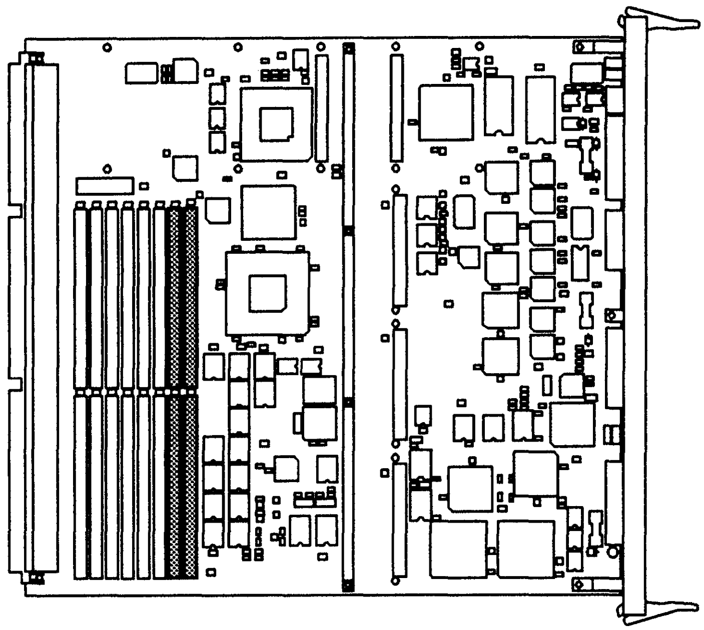


Memory

Up to 16 single in-line memory modules (SIMMs) may be fitted to the MK401 board (JEDEC 36bit SIMMs). The array is constructed of 4 groups of 4 SIMMs, arranged as shown in Figure 3-4. Within each group the SIMMs must be identical, but there is no requirement for the groups to be the same.

SIMMs can use either 4Mbit DRAM or 16Mbit DRAM technology, and be either single or double sided. This gives a minimum memory configuration of 16Mbytes and a maximum of 512Mbytes.

Figure 3-4 Memory slots on an MK401; the shaded region shows one of the 4 groups of 4 memory slots.



Boot ROM and H8 ROM

Both of these ROMs may be upgraded from time to time. They are held in sockets and are readily replaced. Note the position on pin 1 before removing the old device (usually marked by a dot on the packaging).

Realtime Clock and Battery backed RAM

The real time clock and non-volatile RAM device is held in a DIL socket and is easily replaced. Before removing the old device note the position of pin 1 (usually marked by a dot on the packaging). Note that the information within the RAM can only be restored by Meiko's engineers.

Warning - This device contains lithium batteries; never dispose of this device in a fire or attempt to dismantle.

Fuses

There are 3 fuses on the MK401:

Keyboard fuse: 20 × 5mm 1 A quick blow ceramic.

Ethernet fuse: 20 × 5mm 1 A quick blow ceramic.

SCSI bus fuse: 20 × 5mm 500mA quick blow ceramic.

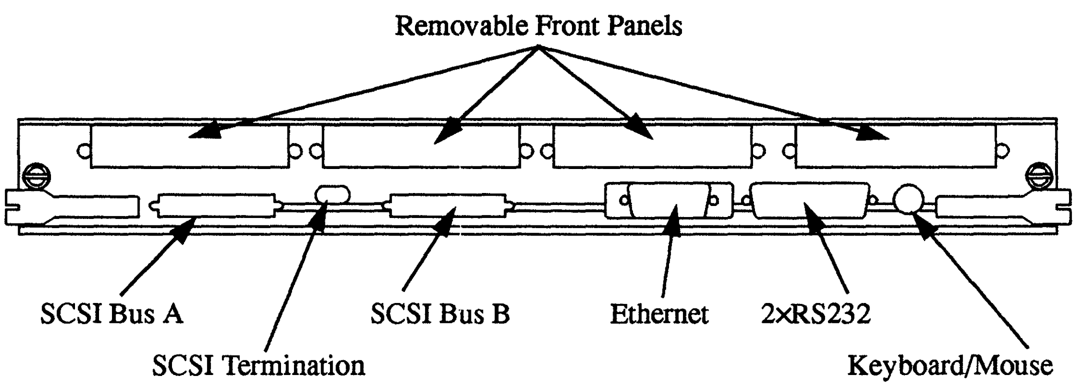
External Connections

External connections are provided for a keyboard/mouse (8 pin circular socket), serial interfaces (2 channels provided by one 25-way D-type connector), Ethernet (15-way D-type connector), and two independent SCSI buses (each via a 50-way miniature connector).

Front Panel Connections

Removable panels provide access to connectors on the optional SBus boards.

Figure 3-5 MK401 Front Panel Connections



RS232 Connections

The two RS232 channels are output via a single 25-way connector. The connections are as shown in the following tables. Signal ground is pin 7, chassis ground in pin 1.

Table 3-1 RS232 Channel A Pinout.

Signal	Input/Output	Pin number
TXD	Out	2
RXD	In	3
RTS	Out	4
CTS	In	5
CSR	In	6
DCD	In	8
DB	In	15
DD	In	17
DA	On	24
DTR	On	20

Table 3-2 RS232 Channel B Pinout.

Signal	Input/Output	Pin number
TXD	Out	14
RXD	In	16
RTS	Out	19
CTS	In	13
DCDB	In	12

Adding SCSI Peripherals

The MK401 includes two SCSI-2 controllers — both SCSI buses are available via connections on the front panel or via the Processor Module's backplane. Up to 3 additional SCSI buses can be added using standard SBus cards.

The disk devices that are optionally fitted into the Processor Modules are connected, via the module's backplane, to the SCSI controllers on your processor boards. The allocation of disks to boards and their controllers, and the number of disk devices is site specific, typical allocations are: 4 disks, one connected to SCSI bus A on each board; 4 disks, 2 connected to SCSI bus A on board's 0 and 1; 4 disks all connected to SCSI bus A on board 0.

SCSI Termination

Two switches on the MK401 front panel allow the SCSI bus termination to be switched; with the switches down termination is on.

If *either* the front panel connections *or* the backplane connections are used, or no devices are attached at all, then the on-board termination must be turned on. Only when both front-panel and backplane connectors are used together should the on-board termination switches be turned off. The end's of a SCSI bus must always be terminated.

External Indicators

Two LEDs (one green, one amber) are included on the board's front panel. The green LED is the heart beat from the board's (H8) CAN controller. The amber light illuminates each time the CAN controller transmits on the CAN bus. Both should flash steadily. These indicators are also displayed on the module's LED display.

The green LED flashes at a slow steady rate (once per second) when operating normally. A quicker flash rate (2×normal) indicates that the board's SPARC processors are not responding; a very quick flash rate (3×normal) indicates that the H8 processor on the module's controller is not responding.

Each processor board within a processor module controls a 4×4 matrix of red LEDs on the module's front panel. The MK401 displays a random pattern on these when running the Boot ROM. When Solaris has been booted a circulating pattern is displayed. The pattern can be changed by user programs and various system commands and daemons.

MBus Address Map A

This section identifies the key devices used on the MK401 and lists their mapping into the MBus physical address space.

For detailed information about any of the devices you should refer to the manufacturer's data sheets.

The following notes are associated with the tables:

1. These locations are byte-wide and are mapped into all 4 bytes of a word. Care should be taken to generate correct byte-wide accesses to the least significant byte of the word in order to maintain future compatibility.
2. These locations are byte-wide memory, mapped into contiguous byte locations. Word or halfword accesses will be automatically mapped into several successive byte-wide accesses.
3. These locations are byte sized registers which are only mapped into the most significant byte of a halfword. To ensure comparability with other boards only byte accesses at the correct address should be used.
4. These locations are byte sized registers which are only mapped into the least significant byte of the word. To conserve MBus bandwidth and ensure comparability with other boards only byte wide accesses at the correct address should be used.

5. These locations are halfword sized registers which are only mapped into the least significant halfword of the word. To conserve MBus bandwidth and ensure compatibility with other boards only halfword wide accesses at the correct address should be used.
6. These locations form a double-word register.

All addresses are in hexadecimal, and all locations are word-wide unless otherwise stated in the notes.

DRAM and SBus Slots

MBus Address	Usage	Read/ Write	Note
000000000	64MB Memory in bank 0.	RW	1
004000000	64MB Memory in bank 1.	RW	1
008000000	64MB Memory in bank 2.	RW	1
00c000000	64MB Memory in bank 3.	RW	1
010000000	64MB Memory in bank 4.	RW	1
014000000	64MB Memory in bank 5.	RW	1
018000000	64MB Memory in bank 6.	RW	1
01c000000	64MB Memory in bank 7.	RW	1
020000000 to dffffffff	Unused (MBus Timeout).		
e00000000 to e0fffffff	SBus Slot 0 (connector J11).	RW	
e10000000 to e1fffffff	SBus Slot 1 (connector J12).	RW	
e20000000 to e2fffffff	SBus Slot 2 (connector J13).	RW	

SBus SCSI and Ethernet

The MBus-to-SBus (M2S) device drives two LSI Logic L64853A SBus DMA devices. One (chip B) drives the SCSI-B controller, the other (chip A) controls both the SCSI-A and Ethernet controllers.

SBus DMA chip B and SCSI

An Emulex FAS101 SCSI controller provides the MK401's SCSI-B bus.

The memory map for the SCSI SBus DMA device and the SCSI controller is:

MBus Address	Usage	Read/ Write	Note
e30000000	SBus DMA_B ID register (= 0xfe810102)	R	
e30000004 to e303fffff	Unused (echoes of above)		
e30400000	SBus DMA_B Control/Status Register	RW	
e30400004	SBus DMA_B (Next)AddressCounter	RW	
e30400008	SBus DMA_B (Next)ByteCount	RW	
e3040000c	Reserved for testing M2S		
e30400010 to e307fffff	Unused (echoes of above)		
e30800000	SCSI_B Transfer Count Low	RW	1
e30800004	SCSI_B Transfer Count Mid	RW	2
e30800008	SCSI_B FIFO Data	RW	2
e3080000c	SCSI_B Command	RW	2
e30800010	SCSI_B Status	R	2
e30800010	SCSI_B Destination Bus ID	W	2

MBus Address	Usage	Read/ Write	Note
e30800014	SCSI_B Interrupt	R	2
e30800014	SCSI_B Select/Reselect Timeout	W	2
e30800018	SCSI_B Sequence Step	R	2
e30800018	SCSI_B Synchronous Period	W	2
e3080001c	SCSI_B FIFO Flags	R	2
e3080001c	SCSI_B Synchronous Offset	W	2
e30800020	SCSI_B Configuration 1	RW	2
e30800024	SCSI_B Clock Conversion Factor	W	2
e30800028	SCSI_B Test mode	W	2
e3080002c	SCSI_B Configuration 2	RW	2
e30800030	SCSI_B Configuration 3	RW	2
e30800034	SCSI_B Reserved		
e30800038	SCSI_B Transfer Count High	RW	2
e3080003c	SCSI_B Reserved		
e30800040 to e30bfffff	Unused (echoes of above)		
e30c00000 to e3fffffff	Reserved (Read Undefined)		

SBus DMA chip A, Ethernet and SCSI

An Emulex FAS101 SCSI controller and an Advanced Micro Devices AM7990 Ethernet controller (LANCE) provide the MK401's SCSI-A bus and Ethernet.

The following table identifies the memory mapping for the SBus DMA device, the SCSI, and the Ethernet controllers.

MBus Address	Usage	Read/ Write	Note
e40000000	SBus DMA_A ID register (=0xfe810102)	R	
e40000004 to e403fffff	Unused (echoes of above)		
e40400000	SBus DMA_A Control/Status Register	RW	
e40400004	SBus DMA_A (Next)AddressCounter	RW	
e40400008	SBus DMA_A (Next)ByteCount	RW	
e4040000c	Reserved for testing		
e40400010 to e407fffff	Unused (echoes of above)		
e40800000	SCSI_A Transfer Count Low	RW	2
e40800004	SCSI_A Transfer Count Mid	RW	2
e40800008	SCSI_A FIFO Data	RW	2
e4080000c	SCSI_A Command	RW	2
e40800010	SCSI_A Status	R	2
e40800010	SCSI_A Destination Bus ID	W	2
e40800014	SCSI_A Interrupt	R	2
e40800014	SCSI_A Select/Reselect Timeout	W	2
e40800018	SCSI_A Sequence Step	R	2
e40800018	SCSI_A Synchronous Period	W	2
e4080001c	SCSI_A FIFO Flags	R	2
e4080001c	SCSI_A Synchronous Offset	W	2
e40800020	SCSI_A Configuration 1	RW	2

MBus Address	Usage	Read/ Write	Note
e40800024	SCSI_A Clock Conversion Factor	W	2
e40800028	SCSI_A Test mode	W	2
e4080002c	SCSI_A Configuration 2	RW	2
e40800030	SCSI_A Configuration 3	RW	2
e40800034	SCSI_A Reserved		
e40800038	SCSI_A Transfer High	RW	2
e4080003c	SCSI_A Reserved		
e40800040 to e40bfffff	Unused (echoes of above)		
e40c00000	LANCE Register Data Port	RW	2
e40c00002	LANCE Register Address Port	RW	2
e40c00004 to e5fffffff	Unused (echoes of above)		
e60000000 to e600000ff	Unused (Invalid TLB Entry) (MBus Error?)		
e60000100 to e600001f0	M2S TLB Slices 0 through 15. (on 16-byte boundaries)	RW	
e60000200 to e6fffffff	Unused (Invalid) (MBus Error?)		
e70000000 to effffffff	Unused (SBus Reserved) (MBus Error?)		

Memory Controller

The MK401 uses an LSI Logic L64860 memory controller.

The following table shows the MBus memory maps for the memory controller's control and diagnostic ports.

MBus Address	Usage	Read/ Write	Note
`f00000000`	Memory Enable	RW	
`f00000004`	Memory Delay	RW	
`f00000008`	Fault Status	R(W)	
`f0000000c`	Video Config.	RW	
`f00000010`	Fault Address 0	R	
`f00000014`	Fault Address 1	R	
`f00000018`	ECC Diagnostics	RW	
`f0000001c` to `f0fffffff`	Unused (Read undefined)		
`f00000000` to `feffffff`	Unused (No response) (MBus Timeout)		

Boot ROM, Serial Ports, Real Time Clock, Miscellaneous

The following table identifies the memory map for various devices on the I/O bus.

MBus Address	Usage	Read/ Write	Note
ff0000000 to ff007ffff	Boot ROM (512KByte).	R(W)	3
ff0080000 to ff00fffff	Unused (Boot ROM Echo)		
ff0100000 to ff0100007	Serial Port Controller		
ff0100000	Control Registers port B	RW	4
ff0100002	Data Buffer port B	RW	4
ff0100004	Control Registers port A	RW	4
ff0100006	Data Buffer port A	RW	4
ff0100008 to ff01fffff	Unused (Serial Port Echoes)		
ff0200000 to ff0200007	Keyboard and Mouse Port Controller		
ff0200000	Control Registers mouse port	RW	4
ff0200002	Data Buffer mouse port	RW	4
ff0200004	Control Registers keyboard port	RW	4
ff0200006	Data Buffer keyboard port.	RW	4
ff0200008 to ff02fffff	Unused (Keyboard and Mouse Port Echoes)		
ff0300000 to ff0301fff	Real Time Clock module and 8KByte SRAM	RW	3
ff0302000 to ff03fffff	Unused (RTC Echoes)		

MBus Address	Usage	Read/ Write	Note
ff0400000 to ff06fffff	Unused (MBus Error)		
ff0700000	Node Reset Request	W	5
ff0700004 to ff07001ff	Unused (echoes)		
ff0700200	MBus Grant readback	R	5
ff0700204 to ff07003ff	Unused (echoes)		
ff0700400	Physical Slot Identifier	R	6
ff0700404 to ff07005ff	Unused (echoes)		
ff0700600	LED Bargraph	RW	6
ff0700604 to ff07007ff	Unused (echoes)		

Control Area Network Interface

The SPARC processors' interface the Control Area Network via Philips PCA82C200 devices, which are mapped into the MBus address space at the following locations:

MBus Address	Usage	Read/ Write	Note
ff0700800	CAN — Control Register	RW	5
ff0700804	CAN — Command Register	W	5
ff0700808	CAN — Status Register	R	5
ff070080c	CAN — Interrupt Register	R	5
ff0700810	CAN — Acceptance Code Register	RW	5
ff0700814	CAN — Acceptance Mask Register	RW	5

MBus Address	Usage	Read/ Write	Note
ff0700818	CAN — Bus Timing Register 0	RW	5
ff070081c	CAN — Bus Timing Register 1	RW	5
ff0700820	CAN — Output Control Register	RW	5
ff0700824	CAN — Test Register		
ff0700828	CAN — TXBuf Identifier	RW	5
ff070082c	CAN — TXBuf RTR Data Length code	RW	5
ff0700830	CAN — TXBuf Data Byte 1	RW	5
ff0700834	CAN — TXBuf Data Byte 2	RW	5
ff0700838	CAN — TXBuf Data Byte 3	RW	5
ff070083c	CAN — TXBuf Data Byte 4	RW	5
ff0700840	CAN — TXBuf Data Byte 5	RW	5
ff0700844	CAN — TXBuf Data Byte 6	RW	5
ff0700848	CAN — TXBuf Data Byte 7	RW	5
ff070084c	CAN — TXBuf Data Byte 8	RW	5
ff0700850	CAN — RXBuf Identifier	RW	5
ff0700854	CAN — RXBuf RTR Data Length code	RW	5
ff0700858	CAN — RXBuf Data Byte 1	RW	5
ff070085c	CAN — RXBuf Data Byte 2	RW	5
ff0700860	CAN — RXBuf Data Byte 3	RW	5
ff0700864	CAN — RXBuf Data Byte 4	RW	5
ff0700868	CAN — RXBuf Data Byte 5	RW	5
ff070086c	CAN — RXBuf Data Byte 6	RW	5
ff0700870	CAN — RXBuf Data Byte 7	RW	5

MBus Address	Usage	Read/ Write	Note
ff0700874	CAN — RXBuf Data Byte 8	RW	5
ff0700878	CAN — Unimplemented		
ff070087c	CAN — Clock Divider Register	RW	5
ff0700880 to ff0700fff	Unused (echoes of above)		

Interrupt Request Control and Status Registers

The MK401 has many sources of interrupts to the SPARCs. These are assigned priority levels and are passed to the Interrupt Controllers (one per SPARC). These enable the handling of interrupts to be shared between the SPARCs. The controller has the following major features:

- Masks for all hardware generated interrupts.
- Software interrupts on six levels using a multi-processor compatible set/clear register structure. This allows one processor to interrupt the other by writing once to a single memory location, without having to do a read-modify-write cycle which is subject to being broken by the other processor. Supported interrupt levels are 1, 4, 6, 12, 13, and 15.
- Latches triggered by the output of the free running timers (the Periodic Interrupt Timers). These latches capture the timer passing through zero event and cause an interrupt to the processor. When the interrupt handler accepts the interrupt and clears the request the latch is reset.
- Priority encoder to generate the 4bit encoded interrupt level that is passed to the SPARC processor module.

The Periodic Interrupt Timers are AMD 82C54 free running timers, one assigned to each of the SPARC processors. Timer 0 generates the lower priority clock ticks signal, issuing a level 10 interrupt when the timer has counted down from a preset value to 0 (the count down is at a rate of one decrement every 3.2us). Timer 1 also counts down from a preset value but issues a level 14 interrupt and decrements at a rate of 0.8us. Both processors share timer 0 from one of the timer devices (to avoid the need for explicit synchronisation).

MBus Address	Usage	Read/ Write	Note
ff0701000	IRQ PAL 0 — Timer Latches	RW	5
ff0701002	IRQ PAL 0 — Mask Register Read/ Clear	RW	5
ff0701006	IRQ PAL 0 — Mask Register Set	RW	5
ff070100a	IRQ PAL 0 — Software Interrupt Reg Read / Clear	RW	5
ff070100e	IRQ PAL 0 — Software Interrupt Reg Set	W	5
ff0701010 to ff07011ff	Unused (echoes)		
ff0701200	Timer 0 Level10 Processor 0 and 1	RW	5
ff0701204	Timer 0 Level14 Processor 0	RW	5
ff0701208	Timer 0 Spare timer	RW	5
ff070120c	Timer 0 Control register	RW	5
ff0701210 to ff07015ff	Unused (echoes)		
ff0701600	Err and Powerfail Latch	RW	5
ff0701604 to ff0701fff	Unused (echoes)		
ff0702000	IRQ PAL 1 — Timer Latches	RW	5
ff0702002	IRQ PAL 1 — Mask Register Read/ Clear	RW	5
ff0702006	IRQ PAL 1 — Mask Register Set	RW	5

MBus Address	Usage	Read/ Write	Note
ff070200a	IRQ PAL 1 — Software Interrupt Reg Read / Clear	RW	5
ff070200e	IRQ PAL 1 — Software Interrupt Reg Set	W	5
ff0702010 to ff07021ff	Unused (echoes)		
ff0702200	Spare timer	RW	5
ff0702204	Timer 1 Level14 Processor 1	RW	5
ff0702208	Timer 1 Spare timer	RW	5
ff070220c	Timer 1 Control register	RW	5
ff0702210 to ff0702fff	Unused (echoes)		
ff0703000 to ff0703fff	Unused (Read Undefined)		
ff0704000 to ff07fffff	Unused (echoes of above)		

MBus to I/O Bus

The MBus to I/O Bus interface is via an LSI Logic L64851 device.

MBus Address	Usage	Read/ Write	Note
ff0800000	Software interrupts and enable.	RW	
ff0800004	Active level of external interrupts.	RW	
ff0800008	Programmable limit register 0.	RW	
ff080000c	Programmable limit register 1.	RW	
ff0800010	I/O bus devices available.	RW	
ff0800018	Latency delay register.	RW	

MBus Address	Usage	Read/ Write	Note
ff080001C	MBus Id register.	R	
ff0800020	Programmable timer.	R	
ff0800024	Programmable timer.	R	

MBus to SBus Chip, Elan, and MBus Slot Slaves

The MBus-to-SBus controller is an LSI Logic L64852C. The Elan Communications Processors is a Meiko device.

MBus Address	Usage	Read/ Write	Note
ff4ffff0	M2S Virtual Address Table Base Address	RW	
ff4ffff4	M2S IO/MMU Control register	RW	
ff4ffff8	M2S Error/Status register	R	
ff4ffffc	M2S — MBus ID Register	R	
ff5000000 to ff6f7ffff	Unused (MBus Timeout)		
ff6f80000 to ff6ffdfff	ELAN Command port area	RW	
ff6ffe000 to ff6ffffbf	ELAN Hush register area	RW	
ff6ffffc0	ELAN Clock Hi	RW	
ff6ffffc4	ELAN Clock Hi	R	
ff6ffffc8	ELAN Clock Lo	RW	
ff6ffffcc	ELAN Clock Lo	R	
ff6ffffd0	ELAN Alarm	RW	
ff6ffffd4	ELAN Alarm	R	

MBus Address	Usage	Read/ Write	Note
ff6ffffd8	ELAN Interrupt	R	
ff6ffffdc	ELAN Interrupt	R	
ff6ffffe0	ELAN Clock Hi	R	7
ff6ffffe4	ELAN Clock Lo (for 64bit accesses)	R	7
ff6ffffe8	ELAN Main Proc. Interrupt Mask	RW	
ff6ffffec	ELAN Main Proc. Interrupt Mask	R	
ff6fffff0	ELAN Control register	RW	
ff6fffff4	ELAN Control register	R	
ff6fffff8	MBus Port ID register for ELAN Chip	R	
ff6fffffc	MBus Port ID register for ELAN Chip	R	
ff7000000 to ff7ffffff	Unused (MBus timeout)		
ff8000000 to ff9ffffff	Used by MBus slave device in MBus Slot 0		
ffa000000 to ffbffffff	Used by MBus slave device in MBus Slot 1		
ffc000000 to fffffffff	Unused (MBus timeout)		

NVRAM Variables — B

The battery-backed RAM in the realtime clock module is used to used to store basic machine start-up and communication options.

These parameters may be queried using the Forth Monitor (i.e. at the `ok` prompt):

`printenv`	Display current variable settings.
`setenv` *variable value*	Assign (or reassign) a value to a variable.
`set-default` *variable*	Restore the variables default value.
`set-defaults`	Restore the default values to all variables.

For example:

```
ok setenv output-device can
```

Alternatively the System Administrator can use the `eeprom(1m)` command to view and change the variables direct from a Unix command shell. For example:

```
root@cs2# eeprom output-device=can
```

Some of the parameters (those marked in the following list) may also be modified using the Set function in Pandora's Network and Configuration Views.

Variable	Default	Description
`sbus-probe-list`	`43012`	Identifies the SBus slots to probe and the probe order.
`keyboard-click?`	`false`	If true, enable keyboard click.
`keymap`	*no default*	Name of custom keymap file.
`output-device` †	`screen`	Power-on output device. One of `screen`, `can`, `ttya`, or `ttyb`. Use `can` to enable console connections to be grabbed by `cancon(1m)` and Pandora.
`input-device` †	`keyboard`	Power-on input device. One of `keyboard`, `can`, `ttya`, or `ttyb`. Use `can` to enable console connections to be grabbed by `cancon(1m)` and Pandora.
`cancon-host`	`4294967295`	Used to record the host of the `cancon(1m)` remote console connection through a reboot of this processor. Do not change.
`elanip-broadcast-high` †	`4096`	Highest Elan Id in network.
`elanip-broadcast-low` †	`0`	Lowest Elan Id in network.
`ep-btxpktlifetime` †	`1000`	Elan packet characteristics.
`ep-btxtimeout` †	`1000`	Elan packet characteristics.
`ep-txpktlifetime` †	`10000`	Elan packet characteristics.
`ep-txtimeout` †	`10000`	Elan packet characteristics.
`ep-bigmsgbcastboxes` †	`4`	Elan packet characteristics.
`ep-bigmsgboxes` †	`32`	Elan packet characteristics.
`ep-bigmsgsize` †	`20416`	Elan packet characteristics.
`ep-smallmsgbcastboxes` †	`4`	Elan packet characteristics.
`ep-smallmsgboxes` †	`32`	Elan packet characteristics.
`ep-smallmsgsize` †	`4032`	Elan packet characteristics.
`elan-boot-id` †	`0`	Elan Id of node that this processor boots from.

Variable	Default	Description
`elan-node-id` †	0	Elan Id of this processor.
`elan-node-level` †	1	The processor's level in the CS-2 network.
`elan-num-levels` †	1	Number of levels in the CS-2 network.
`elan-top-switch` †	0	Specifies the level in the network that the processor sees it's topswitch. Usually this is level 0, the real top of the network.
`elan-switch-plane` †	0	Switch plane that this processor receives it boot code from when booting via the Elan network.
`ttyb-rts-dtr-off`	`false`	If true, Solaris does not assert RTS/DTR on ttyb.
`ttyb-ignore-cd`	`true`	If true, Solaris ignores carrier-detect on ttyb.
`ttya-rts-dtr-off`	`false`	If true, Solaris does not assert RTS/DTR on ttya.
`ttya-ignore-cd`	`true`	If true, Solaris ignores carrier-detect on ttya.
`ttyb-mode`	`9600,8,n,1,-`	ttyb (baud rate, #bits, parity, #stop, handshake). Baud rate is 110, 300, 1200, 2400, 4800, 9600, 19200, or 38400. #bits is 5, 6, 7, or 8. Parity is `n` (none), `e` (even), `o` (odd), `m` (mark), `s` (space). Handshake is – (none), `h` (hardware rts/cts), `s` (software).
`ttya-mode`	`9600,8,n,1,-`	ttyb (baud rate, #bits, parity, #stop, handshake). Baud rate is 110, 300, 1200, 2400, 4800, 9600, 19200, or 38400. #bits is 5, 6, 7, or 8. Parity is `n` (none), `e` (even), `o` (odd), `m` (mark), `s` (space). Handshake is – (none), `h` (hardware rts/cts), `s` (software).
`fcode-debug?`	`false`	If true, include name fields for plug-in device Fcodes.
`diag-file` †	`kadb`	The file and arguments to load from the root filesystem when the `diag-switch?` is true; otherwise use the `boot-file` parameter.
`diag-device` †	`elan`	Device to boot from when the `diag-switch?` is true; one of `disk`, `net`, or `elan`. Specify `elan` to boot over the CS-2 data network.

Variable	Default	Description
boot-file		The file and arguments to load from the root filesystem (e.g. kadb -v, or /kernel/unix -vr). No file implies /kernel/unix.
boot-device †	elan	Device to boot from; one of disk, net, or elan. Specify elan to boot over the CS-2 data network.
auto-boot? †	false	Boot automatically after power-on. Default value is true.
watchdog-reboot?	false	If true, reboot after watchdog reset.
local-mac-address?	false	If true, use the ethernet address taken from the local-mac-address parameter; otherwise use the IdPROM.
screen-#columns	80	Number of on-screen columns.
screen-#rows	34	Number of on-screen rows.
selftest-#megs	1	Megabytes of RAM to test on power-up or memory test.
scsi-initiator-id	7	SCSI bus address of host adapter, range 0–7.
cpu-#mhz	40	CPU clock rate.
use-nvramrc?	false	If true, execute the code stored in the nvramrc parameter when the boot ROM starts up.
nvramrc		Forth code to execute when the boot ROM starts-up (but only if use-nvramrc? is true).
sunmon-compat?	false	If true, come-up with old style prompt '>'.
security-mode	none	System security level for monitor commands; one of none, command, or full. None allows all commands to be executed. Command allows the continue and boot (without parameters) commands to be executed; others require a password. Full requires a password before any commands may be executed.
security-password	*no default*	The password used with security-mode described above.

Variable	Default	Description
`security-#badlogins`	*no default*	System set variable showing the number of times a bad password was specified.
`oem-logo`	*no default*	Byte array OEM logo (enabled by `oem-logo?`). Create a Forth array containing the logo and then copy into the `oem-logo` field.
`oem-logo?`	`false`	Enables OEM logo defined by `oem-logo`.
`oem-banner`	*no default*	Text displayed in the custom OEM banner alongside the OEM logo. Enabled by `oem-banner?`.
`oem-banner?`	`false`	Enables OEM banner text specified in `oem-banner`.
`hardware-revision`	*no default*	Hardware revision of this board (e.g. Rev D).
`last-hardware-update`	*no default*	Date of board's manufacture or last upgrade (e.g. 25May94).
`testarea`	`0`	Unused.
`mfg-switch?`	`false`	If true, perform repeated self tests.
`diag-switch?`	`false`	Run in diagnostic mode.

† These parameters may be changed using the Set function in Pandora's Network and Configuration Views.

Forth Monitor Commands C

The following commands have been added to the Forth Monitor and are in addition to the commands that are normally present on a Solaris system. The additional commands relate to the Control Network (CAN) and Elan network.

CAN Commands

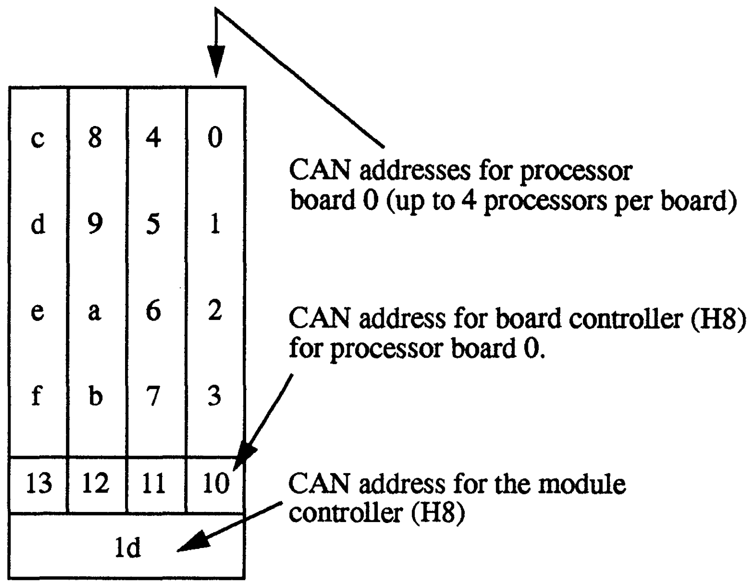
To test and use the CAN bus you need to understand CAN addresses.

Nodes are addressed by their physical position in terms of Cluster, Module, and Node id's (CMN). In CAN packets each of these id's is represented by a 6bit field; the hexadecimal representation of these three 6bit fields is a Node Id.

The module id is derived from the switch at the rear of the module. The numbering of the nodes within a module is shown in Figure C-1.

For example, the Node Id of processor with CMN 0:2:3 (processor 3, module 2, cluster 0) is 00083. The node's controlling H8 has the Node Id 00090. The node's module controller has the Node Id 0009d.

Figure C-1 CAN Addresses within a Module



Testing the CAN Device

Commands are provided to test the SPARC's CAN interface device, to test the board and module controllers, and to monitor activity on the CAN bus.

Testing the CAN Interface Device

The `test` command tests the SPARC's CAN device by writing various values into its test register. The test is repeated using the test registers on the board controller's H8 and the module controller's H8.

```
ok test /can
 Register test 0x00: OK
 Register test 0xff: OK
 Register test 0xaa: OK
 Register test 0x55: OK
 Checking on-board H8: OK.
 Checking module controller: OK.
```

Testing the CAN Bus

You can test the CAN bus connection between nodes by using the `rtest` command. In the following example data is transferred from the current node to node 4:

```
ok 4 rtest
Performing remote write/read test on node 4
Remote node type is MK405
.....................................................
Time taken was 13630mSecs
```

Checking the Board and Module Controllers

You can check that both the board controller and module controller H8 processors are running by using the `ping-h8` and `ping-module` commands. Note that you need to change directory to `/can` before you use these commands.

```
ok cd /can
ok ping-h8
On-Board H8 is MK401.
ok cd ..
```

```
ok cd /can
ok ping-module
Module controller is MK515.
ok cd ..
```

Querying CAN Bus Usage

You can query the utilisation of the CAN bus by using the `perf` command. This command shows the number of CAN packets received since the machine was powered-up, and the number since the last query. You need to change to the `/can` directory before using this command:

```
ok cd /can
ok perf
Total number of messages received since power-up: 259380
No. of messages received per second since the last check: 4
ok cd ..
```

Monitoring CAN Bus Packets

You can *snoop* the CAN bus (monitor that packets on the bus) using the `snoop` command. Note that you cannot use this facility if you are connected to the Forth Monitor via a cancon connection. You need to change to the `can` directory before using this command.

```
ok cd can
ok snoop
Can't can-snoop if you are a cancon slave
ok cd ..
```

CAN Addresses

To determine the CAN address of this node use `.can-id`. This displays the node's address in terms of its CMN, Node Id, and Slot Id. The Slot Id is for Meiko engineering use[1].

```
ok .can-id
SlotId: 0090, CAN Node-id: 00088 [00:02:08]
```

Similarly the CAN address of the board's controlling H8 processor can be obtained with the `.h8-id` command:

```
ok .h8-id
The on-board H8 is node 00092.
```

1. The slot id is the node's physical position in the machine represented by a 5 bit cluster number, a 5 bit module number, a 2 bit slot number, 2 unused (always 0 bits), and a 2 bit processor number; the 2 bits that represent the slot number are transposed.

The CAN address of the H8 that controls the board's module can be determined by the `.module-id` command:

```
ok .module-id
The module controller is node 0009d.
```

You can convert from CAN node id's to Cluster, Module, Node addresses (and vice versa) by using the `canid>cmn` and `cmn>canid` commands respectively. Note that you need to change directory to `can` before you use these commands.

```
ok cd can
ok 4 canid>cmn
0 0 4
ok cd ..
```

```
ok cd can
ok 0 2 8 cmn>canid
Node Id is 00088
ok cd ..
```

Querying CAN Objects

Can packets include a 10 bit address space which, although not sufficient to map into the MBus/H8 physical address space, is adequate to map-in various status and control devices. These are referred to as CAN objects. Reading or writing to these objects allows you to query the status of a processor, board, or module, and to issue control instructions. See the header file `/opt/MEIKOcs2/include/canio/canobj.h` for a list of object addresses and their meanings.

Local CAN objects are those that relate directly to this node. Remote CAN objects maybe those of a board, module controller, or remote SPARC.

You use the `rlo` command to read a local object. You need to pass an object id on the Forth stack; in the following example we request the board type and are returned 191 (an MK401):

```
ok 0 rlo
Read: 191
```

To read a remote object you need to push onto the Forth stack a CAN node id and the object id. In the following example we request the board type of node `c1` (module 3, board 0, node 2), which is an MK405:

```
ok c1 0 rro
Read: 195
```

The following additional example fetches the board type of node `dd`, which is the controller for module 3, cluster 0:

```
ok dd 0 rro
Read: 203
```

Similar commands exist to write to CAN objects, but their direct use is not recommended (they can reconfigure and reset the machine).

Remote Console Connections

You can create a console connection to a remote node by using `cancon`. You need to pass on the Forth stack a CAN node id.

You cannot create a cancon connection from within an existing cancon connection. If you are remotely interacting with a node's Forth monitor via cancon (or Pandora) an attempt to create another cancon connection will fail.

```
ok 4 cancon
Connected to node 00004

cs2-4 console login:
```

If your node is currently serving a remote console connection to someone else you can force it to disconnect that connection by using cancon-dis. In the following example the current connection to node 8 is dropped:

```
ok 8 cancon-dis
Disconnecting node 00008 [00:00:08] ...
```

Elan Commands

The Elan device includes self test code that can be executed by the test-all command (which tests memory, SBus, CAN, Elan and all other devices with self test code) or explicitly by the test /elan command.

```
ok test /elan
Initialising Elan/Selftest software ... OK
Checking threads processor ... OK

Testing from level 1 to level 1.
Generating a route to level 1 ... OK
Ping ... OK
Check-Ping ... OK
Spraying data to top switch ... OK
Testing spray buffer ... OK

Closing down Elan/Selftest software ... OK
```

Computing
Surface

Vector Processing Element (MK403) Users Guide

S1002-10M139.00

meiko

Meiko's address in the US is:

Meiko
130 Baker Avenue
Concord MA01742

508 371 0088
Fax: 508 371 7516

Meiko's address in the UK is:

Meiko Limited
650 Aztec West
Bristol
BS12 4SD

Tel: 01454 616171
Fax: 01454 618188

Issue Status:		
	Draft	
	Preliminary	
	Release	x
	Obsolete	

Circulation Control: *External*

Contents

Overview 1

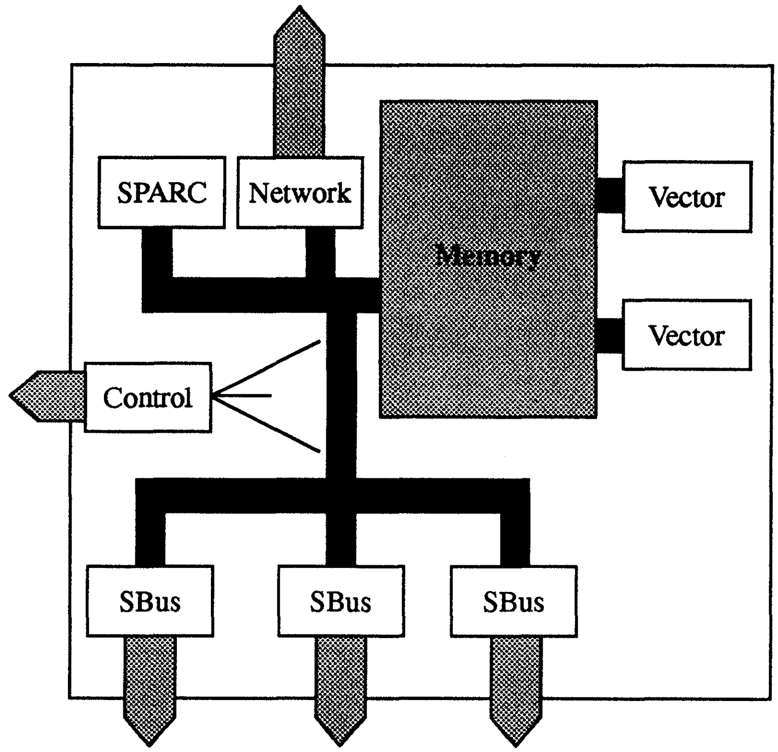
The MK403 Vector Processing Element offers high performance vector computing power and flexible I/O options.

The board design encompasses at the lowest level the principle design objectives for the CS-2, offering a scalable modular construction with easy upgrade options, a reliance on state of the art commodity components, leading edge proprietary network components, and support for system wide fault tolerance.

In outline the Vector Processing Element offers:

- Superscalar SPARC MBus module for scalar computing power and operating system services.
- Meiko Elan Communications Processor offering a high bandwidth, low latency interface to the CS-2 data network.
- Two Fujitsu Vector Processing Units (VPUs) on a plug-in module.
- Up to 512Mbytes of memory with 3 independent ports allowing simultaneous access by the SPARC MBus and the 2 Vector units.
- Three full size SBus slots for SCSI, Ethernet, or other third party options.
- Keyboard, mouse, and dual serial ports.
- Interface to the machine-wide control area network (CAN) offering remote diagnostic control and error logging facilities.

Figure 1-1 MK403 Board Overview



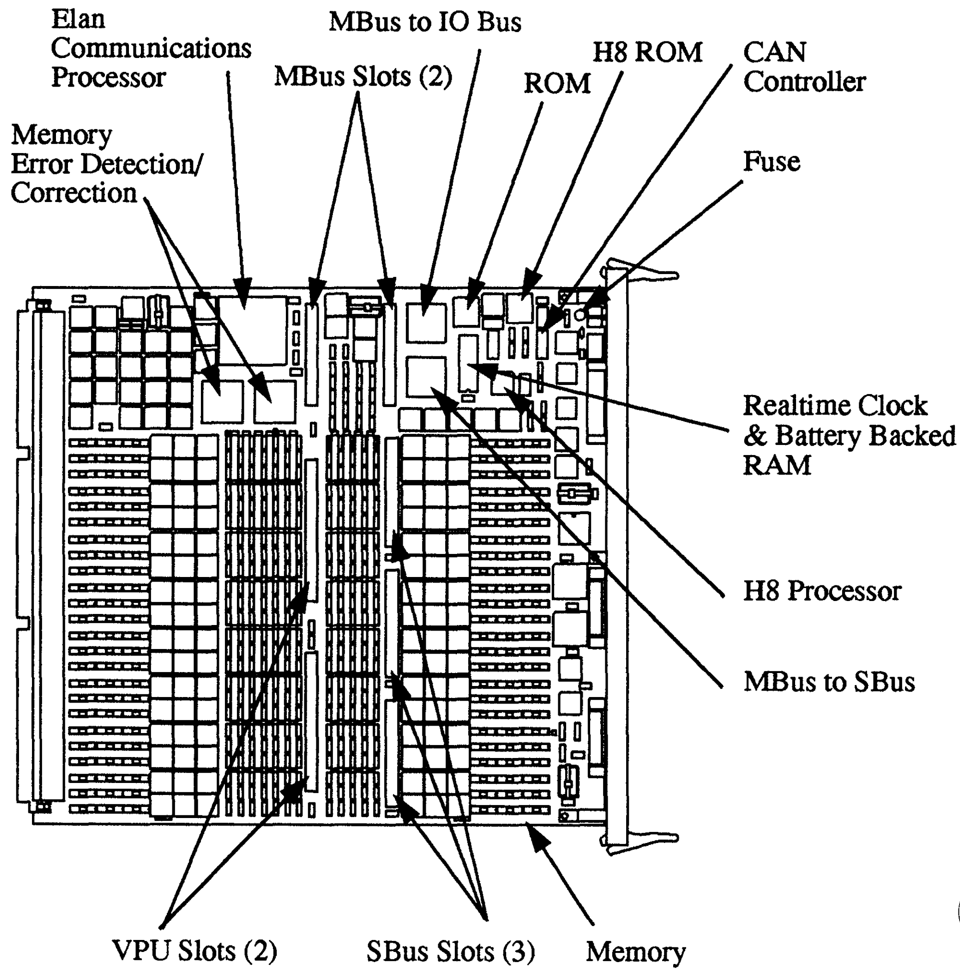
Access by the 3 processors to the memory system is via three independent memory ports. One port of the memory system is connected to an industry standard MBus interface giving access to the SPARC processor, the Elan Communications Processors, the MBus to SBus interface and its 3 SBus slots, and various other minor I/O devices on an I/O bus; in essence a SPARC workstation. Each of the remaining memory ports are used by the Fujitsu micro vector processors.

The 2 vector processing units on the MK403 are mounted on a single plug-in board which offers memory management and cache coherency with the SPARC caches. Cache coherency between the SPARC and vector processors has been achieved by defining a cache directory close to each vector processor; this cache directory attempts to replicate the contents of the SPARC cache, and is therefore potentially updated on every MBus cycle. Memory accesses by the vector processors that conflict with the SPARC cache cause the vector processor to stall until the appropriate MBus cycles have been generated.

Running throughout the whole CS-2 system is a control network (CAN) used to distribute status and configuration information, to provide remote control and diagnostics of all processors, and to create remote console connections to the processors. The MK403 has two interfaces to this network; one connected to the I/O bus (and thus providing a direct interface to the SPARC processors) and one via a dedicated micro-controller which provides board control.

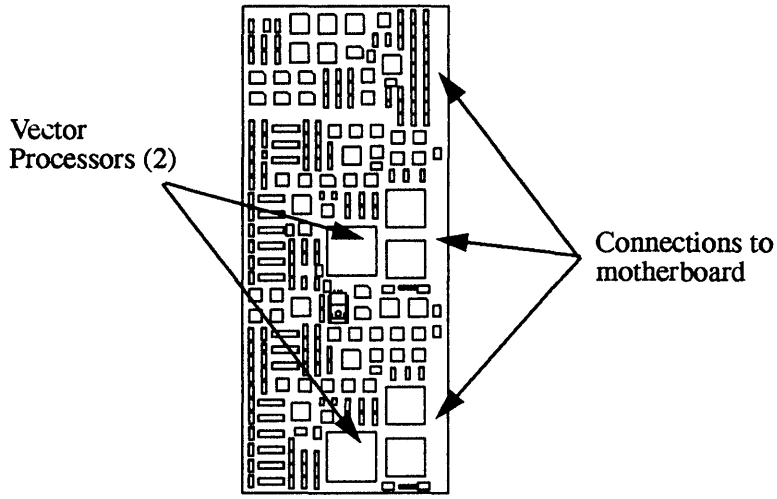
The major components and their placement on the MK403 motherboard are shown in Figure 2-1.

Figure 2-1 Board Schematic Showing Major Components (MK403)



MBus, SBus, and Vector Processor plug-in modules not shown.
Components are also fitted on the reverse of this board.

Figure 2-2 Board Schematic Showing Major Components (MK534)



MMU RAM and Cache RAM
fitted on reverse of board.

MBus

Two full size MBus sites are provided. One of these is used by the plug-in vector processors, the other by a uni-processor SPARC MBus module.

The MBus is fully level 2 compliant and runs at 40MHz. The SPARC processor shares the MBus with the Elan Communications Processor, the MBus-to-SBus interface, the I/O bus controller, and vector processors (but note that the vector board uses the MBus interface solely for memory management and cache coherency with the SPARC, and that direct memory accesses are made via memory ports that are independent of the MBus). The allocation of MBus id's is:

- MBus id 0 is the I/O bus controller.
- MBus id 4 is the MBus to SBus controller.
- MBus id 6 is the Elan Communications Processor.
- MBus id's 8 and 9 are MBus slot 0.
- MBus id's 10 and 11 are MBus slot 1.

MBus slot 0 is always used by the plug-in vector board. MBus slot 1 is always used by a uni-processor SPARC module, currently either a TI Viking (with or without second level cache) or a ROSS Pinnacle.

ROSS Pinnacle Module

The MK403 may be fitted with a uni-processor Pinnacle module which includes second level cache.

The Pinnacle MBus module is built upon a tightly coupled set of three ROSS devices: the RT620 HyperSPARC CPU, the RT625 cache controller, memory management, and tagging unit (CMTU), and the RT627 cache data units (CDUs).

Features of the RT620 CPU are:

- SPARC version 8 conformance.
- 90MHz clock rate.
- 4 execution units offering parallel execution of major instruction types: Load/Store, Branch/Call, integer and floating point units.
- Dual instruction fetch per clock cycle.
- 8Kbyte 2-way set-associative on-chip instruction cache.
- Instruction pipelining including a cache stage to accommodate the latency for second level cache accesses on data. Simultaneous accesses to on-chip and second level cache for each instruction fetch.
- High bandwidth 64bit Intra Module Bus (IMB) provides the interface between the CPU and the second level cache. Use of second level cache decouples the processor clock rate from the lower MBus clock rate.

Key features of the RT625 (CMTU) and RT627 (CDU) devices are:

- Full level 2 cache-coherent MBus compatibility.
- Each CDU has integral 16Kbytes x 32bit SRAM. MBus modules use either 2 or 4 CDUs for 128Kbyte or 256Kbyte second level direct-mapped cache.

- Physical cache tagging with virtual indexing allow the cache coherency logic to determine snoop hits and misses without stalling the CPU's access to the cache.
- Both copy-back and write-through cache modes supported.
- 32 byte read buffer and 64 byte write buffer for buffering the 32 byte cache lines in and out of the second level cache.
- SPARC reference MMU offering 64 entry, fully set-associative TLB with 4096 contexts.

Texas Instruments Viking Module

Two variants of the TI Viking MBus module are available. One contains a Viking SPARC processor with direct connection to the MBus. The second includes a Viking processor with additional Cache Controller and 1Mbyte of second level external cache (E-cache).

Key features of the TMS390Z50 SuperSPARC are:

- SPARC Version 8 conformance.
- 3 instructions per cycle, instruction pipelining, 150MIPs peak performance.
- SPARC Integer Unit.
- SPARC Reference MMU. Cached translation lookaside buffers (TLBs). 32bit virtual addresses, tagged with a 16bit context (65,536 contexts), map to 36bit physical addresses.
- Single and Double precision FPU. Tightly coupled to the integer execution pipeline and allowing one floating-point operation and one memory reference to be issued in each clock cycle. The FPU maintains a 4 entry FIFO queue for FP operations.
- 20Kbyte instruction cache, 16Kbyte data cache. The instruction cache is 5-way set associative, physically addressed, and non-writable. The data cache is 4-way associative and physically addressed. Both cache's are coherent with each other and with optional E-cache or MBus. Without E-cache the instruction and data caches operate in write-through mode; otherwise they are copy-back mode.

- Store buffer. A FIFO queue of 8 entries, each 64bits, decouples the instruction execution pipelines from the E-cache or MBus.
- Multiprocessor cache coherent support; highly pipelined and non-multiplexed VBus interface to optional external cache and the TMS390Z55 Cache Controller, or direct connection to the MBus.
- Prefetch buffer.
- Support for system and software debugging including hardware breakpoint.

The external cache is managed by the TMS390Z55 Cache Controller. Key features of this device are:

- Built in support for cache coherent multiprocessing; multiple Viking modules can share a single MBus and remain fully cache coherent.
- High performance VBus interface with the SPARC processor; decouples the SPARC processor from the MBus clock speed allowing higher processor performance. Reduced MBus traffic reduces contention when multiple processors share a single MBus.
- E-cache is direct mapped, copy-back, and unified: there is a single cache location where a particular byte of the physical address space can reside in the cache (direct mapping); writes by the main processor into the E-cache do not propagate to main memory until the cache is flushed or replaced (copy-back); both instructions and data are supplied to the processor from the same cache (unified).
- 1Mbyte SRAM cache.

Vector Processing Elements

Two vector processors are mounted on a plug-in board that is fitted to MBus slot 0 and the two VPU slots. Included on the vector board is the memory management and cache coherency logic and associated memory. The MBus interface is used by the memory management and cache coherency to remain synchronised with the SPARC processor, whereas direct memory accesses are via the dedicated VPU slots.

Memory Management

The vector processors operate with their address translation units disabled (Real Mode) and instead use an external memory management unit which allows better handling of address translation faults.

Each vector processor references a single-entry TLB and a 64K×16bit MMU RAM (128Kbyte). The MMU RAM contains 16K entries for each of 4 contexts, each entry consists of a 12 bit physical page number, 2 bit read/write flags, and 2 bit reference/modify flags. With a page size of 256Kbytes the 16K entries for each context gives a virtual address space of 4Gbytes. A separate register stores the current context and is used to provide the 2 high order bits into the MMU RAM. The MMU RAM is accessible from the MBus and is kept in-step with the SPARC's page tables by modified page-in code running in the operating system kernel. The single-entry TLB contains a single translation from logical page number to a physical page number, and a two bit reference/modify field.

Memory references by the vector processor are first checked against the TLB. If the logical page number matches, and the page has already been referenced (for a vector process read) or modified (for a vector process write), then the physical page provided by the TLB is used. If the TLB cannot provide the translation the logical page number and the context register are used as an index into the MMU RAM. If the corresponding entry in the MMU RAM has the sufficient access privileges, and the reference/modify bits are set appropriately, then the physical page is read from the MMU RAM and the TLB is updated with this entry. Translations that cannot be resolved by the MMU RAM (access permissions or missing pages) are notified to the SPARC by a level 5 interrupt, which may fault-in a new page or kill the vector process.

Cache Coherency

The vector processing board includes cache coherency logic which allows the vector processors (which have no cache capability themselves) to maintain consistency with the SPARC caches.

Each of the vector processors has an associated cache directory which is used to replicate the state of the SPARC's caches. Entries within the cache directory identify modified cache lines that are held by the SPARC; they are maintained by snooping the activity on the SPARC's MBus. Every memory address generated by the vector processor is checked against the cache directory, and if there is a conflict with the SPARC cache the appropriate MBus operations are generated (one of Invalidate, Coherent-Read and Writeback, or Coherent-Read-with-Invalidate followed by a Writeback).

The vector processor may be used alongside either the ROSS Pinnacle or Texas Instruments SPARC modules, and must therefore accommodate the different caches used by these modules. The TI caches are indexed by physical address, whereas the Pinnacle are indexed by virtual address. The TI E-cache and Pinnacle caches are direct mapped, whereas the TI integral cache is 4-way set associative. The configuration of the vector processors' cache directories is selected by hardware jumpers and is not field configurable.

When used with a TI Viking processor each cache directory is configured as 8Kbyte×10bits of cache tag information, 8Kbytes×8bits of shared and dirty information, and 8Kbytes×1bit of flags used to show that the access was initiated by a set-associative cache. Each cache block consists of 4 lines, each line consists of 32 bytes. Each line has its own shared and dirty bits, but the tag field is shared by the whole block. There is (8K blocks) × (4lines/block) × (32bytes/line) = 1Mbyte of cache represented by this configuration.

When used with the Pinnacle each cache directory is configured as 4Kbytes×10-bits of tag information, 4Kbytes×4bits of shared and dirty information, and 4Kbytes×1bit of flags used to show that the access was initiated by a set-associative cache. Each cache block consists of 2 lines, each line consists of 32bytes. Each line has its own shared and dirty bits, but the tag field is shared by the whole block. There is (4K blocks) × (2 lines/block) × (32bytes/line) = 256Kbytes of cache represented by this configuration.

While the vector processors are idle all MBus operations are snooped into the cache directories. While the vector processors are busy and accessing store all MBus operations are queued in an inbound buffer; the queue can only be serviced when the vector processors are idle, during compute time, or during recovery time between a LDA/STA and the start of the next one. In order to guarantee coherency the vector processors will not start a new series of operations until the queue has been emptied.

Fujitsu MB92831 Micro Vector Processor

The MK403 uses 2 Fujitsu micro vector processors (μVPs) in a co-processor shared memory configuration (the vector processors share the same memory as the SPARC but access it via independent memory ports).

Key features of the Fujitsu μVP are:

- Peak performance 100MFLOPS double precision, 200MFLOPS single precision per processor. External memory bandwidth 400Mbytes/s.
- Rich command set offering 252 vector commands, 57 scalar commands, 9 general control commands. Includes: command load and branch commands; vector load or store commands using stride values or indirect addressing ("scatter-gather"); floating-point, integer, and logical vector and scalar commands; maximum and minimum value search commands; a full range of vector and scalar comparison commands.
- Pipelined memory access allows the vector processor to output memory references before it needs them.
- Compliance with IEEE 754-1985 standard for binary floating-point arithmetic. 32bit and 64bit floating point data types.
- Pipeline execution allows the vector processor to perform the same operation repeatedly for all data items in a vector (or just once for a scalar). Parallel pipelines execute more than one command at a time. Chaining allows the operation in one pipeline to use the result of another.
- Six pipelines: Add (addition, subtraction, comparison, and data conversion), Multiply, Divide, Graphic, Mask, Load-Store.
- Host processor register interface. Commands are written to the 256 command buffer, and parameters to the scalar registers. Program execution is initiated by the start register, and polled for completion via the vector busy register. An abort register allows the host to halt the vector processor.
- 8Kbyte vector registers, 128byte scalar registers, and 64byte mask registers. Vector registers are partitioned into 4 banks, each bank can be read or written on each clock cycle. The banks may be accessed concurrently by 4 pipelines.

SBus Interfaces

The MBus to SBus interface supports up to 5 SBus devices, but on the MK403 only three are used and these map directly to the 3 SBus slots. SBus slot 0 is nearest the MBus processor slots.

The SBus runs at a clock speed of 20MHz.

Memory Configuration

The Superscalar SPARC processors and Elan communication processor are connected to a standard 40MHz MBus. The vector processors and MBus are connected to a 16 bank memory system, each bank providing 64 bits of user data (78 bits including error checking and correction, implemented using 20 by 4 bit DRAMs with two bits unused). Error detection and correction is implemented on each half word (32 bits), allowing write access to 32 bit (ANSI-IEEE 754–1985 `single`) values to be performed at full speed, without requiring a read modify write cycle.

Each bank of memory maintains a currently open DRAM page within which accesses may be performed at full speed. This corresponds to a size within the bank of 8 Kbytes, giving 128 Kbytes total for the 16 banks. When an access is required outside the currently open page a penalty of 6 cycles is incurred to close the previous page, and open the new one.

Refresh cycles are performed on all banks within a few clock cycles of each other, thus allowing the cost of re-opening the banks to be pipelined (since the VP can issue four addresses before stalling for the data from the first), and reducing the overhead of refresh to a few percent of memory bandwidth.

The memory system is clocked at the same speed as the μVP processors (50 MHz), and accesses from the 40 MHz MBus are transferred into the higher speed clock domain. When accessing within an open page each memory bank can accept a new address every two cycles (40ηs), and replies with the data four cycles (80ηs) later, giving a bandwidth of 8 bytes every two cycles (40ηs), that is 200 Mbytes/s. Since there are 16 banks, the total memory system bandwidth is thus 3.2 Gbytes/s.

Each μVP can issue a memory request every cycle (20ηs), and can issue 4 addresses before it requires data to be returned. In the absence of bank contention (which will be discussed below), after a start up latency of four cycles, these requests can be satisfied as fast as they are issued, giving each μVP a steady state bandwidth of 8 bytes every 20ηs, that is 400Mbytes/s.

Since each bank can accept a new address every two cycles (40ηs), but the μVP can generate an address every cycle (20ηs) there is the possibility of bank contention if the μVP generated repeated accesses to the same bank. With a simple linear mapping of addresses to banks, this would occur for all strides which are multiples of 16 (for 64bit double precision accesses). Such an access pattern would then see only one half of the normal bandwidth, that is 200Mbytes/s. All other strides achieve full bandwidth.

To ameliorate this problem as well as allowing the straightforward linear mapping of addresses to banks, Meiko also provide the option (through the choice of the physical addresses which are used to map the memory into user space) of scrambling the allocation of addresses to memory banks. The mapping function has been chosen to guarantee that accesses on "important" strides (1, 2, 4, 8, 16, 32) achieve full performance. Access on other strides may see reduced performance, but there are no strides within the open pages which see the pathological reduction to one half of the available bandwidth.

IO Bus

The IO bus is a slave-only bus used for the connection of minor peripherals to the MBus. The following devices are connected to this bus:

- A 512Kbyte EPROM holding bootstrap and diagnostic programs. The bootstrap code initialises the hardware devices, initial page table construction, and the booting of the Unix kernel.
- A realtime clock module with battery backed SRAM. This holds configuration information and node fault logs, including the log of uncorrectable memory errors. The realtime clock provides year, month, day, hour, minute, and second times.
- Dual DUART devices; one for connection of keyboard and mouse, the other for two general purpose serial ports. In the absence of a keyboard the bootstrap code in the EPROM will usually direct console I/O via serial port A. The serial

ports are clocked at 4.9152MHz, with a working capability of 38.4KBaud. Both serial ports share the same 25-way front panel connection; port A has full synchronous/asynchronous operation and a full complement of modem control lines; port B has a limited set of control lines and is asynchronous only.

- Interrupt controller; this uses registers to mask out certain types of interrupt to relieve the SPARC processor from unnecessary interrupt loading, and to share the handling between the SPARC and the Elan. The programming of the interrupt controllers is handled by the kernel device drivers.
- Periodic interrupt timer; used for maintaining the kernel clock and for kernel profiling.
- A CAN device provides the SPARC with an interface to the machine-wide control area network (CAN). The SPARC processor writes diagnostic information to this bus, and can also act as an X-CAN or G-CAN router. CAN routers transfer data between two levels of the CAN network; an X-CAN router handles transfers between the modules in a Cluster, and a G-CAN router handles transfers between Clusters. The configuration of a SPARC as a router will cause it's CAN device to generate numerous level 2 interrupts which will impact on processor performance.

Board Control Processor

The MK403 board uses an Hitachi H8/534 micro-controller (commonly referred to by Meiko as the H8) to perform basic node control functions. This controller is a single-chip 16bit RISC microcomputer with integral 2KBytes RAM, 32Kbytes EPROM, 16bit RISC CPU, and a number of I/O ports and timers.

The H8 processor runs independently of the other processors on the board and is used solely for control and diagnostic purposes. It has its own interface to the CAN bus via the second of the board's CAN interface devices. This processor receives diagnostic messages, via the CAN bus, from the local SPARC processor, and interprets incoming control messages, such as board reset, console connections, and network configuration.

Using the MK403 3

This chapter describes the usage of the MK403 in terms of its installation, hardware interfaces, and field serviceable components.

Installation

The MK403 is designed for use solely in a CS-2 Processor Module. The Processor Module supplies the board's power, cooling, and connection to the CS-2 data and control networks. The MK403 is fitted into one of the four vertical board slots behind the Processor Module's removable front panel.

Warning - You must disconnect the power from the Processor Module before removing or installing processor boards.

Warning - The board may be fitted with fragile or static sensitive devices. You must handle with care and observe anti-static precautions.

Removing the Module's Front Panel

The module's front panel is held in position by four clips, one in each corner. To remove the panel pull firmly away from the module.

The module's LED display is fitted to the module by two 50-way connectors. To remove the LED display pull firmly away from the module.

Use the reverse procedure to install the LEDs and front panel.

Installing the Processor Board

Insert the board so that it fits into the guide rails at the top and bottom of the module's board rack, ensuring that the component side is to the left (viewed facing the module). Gently push the board squarely on its front panel. Before pushing the board fully into position fold back the levers at each end of the front panel so that they are at 90^{o} to the board; now push the board (while holding the levers) until the base of the two levers is touching the card cage. To lever the board into its final position push both levers until they lie flat on the board's front panel. Secure the board by tightening the two captive screws.

Use the reverse procedure to remove the board.

Warning - You should take care not to damage the connectors at the rear of the board and on the module's backplane. Ensure that the board mates squarely with the module's backplane.

Warning - When removing or installing a board you should take care not to damage the RFI (copper) seals along the edge of the board's front panel.

Warning - To maintain proper circulation of cooling air and to conform to RFI regulations all board slots must be fitted with a processor board or blanking plate.

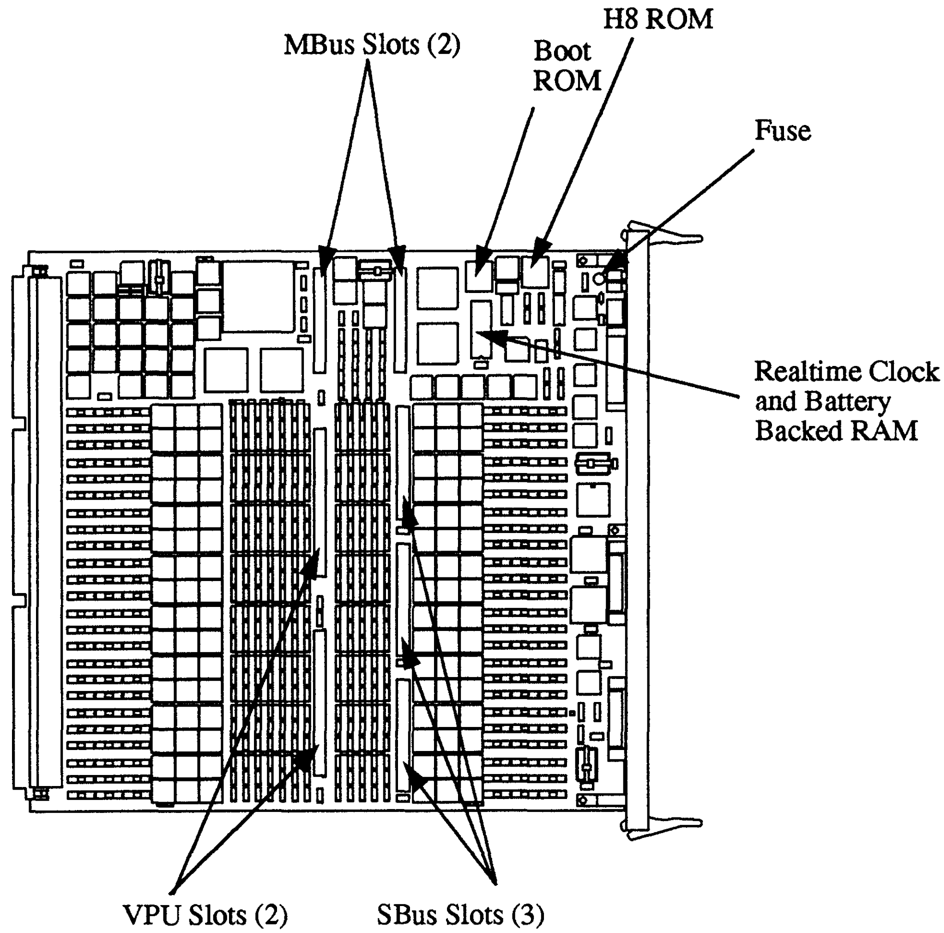
Field Serviceable Components

The MK403 has the following field upgradeable components (see Figure 3-1):

- Superscalar SPARC processor module fitted to MBus slot.
- Three SBus slots.
- Two vector processor slots.
- Boot ROM.

- H8 ROM.
- Realtime clock and non-volatile RAM module.
- Fuses.

Figure 3-1 MK403 Components

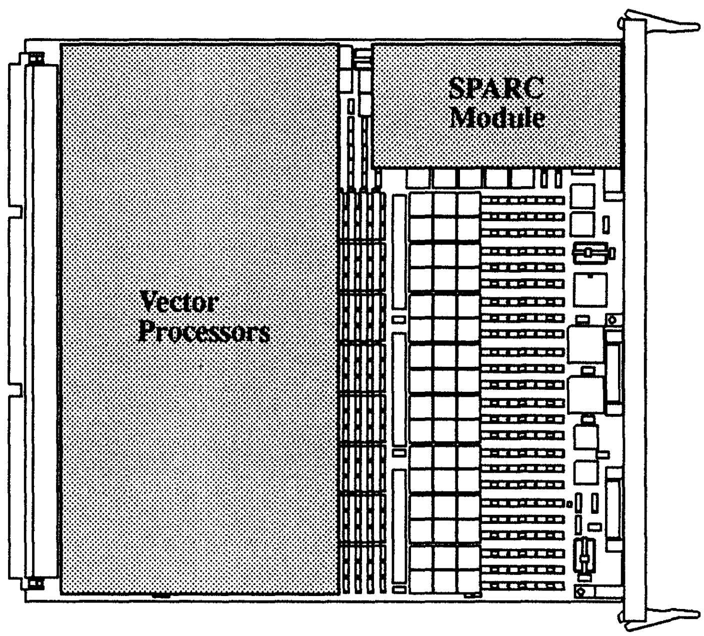


Processor Modules

The 2 vector processor are mounted on a single plug-in board (MK534) which is connected to MBus slot 0 and both of the VPU slots. The board is fixed in place by 4 M3 screws.

A single uni-processor SPARC module is connected to MBus slot 1. The MK403 motherboard includes support for either Texas Instruments Viking or ROSS Pinnacle module (although Pinnacle modules are typically used). The type of SPARC module is set when the MK403 board is manufactured and may not be reconfigured on site.

Figure 3-2 Position of the Vector and SPARC Modules

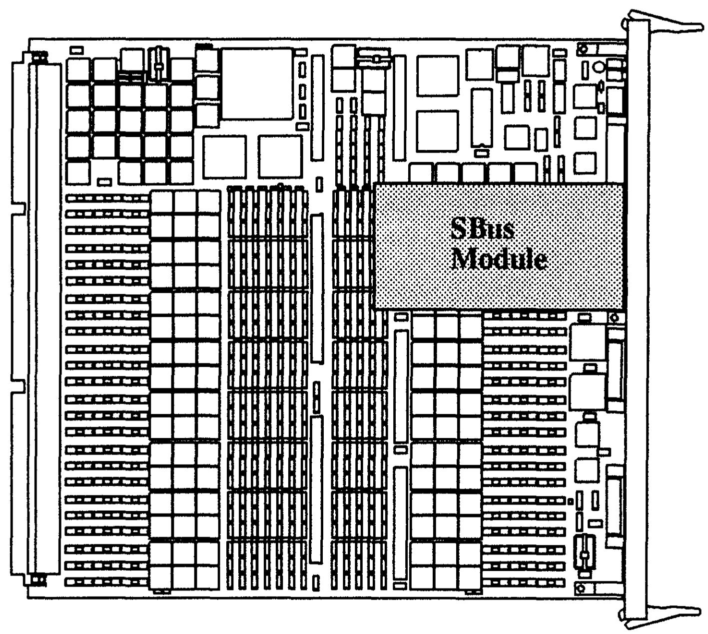


Installing SBus Modules

Three SBus slots are provided and these may be fitted with standard SBus modules; these are plugged into the SBus connectors and secured with two M3 screws. When using SBus cards that have external connections, for example a graphics card, remove the appropriate panel from the front of the MK403 — the panel is held in place by two small screws.

SBus devices are numbered from 0 to 2, device 0 being next to the processor slots (see Figure 3-3).

Figure 3-3 Position of SBus Module 0



SBus SCSI Cards

When using SBus SCSI cards to connect to disk devices within the Processor Module you must connect the front-panel output from the SBus card to the SCSI-A connector on the MK403 motherboard using Meiko cable 60-CA0217-1T.

You should note that the Processor Module's disks may be interconnected in one of three ways: each of the 4 disks connected to a separate processor board via SCSI bus A, disks connected in pairs to SCSI bus A on board's 0 and 1, or all disks connected to SCSI bus A on board 0.

Boot ROM and H8 ROM

Both of these ROMs may be upgraded from time to time. They are held in sockets and are readily replaced. Note the position on pin 1 before removing the old device (usually marked by a dot on the packaging).

Realtime Clock and Battery backed RAM

The real time clock and non-volatile RAM device is held in a DIL socket and is easily replaced. Before removing the old device note the position of pin 1 (usually marked by a dot on the packaging). Note that the information within the RAM can only be restored by Meiko's engineers.

Warning - This device contains lithium batteries; never dispose of this device in a fire or attempt to dismantle.

Fuses

There is one fuse on the MK403 motherboard to protect the keyboard/mouse circuit. This is a 250mA quick blow fuse, Meiko part number 22-FU400-02E250.

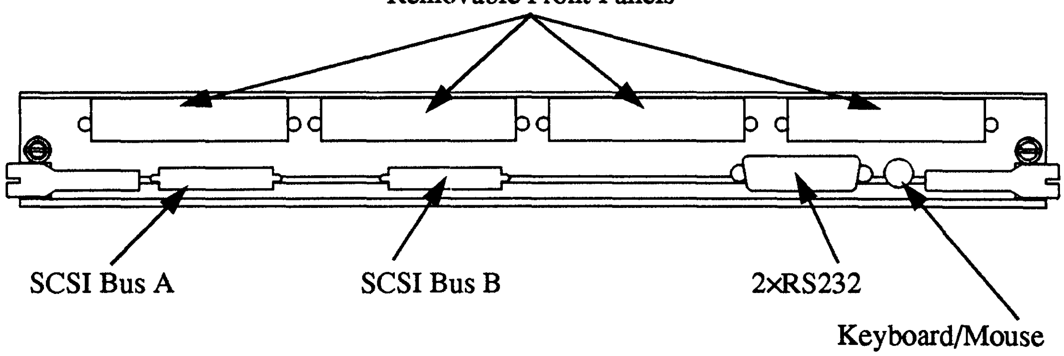
External Connections

External connections are provided for a keyboard/mouse (8 pin circular socket), serial interfaces (2 channels provided by one 25-way D-type connector), and two independent SCSI buses (each via a 50-way miniature connector).

Front Panel Connections

Removable panels provide access to connectors on the optional SBus boards.

Figure 3-4 MK403 Front Panel Connections



RS232 Connections

The two RS232 channels are output via a single 25-way connector. The connections are as shown in the following tables. Signal ground is pin 7, chassis ground in pin 1.

Table 3-1 RS232 Channel A Pinout.

Signal	Input/Output	Pin number
TXD	Out	2
RXD	In	3
RTS	Out	4
CTS	In	5
CSR	In	6
DCD	In	8
DB	In	15

Table 3-1 RS232 Channel A Pinout.

Signal	Input/Output	Pin number
DD	In	17
DA	On	24
DTR	On	20

Table 3-2 RS232 Channel B Pinout.

Signal	Input/Output	Pin number
TXD	Out	14
RXD	In	16
RTS	Out	19
CTS	In	13
DCDB	In	12

External Indicators

Two LEDs (one green, one amber) are included on the board's front panel. The green LED is the heart beat from the board's (H8) CAN controller. The amber light illuminates each time the CAN controller transmits on the CAN bus. Both should flash steadily. These indicators are also displayed on the module's LED display.

The green LED flashes at a slow steady rate (once per second) when operating normally. A quicker flash rate (2×normal) indicates that the board's SPARC processor is not responding; a very quick flash rate (3×normal) indicates that the H8 processor on the module's controller is not responding.

Each processor board within a processor module controls a 4×4 matrix of red LEDs on the module's front panel. The MK403 displays a random pattern on these when running the Boot ROM. When Solaris has been booted a circulating pattern is displayed. The pattern can be changed by user programs and various system commands and daemons.

Address Maps A

MBus Address Maps

This section gives the mapping of memory and peripherals into the MBus physical address space. All addresses are in hexadecimal, all locations are word wide unless otherwise stated in the notes. The following notes are associated with some of the items in the tables:

1. These locations are byte wide and are mapped into all 4 bytes of a word. Care should be taken to generate correct bytewide accesses to the least significant byte of the word in order to maintain future compatibility.
2. These locations are halfword wide and are mapped into both halfwords of the word. Care should be taken to generate correct halfword accesses to the least significant halfword of the word in order to maintain future compatibility.
3. These locations are bytewide memory, mapped into contiguous byte locations. Word or halfword accesses will be automatically mapped into several successive bytewide accesses.
4. These locations are byte sized registers which are only mapped into the least significant byte of the word. Accesses of larger than one byte will be mapped into several IO Bus transactions, but only one (the least significant byte access) will actually select the device. Halfword or word accesses will have no ill effects but will waste MBus bandwidth and should be avoided.

5. These locations are halfword sized registers which are only mapped into the least significant halfword of the word. Accesses of larger than one byte will be mapped into several byte transactions, but only two (the least and next-to-least significant byte) will actually select the device. Word accesses will have no ill effects but will waste MBus bandwidth and should be avoided.
6. These locations form a doubleword register.

MBus Address Map Summary

The following table summarises the MBus memory space usage by the board's principle components:

MBus Address	Usage
`000000000 to 007ffffff`	128MB Memory (normal mapping, no coherency between µVP and SPARC).
`020000000 to 027ffffff`	128MB Memory (scrambled mapping, enforce coherency between µVP and SPARC).
`200000000 to 207ffffff`	128MB Memory (normal mapping, enforce coherency between µVP and SPARC).
`220000000 to 227ffffff`	128MB Memory (scrambled mapping, no coherency between µVP and SPARC).
`100000000 to 10000ffff`	µVP 0 (Supervisor Access).
`100010000 to 10001ffff`	µVP 0 (User Access).
`120000000 to 12000ffff`	µVP 1 (Supervisor Access).
`120010000 to 12001ffff`	µVP 1 (User Access).
`160000000 to 16000ffff`	µVP Broadcast (Supervisor).
`160010000 to 16001ffff`	µVP Broadcast (User).
`1c0000000`	MBus EDC Error Data.

MBus Address	Usage
1c0000008	MBus EDC Error Diagnosis.
1c0000010	MBus EDC Clear.
800000000 to 81fffffff	μVP cache directory.
900000000 to 9A0000018	μVP status, MMU etc.
e00000000 to e0fffffff	SBus Slot 1.
e10000000 to e1fffffff	SBus Slot 2.
e20000000 to e2fffffff	SBus Slot 3.
e60000000 to e600001ff	MBus-to-SBus TLB.
e70000000 to effffffff	SBus Reserved.
ff0000000 to ff00fffff	Boot ROM.
ff0100000 to ff01fffff	Serial Port.
ff0200000 to ff02fffff	Keyboard and Mouse Port.
ff0300000 to ff03fffff	Real Time Clock and 8K SRAM.
ff0700000 to ff07007ff	Node Reset, IRQ.Pal's etc.
ff0700800 to ff0700fff	CAN.
ff0701000	IRQ 0 and Timer 0.
ff0701600	Async. Error Pending 0.
ff0701800	EDC Error Mask 0.

MBus Address	Usage
ff0701a00	SPARC Reset Flag.
ff0702000	IRQ 1 and Timer 1.
ff0702600	Async. Error Pending 1.
ff0702800	EDC Error Mask 1.
ff0703600	LED's.
ff0704000 to ff07042ff	µVP 0 Reset, Busy.
ff0705000 to ff07052ff	µVP 1 Reset, Busy.
ff0800000 to ff08fffff	STDIO Control Registers.
ff4ffffff0 to ff4ffffff	M2S Control Registers.
ff6f80000 to ff6ffffff	ELAN.
ff8000000 to ff9ffffff	MBus Slot 0.
ffa000000 to ffbffffff	MBus Slot 1.

DRAM and SBus Slots

MBus Address	Usage	Rd/Wr	Note
000000000 to 007ffffff	128MB Memory (Normal, no coherency between µVP and SPARC).	RW	
020000000 to 027ffffff	128MB Memory (Scrambled, enforce coherency between µVP and SPARC).	RW	
200000000 to 207ffffff	128MB Memory (Normal, enforce coherency between µVP and SPARC).	RW	
220000000 to 227ffffff	128MB Memory (Scrambled, no coherency between µVP and SPARC)	RW	

MBus Address	Usage	Rd/Wr	Note
`e0000000` to `e0ffffff`	SBus Slot 1.	RW	
`e1000000` to `e1ffffff`	SBus Slot 2.	RW	
`e2000000` to `e2ffffff`	SBus Slot 3.	RW	

µVP Cache Directory Mappings

The cache directories for both of the vector processor are mapped into the MBus address space.

MBus Address	Usage	Rd/Wr	Note
`8uxxy0000 to 8uxxzffc0` [a]	µVP cache directory (write both, read µVP0). x = don't care, u = (binary) `0xxx`, y = (binary) `xx00`, z= (binary) `xx11`. Bits 29,26:18 of address are used as data inputs to the RAMs, not as RAM index.	RW	
`8vxxy0000 to 8vxxzffc0`	µVP1 cache directory (read only). x = don't care, v = (binary) `1xxx`, y = (binary) `xx00`, z= (binary) `xx11`.	R	

a. The cache directory MMU mappings should be such that Logical Address equals Physical Address, or the cache directories should be accessed by bypassing the SPARC MMU. This allows the addresses to behave the same, regardless of whether we are configured as a Pinnacle or a Viking.

μVP Status and MMU Mappings

MBus Address	Usage	Rd/Wr	Note
900000000 to 90001fff8[a]	μVP MMU (write both, read μVP 0)	RW	
910000000 to 91001fff8	μVP MMU (read-only of μVP 1)	R	
920000000	MMU_ADDRESS0	R	
920000000	LD_PF_INBOUND (note same address as above!)	W	
930000000	MMU_ADRESS1	R	
930000000	LD_PF_OUTBOUND (note same address as above!)	W	
980000000	MMU_CONTROL	W	
980000008	MMU_MASK_FAULT	W	
980000010	MMU_CLEAR	W	
980000018	KILL_uVP	W	
980000028	CLEAR_INVALIDATE	W	
9A0000000	uVP Status	R	
9A0000010	SET_GATE_AND_SYNCH	W	
9A0000018	SET_GATE_AND_SEPARATE	W	

a. The MMU RAM mappings should be such that Logical Address equals Physical Address, or the MMU RAMS should be accessed by bypassing the SPARC MMU. This allows the addresses to behave the same, regardless of whether we are configured as a Pinnacle or a Viking.

BootRom, Serial Ports, Miscellaneous

MBus Address	Usage	Rd/Wr	Note
ff0000000 to ff007ffff	BootRom (512KByte).	R(W)	3
ff0040000 to ff007ffff	Unused for 256KByte Flash ROM (Boot ROM Echo)		
ff0100000 to ff010000f	Serial Port Controller		
ff0100000	Control Registers port B	RW	3
ff0100002	Data Buffer port B	RW	3
ff0100004	Control Registers port A	RW	3
ff0100006	Data Buffer port A	RW	3
ff0100010 to ff01fffff	Unused (Serial Port Echos)		
ff0200000 to ff020000f	Keyboard and Mouse Port Controller		
ff0200000	Control Registers mouse port	RW	3
ff0200002	Data Buffer mouse port	RW	3
ff0200004	Control Registers keyboard port	RW	3
ff0200006	Data Buffer keyboard port)	RW	3
ff0200010 to ff02fffff	Unused (Keyboard and Mouse Port Echos)		
ff0300000 to ff0301fff	Real Time Clock module and 8KByte SRAM	RW	3
ff0302000 to ff03fffff	Unused (RTC Echos)		
ff0400000 to ff06fffff	Unused (MBus Error)		
ff0700000	Node Reset Request	RW	4
ff0700004 to ff07001ff	Unused (Echos)		

MBus Address	Usage	Rd/Wr	Note
ff0700200	MBus Grant	R	4
ff0700204 to ff07003ff	Unused (Echos)		
ff0700400	Physical Slot Identifier	R	5
ff0700404 to ff07005ff	Unused (Echos)		
ff0703600	LED Bargraph	RW	5
ff0703604 to ff07007ff	Unused (Echos)		

Control Area Network Interface

MBus Address	Usage	Rd/Wr	Note
ff0700800	CAN - Control Register	RW	4
ff0700804	CAN - Command Register	W	4
ff0700808	CAN - Status Register	R	4
ff070080c	CAN - Interrupt Register	R	4
ff0700810	CAN - Acceptance Code Register	RW	4
ff0700814	CAN - Acceptance Mask Register	RW	4
ff0700818	CAN - Bus Timing Register 0	RW	4
ff070081c	CAN - Bus Timing Register 1	RW	4
ff0700820	CAN - Output Control Register	RW	4
ff0700824	CAN - Test Register		
ff0700828	CAN - TXBuf Identifier	RW	4
ff070082c	CAN - TXBuf RTR Data Length code	RW	4
ff0700830	CAN - TXBuf Data Byte 1	RW	4
ff0700834	CAN - TXBuf Data Byte 2	RW	4
ff0700838	CAN - TXBuf Data Byte 3	RW	4
ff070083c	CAN - TXBuf Data Byte 4	RW	4
ff0700840	CAN - TXBuf Data Byte 5	RW	4

MBus Address	Usage	Rd/Wr	Note
ff0700844	CAN - TXBuf Data Byte 6	RW	4
ff0700848	CAN - TXBuf Data Byte 7	RW	4
ff070084c	CAN - TXBuf Data Byte 8	RW	4
ff0700850	CAN - RXBuf Identifier	RW	4
ff0700854	CAN - RXBuf RTR Data Length code	RW	4
ff0700858	CAN - RXBuf Data Byte 1	RW	4
ff070085c	CAN - RXBuf Data Byte 2	RW	4
ff0700860	CAN - RXBuf Data Byte 3	RW	4
ff0700864	CAN - RXBuf Data Byte 4	RW	4
ff0700868	CAN - RXBuf Data Byte 5	RW	4
ff070086c	CAN - RXBuf Data Byte 6	RW	4
ff0700870	CAN - RXBuf Data Byte 7	RW	4
ff0700874	CAN - RXBuf Data Byte 8	RW	4
ff0700878	CAN - Unimplemented		
ff070087c	CAN - Clock Divider Register	RW	4
ff0700880 to ff0700fff	Unused (Echos of above)		

Interrupt Request Control and Status Registers

MBus Address	Usage	Rd/Wr	Note
ff0701000	IRQ pal 0 - Mask Register Read / Clear	RW	5
ff0701004	IRQ pal 0 - Mask Register Set	RW	5
ff0701008	IRQ pal 0 - Software Interrupt Reg Read / Clear	RW	5
ff070100c	IRQ pal 0 - Software Interrupt Reg Set	W	5
ff0701010 to ff07011ff	Unused (Echos)		
ff0701200	Timer 0 Level10	RW	4

MBus Address	Usage	Rd/Wr	Note
ff0701204	Timer 0 Level14	RW	4
ff0701208	Timer 0 Spare	RW	4
ff070120c	Timer 0 Control register	RW	4
ff0701210 to ff07013ff	Unused (Echos)		
ff0701400	CPU 0 Status and Watchdog Interrupt	RW	4
ff0701404 to ff07015ff	Unused (Echos)		
ff0701600	CPU 0 Async Error Pending	RW	4
ff0701800	CPU 0 EDC Mask	RW	4
ff0701604 to ff0701fff	Unused (Echos of above)		
ff0702000	IRQ pal 1 - Mask Register Read / Clear	RW	5
ff0702004	IRQ pal 1 - Mask Set	RW	5
ff0702008	IRQ pal 1 - Software Interrupt Reg Read / Clear	RW	5
ff070200c	IRQ pal 1 - Software Interrupt Reg Set	W	5
ff0702010 to ff07021ff	Unused (Echos)		
ff0702200	Timer 1 Level10	RW	4
ff0702204	Timer 1 Level14	RW	4
ff0702208	Timer 1 Spare	RW	4
ff070220c	Timer 1 Control register	RW	4
ff0702210 to ff07023ff	Unused (Echos)		
ff0702400	CPU 1 Status and Watchdog Interrupt	RW	4
ff0702404 to ff07025ff	Unused (Echos)		
ff0702600	CPU 1 Async Error Pending	RW	4

MBus Address	Usage	Rd/Wr	Note
ff0702800	CPU 1 EDC Mask	RW	4
ff0702604 to ff0702fff	Unused (Echos of above)		
ff0703000 to ff0703fff	Unused (Read Undefined)		
ff0704000 to ff07fffff	Unused (Echos of above)		

STDIO IO Bus Control Registers

MBus Address	Usage	Rd/Wr	Note
ff0800000	Software Interrupt/Enable	RW	
ff0800004	Active Interrupt Level	RW	
ff0800008	Limit Register 0	RW	
ff080000c	Limit Register 1	RW	
ff0800010	IOBus Devices Available	RW	
ff0800018	Latency Delay Register	RW	
ff080001c	MBusID Register	R	
ff0800020	Timer 0	R	
ff0800024	Timer 1	R	

MBus to SBus, Elan, and MBus Slot Slaves

MBus Address	Usage	Rd/Wr	Note
ff4ffffff0	M2S Virtual Address Table Base Address	RW	
ff4ffffff4	M2S IO/MMU Control register	RW	
ff4ffffff8	M2S Error/Status register	R	
ff4ffffffc	M2S - MBus ID Register	R	
ff5000000 to ff6f7ffff	Unused (MBus Timeout)		

MBus Address	Usage	Rd/Wr	Note
ff6f80000 to ff6ffdfff	ELAN Command port area	RW	
ff6ffe000 to ff6ffffbf	ELAN Hush register area	RW	
ff6ffffc0	ELAN Clock Hi	RW	
ff6ffffc4	ELAN Clock Hi	R	
ff6ffffc8	ELAN Clock Lo	RW	
ff6ffffcc	ELAN Clock Lo	R	
ff6ffffd0	ELAN Alarm	RW	
ff6ffffd4	ELAN Alarm	R	
ff6ffffd8	ELAN Interrupt	R	
ff6ffffdc	ELAN Interrupt	R	
ff6ffffe0	ELAN Clock Hi	R	6
ff6ffffe4	ELAN Clock Lo (For 64-bit accesses)	R	6
ff6ffffe8	ELAN Main Proc. Interrupt Mask	RW	
ff6ffffec	ELAN Main Proc. Interrupt Mask	R	
ff6fffff0	ELAN Control register	RW	
ff6fffff4	ELAN Control register	R	
ff6fffff8	MBus Port ID register for ELAN Chip	R	
ff6fffffc	MBus Port ID register for ELAN Chip	R	
ff7000000 to ff7ffffff	Unused (MBus timeout)		
ff8000000 to ff9ffffff	Used by MBus slave device in MBus Slot 0		
ffa000000 to ffbffffff	Used by MBus slave device in MBus Slot 1		
ffc000000 to fffffffff	Unused (MBus timeout)		

Vector Processors

Each of the two μVP's is mapped into the MBus address space twice, once for supervisor access and once for user access. There are also registers on the IO Bus to control the hardware reset line of each μVP, and to read the hardware busy signal. Only the least significant bit of these registers is used.

MBus Address	Usage	Rd/Wr	Note
`100000000 to 10000ffff`	μVP 0 (Supervisor)	RW	
`100010000 to 10001ffff`	μVP 0 (User)	RW	
`120000000 to 12000ffff`	μVP 1 (Supervisor)	RW	
`120010000 to 12001ffff`	μVP 1 (User)	RW	
`160000000 to 16000ffff`	μVP Broadcast (Super)	W	
`160010000 to 16001ffff`	μVP Broadcast (User)	W	
`9A0000010`	SET_GATE_AND_SYNCH	R	
`9A0000018`	SET_GATE_AND_SEPARATE	R	
`9A0000020`	CLEAR_GATES		
`ff0704000`	μVP 0 Reset	W	4
`ff0704200`	μVP 0 Busy	R	4
`ff0705000`	μVP 1 Reset	W	4
`ff0705200`	μVP 1 Busy	R	4

Within each of these images, the registers can be accessed by aligned 32-bit or 64-bit accesses.

Address Offset	Usage	Rd/Wr	Note
0000 to 03ff	VCB Command Buffer	RW	
0800 to 087f	VSR Scalar Registers	RW	
0a00 to 0bff	VTR Translation Registers	RW(s)	
0e00	VACNT Address Control	RW(s)	
0e08	VMD Mode	RW	
0e10	VLEN Vector Length	RW	
0e14	VCLEN Command Length	RW	
0e18	VSTA Start	RW	
0e20	VCINF Comparison Info	R	
0e28	VSTS Status	RW	
0e30	VEXB Exception Buffer	R	
0e38	VEXA Exception Address	R	
0e44	VABT Abort	W	
0e4c	VBSY Busy	R	

uVP Address Map

µVP Address	Usage
00000000 to 07ffffff	128MB Memory (Normal)
20000000 to 27ffffff	128MB Memory (Scrambled)
40000000 to 47ffffff	128MB Memory (Uncorrected)
60000000 to 67ffffff	128MB Memory (Scrambled Checkbit)

μVP Address	Usage
80000000	μVP EDC Error Data
80000008	μVP EDC Error Diagnosis
80000010	μVP EDC Clear

NVRAM Variables **B**

The battery-backed RAM in the realtime clock module is used to used to store basic machine start-up and communication options.

These parameters may be queried using the Forth Monitor (i.e. at the `ok` prompt):

`printenv`	Display current variable settings.
`setenv` *variable value*	Assign (or reassign) a value to a variable.
`set-default` *variable*	Restore the variables default value.
`set-defaults`	Restore the default values to all variables.

For example:

```
ok setenv output-device can
```

Alternatively the System Administrator can use the `eeprom(1m)` command to view and change the variables direct from a Unix command shell. For example:

```
root@cs2# eeprom output-device=can
```

Some of the parameters (those marked in the following list) may also be modified using the Set function in Pandora's Network and Configuration Views.

Variable	Default	Description
sbus-probe-list	43012	Identifies the SBus slots to probe and the probe order.
keyboard-click?	false	If true, enable keyboard click.
keymap	*no default*	Name of custom keymap file.
output-device †	screen	Power-on output device. One of screen, can, ttya, or ttyb. Use can to enable console connections to be grabbed by cancon(1m) and Pandora.
input-device †	keyboard	Power-on input device. One of keyboard, can, ttya, or ttyb. Use can to enable console connections to be grabbed by cancon(1m) and Pandora.
cancon-host	4294967295	Used to record the host of the cancon(1m) remote console connection through a reboot of this processor. Do not change.
elanip-broadcast-high †	4096	Highest Elan Id in network.
elanip-broadcast-low †	0	Lowest Elan Id in network.
ep-btxpktlifetime †	1000	Elan packet characteristics.
ep-btxtimeout †	1000	Elan packet characteristics.
ep-txpktlifetime †	10000	Elan packet characteristics.
ep-txtimeout †	10000	Elan packet characteristics.
ep-bigmsgbcastboxes †	4	Elan packet characteristics.
ep-bigmsgboxes †	32	Elan packet characteristics.
ep-bigmsgsize †	20416	Elan packet characteristics.
ep-smallmsgbcastboxes †	4	Elan packet characteristics.
ep-smallmsgboxes †	32	Elan packet characteristics.
ep-smallmsgsize †	4032	Elan packet characteristics.
elan-boot-id †	0	Elan Id of node that this processor boots from.

Variable	Default	Description
`elan-node-id` †	`0`	Elan Id of this processor.
`elan-node-level` †	`1`	The processor's level in the CS-2 network.
`elan-num-levels` †	`1`	Number of levels in the CS-2 network.
`elan-top-switch` †	`0`	Specifies the level in the network that the processor sees it's topswitch. Usually this is level 0, the real top of the network.
`elan-switch-plane` †	`0`	Switch plane that this processor receives it boot code from when booting via the Elan network.
`ttyb-rts-dtr-off`	`false`	If true, Solaris does not assert RTS/DTR on ttyb.
`ttyb-ignore-cd`	`true`	If true, Solaris ignores carrier-detect on ttyb.
`ttya-rts-dtr-off`	`false`	If true, Solaris does not assert RTS/DTR on ttya.
`ttya-ignore-cd`	`true`	If true, Solaris ignores carrier-detect on ttya.
`ttyb-mode`	`9600,8,n,1,-`	ttyb (baud rate, #bits, parity, #stop, handshake). Baud rate is 110, 300, 1200, 2400, 4800, 9600, 19200, or 38400. #bits is 5, 6, 7, or 8. Parity is `n` (none), `e` (even), `o` (odd), `m` (mark), `s` (space). Handshake is – (none), `h` (hardware rts/cts), `s` (software).
`ttya-mode`	`9600,8,n,1,-`	ttyb (baud rate, #bits, parity, #stop, handshake). Baud rate is 110, 300, 1200, 2400, 4800, 9600, 19200, or 38400. #bits is 5, 6, 7, or 8. Parity is `n` (none), `e` (even), `o` (odd), `m` (mark), `s` (space). Handshake is – (none), `h` (hardware rts/cts), `s` (software).
`fcode-debug?`	`false`	If true, include name fields for plug-in device Fcodes.
`diag-file` †	`kadb`	The file and arguments to load from the root filesystem when the `diag-switch?` is true; otherwise use the `boot-file` parameter.
`diag-device` †	`elan`	Device to boot from when the `diag-switch?` is true; one of `disk`, `net`, or `elan`. Specify `elan` to boot over the CS-2 data network.

Variable	Default	Description
`boot-file`		The file and arguments to load from the root filesystem (e.g. `kadb -v`, or `/kernel/unix -vr`). No file implies `/kernel/unix`.
`boot-device` †	`elan`	Device to boot from; one of `disk`, `net`, or `elan`. Specify `elan` to boot over the CS-2 data network.
`auto-boot?` †	`false`	Boot automatically after power-on. Default value is true.
`watchdog-reboot?`	`false`	If true, reboot after watchdog reset.
`local-mac-address?`	`false`	If true, use the ethernet address taken from the `local-mac-address` parameter; otherwise use the IdPROM.
`screen-#columns`	`80`	Number of on-screen columns.
`screen-#rows`	`34`	Number of on-screen rows.
`selftest-#megs`	`1`	Megabytes of RAM to test on power-up or memory test.
`scsi-initiator-id`	`7`	SCSI bus address of host adapter, range 0–7.
`cpu-#mhz`	`40`	CPU clock rate.
`use-nvramrc?`	`false`	If true, execute the code stored in the nvramrc parameter when the boot ROM starts up.
`nvramrc`		Forth code to execute when the boot ROM starts-up (but only if `use-nvramrc?` is true).
`sunmon-compat?`	`false`	If true, come-up with old style prompt '>'.
`security-mode`	`none`	System security level for monitor commands; one of `none`, `command`, or `full`. None allows all commands to be executed. Command allows the continue and boot (without parameters) commands to be executed; others require a password. Full requires a password before any commands may be executed.
`security-password`	*no default*	The password used with security-mode described above.

Variable	Default	Description
`security-#badlogins`	*no default*	System set variable showing the number of times a bad password was specified.
`oem-logo`	*no default*	Byte array OEM logo (enabled by `oem-logo?`). Create a Forth array containing the logo and then copy into the `oem-logo` field.
`oem-logo?`	`false`	Enables OEM logo defined by `oem-logo`.
`oem-banner`	*no default*	Text displayed in the custom OEM banner alongside the OEM logo. Enabled by `oem-banner?`.
`oem-banner?`	`false`	Enables OEM banner text specified in `oem-banner`.
`hardware-revision`	*no default*	Hardware revision of this board (e.g. Rev D).
`last-hardware-update`	*no default*	Date of board's manufacture or last upgrade (e.g. 25May94).
`testarea`‡		Set bit 1 to split the memory system (i.e memory test only sees half of what is available). Other bits indicate VPU silicon revision, board revision etc. and are not documented.
`mfg-switch?`	`false`	If true, perform repeated self tests.
`diag-switch?`	`false`	Run in diagnostic mode.
`meiko-sbus-slot`‡	*no default*	Identifies the SBus slot that will be used as the boot device (fitted with either a SCSI or Ethernet board). "net" is aliased to slot 2; net0, net1, net2 exist for booting off an explicit slot.
`meiko-sbus-use`‡	*no default*	Used with `meiko-sbus-slot` (above) to identify the device type; one of 0 (none), 1 (Ethernet), 2 (SCSI), or 3(Ethernet and SCSI).

† These parameters may be changed using the Set function in Pandora's Network and Configuration Views.

‡ The definition of these variables is unique to the MK403 Vector Processing Element (all others are the same for all CS-2 boards).

Forth Monitor Commands C

The following commands have been added to the Forth Monitor and are in addition to the commands that are normally present on a Solaris system. The additional commands relate to the Control Network (CAN) and Elan network.

CAN Commands

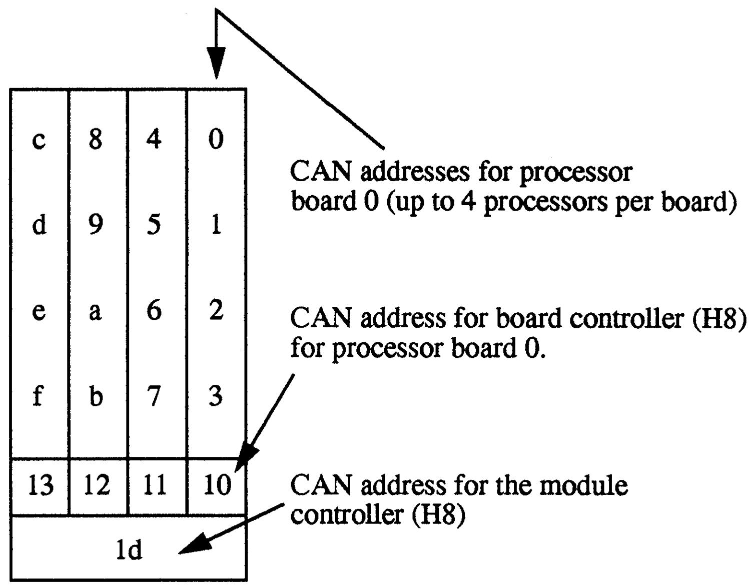
To test and use the CAN bus you need to understand CAN addresses.

Nodes are addressed by their physical position in terms of Cluster, Module, and Node id's (CMN). In CAN packets each of these id's is represented by a 6bit field; the hexadecimal representation of these three 6bit fields is a Node Id.

The module id is derived from the switch at the rear of the module. The numbering of the nodes within a module is shown in Figure C-1.

For example, the Node Id of processor with CMN 0:2:3 (processor 3, module 2, cluster 0) is 00083. The node's controlling H8 has the Node Id 00090. The node's module controller has the Node Id 0009d.

Figure C-1 CAN Addresses within a Module



Testing the CAN Device

Commands are provided to test the SPARC's CAN interface device, to test the board and module controllers, and to monitor activity on the CAN bus.

Testing the CAN Interface Device

The `test` command tests the SPARC's CAN device by writing various values into its test register. The test is repeated using the test registers on the board controller's H8 and the module controller's H8.

```
ok test /can
 Register test 0x00: OK
 Register test 0xff: OK
 Register test 0xaa: OK
 Register test 0x55: OK
 Checking on-board H8: OK.
 Checking module controller: OK.
```

Testing the CAN Bus

You can test the CAN bus connection between nodes by using the `rtest` command. In the following example data is transferred from the current node to node 4:

```
ok 4 rtest
Performing remote write/read test on node 4
Remote node type is MK405
.....................................................
Time taken was 13630mSecs
```

Checking the Board and Module Controllers

You can check that both the board controller and module controller H8 processors are running by using the `ping-h8` and `ping-module` commands. Note that you need to change directory to `/can` before you use these commands.

```
ok cd /can
ok ping-h8
On-Board H8 is MK401.
ok cd ..
```

```
ok cd /can
ok ping-module
Module controller is MK515.
ok cd ..
```

Querying CAN Bus Usage

You can query the utilisation of the CAN bus by using the `perf` command. This command shows the number of CAN packets received since the machine was powered-up, and the number since the last query. You need to change to the `/can` directory before using this command:

```
ok cd /can
ok perf
Total number of messages received since power-up: 259380
No. of messages received per second since the last check: 4
ok cd ..
```

Monitoring CAN Bus Packets

You can *snoop* the CAN bus (monitor that packets on the bus) using the `snoop` command. Note that you cannot use this facility if you are connected to the Forth Monitor via a cancon connection. You need to change to the `can` directory before using this command.

```
ok cd can
ok snoop
Can't can-snoop if you are a cancon slave
ok cd ..
```

CAN Addresses

To determine the CAN address of this node use `.can-id`. This displays the node's address in terms of its CMN, Node Id, and Slot Id. The Slot Id is for Meiko engineering use[1].

```
ok .can-id
SlotId: 0090, CAN Node-id: 00088 [00:02:08]
```

Similarly the CAN address of the board's controlling H8 processor can be obtained with the `.h8-id` command:

```
ok .h8-id
The on-board H8 is node 00092.
```

1. The slot id is the node's physical position in the machine represented by a 5 bit cluster number, a 5 bit module number, a 2 bit slot number, 2 unused (always 0 bits), and a 2 bit processor number; the 2 bits that represent the slot number are transposed.

The CAN address of the H8 that controls the board's module can be determined by the `.module-id` command:

```
ok .module-id
The module controller is node 0009d.
```

You can convert from CAN node id's to Cluster, Module, Node addresses (and vice versa) by using the `canid>cmn` and `cmn>canid` commands respectively. Note that you need to change directory to `can` before you use these commands.

```
ok cd can
ok 4 canid>cmn
0 0 4
ok cd ..
```

```
ok cd can
ok 0 2 8 cmn>canid
Node Id is 00088
ok cd ..
```

Querying CAN Objects

Can packets include a 10 bit address space which, although not sufficient to map into the MBus/H8 physical address space, is adequate to map-in various status and control devices. These are referred to as CAN objects. Reading or writing to these objects allows you to query the status of a processor, board, or module, and to issue control instructions. See the header file `/opt/MEIKOcs2/include/canio/canobj.h` for a list of object addresses and their meanings.

Local CAN objects are those that relate directly to this node. Remote CAN objects maybe those of a board, module controller, or remote SPARC.

You use the `rlo` command to read a local object. You need to pass an object id on the Forth stack; in the following example we request the board type and are returned 191 (an MK401):

```
ok 0 rlo
Read: 191
```

To read a remote object you need to push onto the Forth stack a CAN node id and the object id. In the following example we request the board type of node `c1` (module 3, board 0, node 2), which is an MK405:

```
ok c1 0 rro
Read: 195
```

The following additional example fetches the board type of node `dd`, which is the controller for module 3, cluster 0:

```
ok dd 0 rro
Read: 203
```

Similar commands exist to write to CAN objects, but their direct use is not recommended (they can reconfigure and reset the machine).

Remote Console Connections

You can create a console connection to a remote node by using `cancon`. You need to pass on the Forth stack a CAN node id.

You cannot create a cancon connection from within an existing cancon connection. If you are remotely interacting with a node's Forth monitor via cancon (or Pandora) an attempt to create another cancon connection will fail.

```
ok 4 cancon
Connected to node 00004

cs2-4 console login:
```

If your node is currently serving a remote console connection to someone else you can force it to disconnect that connection by using cancon-dis. In the following example the current connection to node 8 is dropped:

```
ok 8 cancon-dis
Disconnecting node 00008 [00:00:08] ...
```

Elan Commands

The Elan device includes self test code that can be executed by the test-all command (which tests memory, SBus, CAN, Elan and all other devices with self test code) or explicitly by the test /elan command.

```
ok test /elan
Initialising Elan/Selftest software ... OK
Checking threads processor ... OK

Testing from level 1 to level 1.
Generating a route to level 1 ... OK
Ping ... OK
Check-Ping ... OK
Spraying data to top switch ... OK
Testing spray buffer ... OK

Closing down Elan/Selftest software ... OK
```

Computing
Surface

Quad SPARC/Compute Processing Elements (MK405) Users Guide

S1002-10M138.00

meiko

Meiko's address in the US is:

Meiko
130 Baker Avenue
Concord MA01742

508 371 0088
Fax: 508 371 7516

Meiko's address in the UK is:

Meiko Limited
650 Aztec West
Bristol
BS12 4SD

Tel: 01454 616171
Fax: 01454 618188

Issue Status:		
	Draft	
	Preliminary	
	Release	x
	Obsolete	

Circulation Control: *External*

Contents

Overview 1

The MK405 SuperSPARC/Compute board offers a scalar multi-processor capability for computationally intensive parallel workloads.

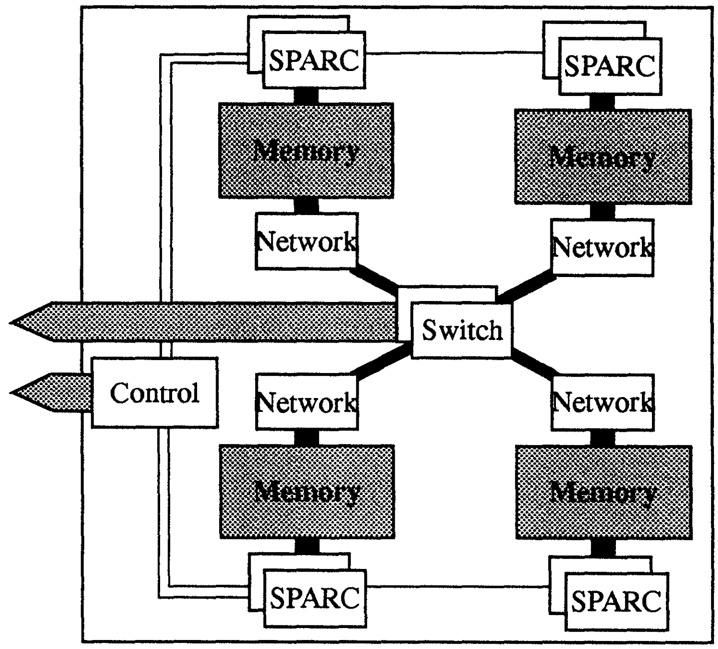
The board design encompasses at the lowest level the principle design objectives for the CS-2, offering a scalable modular construction with easy upgrade options, a reliance on state of the art commodity components, leading edge proprietary network components, and support for system wide fault tolerance.

In outline the Superscalar SPARC/Compute board offers:

- Four independent MBus interfaces, each taking a uni-processor or dual processor SPARC processor module. Easy upgrade to next generation technology protects investment and extends system life.
- Two Meiko Elan Communication processors, one per MBus.
- Up to 512Mbytes of field configurable memory (up to 128Mbytes per MBus).
- Two Elite network switches provide first level switch network.
- Fully compliant with the SPARC Compliance Definition (SCD); applications can be ported between SCD environments without change.

This densely packed compute board has connections to the CS-2 data network but no other I/O capability; disk and other I/O services must be provided by other board types.

Figure 1-1 Board Overview



MK405 Board Description 2

The MK405 SuperSPARC/Compute board includes four independent processing elements, each with their own processor, memory, and connection to the CS-2 data network. The interconnection of the devices in each processing element is via the industry standard SPARC MBus architecture using a 40MHz, level 2 cache coherent implementation. Interfacing the MBus is an I/O bus, which connect minor devices (such as the real time clock).

Running throughout the whole CS-2 is a control network (CAN) used to distribute status and configuration information, to provide remote control and diagnostics of all processors, and to create console connections to the processors. The MK405 has 5 interfaces to this network; one connected to each of the 4 I/O buses (which provides a direct interface to the SPARC processors), and one via a dedicated micro-controller which is used for board control.

The MK405 includes two Elite network switches which provide the first level switch network for the 4 processing elements. Each Elite connects one link from each of the four Elan Communications Processors, i.e. each Elite serves a separate network plane.

MBus

Each processing element has one MBus site allowing two SPARC processors to be supported on a single plug-in daughter board. There are four processing elements, giving a maximum board capacity of 8 processors. The MBus boards simply plug into the slots provided.

The MBus is fully level 2 compliant and runs at 40MHz. The two SPARC processors on each plug-in card share the MBus with an Elan Communications Processor and the I/O bus controller; the allocation of MBus Id's is:

- MBus id 0 is the I/O bus controller.
- MBus id 6 is the Elan Communications Processor.
- MBus id's 8 and 9 are the MBus processor module.

The use of MBus sockets on the motherboard allows PSARC modules to be easily upgraded as SPARC technology develops. The TI Viking and ROSS Pinnacle SPARC technologies are current options.

Note that each of the 4 MBus interfaces on the MK405 are independent; there is no requirement that the SPARC modules are the same.

ROSS Pinnacle Module

Two variants of the Pinnacle MBus module are available; you may use either the single processor modules or the dual processor module. Both variants include external second level cache.

The Pinnacle MBus modules are built upon a tightly coupled set of three ROSS devices: the RT620 HyperSPARC CPU, the RT625 cache controller, memory management, and tagging unit (CMTU), and the RT627 cache data units (CDUs).

Features of the RT620 CPU are:

- SPARC version 8 conformance.
- 90MHz clock rate.
- 4 execution units offering parallel execution of major instruction types: Load/Store, Branch/Call, integer and floating point units.
- Dual instruction fetch per clock cycle.
- 8Kbyte 2-way set-associative on-chip instruction cache.
- Instruction pipelining including a cache stage to accommodate the latency for second level cache accesses on data. Simultaneous accesses to on-chip and second level cache for each instruction fetch.

- High bandwidth 64bit Intra Module Bus (IMB) provides the interface between the CPU and the second level cache. Use of second level cache decouples the processor clock rate from the lower MBus clock rate.

Key features of the RT625 (CMTU) and RT627 (CDU) devices are:

- Full level 2 cache-coherent MBus compatibility.
- Each CDU has integral 16Kbytes x 32bit SRAM. MBus modules use either 2 or 4 CDUs for 128Kbyte or 256Kbyte second level direct-mapped cache.
- Physical cache tagging with virtual indexing allow the cache coherency logic to determine snoop hits and misses without stalling the CPU's access to the cache.
- Both copy-back and write-through cache modes supported.
- 32 byte read buffer and 64 byte write buffer for buffering the 32 byte cache lines in and out of the second level cache.
- SPARC reference MMU offering 64 entry, fully set-associative TLB with 4096 contexts.

Texas Instruments Viking Module

Two variants of the TI Viking MBus module are available. One contains a Viking SPARC processor with direct connection to the MBus. The second includes a Viking processor with additional Cache Controller and 1Mbyte of second level external cache (E-cache).

Key features of the TMS390Z50 SuperSPARC are:

- SPARC Version 8 conformance.
- 3 instructions per cycle, instruction pipelining, 150MIPs peak performance.
- SPARC Integer Unit.
- SPARC Reference MMU. Cached translation lookaside buffers (TLBs). 32bit virtual addresses, tagged with a 16bit context (65,536 contexts), map to 36bit physical addresses.

- Single and Double precision FPU. Tightly coupled to the integer execution pipeline and allowing one floating-point operation and one memory reference to be issued in each clock cycle. The FPU maintains a 4 entry FIFO queue for FP operations.
- 20Kbyte instruction cache, 16Kbyte data cache. The instruction cache is 5-way set associative, physically addressed, and non-writable. The data cache is 4-way associative and physically addressed. Both cache's are coherent with each other and with optional E-cache or MBus. Without E-cache the instruction and data caches operate in write-through mode; otherwise they are copy-back mode.
- Store buffer. A FIFO queue of 8 entries, each 64bits, decouples the instruction execution pipelines from the E-cache or MBus.
- Multiprocessor cache coherent support; highly pipelined and non-multiplexed VBus interface to optional external cache and the TMS390Z55 Cache Controller, or direct connection to the MBus.
- Prefetch buffer.
- Support for system and software debugging including hardware breakpoint.

The external cache is managed by the TMS390Z55 Cache Controller. Key features of this device are:

- Built in support for cache coherent multiprocessing; multiple Viking modules can share a single MBus and remain fully cache coherent.
- High performance VBus interface with the SPARC processor; decouples the SPARC processor from the MBus clock speed allowing higher processor performance. Reduced MBus traffic reduces contention when multiple processors share a single MBus.
- E-cache is direct mapped, copy-back, and unified: there is a single cache location where a particular byte of the physical address space can reside in the cache (direct mapping); writes by the main processor into the E-cache do not

propogate to main memory until the cache is flushed or replaced (copy-back); both instructions and data are supplied to the processor from the same cache (unified).

- 1Mbyte SRAM cache.

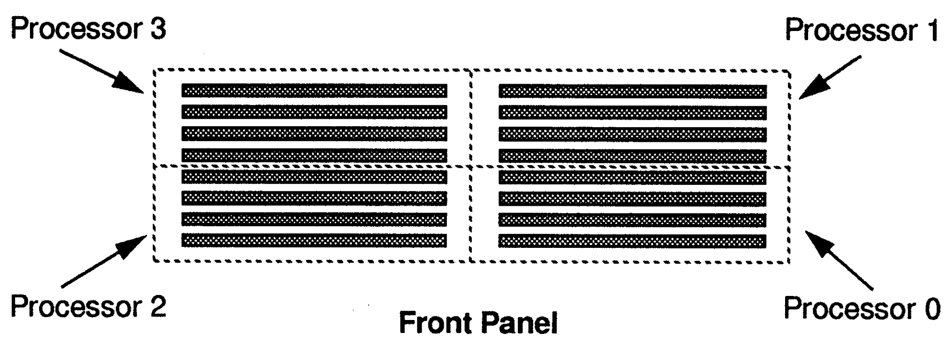
Memory Configurations

The memory controller used by each processing element offers a 128bit wide data bus for fast memory access, with single bit error correction, 2 bit error detection, and multiple bits within nibble detection. Memory errors are logged by kernel software and propagated via the CAN bus.

The MK405 includes 16 memory slots, of which 4 are allocated to each processing element. Both single and double sided, 4Mbit and 16Mbit SIMMS may be used, giving a total maximum on-board memory of 512Mbytes, or 128Mbytes per processing element. Within each group of 4 slots the SIMMs must be of the same type, but there is no requirement for the groups to be the same.

Memory is mapped into each processing elements address space as two banks of 64Mbytes. Bank 0 maps onto side 0 of the 4 SIMMs; bank 1 maps onto side 1. If less than 64Mbytes of memory is present in a bank it is echoed throughout the 64Mbytes. For each 128bit MBus data access 4 bytes are read from each SIMM.

Figure 2-1 Allocation of Memory Slots to Processing Elements



IO Bus

The IO bus is a slave-only bus used for the connection of minor peripherals to the MBus. There is one I/O bus for each of the 4 processing elements. The following devices are connected to this bus:

- A 512Kbyte EPROM holding bootstrap and diagnostic programs. The bootstrap code initialises the hardware devices, initial page table construction, and the booting of the Unix kernel. Normally only processor 0 in a dual-processor configuration handles the hardware initialisation; processor 1 sleeps until it is woken by processor 0 with a software interrupt.
- A realtime clock module with battery backed SRAM. This holds configuration information and node fault logs, including the log of uncorrectable memory errors. The realtime clock provides year, month, day, hour, minute, and second times.
- Interrupt controllers, one per SPARC processor. These use registers to mask out certain types of interrupt to relieve the SPARC processor from unnecessary interrupt loading, and to share the handling between the two processors. In addition some levels of interrupt are also presented direct to the Elan allowing them to be handled there and thus masked out of both the SPARC processors. The programming of the interrupt controllers is handled by the kernel device drivers.
- Periodic interrupt timers; used for maintaining the kernel clock and for kernel profiling. Two clock devices are used, one per SPARC processor. The lower level (kernel clock) interrupts from one of the clock devices are shared by both SPARC processors to remove the need for the two processor clocks to synchronise.
- A single CAN device provides both SPARCs with an interface to the machine-wide control area network (CAN). The SPARC processors write diagnostic information to this bus, and can also act as X-CAN or G-CAN routers. CAN routers transfer data between two levels of the CAN network; an X-CAN router handles transfers between the modules in a Cluster, and a G-CAN router handles transfers between Clusters. The configuration of a SPARC as a router will cause it's CAN device to generate numerous level 2 interrupts which will impact on processor performance.

Board Control Processor

The MK405 board uses an Hitachi H8/534 micro-controller (commonly referred to by Meiko as the H8) to perform basic node control functions. This controller is a single-chip 16bit RISC microcomputer with integral 2KBytes RAM, 32Kbytes EPROM, 16bit RISC CPU, and a number of I/O ports and timers.

The H8 processor runs independently of the other processors on the board and is used solely for control and diagnostic purposes. It has its own interface to the CAN bus via the second of the board's CAN interface devices. This processor receives diagnostic messages, via the CAN bus, from the local SPARC processors, and interprets incoming control messages, such as board reset, console connections, and network configuration.

Using the MK405 3

This chapter describes the usage of the MK405 in terms of its installation, hardware interfaces, and field serviceable components.

Installation

The MK405 is designed for use solely in a CS-2 Processor Module. The Processor Module supplies the board's power, cooling, and connection to the CS-2 data and control networks. The MK405 is fitted into one of the four vertical board slots behind the Processor Module's removable front panel.

The MK405 has external connection to the CS-2 data and control networks but has no other I/O interfaces. The MK405 must be installed alongside other board types which can provide disk and other I/O services.

Warning - You must disconnect the power from the Processor Module before removing or installing processor boards.

Warning - The board may be fitted with fragile or static sensitive devices. You must handle with care and observe anti-static precautions.

Removing the Module's Front Panel

The module's front panel is held in position by four clips, one in each corner. To remove the panel pull firmly away from the module.

The module's LED display is fitted to the module by two 50-way connectors. To remove the LED display pull firmly away from the module.

Use the reverse procedure to install the LEDs and front panel.

Installing the Processor Board

Insert the board so that it fits into the guide rails at the top and bottom of the module's board rack, ensuring that the component side is to the left (viewed facing the module). Gently push the board squarely on its front panel. Before pushing the board fully into position fold back the levers at each end of the front panel so that they are at 90^{0} to the board; now push the board (while holding the levers) until the base of the two levers is touching the card cage. To lever the board into its final position push both levers until they lie flat on the board's front panel. Secure the board by tightening the two captive screws.

Use the reverse procedure to remove the board.

Warning - You should take care not to damage the connectors at the rear of the board and on the module's backplane. Ensure that the board mates squarely with the module's backplane.

Warning - When removing or installing a board you should take care not to damage the RFI (copper) seals along the edge of the board's front panel.

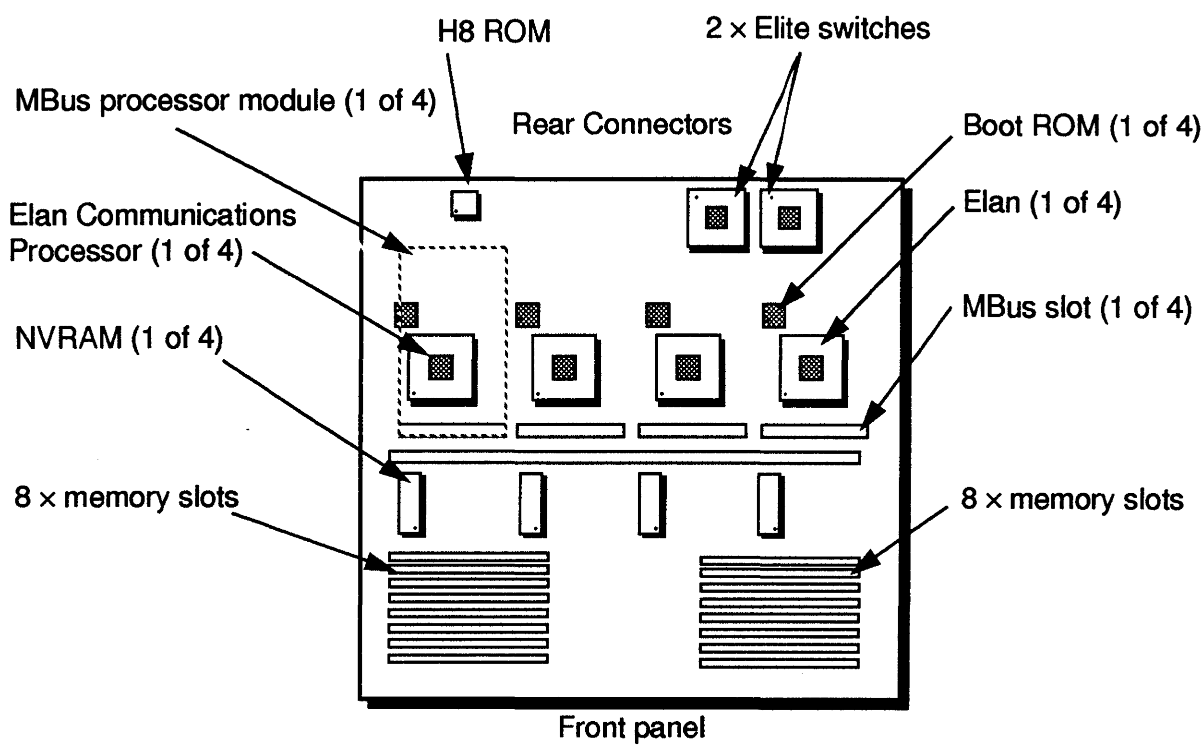
Warning - To maintain proper circulation of cooling air and to conform to RFI regulations all board slots must be fitted with a processor board or blanking plate.

Field Serviceable Components

The MK405 has the following field serviceable components.

- Superscalar SPARC processor modules fitted to MBus slots.
- 16 memory slots.
- Boot ROMs.
- H8 ROMs.
- Non-volatile RAM (NVRAM).

Figure 3-1 MK405 Components

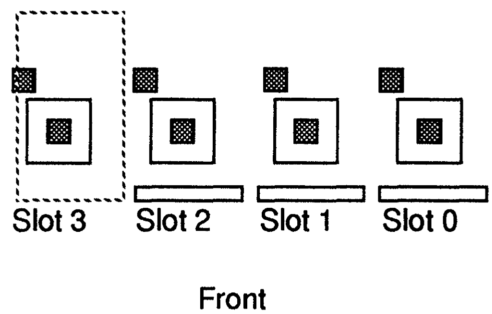


MBus Processor Modules

Up to 4 Superscalar SPARC modules may be installed in the MBus slots. Installation is simple — push the processor into the connector on the MK405 and secure into position with a screw at each corner.

The processor slots are numbered from 0 to 3, from right to left (when viewed from the front) as shown in Figure 3-2.

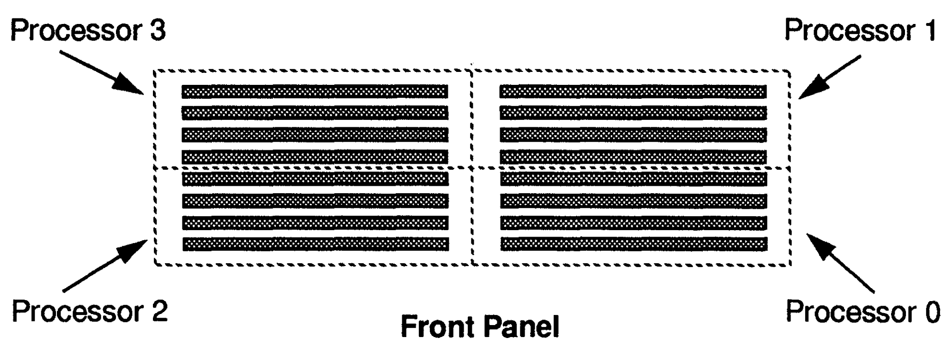
Figure 3-2 MK405 Processor Numbering



Memory

Up to 16 single in-line memory modules (SIMMs) may be fitted to the MK405 board (JEDEC 36 bit SIMMs). The array is constructed of 4 groups of 4 SIMMs, one group for each processor — see Figure 3-3. Within each group the SIMMs must be identical but there is no requirement for the groups to be the same.

Figure 3-3 Memory slots on an MK405 showing the allocation of slots to each of the 4 processors.



SIMMs can be either 4 Mbit DRAM or 16 Mbit DRAM technology, either single or double sided. This gives a total minimum memory configuration of 64 Mbytes and a maximum of 512 Mbytes (16 - 128 Mbytes per processor).

Boot ROM and H8 ROM

Both of these ROMs may be upgraded from time to time. They are held in sockets and are readily replaced. Note the position of pin 1 before removing the old device (usually marked by a dot on the packaging).

Realtime Clock and Battery Backed RAM

The realtime clock and non-volatile RAM device is held in a DIL socket and is easily replaced. Before removing the old device note the position of pin 1 (usually marked by a dot on the packaging). Note that the information within the NVRAM can only be restored by Meiko's engineers.

Warning - This device contains Lithium batteries; never dispose of this device in a fire or attempt to dismantle.

External Indicators

The MK405 front panel includes two pairs of LEDs. In the middle are two amber LEDs — illumination of these indicates an error with the Elite switches. The green/amber LED pair are the board's CAN bus indicators — green is the board's heartbeat, and amber indicates a transmission onto the bus.

The green LED flashes at a slow steady rate (once per second) when operating normally. A quicker flash rate (2×normal) indicates that one of the SPARC processors is not responding; a very quick flash rate (3×normal) indicates that the H8 processor on the module controller is not responding.

Each processor board within a processor module controls a 4×4 matrix of red LEDs on the module's front panel. Each of the four processors on an MK405 controls one row of 4 lights. Processor 0 controls the top row, processor 3 the bottom.

MBus Address Maps A

This section identifies the key devices used on the MK405 and lists their mapping into the MBus physical address space.

For detailed information about any of the devices you should refer to the manufacturer's data sheets.

The following notes are associated with the tables:

1. Memory banks are spaced 64MBytes apart. If less than 64MBytes of memory is present in each bank, it will appear echoed throughout the 64MBytes.
2. These locations are byte wide and are mapped into all 4 bytes of a word. Care should be taken to generate correct byte-wide accesses to the least significant byte of the word in order to maintain future compatibility.
3. These locations are byte-wide memory, mapped into contiguous byte locations. Word or halfword accesses will be automatically mapped into several successive byte-wide accesses.
4. These locations are byte sized registers which are only mapped into the most significant byte of a halfword. To ensure compatibility with other boards only byte accesses at the correct address should be used.
5. These locations are byte sized registers which are only mapped into the least significant byte of the word. To conserve MBus bandwidth and ensure compatibility with other boards only byte wide accesses at the correct address should be used.

6. These locations are halfword sized registers which are only mapped into the least significant halfword of the word. To conserve MBus bandwidth and ensure compatibility with other boards only halfword wide accesses at the correct address should be used.
7. These locations form a double-word register.

DRAM Slots

MBus Address	Usage	Rd/Wr	Note
`000000000`	64MB Memory in bank 0	RW	1
`004000000`	64MB Memory in bank 1	RW	1
`008000000 to effffffff`	Unused (MBus Timeout)		

Memory Controller

The MK405 uses LSI Logic L64860 memory controllers.

The following table shows the MBus memory maps for the memory controller's control and diagnostic ports.

MBus Address	Usage	Rd/Wr	Note
`f00000000`	Memory Enable	RW	
`f00000004`	Memory Delay	RW	
`f00000008`	Fault Status	R(W)	
`f0000000c`	Video Config	RW	
`f00000010`	Fault Address 0	R	
`f00000014`	Fault Address 1	R	
`f00000018`	ECC Diagnostics	RW	
`f0000001c to f0fffffff`	Unused (Read undefined)		
`f00000000 to fefffffff`	Unused (No response) (MBus Timeout)		

Boot ROM, Real Time Clock, Miscellaneous

The following table identifies the memory map for various devices on the I/O bus.

MBus Address	Usage	Rd/Wr	Note
`ff0000000 to ff007ffff`	Boot ROM (512KByte).	R	3
`ff0080000 to ff00fffff`	Unused (Boot ROM Echo)		
`ff0100000 to ff02fffff`	Unused (MBus Error)		
`ff0300000 to ff0301fff`	Real Time Clock module and 8KByte SRAM	RW	3
`ff0302000 to ff03fffff`	Unused (RTC Echoes)		
`ff0400000 to ff06fffff`	Unused (MBus Error)		
`ff0700000`	Node Reset Request	W	5
`ff0700004 to ff07001ff`	Unused (Echoes)		
`ff0700200`	MBus Grant readback	R	5
`ff0700204 to ff07003ff`	Unused (Echoes)		
`ff0700400`	Physical Slot Identifier	R	6
`ff0700404 to ff07005ff`	Unused (Echoes)		
`ff0700600`	LED Bargraph	RW	6
`ff0700604 to ff07007ff`	Unused (Echoes)		

Control Area Network Interface

The SPARC processors' interface the Control Area Network is via Philips PCA82C200 devices, which are mapped into the MBus address space at the following locations:

MBus Address	Usage	Rd/Wr	Note
ff0700800	CAN — Control Register	RW	5
ff0700804	CAN — Command Register	W	5
ff0700808	CAN — Status Register	R	5
ff070080c	CAN — Interrupt Register	R	5
ff0700810	CAN — Acceptance Code Register	RW	5
ff0700814	CAN — Acceptance Mask Register	RW	5
ff0700818	CAN — Bus Timing Register 0	RW	5
ff070081c	CAN — Bus Timing Register 1	RW	5
ff0700820	CAN — Output Control Register	RW	5
ff0700824	CAN — Test Register		
ff0700828	CAN — TXBuf Identifier	RW	5
ff070082c	CAN — TXBuf RTR Data Length code	RW	5
ff0700830	CAN — TXBuf Data Byte 1	RW	5
ff0700834	CAN — TXBuf Data Byte 2	RW	5
ff0700838	CAN — TXBuf Data Byte 3	RW	5
ff070083c	CAN — TXBuf Data Byte 4	RW	5
ff0700840	CAN — TXBuf Data Byte 5	RW	5
ff0700844	CAN — TXBuf Data Byte 6	RW	5
ff0700848	CAN — TXBuf Data Byte 7	RW	5
ff070084c	CAN — TXBuf Data Byte 8	RW	5
ff0700850	CAN — RXBuf Identifier	RW	5
ff0700854	CAN — RXBuf RTR Data Length code	RW	5

MBus Address	Usage	Rd/Wr	Note
ff0700858	CAN — RXBuf Data Byte 1	RW	5
ff070085c	CAN — RXBuf Data Byte 2	RW	5
ff0700860	CAN — RXBuf Data Byte 3	RW	5
ff0700864	CAN — RXBuf Data Byte 4	RW	5
ff0700868	CAN — RXBuf Data Byte 5	RW	5
ff070086c	CAN — RXBuf Data Byte 6	RW	5
ff0700870	CAN — RXBuf Data Byte 7	RW	5
ff0700874	CAN — RXBuf Data Byte 8	RW	5
ff0700878	CAN — Unimplemented		
ff070087c	CAN — Clock Divider Register	RW	5
ff0700880 to ff0700fff	Unused (Echoes of above)		

Interrupt Request Control and Status Registers

The MK405 has many sources of interrupts to the SPARCs. These are assigned priority levels and are passed to the Interrupt Controllers (one per SPARC). These enable the handling of interrupts to be shared between the SPARCs. The controller has the following major features:

- Masks for all hardware generated interrupts.
- Software interrupts on six levels using a multi-processor compatible set/clear register structure. This allows one processor to interrupt the other by writing once to a single memory location, without having to do a read-modify-write cycle which is subject to being broken by the other processor. Supported interrupt levels are 1, 4, 6, 12, 13, and 15.
- Latches triggered by the output of the free running timers (the Periodic Interrupt Timers). These latches capture the timer passing through zero event and cause an interrupt to the processor. When the interrupt handler accepts the interrupt and clears the request the latch is reset.
- Priority encoder to generate the 4bit encoded interrupt level that is passed to the SPARC processor module.

The Periodic Interrupt Timers are AMD 82C54 free running timers, one assigned to each of the SPARC processors. Timer 0 generates the lower priority clock ticks signal, issuing a level 10 interrupt when the timer has counted down from a preset value to 0 (the count down is at a rate of one decrement every 3.2us). Timer 1 also counts down from a preset value but issues a level 14 interrupt and decrements at a rate of 0.8us. Both processors share timer 0 from one of the timer devices (to avoid the need for explicit synchronisation).

MBus Address	**Usage**	**Rd/Wr**	**Note**
`ff0701000`	IRQ PAL 0 — Timer Latches	RW	5
`ff0701002`	IRQ PAL 0 — Mask Register Read / Clear	RW	5
`ff0701006`	IRQ PAL 0 — Mask Register Set	RW	5
`ff070100a`	IRQ PAL 0 — Software Interrupt Reg Read / Clear	RW	5
`ff070100e`	IRQ PAL 0 — Software Interrupt Reg Set	W	5
`ff0701010 to ff07011ff`	Unused (Echoes)		
`ff0701200`	Timer 0 Level10 Processor 0 and 1	RW	5
`ff0701204`	Timer 0 Level14 Processor 0	RW	5
`ff0701208`	Timer 0 Spare timer	RW	5
`ff070120c`	Timer 0 Control register	RW	5
`ff0701210 to ff07015ff`	Unused (Echoes)		
`ff0701600`	Err and Powerfail Latch	RW	5
`ff0701604 to ff0701fff`	Unused (Echoes)		
`ff0702000`	IRQ PAL 1 — Timer Latches	RW	5
`ff0702002`	IRQ PAL 1 — Mask Register Read / Clear	RW	5
`ff0702006`	IRQ PAL 1 — Mask Register Set	RW	5

MBus Address	Usage	Rd/Wr	Note
ff070200a	IRQ PAL 1 — Software Interrupt Reg Read / Clear	RW	5
ff070200e	IRQ PAL 1 — Software Interrupt Reg Set	W	5
ff0702010 to ff07021ff	Unused (Echoes)		
ff0702200	Spare timer	RW	5
ff0702204	Timer 1 Level14 Processor 1	RW	5
ff0702208	Timer 1 Spare timer	RW	5
ff070220c	Timer 1 Control register	RW	5
ff0702210 to ff0702fff	Unused (Echoes)		
ff0703000 to ff0703fff	Unused (Read Undefined)		
ff0704000 to ff07fffff	Unused (Echoes of above)		

MBus to I/O Bus

The MBus to I/O Bus interface is via an LSI Logic L64851 device.

MBus Address	Usage	Read/ Write	Note
ff0800000	Software interrupts and enable.	RW	
ff0800004	Active level of external interrupts.	RW	
ff0800008	Programmable limit register 0.	RW	
ff080000c	Programmable limit register 1.	RW	
ff0800010	I/O bus devices available.	RW	
ff0800018	Latency delay register.	RW	

MBus Address	Usage	Read/ Write	Note
ff080001C	MBus Id register.	R	
ff0800020	Programmable timer.	R	
ff0800024	Programmable timer.	R	

Elan Communications Processor and MBus Slot Slaves

The Elan Communications Processors is a Meiko device.

MBus Address	Usage	Rd/Wr	Note
ff6f80000 to ff6ffdfff	ELAN Command port area	RW	
ff6ffe000 to ff6ffffbf	ELAN Hush register area	RW	
ff6ffffc0	ELAN Clock Hi	RW	
ff6ffffc4	ELAN Clock Hi	R	
ff6ffffc8	ELAN Clock Lo	RW	
ff6ffffcc	ELAN Clock Lo	R	
ff6ffffd0	ELAN Alarm	RW	
ff6ffffd4	ELAN Alarm	R	
ff6ffffd8	ELAN Interrupt	R	
ff6ffffdc	ELAN Interrupt	R	
ff6ffffe0	ELAN Clock Hi	R	7
ff6ffffe4	ELAN Clock Lo (For 64bit accesses)	R	7
ff6ffffe8	ELAN Main Proc. Interrupt Mask	RW	
ff6ffffec	ELAN Main Proc. Interrupt Mask	R	
ff6fffff0	ELAN Control register	RW	
ff6fffff4	ELAN Control register	R	
ff6fffff8	MBus Port ID register for ELAN Chip	R	

MBus Address	Usage	Rd/Wr	Note
ff6ffffc	MBus Port ID register for ELAN Chip	R	
ff7000000 to ff7ffffff	Unused (MBus timeout)		
ff8000000 to ff9ffffff	Used by MBus slave device in MBus Slot		
ffa000000 to fffffffff	Unused (MBus timeout)		

NVRAM Variables **B**

The battery-backed RAM in the realtime clock module is used to used to store basic machine start-up and communication options.

These parameters may be queried using the Forth Monitor (i.e. at the `ok` prompt):

`printenv`	Display current variable settings.
`setenv` *variable value*	Assign (or reassign) a value to a variable.
`set-default` *variable*	Restore the variables default value.
`set-defaults`	Restore the default values to all variables.

For example:

```
ok setenv output-device can
```

Alternatively the System Administrator can use the `eeprom(1m)` command to view and change the variables direct from a Unix command shell. For example:

```
root@cs2# eeprom output-device=can
```

Some of the parameters (those marked in the following list) may also be modified using the Set function in Pandora's Network and Configuration Views.

Variable	Default	Description
sbus-probe-list	43012	Identifies the SBus slots to probe and the probe order.
keyboard-click?	false	If true, enable keyboard click.
keymap	*no default*	Name of custom keymap file.
output-device †	screen	Power-on output device. One of screen, can, ttya, or ttyb. Use can to enable console connections to be grabbed by cancon(1m) and Pandora.
input-device †	keyboard	Power-on input device. One of keyboard, can, ttya, or ttyb. Use can to enable console connections to be grabbed by cancon(1m) and Pandora.
cancon-host	4294967295	Used to record the host of the cancon(1m) remote console connection through a reboot of this processor. Do not change.
elanip-broadcast-high †	4096	Highest Elan Id in network.
elanip-broadcast-low †	0	Lowest Elan Id in network.
ep-btxpktlifetime †	1000	Elan packet characteristics.
ep-btxtimeout †	1000	Elan packet characteristics.
ep-txpktlifetime †	10000	Elan packet characteristics.
ep-txtimeout †	10000	Elan packet characteristics.
ep-bigmsgbcastboxes †	4	Elan packet characteristics.
ep-bigmsgboxes †	32	Elan packet characteristics.
ep-bigmsgsize †	20416	Elan packet characteristics.
ep-smallmsgbcastboxes †	4	Elan packet characteristics.
ep-smallmsgboxes †	32	Elan packet characteristics.
ep-smallmsgsize †	4032	Elan packet characteristics.
elan-boot-id †	0	Elan Id of node that this processor boots from.

Variable	Default	Description
`elan-node-id` †	`0`	Elan Id of this processor.
`elan-node-level` †	`1`	The processor's level in the CS-2 network.
`elan-num-levels` †	`1`	Number of levels in the CS-2 network.
`elan-top-switch` †	`0`	Specifies the level in the network that the processor sees it's topswitch. Usually this is level 0, the real top of the network.
`elan-switch-plane` †	`0`	Switch plane that this processor receives it boot code from when booting via the Elan network.
`ttyb-rts-dtr-off`	`false`	If true, Solaris does not assert RTS/DTR on ttyb.
`ttyb-ignore-cd`	`true`	If true, Solaris ignores carrier-detect on ttyb.
`ttya-rts-dtr-off`	`false`	If true, Solaris does not assert RTS/DTR on ttya.
`ttya-ignore-cd`	`true`	If true, Solaris ignores carrier-detect on ttya.
`ttyb-mode`	`9600,8,n,1,-`	ttyb (baud rate, #bits, parity, #stop, handshake). Baud rate is 110, 300, 1200, 2400, 4800, 9600, 19200, or 38400. #bits is 5, 6, 7, or 8. Parity is `n` (none), `e` (even), `o` (odd), `m` (mark), `s` (space). Handshake is – (none), `h` (hardware rts/cts), `s` (software).
`ttya-mode`	`9600,8,n,1,-`	ttyb (baud rate, #bits, parity, #stop, handshake). Baud rate is 110, 300, 1200, 2400, 4800, 9600, 19200, or 38400. #bits is 5, 6, 7, or 8. Parity is `n` (none), `e` (even), `o` (odd), `m` (mark), `s` (space). Handshake is – (none), `h` (hardware rts/cts), `s` (software).
`fcode-debug?`	`false`	If true, include name fields for plug-in device Fcodes.
`diag-file` †	`kadb`	The file and arguments to load from the root filesystem when the `diag-switch?` is true; otherwise use the `boot-file` parameter.
`diag-device` †	`elan`	Device to boot from when the `diag-switch?` is true; one of `disk`, `net`, or `elan`. Specify `elan` to boot over the CS-2 data network.

Variable	Default	Description
`boot-file`		The file and arguments to load from the root filesystem (e.g. `kadb -v`, or `/kernel/unix -vr`). No file implies `/kernel/unix`.
`boot-device` †	`elan`	Device to boot from; one of `disk`, `net`, or `elan`. Specify `elan` to boot over the CS-2 data network.
`auto-boot?` †	`false`	Boot automatically after power-on. Default value is true.
`watchdog-reboot?`	`false`	If true, reboot after watchdog reset.
`local-mac-address?`	`false`	If true, use the ethernet address taken from the `local-mac-address` parameter; otherwise use the IdPROM.
`screen-#columns`	`80`	Number of on-screen columns.
`screen-#rows`	`34`	Number of on-screen rows.
`selftest-#megs`	`1`	Megabytes of RAM to test on power-up or memory test.
`scsi-initiator-id`	`7`	SCSI bus address of host adapter, range 0–7.
`cpu-#mhz`	`40`	CPU clock rate.
`use-nvramrc?`	`false`	If true, execute the code stored in the nvramrc parameter when the boot ROM starts up.
`nvramrc`		Forth code to execute when the boot ROM starts-up (but only if `use-nvramrc?` is true).
`sunmon-compat?`	`false`	If true, come-up with old style prompt '>'.
`security-mode`	`none`	System security level for monitor commands; one of `none`, `command`, or `full`. None allows all commands to be executed. Command allows the continue and boot (without parameters) commands to be executed; others require a password. Full requires a password before any commands may be executed.
`security-password`	*no default*	The password used with security-mode described above.

Variable	Default	Description
`security-#badlogins`	*no default*	System set variable showing the number of times a bad password was specified.
`oem-logo`	*no default*	Byte array OEM logo (enabled by `oem-logo?`). Create a Forth array containing the logo and then copy into the `oem-logo` field.
`oem-logo?`	`false`	Enables OEM logo defined by `oem-logo`.
`oem-banner`	*no default*	Text displayed in the custom OEM banner alongside the OEM logo. Enabled by `oem-banner?`.
`oem-banner?`	`false`	Enables OEM banner text specified in `oem-banner`.
`hardware-revision`	*no default*	Hardware revision of this board (e.g. Rev D).
`last-hardware-update`	*no default*	Date of board's manufacture or last upgrade (e.g. 25May94).
`testarea`	`0`	Unused.
`mfg-switch?`	`false`	If true, perform repeated self tests.
`diag-switch?`	`false`	Run in diagnostic mode.

† These parameters may be changed using the Set function in Pandora's Network and Configuration Views.

Forth Monitor Commands C

The following commands have been added to the Forth Monitor and are in addition to the commands that are normally present on a Solaris system. The additional commands relate to the Control Network (CAN) and Elan network.

CAN Commands

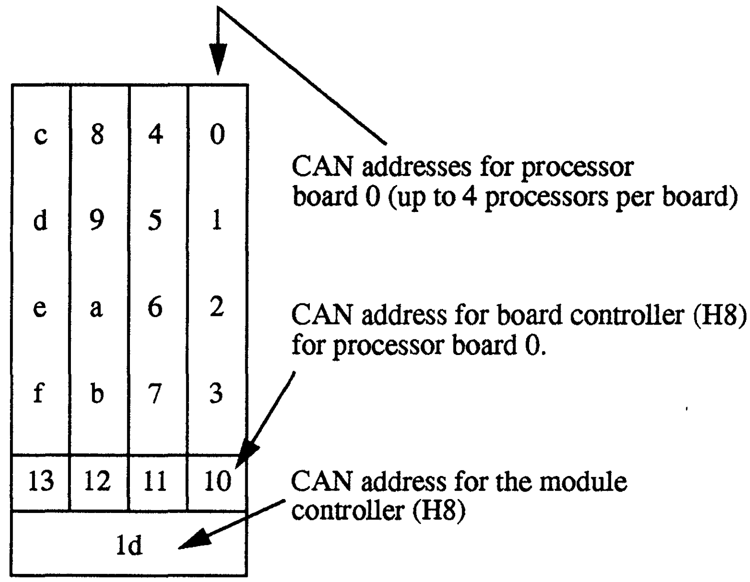
To test and use the CAN bus you need to understand CAN addresses.

Nodes are addressed by their physical position in terms of Cluster, Module, and Node id's (CMN). In CAN packets each of these id's is represented by a 6bit field; the hexadecimal representation of these three 6bit fields is a Node Id.

The module id is derived from the switch at the rear of the module. The numbering of the nodes within a module is shown in Figure C-1.

For example, the Node Id of processor with CMN 0:2:3 (processor 3, module 2, cluster 0) is 00083. The node's controlling H8 has the Node Id 00090. The node's module controller has the Node Id 0009d.

Figure C-1 CAN Addresses within a Module



Testing the CAN Device

Commands are provided to test the SPARC's CAN interface device, to test the board and module controllers, and to monitor activity on the CAN bus.

Testing the CAN Interface Device

The `test` command tests the SPARC's CAN device by writing various values into its test register. The test is repeated using the test registers on the board controller's H8 and the module controller's H8.

```
ok test /can
 Register test 0x00: OK
 Register test 0xff: OK
 Register test 0xaa: OK
 Register test 0x55: OK
 Checking on-board H8: OK.
 Checking module controller: OK.
```

Testing the CAN Bus

You can test the CAN bus connection between nodes by using the `rtest` command. In the following example data is transferred from the current node to node 4:

```
ok 4 rtest
Performing remote write/read test on node 4
Remote node type is MK405
....................................................
Time taken was 13630mSecs
```

Checking the Board and Module Controllers

You can check that both the board controller and module controller H8 processors are running by using the `ping-h8` and `ping-module` commands. Note that you need to change directory to `/can` before you use these commands.

```
ok cd /can
ok ping-h8
On-Board H8 is MK401.
ok cd ..
```

```
ok cd /can
ok ping-module
Module controller is MK515.
ok cd ..
```

Querying CAN Bus Usage

You can query the utilisation of the CAN bus by using the `perf` command. This command shows the number of CAN packets received since the machine was powered-up, and the number since the last query. You need to change to the `/can` directory before using this command:

```
ok cd /can
ok perf
Total number of messages received since power-up: 259380
No. of messages received per second since the last check: 4
ok cd ..
```

Monitoring CAN Bus Packets

You can *snoop* the CAN bus (monitor that packets on the bus) using the `snoop` command. Note that you cannot use this facility if you are connected to the Forth Monitor via a cancon connection. You need to change to the `can` directory before using this command.

```
ok cd can
ok snoop
Can't can-snoop if you are a cancon slave
ok cd ..
```

CAN Addresses

To determine the CAN address of this node use `.can-id`. This displays the node's address in terms of its CMN, Node Id, and Slot Id. The Slot Id is for Meiko engineering use[1].

```
ok .can-id
SlotId: 0090, CAN Node-id: 00088 [00:02:08]
```

Similarly the CAN address of the board's controlling H8 processor can be obtained with the `.h8-id` command:

```
ok .h8-id
The on-board H8 is node 00092.
```

1. The slot id is the node's physical position in the machine represented by a 5 bit cluster number, a 5 bit module number, a 2 bit slot number, 2 unused (always 0 bits), and a 2 bit processor number; the 2 bits that represent the slot number are transposed.

The CAN address of the H8 that controls the board's module can be determined by the `.module-id` command:

```
ok .module-id
The module controller is node 0009d.
```

You can convert from CAN node id's to Cluster, Module, Node addresses (and vice versa) by using the `canid>cmn` and `cmn>canid` commands respectively. Note that you need to change directory to `can` before you use these commands.

```
ok cd can
ok 4 canid>cmn
0 0 4
ok cd ..
```

```
ok cd can
ok 0 2 8 cmn>canid
Node Id is 00088
ok cd ..
```

Querying CAN Objects

Can packets include a 10 bit address space which, although not sufficient to map into the MBus/H8 physical address space, is adequate to map-in various status and control devices. These are referred to as CAN objects. Reading or writing to these objects allows you to query the status of a processor, board, or module, and to issue control instructions. See the header file `/opt/MEIKOcs2/include/canio/canobj.h` for a list of object addresses and their meanings.

Local CAN objects are those that relate directly to this node. Remote CAN objects maybe those of a board, module controller, or remote SPARC.

You use the `rlo` command to read a local object. You need to pass an object id on the Forth stack; in the following example we request the board type and are returned 191 (an MK401):

```
ok 0 rlo
Read: 191
```

To read a remote object you need to push onto the Forth stack a CAN node id and the object id. In the following example we request the board type of node `c1` (module 3, board 0, node 2), which is an MK405:

```
ok c1 0 rro
Read: 195
```

The following additional example fetches the board type of node `dd`, which is the controller for module 3, cluster 0:

```
ok dd 0 rro
Read: 203
```

Similar commands exist to write to CAN objects, but their direct use is not recommended (they can reconfigure and reset the machine).

Remote Console Connections

You can create a console connection to a remote node by using `cancon`. You need to pass on the Forth stack a CAN node id.

You cannot create a cancon connection from within an existing cancon connection. If you are remotely interacting with a node's Forth monitor via cancon (or Pandora) an attempt to create another cancon connection will fail.

```
ok 4 cancon
Connected to node 00004

cs2-4 console login:
```

If your node is currently serving a remote console connection to someone else you can force it to disconnect that connection by using cancon-dis. In the following example the current connection to node 8 is dropped:

```
ok 8 cancon-dis
Disconnecting node 00008 [00:00:08] ...
```

Elan Commands

The Elan device includes self test code that can be executed by the test-all command (which tests memory, SBus, CAN, Elan and all other devices with self test code) or explicitly by the test /elan command.

```
ok test /elan
Initialising Elan/Selftest software ... OK
Checking threads processor ... OK

Testing from level 1 to level 1.
Generating a route to level 1 ... OK
Ping ... OK
Check-Ping ... OK
Spraying data to top switch ... OK
Testing spray buffer ... OK

Closing down Elan/Selftest software ... OK
```

Computing

Surface

Overview of the Control Area Network (CAN)

S1002-10M140.00

meiko

Meiko's address in the US is:

Meiko
130 Baker Avenue
Concord MA01742

508 371 0088
Fax: 508 371 7516

Meiko's address in the UK is:

Meiko Limited
650 Aztec West
Bristol
BS12 4SD

Tel: 01454 616171
Fax: 01454 618188

Issue Status:		
	Draft	
	Preliminary	
	Release	x
	Obsolete	

Circulation Control: *External*

Contents

The Control Area Network 1

Introduction

The Control Area Network (CAN) is a low bandwidth serial network. It is used by the CS-2 to carry control, diagnostic, and remote console traffic between processors. The CAN is independent of the CS-2 data network and does not therefore impact on its performance.

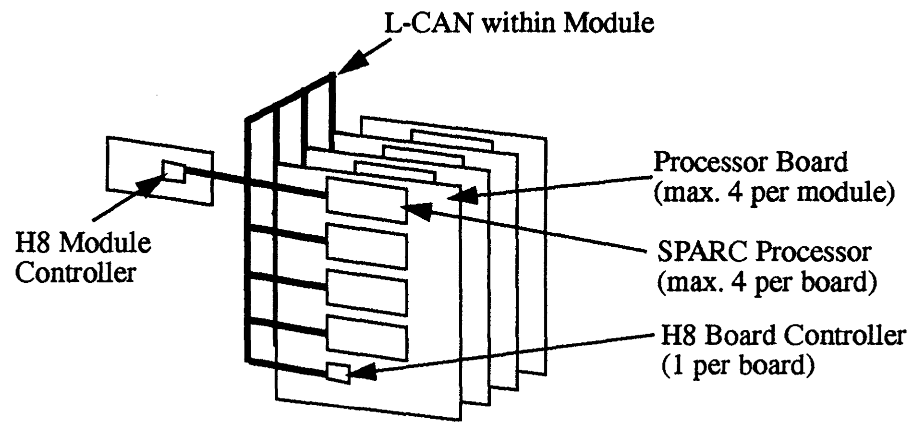
An understanding of the CAN is not required for normal operation of the CS-2. It is typically used by the resource management system to maintain the machine database, and by Pandora to create remote console connections and to gather component operating status. The information in this document is therefore provided for information only.

Network Hierarchy

The CAN is hierarchical, with the number of nodes on each network limited (by the electrical characteristics of the CAN transceivers) to around 30 nodes.

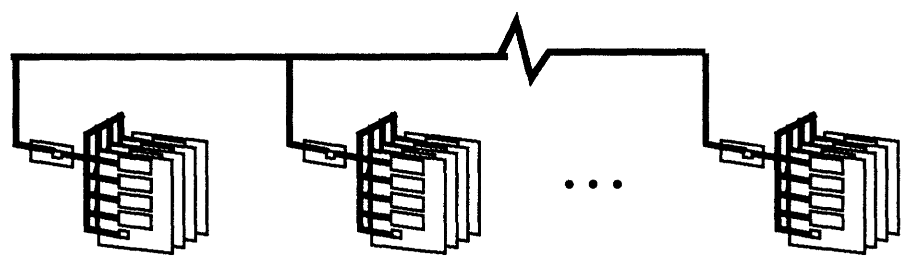
At the lowest level is the L-CAN, a network interconnecting the processors in each module. This connects up to 16 SPARC processors (allowing for four boards each with four SPARC processors), 4 board controller processors (H8 processors), and 1 module controller (also an H8 processor). The interface between the processor and the CAN is handled in each case by dedicated CAN transceiver devices.

Figure 1-1 L-CAN Connections within a Processor Module



The modules within a Cluster (3 bays, up to 24 modules) are interconnected by the X-CAN, and the interconnection of Clusters is via the G-CAN. The transfer of network traffic from the L-CAN to the X-CAN is handled by each module's controller, whereas the transfer from X-CAN to G-CAN is via nominated routers (selectable from Pandora).

Figure 1-2 X-CAN Connections within a Cluster



CAN Messages

A node requests status information from, or sends control requests to, another CAN node by sending a message to it. The information or function that is required by the sender is specified by addressing an *object* at the recipient. Objects are either hardware devices or software functions, the mapping from an object id to device or function being performed at the recipient. The operation that is to be performed on the object, such as a read or write, and any associated object-specific data is also included in the CAN message.

The CAN packet consists of up to 10 bytes, of which 2 bytes are the packet header, 4 bytes are the object address, and 4 bytes are the optional object-specific data. The header defines the source and destination nodes for the packet; the source and destination will always be on the same L-CAN. The object address define the absolute address (in terms of cluster, module, node, and object) of the target object, and also define the operation that is to be performed on that object (e.g. a read, write, or a signal). Absolute addressing of the object is required by the routing software when the source and destination nodes are on different L-CAN networks; in this case the source node sends a message to its module controller, and the software running on the module controller determines from the full object address the routing of the message, via the network hierarchy, to the remote node.

Network Protocol

The CAN uses a master/slave protocol in which most SPARC processors are configured as slaves, and most board and module controllers (i.e the H8 processors) are configured as masters.

Masters have the capability to read from and write to objects over the CAN, and thus have the capability to control and reconfigure the machine. Slaves can signal to masters that they have information for their attention by sending a signal message. A slave can be given the capability of a master by another master for specific purposes, for example to enable a SPARC host to read/write the console object of another SPARC. By assigning the role of masters to the H8 processors, which are themselves controlled by firmware programme code, interference of the CAN by user processes is prevented.

By default the masters (the H8 controllers) receive all messages from the CAN and use filters on the CAN controllers to extract the messages appropriate to each processor; this allows a broadcast capability to be defined for the module and board controllers. Slave processors (the SPARCs) are configured only to receive messages that are explicitly targeted at them, which reduces CAN workload on the SPARC processors.

Prioritisation

Message priorities are used at two stages; during arbitration between CAN devices on the L-CAN, and during routing between network levels by a routing processor. The two stages of prioritisation are represented by the two priority bits in the CAN message packet; the header defines the priority between CAN devices (either 0 for high priority, or 1 for low), and address data specifies the prioritisation for routed messages. A high priority message might indicate a power supply failure, whereas less urgent messages (switch errors) are recorded by low priority messages. During routing the priority in the message's address field is used to reduce congestion at the routers and on the X-CAN/G-CAN networks, and may be modified by the router processors to give highest priority to message responses and those making their way down the CAN hierarchy. Note that a low priority message sent to a node can be overwritten by a high priority message, causing the low priority message to be scrapped.

Network Error Detection and Recovery

Packets are acknowledged (ACK'ed) if they reach their destination and are interpreted correctly. A not-acknowledge (NACK) is sent if they fail to be correctly interpreted at their destination.

Reasons for failure are:

- Bad message. Perhaps the sender attempted to write to a read-only object, or an object that doesn't exist.
- Hardware errors. Either the message or the acknowledgement failed.
- Hardware overruns. No spare input buffer at the transceiver.

Bad messages, or messages to non-existent objects, are signalled to the sender by the return of a not-acknowledge packet.

Hardware errors are detected using a timeout at the message sender. The expiry of the timeout period indicates that the message and/or its acknowledgement failed. The behaviour of the sender in this case is function specific, but may be to resend the message or to give-up.

Hardware overruns occur when a transceiver's two receive buffers are full and a message on the network is targeted at the device. The incoming message is ignored, and an interrupt is issued to the CAN device's controlling processor. The sender of the message detects the failed transmission by the absence of an acknowledgement within its timeout period.

The CAN transceivers maintain a count of input and output errors using two counters. The transceivers are initialised with two threshold values called the *warning* limit and the *bus-off* limit at which the number of errors becomes serious and requires attention. When a count reaches the warning limit an error flag is set and the attached processor is interrupted, but the transceiver continues to operate. On reaching the bus-off limit the transceiver goes into a permanent reset state and is no longer operational. The accumulation of not-acknowledge errors cannot force the transceiver into a bus-off state, to accommodate the case at system start-up where a transceiver may become operational in advance of its peers.

Example — Snooping the CAN

The `cansnoop(1m)` command can be used by the root user to monitor the activity on the local CAN. The following example output shows some of the CAN traffic that is generated when a console connection is created via the CAN with either Pandora or `cancon(1m)`. The following example has been edited (for clarity) to remove the numerous heartbeat signals which are sent between nodes to signal their continuing operation.

In the following output the initial 2 digit field is a feature of `cansnoop(1m)` and represents a microsecond time stamp; subsequent fields represent the contents of the CAN packets. The second field identifies the source and destination CAN nodes (the packet header), the third field identifies the message type (WO= write, ACK = acknowledge, D = data), the next group of 4 fields is the full object address (with key object id's replaced by an ASCII mnemonic), and the remaining fields being the optional 4 bytes of data and their ASCII representation.

The output shows node 008 sending a packet to the console connection object at node 004; the data field identifies the initiator of the connection and the object that is to be used for subsequent communications. The connection request is acknowledged by node 004.

Following the initial handshaking characters typed at the keyboard are sent via the CAN to the remote console port. Each typed character is acknowledged by node 004 and echoed back to node 008. The connection is dropped by a write from node 008 to node 004's console disconnect object.

```
111.870 008->004  WO  00,02,04 CONSOLECONN 30 02 23 82 '0.#.'
111.897 004->008  ACK 00,02,04 03f0        ff ff ff ff '....'
122.335 008->004  D   00,02,04 0000        0d          '.'
122.335 004->008  ACK 00,02,04 0000
122.336 004->008  D   00,02,08 0382        0d 0a       '..'
122.336 008->004  ACK 00,02,08 0382
122.340 004->008  D   00,02,08 0382        6e 6f 76 61 'nova'
122.341 008->004  ACK 00,02,08 0382
122.341 004->008  D   00,02,08 0382        31 20 63 6f '1 co'
122.341 008->004  ACK 00,02,08 0382
122.341 004->008  D   00,02,08 0382        6e 73 6f 6c 'nsol'
122.341 008->004  ACK 00,02,08 0382
122.342 004->008  D   00,02,08 0382        65 20 6c 6f 'e lo'
122.342 008->004  ACK 00,02,08 0382
122.342 004->008  D   00,02,08 0382        67 69 6e 3a 'gin:'
122.342 008->004  ACK 00,02,08 0382


138.993 004->008  D   00,02,08 0382        36 5d 23 20 '6]# '
138.993 008->004  ACK 00,02,08 0382
138.993 004->008  D   00,02,08 0382        65 78 69 74 'exit'
138.993 008->004  ACK 00,02,08 0382
138.994 004->008  D   00,02,08 0382        0d 0a       '..'
138.994 008->004  ACK 00,02,08 0382
144.698 004->008  D   00,02,08 0382        0d 0d 0a    '...'
144.698 008->004  ACK 00,02,08 0382
144.702 004->008  D   00,02,08 0382        6e 6f 76 61 'nova'
144.702 008->004  ACK 00,02,08 0382
144.702 004->008  D   00,02,08 0382        31 20 63 6f '1 co'
144.702 008->004  ACK 00,02,08 0382
144.703 004->008  D   00,02,08 0382        6e 73 6f 6c 'nsol'
144.703 008->004  ACK 00,02,08 0382
144.703 004->008  D   00,02,08 0382        65 20 6c 6f 'e lo'
144.703 008->004  ACK 00,02,08 0382
144.704 004->008  D   00,02,08 0382        67 69 6e 3a 'gin:'
144.704 008->004  ACK 00,02,08 0382
144.704 004->008  D   00,02,08 0382        20          ' '
144.704 008->004  ACK 00,02,08 0382
149.177 008->004  WO  00,02,04 CONSOLEDISC 30 02 23 82 '0.#.'
149.203 004->008  ACK 00,02,04 03f1        30 02 23 82 '0.#.'
```

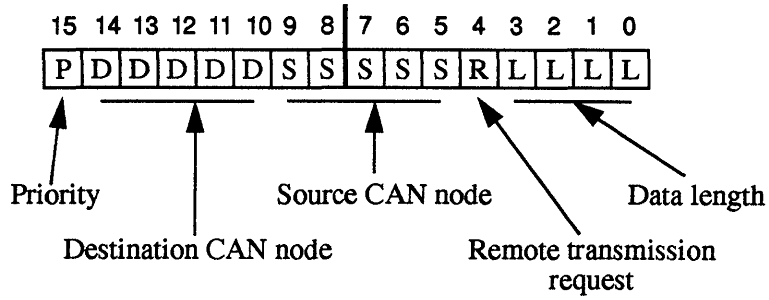
Appendix A — Packet Format

CAN packets are 10 bytes in length consisting of 2 bytes of header information, 4 bytes address data, and an optional 4 bytes of message data.

Header Data

The 2 byte message header identifies the source node, destination node, and message priority.

Figure 1-3 Message Header (2 bytes)



The fields in the packet header have the following meanings:

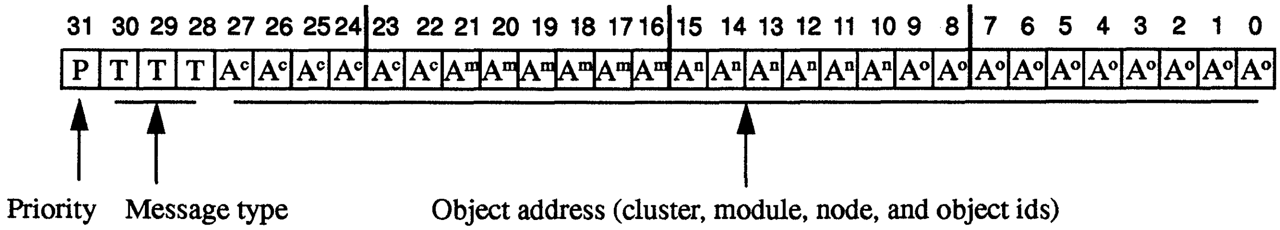
Bit(s)	Meaning
15	Message priority (for arbitration on the L-CAN). 0 is high, 1 is low.
14-10	Destination CAN node id. (in the range 0–29).
9-5	Source CAN node id. (in the range 0–29).
4	Remote transmission request (always 1 for the CS-2).
3-0	Length of following data. This will be either 4 or 8 for the CS-2; 4 bytes are required for the address data, and an optional 4 bytes for object-specific data.

Within a module CAN node ids are allocated as shown in Figure 1-5.

Address Data

The action performed at the destination node is specified by the address data. This defines the message type (Write to object, read from object etc.), and the address of the object that is to be targeted.

Figure 1-4 Object Addressing (4 bytes)



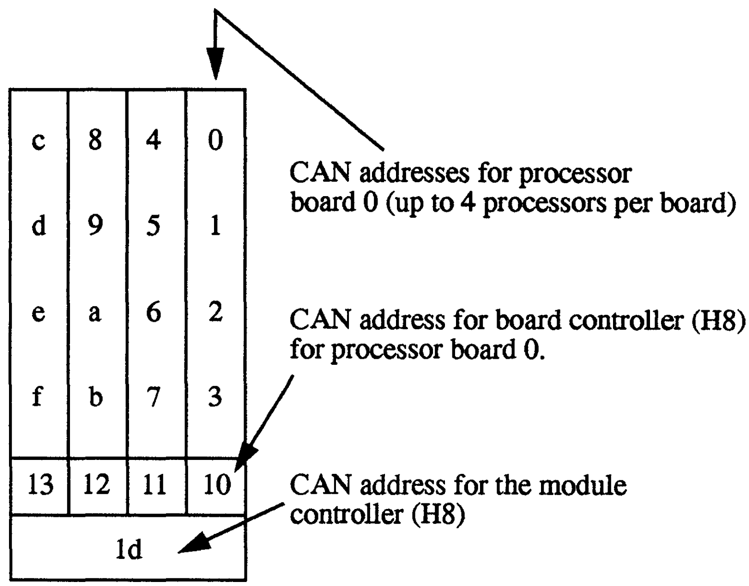
The fields in the address data have the following meanings:

Bit(s)	Meaning
31	Message priority for arbitration by the X-CAN/G-CAN routers. 0 is high, 1 is low.
30-28	Message type (see below).
27-0	The address of the object that the transaction is to apply to. This if a full machine address consisting of a 6 bit cluster id, 6 bit module id, a 6 bit node id, and a 10 bit object id. A broadcast is specified by 111110; 111111 means never route to this level. Object id's are listed in the header files `/usr/include/sys/cankobj.h` and `/opt/MEIKOcs2/etc/include/canio/canobj.h`.

Message types are:

Id	Meaning
000	Read request; the address identifies an object to be read. For use by master CAN nodes only (typically the H8 processors).
001	Write request; the address identifies the object to be written to. For use by master CAN nodes only (typically the H8 processors).
010	Write request without acknowledge from destination; the address identifies the object to be written to. For use by master CAN nodes only (typically the H8 processors).
011	Data.
100	Write acknowledge.
101	Unused.
110	Write a not-acknowledgement.
111	Send a signal. The data field defines the object that has changed (using the least significant 10bits of the 4byte data field); the address field includes the broadcast code (111110) in the node, and/or module and/or cluster fields. This is used by CAN slave nodes to notify CAN masters that an object has changed. At least one master should read the changed object. The signal is repeated at regular intervals (the timeout period) until at least one master queries the object. Signals are not directly acknowledged.

Figure 1-5 CAN Addresses within a Module



MAN-005000-000
Rev. 1.1
August 1993

CRD-5000

SCSI RAID Controller User's Manual

CMD Technology, Inc.
1 Vanderbilt
Irvine, California 92718
(714) 454-0800

Copyright

Disclaimer

CMD reserves the right to make changes to this manual and the equipment described herein without notice. CMD has made all reasonable efforts to insure that the information in this manual is accurate and complete. However, CMD shall not be liable for any technical or editorial errors or omissions made herein or for incidental, special, or consequential damage of whatsoever nature resulting from the furnishing of this manual, or operation and performance of equipment in connection with this manual.

FCC Notice

Class A Computing Device: This equipment has been tested and found to comply with the limits for a Class A digital device pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. This equipment generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the instruction manual, may cause harmful interference to radio communications. Operation of this equipment in a residential area is likely to cause harmful interference in which case the user will be required to correct the interference at his own expense.

Warranty

BASIC WARRANTY—In the absence of any optional warranty or continuing provisions by formal agreement, CMD warrants its products in accordance with the schedules listed below. Purchaser hereafter mentioned refers at all times to the customer who purchased CMD product(s).

RAID CONTROLLER WARRANTY—CMD warrants RAID Controller products of its manufacture to be free from defect in material and workmanship for a period of one year from the date of shipment. During this period, if the customer experiences difficulties with a CMD RAID Controller and is unable to resolve the problem via phone with CMD Technical Support, a Return Material Authorization (RMA) will be issued. Following receipt of an RMA, the Purchaser is responsible for returning the product to CMD, freight prepaid. CMD, upon verification of warranty, will repair or replace at its option the RAID Controller in question, and will then return the product to the Purchaser, freight prepaid.

CABLE WARRANTY—All CMD provided cables are warranted for ninety (90) days from the time of shipment. Questionable cables should be returned to CMD, freight prepaid, where they will be repaired or replaced by CMD at its option and returned to the Purchaser, freight prepaid.

GENERAL TERMS—The above warranties shall not apply to expendable components such as fuses, bulbs, and the like, nor to connectors, adapters, and other items not a part of the basic product. CMD shall have no obligation to make repairs or to cause replacement required through normal wear and tear or necessitated in whole or in part by catastrophe, fault or negligence of the user, improper or unauthorized use of the product, or use of the product in such a manner for which it was not designed, or by causes external to the product, such as, but not limited to, power failure or air conditioning. CMD's sole obligation hereunder shall be to repair or replace any defective product, and, unless stated, pay return transportation costs within the United States of America for such replacement. Purchaser shall provide labor for removal of the defective product, shipping charges for return to CMD and installation of its replacement. On-site services are not a part of this warranty. Above warranties are subject to change without notice.

RETURNED MATERIAL—Warranty claims must be received by CMD within the applicable warranty period. A replaced product, or part thereof, shall become the property of CMD and shall be returned to CMD at Purchaser's expense. All returned material must be accompanied by a Return Materials Authorization (RMA) number assigned by CMD. For RMA numbers call CMD at (714) 454-0800.

THE EXPRESSED WARRANTIES SET FORTH IN THIS AGREEMENT ARE IN LIEU OF ALL OTHER WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING WITHOUT LIMITATION, ANY WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE, AND ALL SUCH OTHER WARRANTIES ARE HEREBY DISCLAIMED AND EXCLUDED BY CMD. THESE STANDARD EXPRESS WARRANTIES ARE IN LIEU OF ALL OBLIGATIONS OR LIABILITIES ON THE PART OF CMD FOR DAMAGES, INCLUDING BUT NOT LIMITED TO SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THE PRODUCT.

Return and Repair Policy

Warranty Period

The following warranty period is from the date of shipment:

CMD RAID Controller	One year
Cable	90 days
Drives	Manufacturer's warranty

Return For Credit

The allowable period of return for credit from the date of shipment is as follows

CMD RAID Controller	Less than 30 days
Cable	Less than 30 days
Drives	Not applicable

Return for Repair

CMD RAID Controller	
In-Warranty (Less than 1 year)	CMD offers a *15 working day turnaround repair service* at the cost of parts only. Defective boards will be repaired and returned to the customer within 15 working days from that date that CMD receives the board. CMD also offers two *in-warranty 24 hour expediting services:*

CMD RAID Controller	
	24 Hour Turnaround Loaner Service: Under this policy, CMD will ship a loaner in 24 hours during regular working days to the customer for a charge of $100.00 per loaner. Upon receiving the loaner, customer must return the defective board to CMD within seven (7) days for repair. CMD will repair the defective board and return the board to the customer. Customer must then return the loaner in seven (7) days after the receipt of the repaired board. Approval for loaner service is based on credit verification. *24 Hour Turnaround Swap Service:* In the case that the defective board is within the first six (6) months of the warranty, CMD, at its own option, offers a 24 hour turnaround swap service. CMD will ship the same model of the board to customer within 24 hours during working days in exchange for the defective board. CMD will swap with a new board if board is not functional upon arrival. For all other cases, swap will occur with either a new or refurbished board for a charge of $200.00. CMD does not offer swap services for boards that are purchased more than six months from the date of shipment. Customer is responsible for returning the defective board to CMD within seven (7) days after receipt of the swapped board. The remaining warranty period shall apply to the repaired or swapped board.
Out-of-Warranty (more than 1 year)	CMD offers a *15 working day turnaround repair service* at a rate of $300.00 plus parts and freight for all out-of-warranty RAID Controllers. CMD will repair and return defective controllers within 15 working days from the date that CMD receives the controller. CMD also offers an *Out-of-Warranty 24 Hour Turnaround Loaner Service:* Under this policy, CMD will ship the same model loaner in the 24 hour time frame of working days to customer for an additional charge of $100.00 plus freight per loaner. The loaner is for use by the customer during the period that the defective board is being repaired. Customer is responsible for returning the defective board to CMD within seven days after the receipt of loaner and returning the loaner in seven (7) days once the defective board is repaired and received. The approval of the loaner service is at CMD's option and based upon customer credit verification. CMD will extend warranty for a period of six (6) months on any out-of-warranty repaired board.
Cable	
In-Warranty (90 days)	Free swap
Out-of-Warranty (90 days)	Not applicable
Drives	
In-Warranty (per manufacturer)	Manufacturer charge only.

CMD RAID Controller	
Out-of-Warranty (per manufacturer)	manufacturer charge plus $100 CMD handling.

Return for Upgrade / Update

CMD RAID Controller	
In-Warranty (less than 1 year)	CMD offers a *15 working day turnaround different function upgrade service* for boards that can be upgraded to a higher function; and a *free 15 working day turnaround ECO Field Upgrade* for all its boards. CMD will *upgrade* the hardware of its board to a higher function for a charge of the difference of list prices of the original and upgraded functions. CMD will also update its board to its latest firmware release at no charge to the customer. Boards will be upgraded/updated and returned to the customer within 15 working days from the date that CMD receives the board. CMD also offers *24 hour turnaround loaner service* as stated in "Return for Repair." The remaining warranty period shall apply to the updated board. For upgraded boards, CMD will extend warranty for a period of six months.
Out-of-Warranty (More than 1 year)	CMD offers a *15 working day turnaround different function upgrade service* for boards that can be upgraded to a higher function at a charge of the difference of list prices of two functions. CMD also offers a *free 15 working day turnaround ECO Field Upgrade* for all its boards. Boards will be upgraded/updated and returned to customer within 15 working days from the date that CMD receives the board. CMD also offers *24 hours turnaround Loaner Service* as stated in "Return for Repair." There will be no warranty extension for same function firmware update. For different function Hardware upgrade, CMD will extend warranty for a period of six (6) months.
Drives	Same as in "Return for Repair."

Shipping Charges

The following shipping charges apply to all REPAIR, SWAP, LOANER, and UPGRADE UNITS.

In-Warranty	
Domestic	Freight from CMD to customer is to be paid by CMD; freight from customer to CMD is to be paid by customer.
International	All fees are to be paid by customer (including custom duty and broker fees).
Out-of Warranty	
Domestic	All fees are to be paid by customer.

In-Warranty	
International	All fees are to be paid by customer (including custom duty and broker fees).

General Conditions

All goods returned to CMD including returns for credit, swap returns, loaner returns, and evaluation returns shall remain in good condition. Any damage or alteration done by the customer will result in a rejection or additional charge to the customer.

A customer who receives a board that is defective upon arrival must contact CMD Technical Support and obtain authorization before returning the board for repair or replacement. CMD Sales personnel must be consulted for authorization of returned goods for credit and/or evaluation.

Table of Contents

Chapter 1

Introduction

This manual provides instructions on how to install, configure and use the CRD-5000 Disk Array Controller. Step-by-step instructions in the chapters titled "Installation" and "Configuration" will help you get your CRD-5000 up and running in just a few minutes. Other chapters go into greater detail on other features of the unit and possible configuration alternatives. For those of you who are new to RAID, this chapter presents a quick overview of the concept.

The Need for RAID

Fixed disk manufacturers have made great strides in designing faster and more capacious devices. Yet advancements in processor technology have far outpaced their efforts. As a result, fixed disks still act as brakes on today's turbocharged computer systems. Modern processors sometimes operate at less than full capacity or even sit idle while waiting for the fixed disk to write or read data.

Improvements in fixed disk technology have been relatively modest, because, at heart, fixed disks are mechanical devices: an actuator must physically move the read/write head across the surface of the spinning disk. Manufacturers can do only so much to accelerate the movement of this mechanical relationship.

The constraints of fixed disk technology could be tolerated if disk accessing occurred infrequently. But the trend is heading towards more disk accessing, not less. Networks and applications such as imaging and graphics are placing heavy, new burdens on fixed disk drives, emphasizing the need to bridge the gap between processor performance and input/output rates.

Data availability is another trend that is driving innovation in fixed disk technology. In many instances today, computers are the only repositories for critical business or scientific data. When a hard disk crashes, restoring the data from tape may take hours, and even then data saved since the last backup will be lost. Disk mirroring provides a high degree of data availability—but at double the cost of a non-mirrored system.

RAID (redundant array of independent disks) offers significant advantages on both fronts. If offers increased speed, because, in effect, it turns several inexpensive drives into one, big drive with multiple actuators. The controller can manipulate these actuators either in parallel to share the work on reads or writes of large files or independently to perform multiple, simultaneous reads or writes of small files. RAID improves data availability by performing parity calculations on every bit of data written to the array. This parity information is saved on the array and can be used to reconstruct the data on any drive that goes bad. Unlike mirroring, where the ratio between data drives and redundant drives is one to one, the CRD-5000 RAID controller can support a ratio of up to six data drives for each parity drive.

What Is RAID?

In simple terms, whenever the CRD-5000 RAID controller receives instructions from the host to write a block of data to the array, it "stripes" or breaks that block of data into smaller pieces and apportions the pieces among the drives in the array. In this way, each drive's actuator gets into the act, decreasing the time needed to complete the write operation. Similarly, when the host asks to read a block of data from the array, the CRD-5000 determines which drives are involved and simultaneously fetches the data from each drive, again making use of the multiple actuators at its disposal.

Several different RAID methods exist. The chief differences lie in the way data is striped and if and how parity is implemented. The industry has tacitly agreed to call these different methods "RAID Levels." The CRD-5000 supports levels 0, 3 and 5, the three most popular levels. Each level is aimed at a certain type of user, as the following table indicates.

Which RAID Level is right for you?

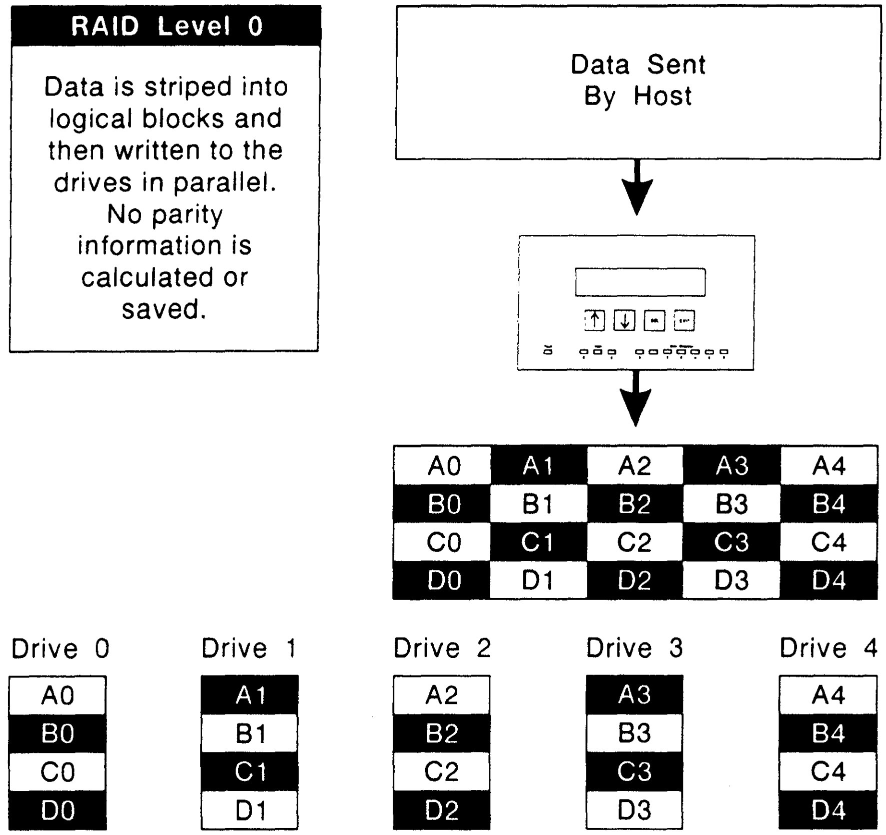
RAID Level	Best Use
0	Applications such as imaging or graphics
3	Applications such as imaging or graphics
5	Applications such as transaction processing

The next three sections will describe these RAID levels in more detail.

RAID 0

RAID 0 is somewhat of a misnomer, because this level has no redundancy at all. This means that if a drive in a RAID 0 array goes bad, all the data on the array, even the data on the remaining good drives, is useless. RAID 0's claim to fame is its ability to read and write large data files at great speed, which may make it attractive to some users who need all the speed they can get and either don't mind losing access to their data occasionally or have other ways to guarantee access.

The one RAID tool employed by RAID 0 is striping. When the host requests that data be written to the array, the RAID 0 controller breaks up the data into smaller pieces. The size of each piece is equal to the size of a logical block on the SCSI disks connected to the array. (A logical block is equal to 512 bytes on the CRD-5000.) Then the controller simultaneously writes one block to each data disk and repeats the process until all the blocks have been written to disk.



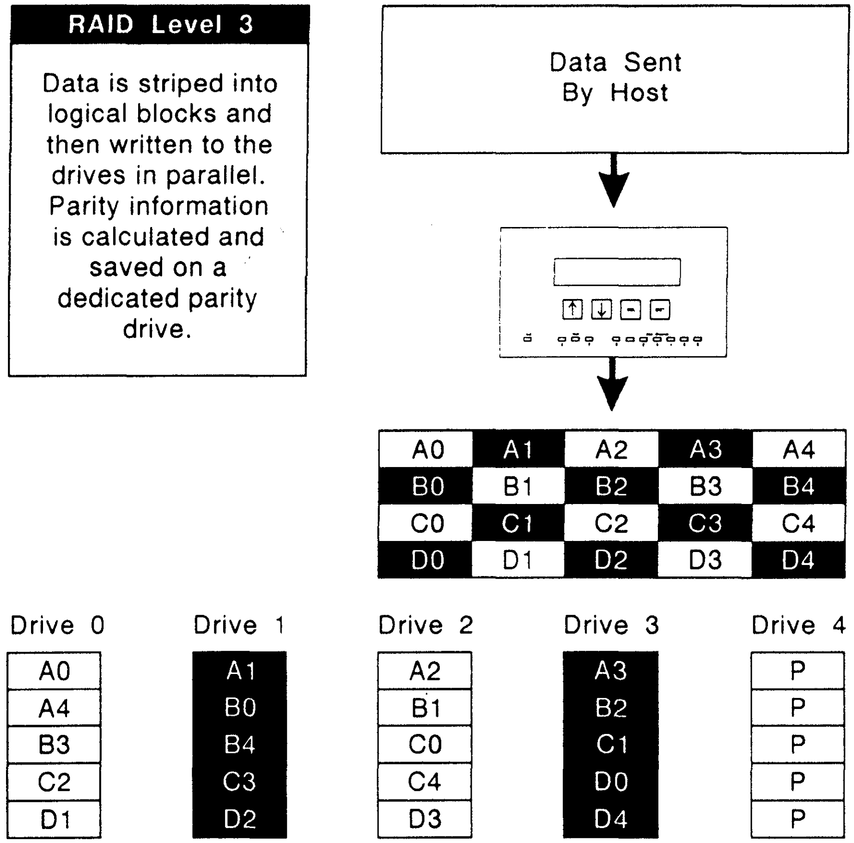
Similarly, when the host issues a read request, the CRD-5000 fetches the data from each drive in the array in parallel.

RAID 0 is designed to employ every disk drive on every input/output operation for all but the smallest data transfers. This enables the array's multiple actuators to gang up on reading or writing large files, resulting in very high transfer rates. Applications that produce very large files, such as imaging or graphics, will benefit from RAID 0's high throughput levels.

RAID 3

RAID 3 stripes data in the same way that RAID 0 does. It breaks up large chunks of data into pieces the size of the logical blocks on the SCSI disks. Then it simultaneously writes one piece to each data disk and repeats the process until all the pieces have been written to disk. The crucial difference between RAID 3 and RAID 0 is that RAID 3 calculates parity information on every bit of data written to the data disks and writes this information to a dedicated parity disk.

The parity information gives RAID 3 the ability to recover from the failure of any single disk drive in the array. When a data drive fails, the CRD-5000 uses the parity information to rebuild that drive's data onto a spare drive. Rebuilding can take place automatically or at the operator's direction. It can also take place in the background, while the controller concurrently handles all I/O requests.



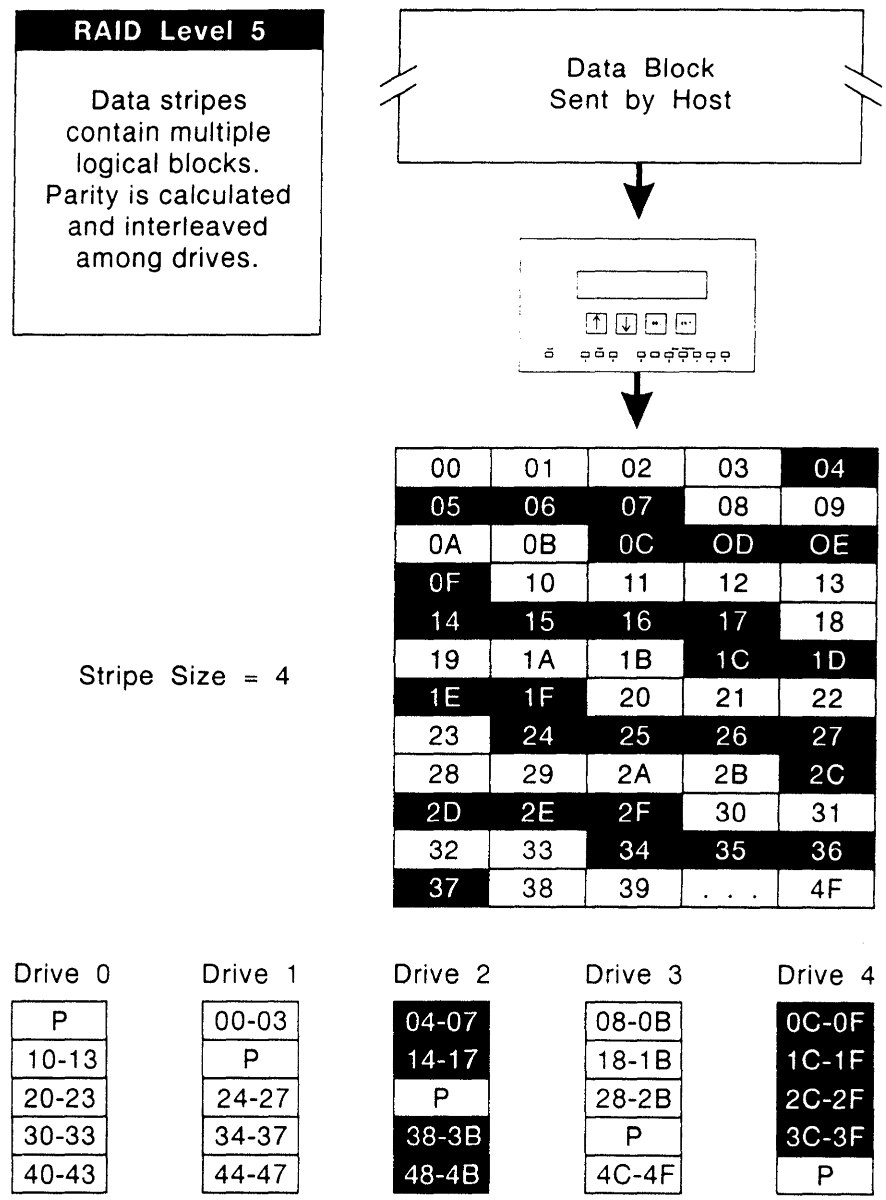
The parity drive virtually guarantees that your data will always be available. RAID cannot make a total guarantee of data availability, however. No RAID level employing parity can survive the failure of a second drive before the first failed drive has been rebuilt. The probability of a second drive failing in the time it takes to rebuild the first drive is very low, however.

This ability to rebuild data comes at a price. Because one drive in the array is set aside for parity information, the total capacity of the array is reduced by the capacity of the parity drive. For instance, in a four-drive array, the price of parity is 25 percent. In other words, only 75 percent of a RAID 3 array's total capacity is available for data storage. Since there is never more than one parity drive, the price declines as the total number of drives in the array grows. The CRD-5000 supports arrays of up to seven disk drives, which lowers the parity price to a little over 14 percent.

Like RAID 0, RAID 3 is designed to employ every disk drive on every input/output operation and take advantage of the high transfer rates possible when multiple actuators work in parallel. Unlike RAID 0, RAID 3 is truly a redundant system that can withstand the failure of one of its drives with hardly any interruption to data access. RAID 3 is best suited for inputting and outputting large files, in situations that also require a high degree of data availability.

RAID 5

RAID 5 takes a different approach towards striping and parity storage to better handle applications that barrage the system with many, small input/output operations. While RAID 0 and 3 are designed to engage all drives in the array at the same time on the *same* read or write, RAID 5 is designed to engage as many drives as possible at the same time on *different* reads and writes. This enables RAID 5 to outperform RAID levels 0 and 3 in the category of I/Os per second.



Rather than striping in units of single logical blocks, RAID 5 permits the user to select a stripe size. (The CRD-5000 supports stripe sizes from one to 255 logical blocks.) When the host sends a chunk of data to be written, the CRD-5000 breaks it up according to this stripe size and writes the pieces to one or more disks in parallel. The goal is to confine most read or write operations to one disk, with as little overlap to other disks as possible. If the system can satisfy a request for one block of data by reading drive one and for another block of data by reading drive four, then it can perform both reads simultaneously.

The parity scheme used by RAID 5 also furthers the goal of multiple, simultaneous I/Os. Instead of storing parity information on a single, dedicated drive, RAID 5 interleaves the information throughout the array. Parity information for the first stripe of data is stored on drive 0. The second stripe's parity information is stored on drive one, and so on. Ideally, then, most writes would involve just two drives—the data drive and the drive where the parity information for that data's stripe is stored. In this way, even many writes can occur simultaneously, if the system has four or more drives.

Due to its ability to perform multiple, simultaneous I/Os, RAID 5 is ideal for such applications as networks, databases and transaction processing systems, which often call for many, small I/O operations.

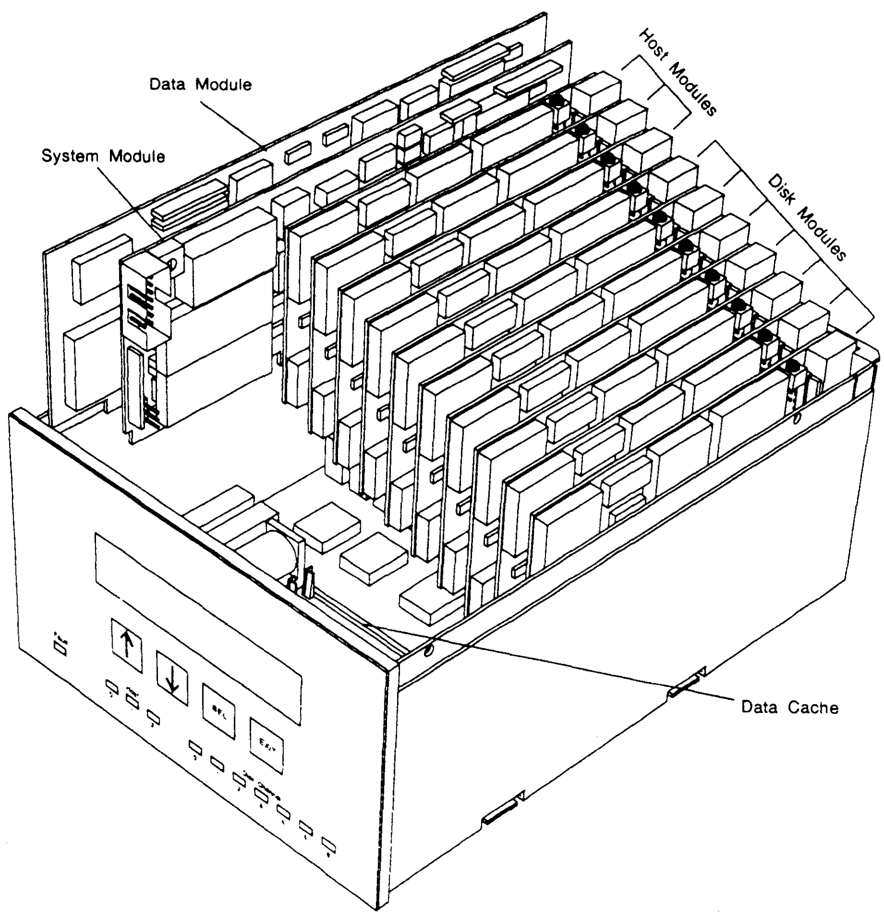
Chapter 2

Technical Specifications

CRD-5000 Disk Array Controller

Supported RAID Levels	0, 3 & 5
Host Bus Interface	SCSI-2 "Fast" compatible
SCSI Transfer Rate	10MBytes/second (synchronous) 7MBytes/second (asynchronous) (per host port)
Tag Queuing	SCSI Tag Queuing supported
Number of SCSI Devices Supported	Host Channels: Up to 21 hosts on three separate SCSI host channels Disk Channels: Up to seven SCSI disk channels, with up to four disk per channel (28 disk total)
System Architecture	CMD Distributed Processor Architecture (passive bus with multiple functional processors)
Cache Size	Up to 32MBytes (8MByte increments)
Operating Systems Supported	DEC VMS and ULTRIX, IBM AIX, SunOS/Solaris, SGI IRIX, MS-DOS, HP/UX, SCO/Unix, Novell NetWare, Windows NT
Configuration and Display	On-board utilities accessed via front-panel LCD/keypad and/or RS-232 port. Module status LEDs on front and top of unit.
MTBF/MTTR	150,000 POH, 20 minutes (LRU level)

Controller Size	8.125" x 5.75" x 3.25" (Full height 5.25" form factor with standard SCSI disk mount spacing) LCD display/keypad mounted to front of unit. 8 lbs. controller weight.
Power Required	5.0 VDC, 6 Amps (max) 12.0 VDC, 0.25 Amp (max)
Operating Temperature	5C to 50C
Relative Humidity	10% to 90%, Non-Condensing
Connector Type	SCSI: Dual row, 50-pin flat header Power: 4-pin Molex Serial: TBD



Components

All CRD-5000 units contain one data module and one system module. The standard configuration includes one host module and five disk modules. The CRD-5000 can accommodate up to three host modules and up to seven disk modules.

DATA MODULE: The data module controls all data flow within the CRD-5000's internal bus architecture, as well as performing cache management and data reconstruction. The data module also contains a battery-backed write data buffer to speed disk writes and assure data integrity in case of power or host interface failure during write operations.

SYSTEM MODULE: The system module acts as the CRD-5000's intelligent system manager and overall command and status processor. The system module performs built-in test and downloads microcode for each functional module at power-up as well as controlling the LCD and serial port. The system module is capable of dynamically reconfiguring the CRD-5000 to maintain maximum reliability and performance in case of SCSI disk or channel failure.

SCSI MODULES: The CRD-5000's host channel and disk channel SCSI modules are identical. At power-up, the system module downloads microcode which determines which task each SCSI module is to perform. When inserted in a host channel slot, the SCSI module manages all interface tasks unique to that host channel. When inserted in a disk channel slot, the SCSI module manages all command and data processing tasks unique to that channel. Each SCSI module acts as a standalone functional module with its own CPU and memory.

DATA CACHE: The CRD-5000 uses industry standard 4MByte SIMMs for data cache. Each CRD-5000 can accommodate up to 32MBytes of data cache, expandable in 8MByte increments. The SIMMs should be rated at 80 nanoseconds or faster.

Features

The flexibility of the CRD-5000 permits you to design a RAID system tailored to your particular needs today, while maintaining an easy upgrade path to enhance the capacity and performance of the system in the future.

- The CRD-5000 supports as many as 28 disks arranged in four ranks, each with a separate logical unit number.
- Ranks are independent of each other and can be set to different RAID levels.
- Disk drives can have different capacities, speeds and manufacturers.
- Spares, whether they are warm or hot, can be connected to any disk channel in the system.
- The controller can be connected to three different host channels at the same time.
- The data cache memory employs a highly efficient cache algorithm and is expandable to 32 megabytes, using commonly available 4MB SIMM memory modules.
- Logical channels are not slot-specific. When power is off, you can change the *physical* channels by swapping a disk channel module from one slot to another or by moving a cable from one disk channel module to another. The CRD-5000 will automatically realign its logical channels at power-up to reflect these changes.

Chapter 3

System Integration

The CRD-5000 is designed to give you wide latitude in constructing, maintaining and upgrading a high-performance RAID system. Rather than bind you to a single configuration with specific parts, the CRD-5000 can be installed in dozens of different configurations with disk drives from any manufacturer. The CRD-5000 is also highly scaleable. As your storage demands grow, you can add drive ranks (up to a maximum of four, containing as many as 28 disk drives). Or add host channels (up to a maximum of three) to permit additional computers to share the CRD-5000. You can add a hot or warm spare to an already functioning rank or increase the size of the data cache by inserting industry-standard SIMMs. But to take full advantage of this flexibility, the CRD-5000 and the drives it controls must be part of a carefully designed system—including an enclosure—which takes into account such issues as cooling, cabling, termination and SCSI bus isolation.

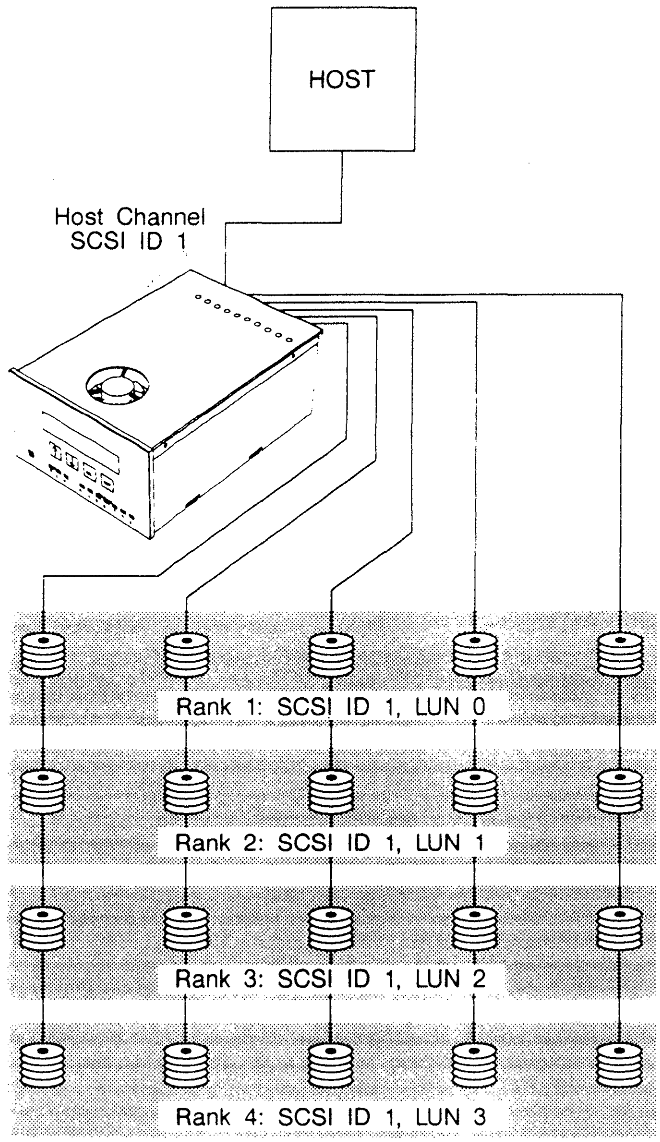
System Configuration Issues

This section discusses some, but certainly not all, of the many RAID system configuration possibilities associated with the CRD-5000. Please call CMD Technology if you have questions about specific applications not covered in this section.

Rank, LUN and SCSI ID

An understanding of the relationships among rank, logical unit number (LUN) and SCSI ID is key to properly configuring a RAID system. A rank is a group of drives, each on a separate disk channel, that are either part of a RAID set or designated as spares for that RAID set. All the drives on a rank, whether they are active members of a RAID set or spares, must have the same SCSI ID. Each rank connected to the CRD-5000 is assigned a unique logical unit number (LUN).

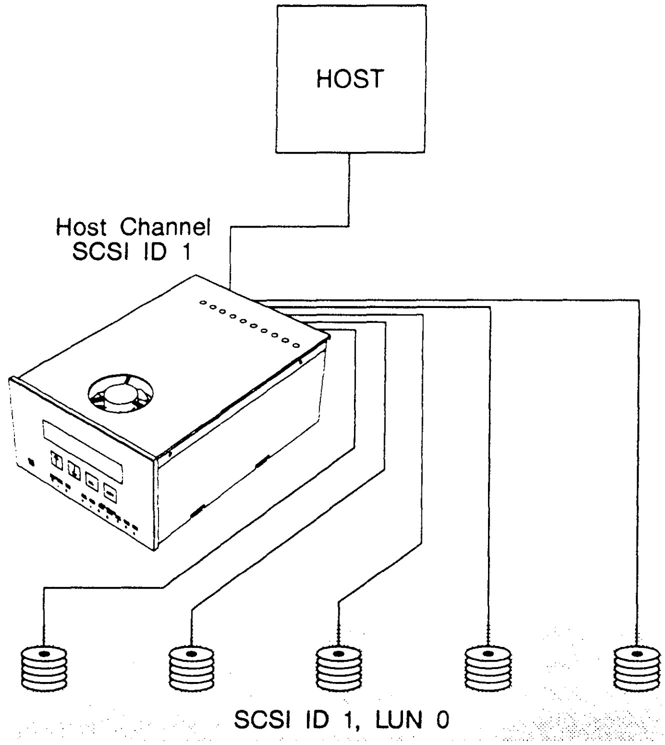
From a host's point of view, a rank appears to be a single, large disk drive. A host would access a rank by first addressing the SCSI ID of its CRD-5000 host channel, then the LUN associated with the desired rank. The following figure illustrates the concept of rank, LUN and SCSI ID on a single-host, multiple-rank configuration.



The rank addresses are from the host's point of reference.

Single Host, Single Rank

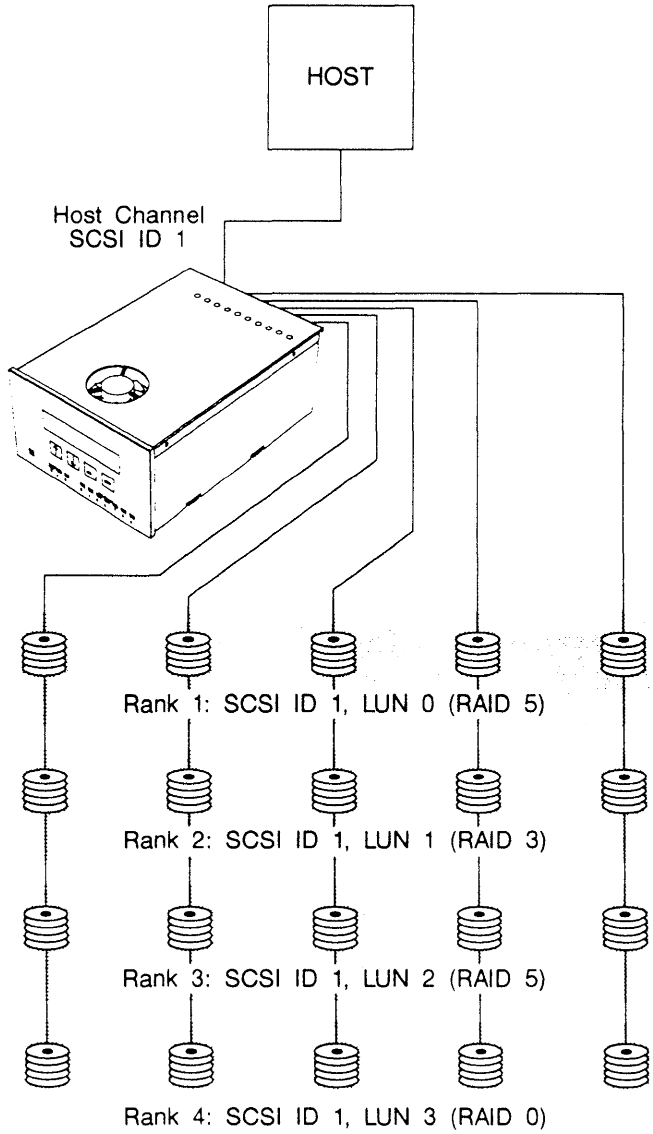
The following diagram shows a basic one host, one rank configuration.



Each drive in the rank must be connected to a separate disk channel. While this configuration is for a five-drive rank, the CRD-5000 will accept as many as seven disk drives on each rank—as long as there is a disk channel module for each drive. All the drives on the rank must have the same SCSI ID. In this case, the host computer is connected to a CRD-5000 host module that has a SCSI ID of one; therefore, the host will recognize the rank by the address SCSI ID 1, LUN 0.

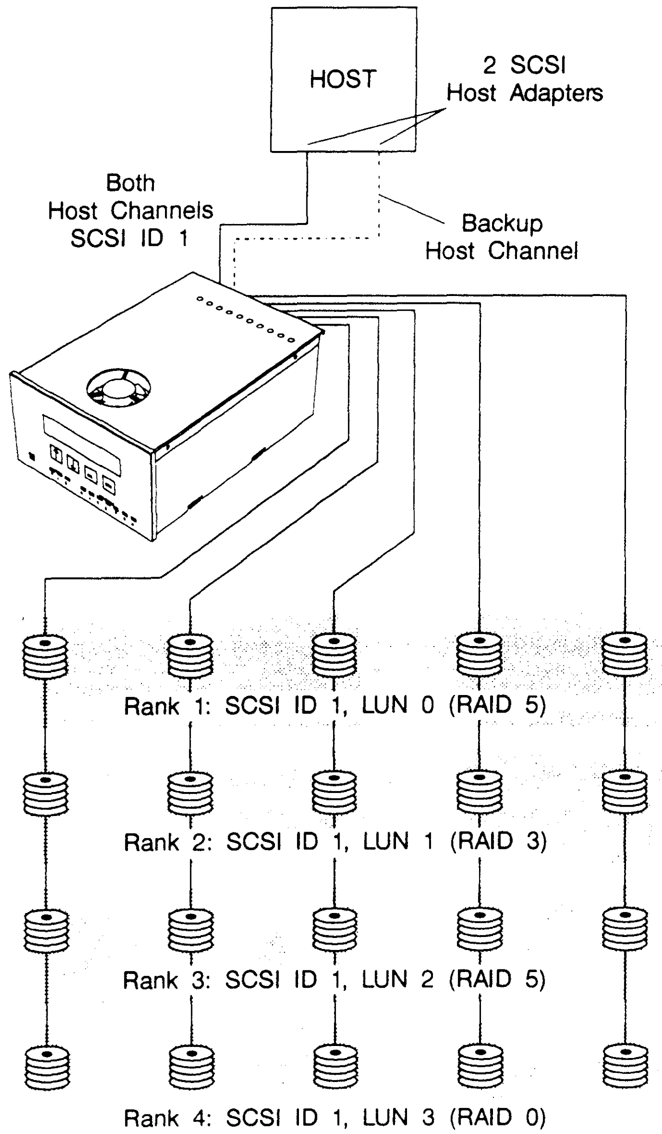
Single Host, Multiple Ranks

This diagram illustrates a configuration with a one host accessing four ranks. Each rank appears to the host as a separate device. Since each "device" has the same SCSI ID, they are differentiated by logical unit number. Notice that each rank can be configured for a different RAID level.



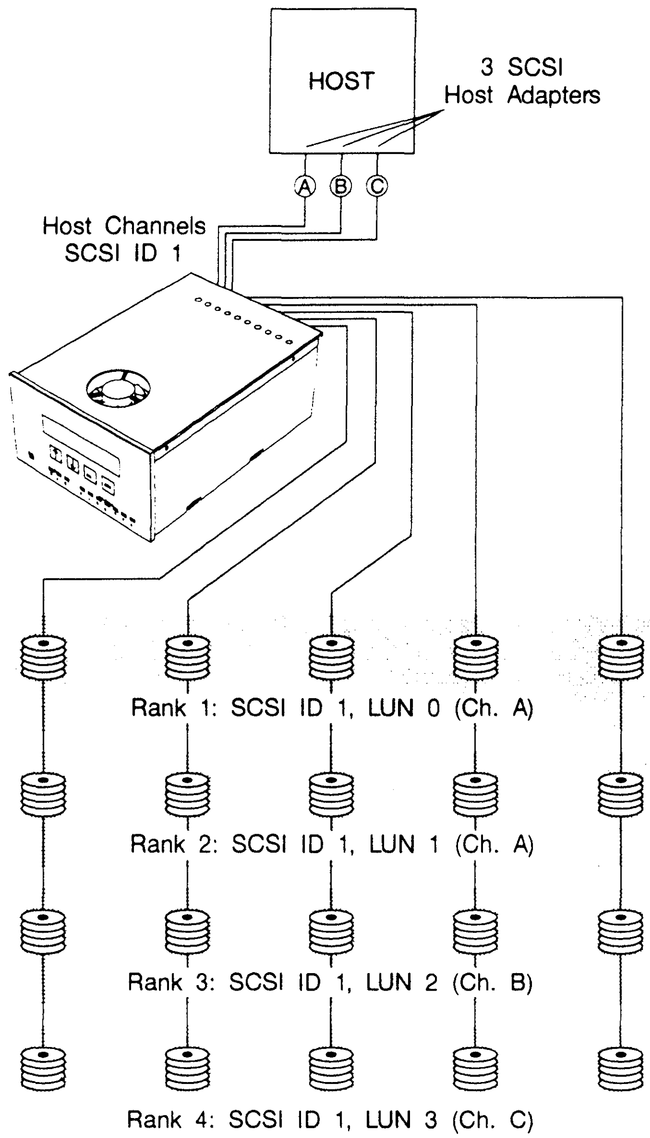
Single Host, Multiple Host Channels for Reliability

In this configuration, the host is equipped with two SCSI host adapters, each of which is connected to the CRD-5000 on a separate host channel. One host channel is idle until the other host channel fails, at which time a script or driver written by the user would shunt data transfers to the backup channel.



Single Host, Multiple Host Channels for Performance

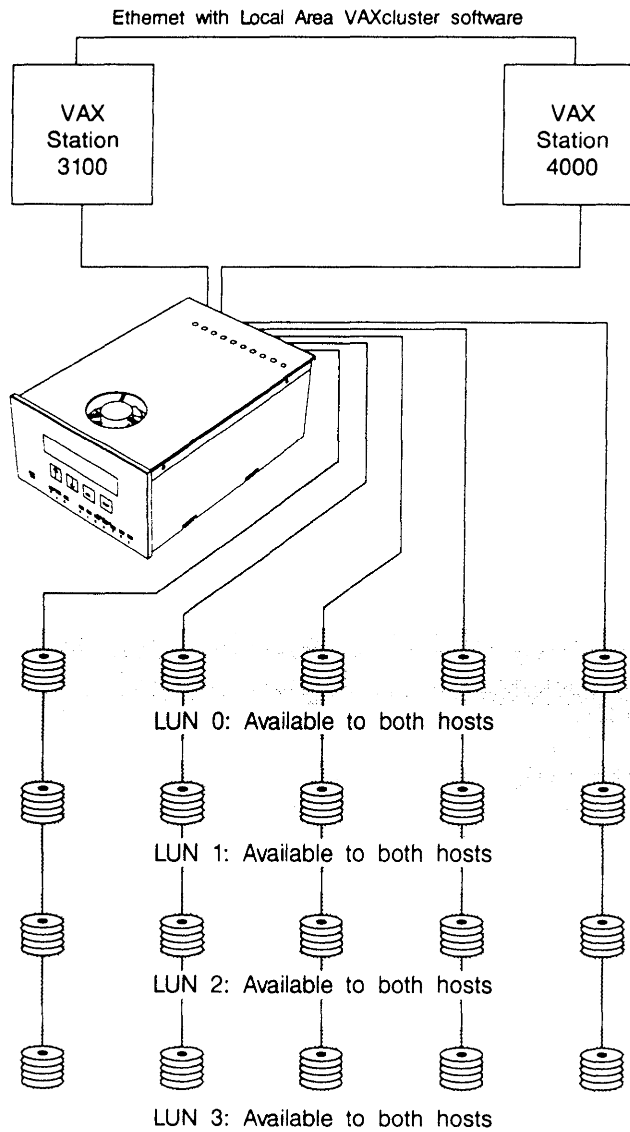
In this example, the host is equipped with multiple SCSI host adapters, each of which is connected to a separate host channel as in the previous example. The difference in this configuration is that all the host channels are operational and assigned to a specific rank. Host channel A transfers data from ranks one and two only. Host channel B transfers data from rank three only, and so on.



The purpose of this configuration is to break up the data transfer bottleneck that can develop when you have multiple disk channels transferring data from *multiple* ranks to the CRD-5000, which must turn around and transfer all this data to the host on *one* host channel. By dedicating a host channel to no more than one or two ranks, you can lighten the data transfer load that each channel must bear, effectively speeding up performance on the system as a whole.

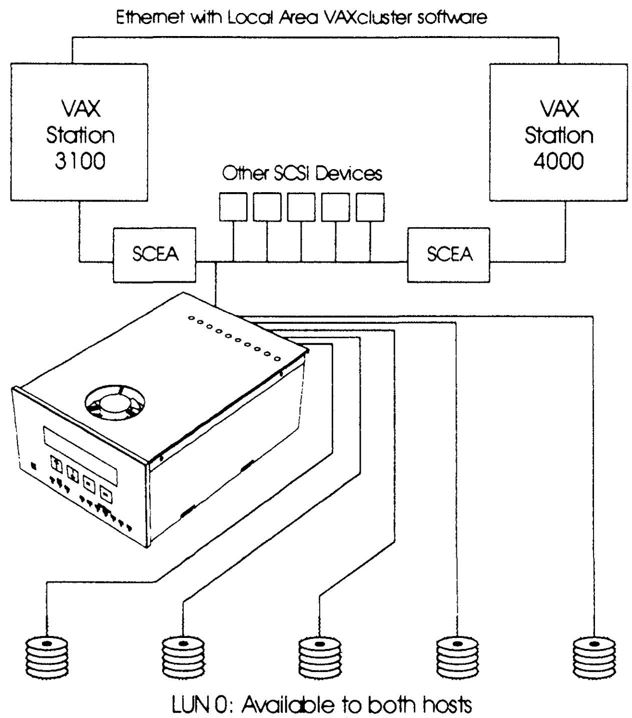
Multiple VAX Stations, Channels and Ranks

The next diagram shows a configuration with two VAX stations, each connected to the CRD-5000 on a separate host channel and to each other on an Ethernet. Because of the unique capabilities of the VMS operating system, the two hosts in this configuration can share data on every rank. VMS makes sure that the hosts do not use the same data at the same time.



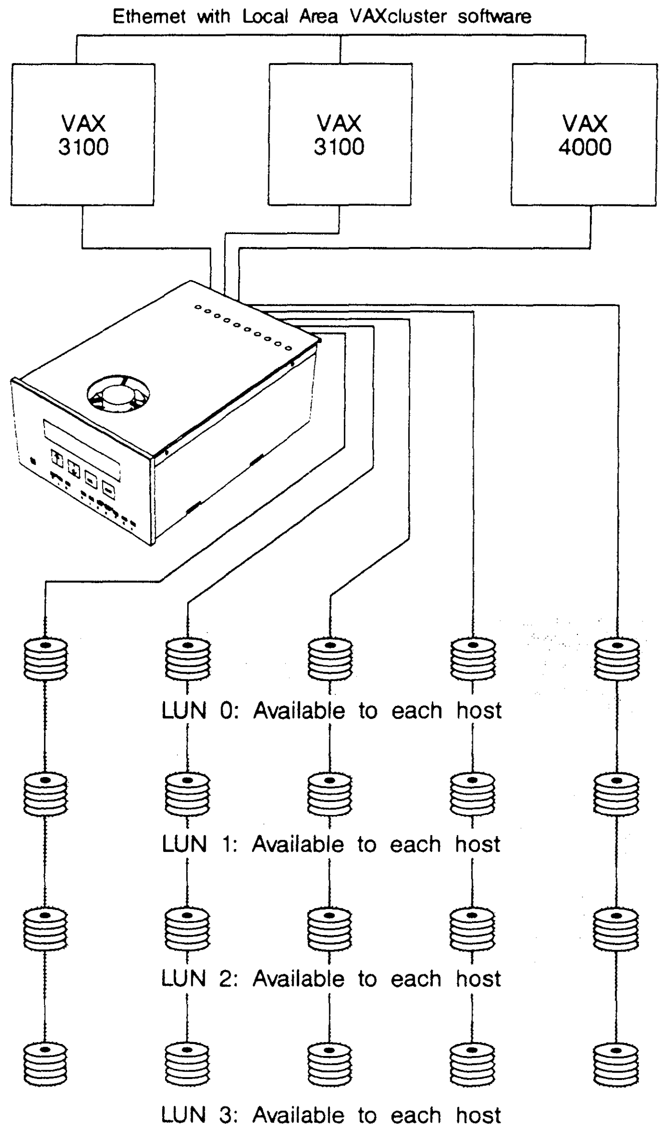
Multiple VAX Stations with SCEA Boards

Like the previous example, this configuration permits two VAX stations to share data on a rank. The difference here is the introduction of CMD Technology SCEA (SCSI Cluster Expansion Adapter) boards. The SCEA boards permit both VAX station hosts to share one host channel to the CRD-5000. More importantly, up to five additional SCSI devices may be connected to the SCSI bus between the two SCEA boards and shared by the hosts. Note that the CRD-5000 will support only one rank per host channel in this configuration. You may increase the number of shared ranks in this configuration to three by adding two host channels to the CRD-5000 and removing two of the other SCSI devices between the SCEA boards. Please consult the SCEA manual for more information on configuring such a system.



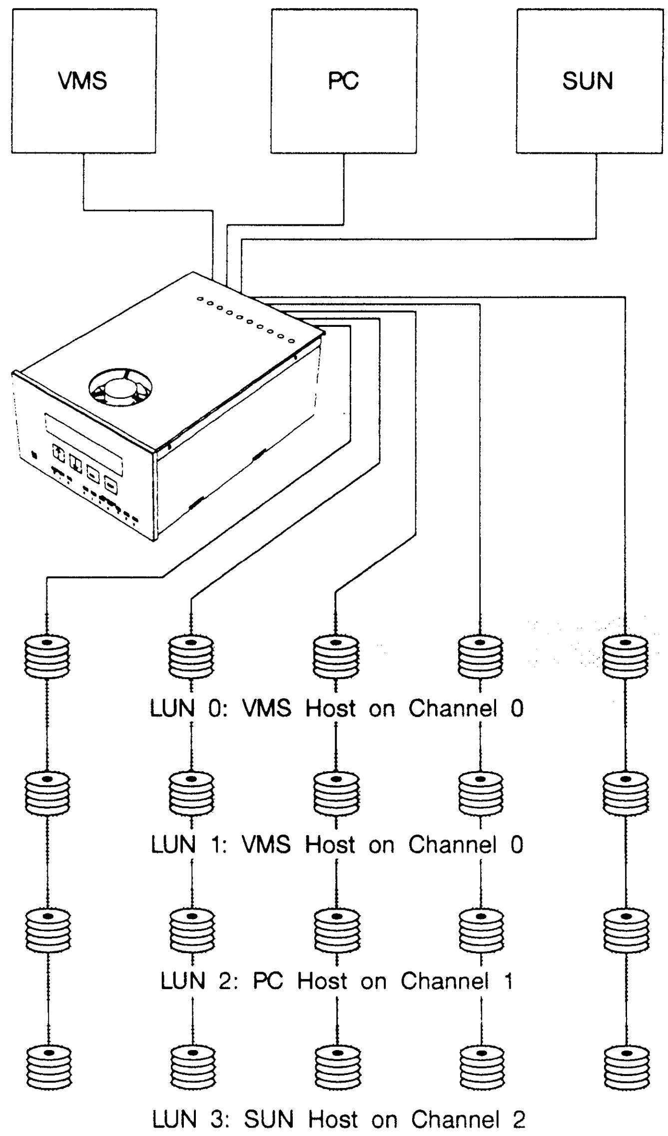
Multiple VAX Processors, Multiple Ranks

As this diagram illustrates, you can connect up to three VAX processors, each on its own host channel, to the CRD-5000. The host computers can be any combination of QBUS, VAXBI bus, or UNIBUS, as long as each contains a SCSI host adapter. Similarly, a DSSI computer would need a DSSI to SCSI adapter. As long as the ranks controlled by the CRD-5000 are of the same allocation class, the hosts will be able to share data from all of them. This sharing of data is made possible by the VMS operating system.



Multiple Hosts, Unlike Operating Systems

Hosts do not need to have the same operating system. In this configuration example, there are three hosts, each with a different operating system. They must be connected to separate host channels and cannot share data. Each rank is dedicated to a particular host.

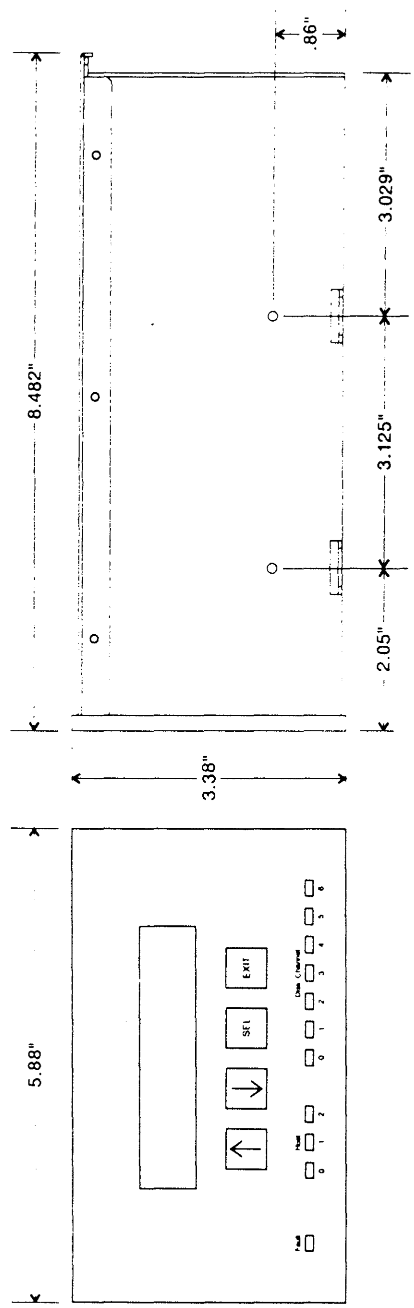


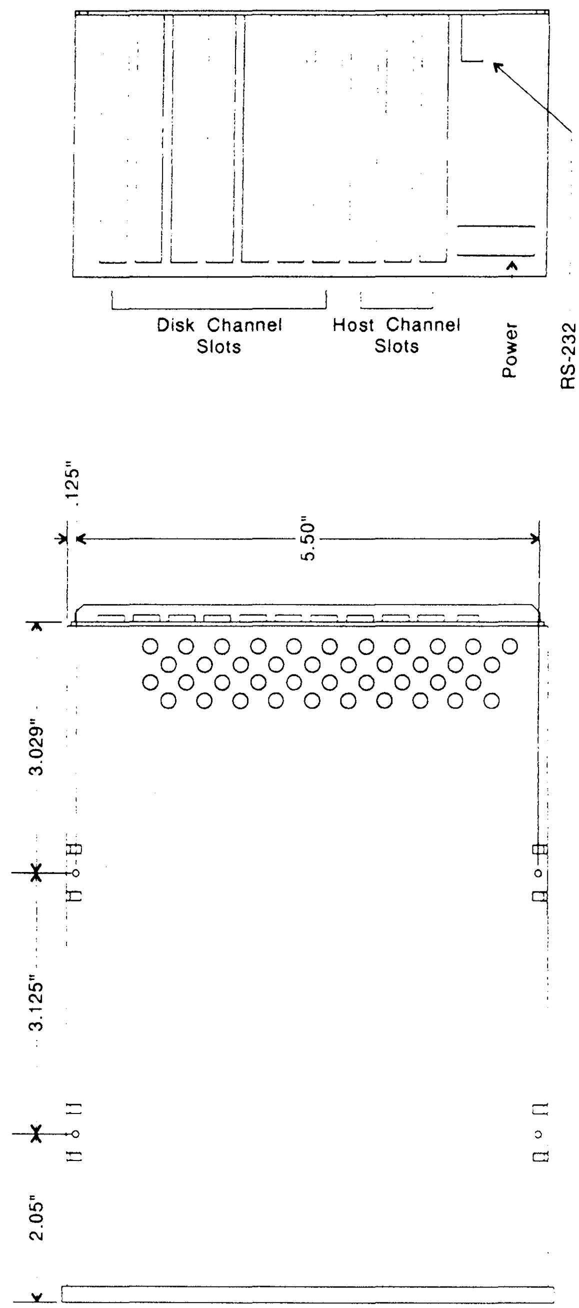
Installation Issues

This remainder of this chapter covers the issues that you must consider when designing a RAID system enclosure.

Dimensions

The following outline and mounting drawings show the size of the CRD-5000 and the location of mounting screw holes.

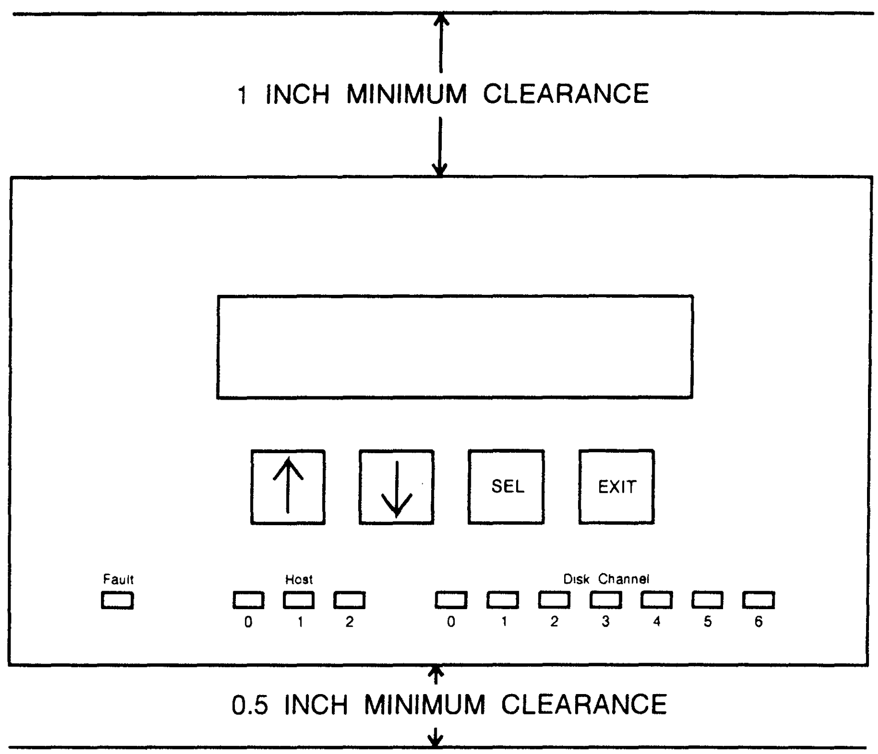




Disk Channel Slots
Host Channel Slots
Power
RS-232
.125"
5.50"
3.029"
3.125"
2.05"

Cooling

The CRD-5000's fan is more than sufficient to maintain a safe operating temperature—provided there is adequate ventilation space both below and above the unit. CMD recommends a minimum vertical clearance of 0.5 inch below and 1.0 inch above the box.

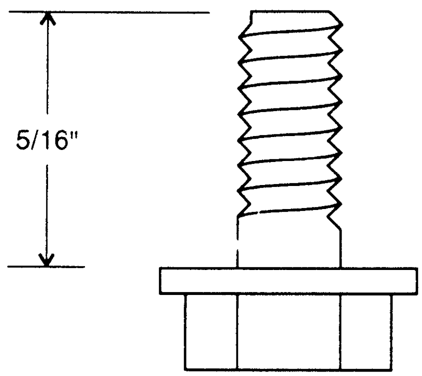


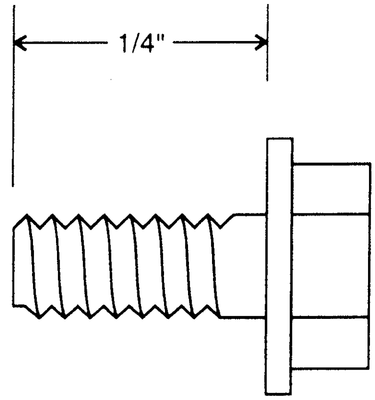
If the CRD-5000 and the disk drives it controls are housed in the same enclosure, make sure the enclosure has its own ventilation system, preferably a fan and vent system that pulls fresh air through the front and exhausts hot air through the back.

A two-level temperature sensor will monitor any heat build-up inside the CRD-5000. When the temperature reaches the the first level, the alarm will sound every two seconds. In such an event, you should take immediate steps to cool the unit or shut the system down in an orderly fashion. If the temperature climbs to the second level, the alarm will sound twice a second. The CRD-5000 will complete any pending I/O activity and then accept no more instructions from the host. You may silence the alarm by simultaneously pressing the "↑" and "↓" buttons on the front panel.

Screw Length

To secure CRD-5000 secured within an enclosure, drive screws through the enclosure's rails and into the four holes on the bottom of the CRD-5000 or the four holes drilled on the side. Screws driven into the bottom should be no longer than 5/16 inch. Screws driven into the sides should be no longer than 1/4 inch. Longer screws can damage the circuitry inside the box.





Power Requirements

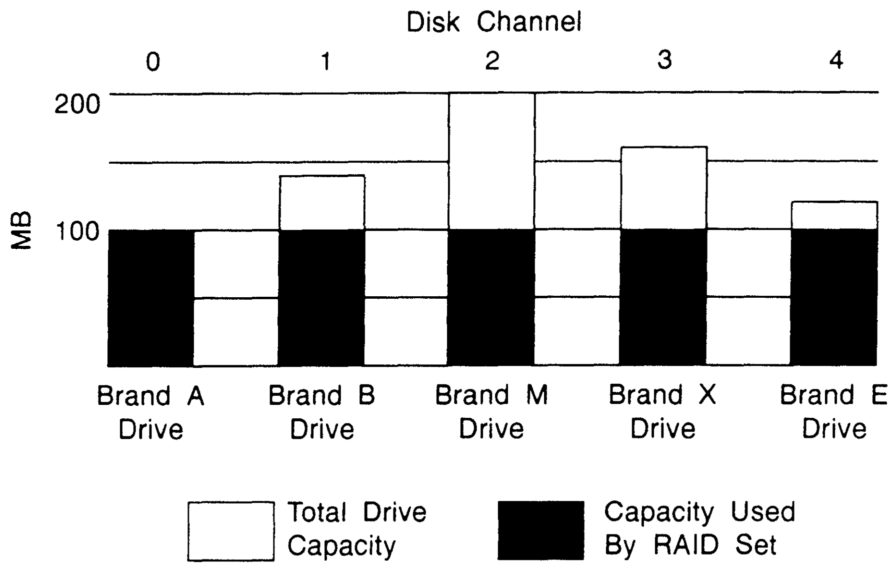
The CRD-5000 is equipped with a standard power connector, which is accessible from the back of the unit. The connector is keyed to assure proper pin alignment. The following chart shows the pin assignments and power specifications:

Pin	DC Voltage	Current
1	+12V ±5%	0.25 Amp Max
2	+12V Ret	
3	+5V Ret	
4	+5V ±5%	6 Amp Max

While the redundancy of a RAID system assures data availability in the case of a disk drive failure, RAID offers no protection against power failures. For this reason, a truly reliable RAID system should incorporate special power sources, such as redundant power supplies or a battery backup.

Disk Drives

The CRD-5000 is designed to accept SCSI disk drives of any size and from any manufacturer. While drives may be mixed and matched on the same rank, the capacity of the resulting RAID set will be dependent on the smallest disk drive. In other words, if a 100MB hard drive is combined with four 150MB drives, the RAID set will treat all the drives as if they had 100MB capacities.



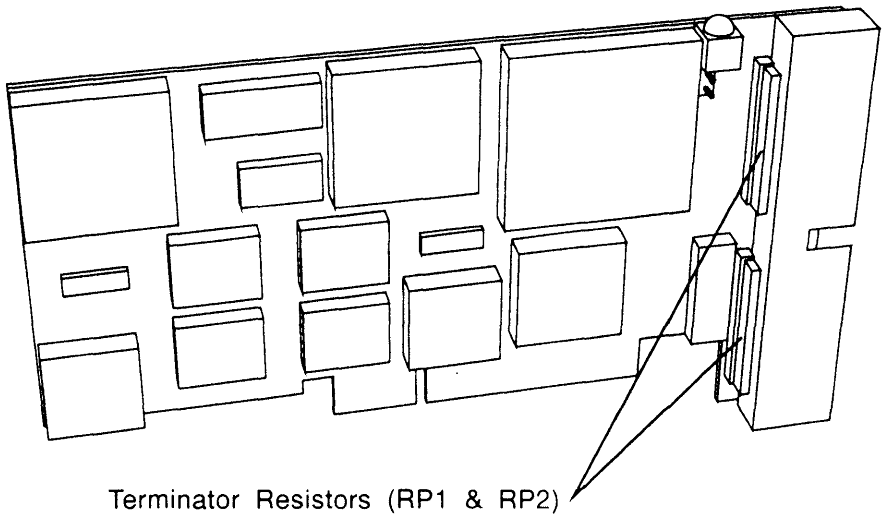
If you wish to enable tag queuing, all drives on the rank must be identical. The CRD-5000 will not support tag queuing, if the drives in a rank are made by different manufacturers. This is because all drive manufacturers do not implement tag queuing in the same way.

Terminator Resistors

SCSI signals must be terminated at both ends of the SCSI bus. There are several SCSI buses involved in the CRD-5000: one bus for each host channel and one bus for each disk channel. Each bus must be properly terminated with terminator resistors. The CRD-5000 is always shipped with terminator resistors installed on all host and disk modules. For a typical installation—with the CRD-5000 at one end of each SCSI bus to the host and at one end of each SCSI bus to the disk drives—all you need to do is see that the host and disk buses are terminated at the *other* end.

There are some possible installations that will require the removal of terminator resistors on one or more modules in the CRD-5000. This is particularly true of host modules. From the host's point of view, the CRD-5000 is just like any other SCSI device that can be daisy-chained on a SCSI bus. If the CRD-5000 shares a SCSI bus with other devices and is not at the end of the bus, the terminator resistors on the host module connected to that bus must be removed.

The CRD-5000's terminator resistors are always located at RP1 and RP2, whether the board is a host module or a disk module.



CMD has used active terminators in the CRD-5000. When running in synchronous mode at seven megabytes per second or faster, the SCSI device at the other end of the bus may require special terminators.

Removable Drives

At some point during the life of your RAID system, you probably will need to replace a disk drive on a rank. In a simple RAID enclosure built to accommodate only one rank, the disk drives might be attached directly to the enclosure chassis. To replace a drive, you would shut off power to the whole system, open the enclosure, remove the drive with a screwdriver and replace it with another drive. While this is inconvenient, it does minimize the risk to your data.

A more sophisticated RAID enclosure will accommodate multiple ranks of drives and permit "hot plugging" or "hot swapping" (the ability to replace and rebuild a drive while power is turned on). For such an enclosure, the question of designing a scheme for removing drives becomes more complicated. First of all, you should use external termination rather than the terminator resistors on the last drive of each drive channel. In other words, extend the SCSI cable beyond the last drive of the disk channel and terminate with an external terminator. In fact, CMD recommends that external termination be the rule, not the exception, in a RAID environment.

To facilitate drive removal, the enclosure should house drives in special containers that slide into backplane power and SCSI connectors. Two issues must be considered when designing such a system: SCSI bus isolation and stub length.

SCSI bus isolation becomes an issue, when you intend to permit "hot plugging" or "hot swapping" on a multiple rank system. The CRD-5000's disk channels are SCSI buses, and like all SCSI buses, electrical signals carrying data or SCSI communications travel up and down the bus past each device connected to it. If you remove a device from the bus while the bus is active—as you would in a "hot swap"—you disrupt the signal along the entire bus, resulting in a bus failure. Such a scenario is particularly catastrophic in a multiple-rank RAID environment, because a disk channel failure will cause every rank that includes an active drive on that channel to be degraded.

To isolate the SCSI bus from such disruptions the RAID enclosure should have power switches at each drive connection so that an individual drive can be powered down before removing it from the bus. Another solution would be to incorporate a tri-state buffer between each drive container and the bus.

A drive container scheme will entail branching off of the SCSI bus at the backplane to the drive's SCSI connector. Such a branch is known in SCSI parlance as a stub and must be kept as short as possible. The SCSI-2 standard specifies that stub lengths should not exceed 0.1 meters (3.9 inches).

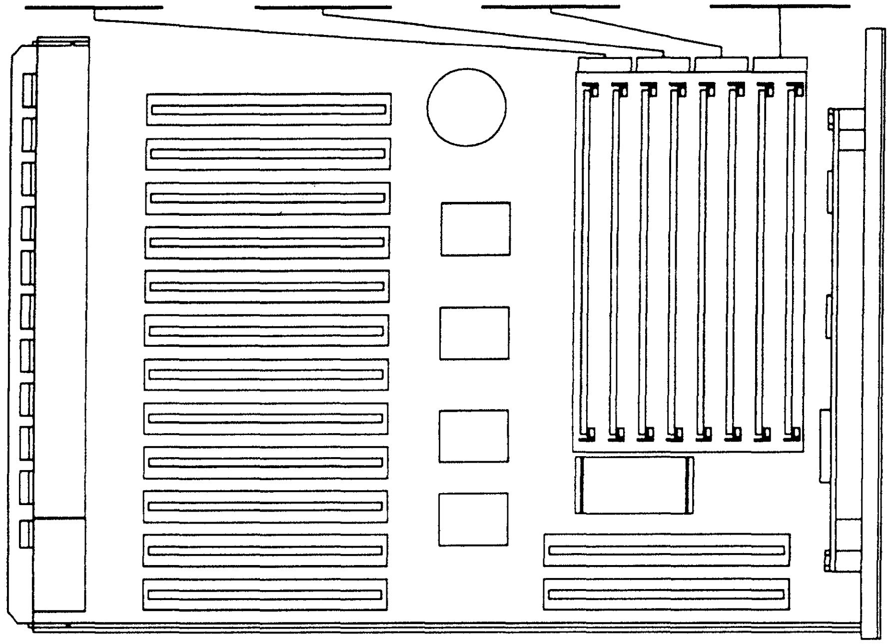
Cache

The CRD-5000 will accept up to 32 megabytes of data cache memory. The cache stores data read from the disk in case the host requests it again, in which case the CRD-5000 can fetch the data from high-speed memory rather than from disk. In designing the cache, CMD has employed a set associative algorithm, featuring as many as 128 sets and a granularity of one. This produces a highly efficient cache that avoids cluttering itself with superfluous data and retains useful data longer.

To activate caching, the CRD-5000 needs at least 8MB of RAM. Additional memory may be added in 8MB increments. The RAM should be in the form of 30-pin, 9-bit, 4MB SIMMs (single in-line memory modules), which are readily available from a variety of third party sources. The SIMMs should be rated at 80 nanoseconds or faster. Two SIMMs are necessary for each upgrade increment.

Install cache modules in 8MB increments, from back to front.

When installing cache, be sure to power down the system. Then insert the SIMMs two at a time, from back to front. For the first 8MB of cache, install SIMMs in slots J22 and J21. For 16MB of cache, install the second two SIMMs in slots J20 and J19. Continue filling the cache bank in the same direction, from back to front, for cache sizes of 24MB and 32MB.



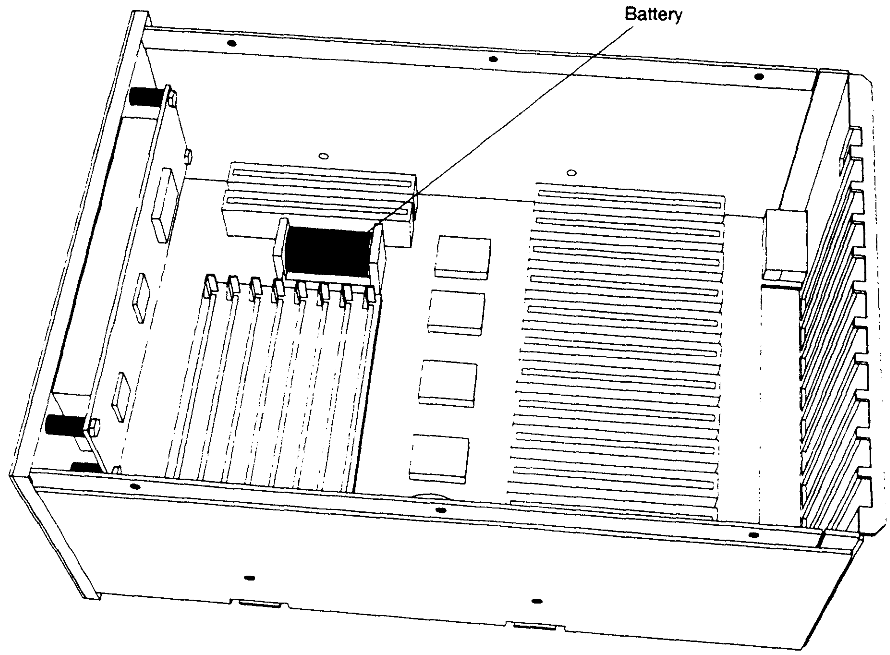
The CRD-5000 is "cache-ready," meaning that the unit will configure itself automatically to use cache as SIMMs are installed.

Battery

RAID disk writes are a cumbersome and time-consuming process, because they involve a *read* from two disks before the write can take place. Whenever the host requests that data be written to the array, the CRD-5000 must read the old version of the data from one disk and its associated parity information from another disk. Updated parity information is calculated using the new data and written to the parity disk, while the new data is written to the data disk. To speed up this process—at least from the host's point of view—the CRD-5000 will put data for a write operation directly into its buffer and immediately tell the host that the write operation has been completed. This frees the host to perform other tasks, while the CRD-5000 processes the data in its buffer and completes the write operation.

The system will function with a weak or missing battery, but performance will suffer noticeably.

The CRD-5000's battery backup protects any data in the buffer from sudden power interruptions. If system power is shut down before the CRD-5000 can complete a write operation, the battery will supply auxiliary power to the buffer's RAM until primary power is restored and the write operation can be completed.



The CRD-5000 drains current from the battery only when primary power is shut off. So for systems that are powered up nearly all the time, the useful life of the battery should approach its shelf life: approximately ten years. The battery on a system that is powered down more often will have a shorter life. The unit will periodically run a battery test and flash an error message on the front-panel LCD if the current is getting weak.

When necessary, replace the battery with a 3.6 volt, ½-AA-size, lithium battery. This is a common battery, which should be available at any electronics supply store.

Cabling

Proper SCSI cabling in all disk drive and host channels is essential to the operation of the CRD-5000. The SCSI-3 proposal recommends that the total length of a SCSI cable on any single bus should be no longer than three (3) meters or (9.9 feet) when the transfer rate is 10 MBytes/second. In the case of the CRD-5000, this means that the total length of any host or disk drive channel cable must be three meters or less. When the CRD-5000 is housed in an enclosure, be sure to include the distance of any internal cable in your calculation. Stubs, which are branches from the SCSI cable, must be kept shorter than 0.1 meters (3.9 inches).

In addition to meeting the length limitation, cables should adhere to the following specifications.

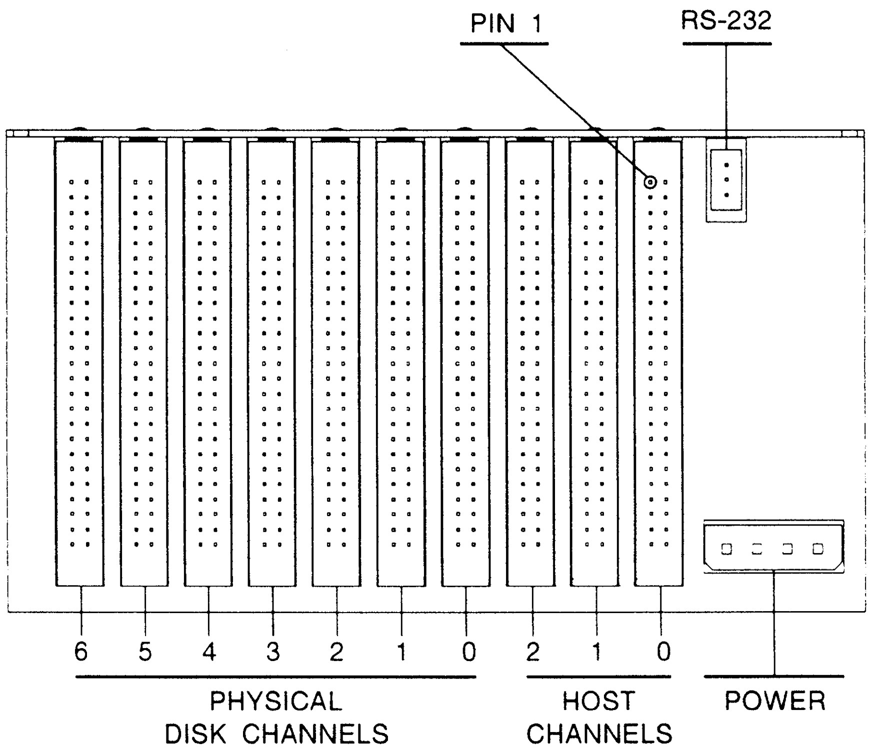
Cable Specifications

Characteristic Impedance	No less than 90 ohms and no greater than 140 ohms
Signal Attenuation	.095 dB maximum per meter at 5 MHz
Pair-to-pair Propagation Delay Delta	.15 nanoseconds maximum per meter
DC Resistance	.23 ohms maximum per meter at 20 degrees C

Making the Connections

The CRD-5000 comes with a maximum of seven disk channels and three host channels. The disk channel ports are the first seven slots from the left, as viewed from behind the box. The first slot on the left is for disk channel six. Channel five is the second from the left and so on to channel zero. (The CRD-5000's numbering system begins with zero rather than one.) Host channel number two is the eighth slot from the left. Host channel number one is slot nine, and host channel zero is slot 10.

The CRD-5000 distinguishes between physical and logical disk channels. This means that you can swap cables and disk modules (the physical channels), and the CRD-5000 will automatically recreate its internal route map (the logical channels) at power up. Make sure power is turned off before you do any swapping.



Remember, the disk channel numbers assigned to the slots are just *physical* designations. Once you cable the drives up and create a RAID set, the CRD-5000 will assign *logical* channel numbers to each drive. It is these logical numbers that really mean something to the system. Initially, the logical and physical numbers will be the same. But the CRD-5000 permits you to swap cables and disk modules. You could, for instance, take the cable from disk drive four and connect it to disk drive five, while moving drive five's cable to drive four. Or you could move disk module one to slot six.

CAUTION: **Make sure the power to the CRD-5000 is turned off before you swap cables or disk modules. Failure to turn the power off, could result in lost data.**

As a result of all this swapping, the physical and logical channel numbers no longer will be the same. Continuing our example, logical drive four will be on channel five. Logical drive five will be on physical channel four. Logical drive one will be on physical channel six. Fortunately, you don't have to keep track of which logical drive is on what physical channel. The CRD-5000 does this for you automatically. You can always ascertain the location of each logical drive on a monitor attached to the RS-232 port.

All the drives in each rank must have the same SCSI ID number. Follow the drive manufacturer's directions for setting the SCSI ID. Set all the drives in the first logical unit (LUN 0) to SCSI ID zero. If you are going to have more than one logical unit (rank), set all the drives in the next logical unit (LUN 1) to SCSI ID one, and so on up to four LUNs. The CRD-5000's disk modules themselves are set at the factory to SCSI ID seven. Use the following table as a guide to setting SCSI IDs.

LUN Number	SCSI ID
LUN 0	0
LUN 1	1
LUN 2	2
LUN 3	3

Next, use 50-pin, ribbon cables to connect each drive in the logical unit to a disk module. The colored stripe on the cable should align with pin one at the top of each CRD-5000 socket (the sockets are keyed to assure a pin 1 match). Make sure the terminator resistors are installed on these drives if there is to be only one logical unit. For additional logical units, attach additional drives to each SCSI cable. You can attach up to four drives on each cable, and you can have different numbers of drives on each logical unit. Remove the terminator resistors from all but the last drive on each cable.

Finally, connect power cables to each drive and to the CRD-5000. You are now ready to configure the system and create a RAID set.

Chapter 4

Configuration

After connecting the cables and supplying power to the CRD-5000 and its drives, the next step is to configure the system in preparation for creating a RAID set on each logical unit. The CRD-5000 will prompt you via the front-panel display for information about the RAID level of the logical unit, the stripe size if you select RAID level five and the status of each physical channel in the array.

This chapter covers only those front-panel functions that are essential to configuring a RAID set. Chapter 5 gives a full description of the many other functions available through the front-panel display.

You may use the CRD-5000's default RAID set parameters and proceed immediately to creating your RAID set.

If you are satisfied with the CRD-5000's default RAID set parameters, you can skip to the "Creating a RAID Set" on page 4-4. The default settings are RAID level five, 128-block stripe size, all occupied disk channels included as online members of the RAID set, and no spares.

If you have a monitor at your disposal, you can connect it to the CRD-5000's RS-232 port and view an expanded version of the information presented on the front-panel LCD. You can also use a keyboard to move around the configuration menus and input your selections. For more information about the monitor utility, see Chapter 6.

The Configuration Menu

A menu of configuration options will appear on the display when you press the "SEL" button after the completion of the self-test. By using the "↑" "↓" "SEL" and "EXIT" buttons, you can move through the various menu levels and select desired parameters. The buttons work as follows:

The up arrow button moves up within the same menu level. This button also moves up through the possible values for stripe size; holding the button down will cause the value to increase more rapidly.

The down arrow button moves down within the same menu level. This button also moves down through the possible values for stripe size; holding the button down will cause the values to decrease more rapidly.

The "SEL" button selects a menu level and causes the display to branch to that level's submenu. The "SEL" button saves any changes made at a particular level before moving on to the next.

The "EXIT" button exits a menu level and causes the display to branch to the next level higher in the menu hierarchy. The "EXIT" button cancels any changes made at a particular level before moving on to the previous level.

To begin configuring the RAID set, press either the "↑" or "↓" button until "RAID SET PARAMS" appears in the display. Press "SEL" to drop down to the first submenu, which asks you to enter the logical unit number of the RAID set. Use the "↑" or "↓" button to cycle through the possible logical unit numbers until the display reaches the number you wish to configure. Press "SEL."

RAID Level

Selecting the appropriate RAID level for your needs is crucial to maximizing performance.

The first submenu asks you to select a RAID level. The appropriate level will depend on how you intend to use the computer system. The following table gives a thumbnail sketch of each level. For a more complete description, see Chapter 1 "Introduction."

RAID Level	Description
RAID 0	Offers very high throughput but no redundancy, leaving the data vulnerable to disk failures. Use only when speed is all important.
RAID 3	Large chunks of data are split into small pieces, which are written in parallel to the drives in the array. One drive is dedicated to storing parity information. Similar to RAID 0, except the parity drive will protect the data if a disk fails. Best suited for applications such as imaging and graphics, which call for sequential reads of large files.
RAID 5	Data is striped in blocks defined by the user and written to the drives in parallel. Parity information is spread evenly across the drives, rather than concentrated on a single drive. This permits multiple I/Os to occur simultaneously. Ideal for applications requiring many simultaneous random reads and writes, such as databases and transaction processing systems.

Cycle through the RAID levels with the arrow keys and press "SEL" when the level you desire appears in the display.

Stripe Size

One rule of thumb for setting the RAID 5 stripe size is three times your average QIO size.

This submenu applies only to RAID level five, which requires the user to select a stripe size. The optimal stripe size (which is measured in logical blocks) should be large enough to enable most read operations to be satisfied on one drive and small enough to prevent excessive amounts of data to be concentrated on one drive. A good rule-of-thumb is to select a stripe size that is equal to three times the average queued input/output (QIO) size. If you do not know your average QIO size, the I/O Statistics screen of the CRD-5000 monitor utility displays information that will help you determine if your stripe size is set for optimal performance (see page 6-28).

Use the arrow buttons to increase or decrease the stripe size and press "SEL" when you reach the desired value. The minimum stripe size is one logical block; the maximum is 255 blocks.

Channel Designation

The purpose of each drive in the RAID set must be communicated by the user to the CRD-5000. The unit permits you to designate any drive in the RAID set as either offline or online, or as a warm or hot spare.

EMPTY: This designates that no disk drive exists on this particular LUN and disk channel. If you later attach a drive to this disk channel, the CRD-5000 will detect its presence at power up and report "Unknown" rather than "Empty."

OFFLINE: This designates that the drive will not be used by the RAID set. No data will be written to the disk.

ONLINE: This designates that the drive will be a fully functioning member of the RAID set. Data will be written to and read from the disk.

WARM SPARE: This drive will be connected to the CRD-5000, but it will not be permitted to spin up until another drive in the array fails and the system *automatically* brings the warm spare online as a substitute.

HOT SPARE: This drive will be connected to the CRD-5000 and will be kept spinning whenever the RAID set is operating. The CRD-5000 will *automatically* substitute the hot spare for a failed drive. Because the hot spare is kept spinning at all times, it is subject to failure even before it is brought online.

Use the arrow buttons to cycle through the disk channels. Press "SEL" when you arrive at the channel you wish to designate. Then cycle through the possible designations and press "SEL" again to confirm you choice. Repeat these steps for each drive channel in the RAID set.

Exiting RAID Set Params

Once you have finished selecting the RAID set parameters for each logical unit, return to the top menu level by pressing "EXIT" until "RAID SET PARAMS" appears on the display. Then use the arrow buttons to scroll the display to "SYSTEM FUNCTIONS." This is where you will perform the final step in configuring the CRD-5000.

Creating a RAID Set

Before commencing RAID operations, the CRD-5000 must configure its internal logic to conform to the parameters you selected under the "CREATE RAID PARAMS" menu. The unit also writes zeros to every bit in the array to "clean the slate" for parity calculations. Finally, the system writes identifying information to each drive in the array. Every time the unit is powered up, it queries this information and automatically adjusts if it finds that a drive is connected to a different physical disk channel.

This process is called "creating the RAID set" and is initiated by selecting the "CREATE RAID SET" menu item under "RAID SET FUNCTIONS." Use the arrow buttons to display the logical unit number of the RAID set. The system will ask you if you want to use the default RAID set parameter values (see page 4-1). To use the defaults select "YES." To use parameters that you have set, select "NO." Then press "SEL" to begin creating the RAID set. The system will prompt you to confirm your choice.

CAUTION: **Creating a RAID set will destroy any data on the disk drives. If you don't want to lose this data, be sure to back it up before creating the RAID set.**

The RAID set creation will take a few minutes, with larger capacity arrays requiring more time than smaller capacity arrays. A message on the front-panel LCD will indicate the percentage of work completed. Creating a RAID set on an array with five 340MB disk drives takes a little over four and a half minutes.

The front-panel LCD will indicate when the RAID set creation has been completed. If your system has multiple logical units, you must run a separate RAID set creation on each LUN.

If the CREATE RAID SET operation detects any problems with any of the drives on the logical unit, it will abort and display one of the following error messages:

NO DRIVE ON CH X: Where X represents the number of the physical disk channel that according to the configuration parameters was supposed to have a drive connected to it but none was detected, or the drive would not spin up.

TOO FEW DRIVES: There were not enough drives to make up a valid RAID set. At least two drives are required to create a RAID set.

BAD DRIVE ON CH X: Where X represents the number of the physical channel that is inoperable or is connected to a drive that is inoperable.

Once a RAID set has been created, it is ready for use. Press the "EXIT" button several times in succession until you are completely out of the configuration menu tree. The RAID set should be fully functional at this point.

Modifying a RAID Set

Once you have created a RAID set, the CRD-5000 permits you to add or delete spare drives, or change one type of spare to another. This is done with the MODIFY RAID SET function, which, unlike the CREATE RAID SET function, does not destroy the data on the drives in the RAID set. You also can use the MODIFY RAID SET function to delete one drive from a good RAID set to make it degraded if you think that drive is showing early signs of failing or if you want to upgrade the drive.

To modify a RAID set, you first must go to the RAID SET PARAMS menu and change the designation for the channel you wish to modify. The channel designation is the only parameter that the CRD-5000 will permit you to modify without recreating the RAID set. Furthermore, the CRD-5000 will not permit you to change a drive's designation to "Online" from any other designation: you cannot increase the number of drives on array without recreating the RAID set, and the Rebuild function takes care of bringing a drive online to replace another drive.

If you try to add a spare drive with a capacity smaller than the smallest capacity drive already part of the RAID set, the CRD-5000 will reject the drive and display an error message.

To remove a drive from a good RAID set, change its designation to "Offline" or "Empty" if it is to be left cabled up and powered on. Change its designation to "Empty" if you intend to uncable it from the CRD-5000.

The Modify function will generate the following error messages if it encounters a problem:

NO DRIVE ON CH X: The CRD-5000 could not find a drive on channel number X.

DRIVE TOO SMALL CH X: The drive on channel number X is too small to be used with this RAID set.

RAID SET TOO SMALL: The drive could not be deleted because the RAID set is already degraded.

ILLEGAL CHOICE: The CRD-5000 will not permit you to change a drive's status to "Online" from any other designation, or to redesignate an "Online" drive as a spare.

Chapter 5

Front Panel Display

The two-line, 20-character LCD display on the front panel provides all the information you need to identify system failures and configure RAID parameters. An external monitor may be connected to the CRD-5000's RS-232 serial port to obtain more information about the system but is not necessary. For information about connecting a monitor to the CRD-5000, see the the next chapter "Monitor Utility."

All information displayed on the front panel may also be viewed on a monitor connected to the CRD-5000's RS-232 port. Similarly, all configuration data may be input through the front panel or a from a keyboard attached to the monitor.

The display operates in three modes: self-test, configuration and dynamic. Self-test mode is engaged automatically at power up. No user input is permitted in self-test mode. This mode simply reports the status of the system's components and identifies errors.

Configuration mode may be selected at any time except during the self-test. This mode is completely menu-driven and presents you with a number of configuration options.

Dynamic mode continuously reports the status of the system and the drives connected to it during RAID operation. The remainder of this chapter discusses the three modes in greater detail.

Self-test Mode

To assure data integrity and reliable performance, the CRD-5000 contains a built-in, self-test, which runs automatically every time the unit is powered up.

The self-test first performs a survey of the system to identify the components present. The results of this survey are shown on the top line of the display from left to right. The first seven characters from the left represent disk modules zero through six (in that order, from left to right). The self-test will display the symbol ">" to indicate that it has found a module installed on a particular channel. The letter "X" indicates that the self-test has found no module on that channel. One or more "X's" are normal if the CRD-5000 is equipped with fewer than its full complement of seven disk modules. The following illustration shows the results of a self-test on a system with five disk channels.

Disk Modules
>>>>>XX

The next three characters on the top line represent host channels zero through three (in that order from left to right). A ">" symbol means that the self-test has found a module on that host channel, while the letter "X" means that no module has been found. An "E" indicates that the self-test has detected an error. The CRD-5000 will work with up to three host channels. The following is an example of a typical self-test screen for a single-host system.

	Host Modules
>>>>>XX	>XX

The last character on the top line represents the data module. If the self-test finds that the data module is in place, it will report a ">" symbol. The following simulation shows the full top line from the self-test of a typical system.

		Data Module
>>>>>XX	>XX	>

After identifying the components of the system, the self-test verifies their integrity, beginning with disk module 0. The results are displayed from left to right on the second line, directly below the corresponding results from the identification test. Components passing the test are indicated by a ">" symbol. The self-test will skip empty slots. This will be represented by an "X" on the second line directly below an "X." If an "E" appears on the second line directly below a ">" then the self-test has found an error with that module.

The final character on the second line represents the data module. The self-test of the data module is more complicated than the other modules and the time required to complete the test will vary according to the size of the cache and the number of SCSI modules. To help you confirm that the self-test is active, the display will show the hex numbers from zero through E as the self-test progresses through its steps. Not every number or letter will be visible, since some steps require only a fraction of a second.

At the conclusion of a successful self-test, the screen for a unit with five disk modules and one host will look like the following simulation. The screen will immediately display a company identification message and then switch to dynamic mode.

Disk Modules	Host Modules	Data Module
>>>>>XX	>XX	>
>>>>>XX	>XX	0

If the self-test fails to find any host modules or a data module, it will immediately terminate and display an error message. Other error messages are displayed at the conclusion of the self-test.

Dynamic Mode

During normal operation of the CRD-5000, the LCD display shows the status of each disk channel, one LUN (logical unit number) at a time. The display automatically cycles through all logical units that contain (or recently contained) a valid RAID sets, updating the information every few seconds. Logical units that do not contain valid RAID sets will not show up on the display, unless you manually instruct the display to show that LUN by pressing the "↑" or "↓" buttons. You also may skip to a specific LUN at any time by pressing either the "↑" or "↓" buttons.

The system will display the current firmware version number when you press the "EXIT" button at any time while the dynamic display is in effect.

A typical dynamic display looks like this:

```
LUN 1        STAT NNNNN__-N
       CACHE HITS 0%
```

"LUN 1" in the top, left corner refers to the number of the logical unit being examined. This number will cycle from LUN 0 to LUN 3, whether or not the RAID system has more than one logical unit. The first seven characters following "STAT" in the top, right corner refer to the status of the seven possible disk channels.

Channel Status Characters

O	Offline
N	Online
H	Hot Spare
W	Warm Spare
R	Rebuilding
U	Unknown
_	Empty

When you add a new disk drive to a rank, it will register as "Unknown" until you either give it a channel designation or perform the rebuild function if it is replacing a failed drive. A drive can also show up as "Unknown" if you place it on a channel that was previously designated as "Empty." The final character on the top line is an evaluation of the RAID set status on the logical unit.

RAID Set Status Characters

O	Offline
N	Non-Degraded
D	Degraded
C	Creating
_	Empty Set

A logical unit can survive the loss of one disk channel with no effect on data availability. The CRD-5000 will go into degraded mode and respond to all read and write requests—though at a diminished performance level and without the protection of redundancy. A bad disk drive should be replaced or a spare substituted as soon as possible to return the system to normal operation. See Chapter 7 for more information about disk channel failures and degraded mode.

"CACHE HITS" on the second line refers to the percentage of read commands that have been satisfied by accessing data from cache rather than from disk. The CRD-5000 must be equipped with optional, industry-standard SIMM modules to enable caching.

Configuration Mode

An important feature of the CRD-5000 is its flexibility. The unit can control up to 28 SCSI disks in four separate arrays. The disks can have different capacities and speeds. They even can be from different manufacturers. Once the RAID set has been created, disk modules can be moved from slot to slot and cables swapped among drives (*Make sure the power is off before doing this!*). The CRD-5000 will recognize and automatically adjust to changes to the system, with no deterioration in performance or data reliability.

To make this flexibility possible, the CRD-5000 requires you to provide a little information about the system and make a few decisions about how the system will be configured. This is accomplished through the front panel display.

A menu of configuration options will appear on the display when you press the "SEL" button after the completion of the self-test. (If password checking is enabled, the system will not permit you to go to the configuration menus until you correctly enter your password—see pages 5-9 and 5-10 for more information.) By using the "↑" "↓" "SEL" and "EXIT" buttons, you can move through the various menu levels and select desired parameters. The buttons work as follows:

The up arrow button moves up within the same menu level. This button also moves up through the possible values for stripe size; holding the button down will cause the value to increase more rapidly.

The down arrow button moves down within the same menu level. This button also moves down through the possible values for stripe size; holding the button down will cause the values to decrease more rapidly.

The "SEL" button selects a menu level and causes the display to branch to that level's submenu. The "SEL" button saves any changes made at a particular level before moving on to the next.

The "EXIT" button exits a menu level and causes the display to branch to the next level higher in the menu hierarchy. The "EXIT" button cancels any changes made at a particular level before moving on to the previous level.

The following diagram shows the branches of the configuration menu.

RAID SET PARAMS

- LUN [0-3]
  - RAID LEVEL
    - [0,3,5]
  - STRIPE SIZE
    - [1-255]
  - CHANNEL [0-6] DESIGNATION
    - EMPTY
    - OFFLINE
    - ONLINE
    - WARM SPARE
    - HOT SPARE

◆

SYSTEM FUNCTIONS

- CHANGE PASSWORD
  - PRESS SEL TO CHANGE PASSWORD
- RESTART SYSTEM
  - PRESS SEL TO RESTART

◆

DISK PARAMS

- LUN [0-3]
  - TAG QUEUING [OFF/ON]
  - SYNC MODE [ON/OFF]
  - SYNC RATE [5-10]

◆

SYSTEM PARAMS

- PASSWORD CHECKING [OFF/ON]
- DATE
- TIME

RAID SET FUNCTIONS

- REBUILD DISK
  - SPECIFY DISK [NO]
    - REBUILD LUN X, CH X [NO/YES]
      - PRESS SEL TO REBUILD DISK AT LUN X, CH X
  - SPECIFY DISK [YES]
    - LUN [0-3]
      - CHANNEL [0-6]
        - PRESS SEL TO REBUILD DISK AT LUN X, CH X
- CREATE RAID SET
  - LUN [0-3]
    - USE DEFAULT VALUES [NO/YES]
      - PRESS SEL TO CREATE RAID SET ON LUN X
- MODIFY RAID SET
  - LUN [0-3]
    - PRESS SEL TO MODIFY RAID SET AT LUN X

◆

HOST PARAMS

- CHANNEL [0-2]
  - SCSI ID [0-7]
  - TAG QUEUING [OFF/ON]
  - SYNC MODE [ON/OFF]
  - SYNC RATE [5-10]

- RAID SET PARAMS
  - LUN [0-3]
    - RAID LEVEL
      - [0,3,5]
    - STRIPE SIZE
      - [1-255]
    - CHANNEL [0-6] DESIGNATION
      - EMPTY
      - OFFLINE
      - ONLINE
      - WARM SPARE
      - HOT SPARE

Once you create a RAID set, channel designation is the only parameter that may be modified without recreating the RAID set.

RAID SET PARAMS

The RAID set parameters menu provides the system with the basic information it needs to create a RAID set. This menu is typically the first step in configuring a RAID system.

LUN [0-3]: Select the desired logical unit number. Each logical unit can have up to seven disk drives (or as many drives as there are disk modules in the CRD-5000). From the perspective of the host, a logical unit appears to be a single drive. The CRD-5000 creates a separate RAID set for each logical unit, and each logical unit can have different RAID parameters.

RAID LEVEL: The CRD-5000 supports three RAID levels:

LEVEL 0: Data is striped logical block by logical block, with each block being written to a different drive in parallel. Offers very high throughput but no redundancy, leaving the data vulnerable to disk failures.

LEVEL 3: Data is striped logical block by logical block and written to the drives in the array in parallel. One disk in the array is set aside for storing parity information. Level 3 features high data throughput, but only one write operation can be processed at a time, because the parity drive is involved in every write. This level is best suited for applications such as imaging and graphics, which call for sequential reads of large files.

LEVEL 5: Data is striped in multiple blocks and written to the drives in parallel. Parity information is spread evenly across the drives, rather than stored on a single, dedicated drive. For this reason, Level 5 permits multiple write operations to occur in parallel. This level is ideal for applications requiring many simultaneous random reads and writes, such as databases and transaction processing systems.

STRIPE SIZE: RAID systems "stripe" data when writing to the array. This means that data from the host is divided into smaller units, with each unit then being written simultaneously to a disk in the array. The size of these units (measured in logical blocks) is determined by the stripe size parameter. This parameter does not apply to RAID Levels 0 and 3, which always use a stripe size of one logical block.

The size of the stripe is particularly important in RAID Level 5, where the goal is to confine a read or write task to a single disk and enable multiple, parallel I/O operations. A stripe size that is too small will cause more data to be broken up across two or more drives in the array. A stripe size that is too large will create a "hot spot" on a disk, which will receive more than its share of I/O activity, while the other disks in the array are underused. In general, the stripe size should be three times the average QIO size. If you do not know your average QIO size, the I/O Statistics screen of the CRD-5000 monitor utility displays information that will help you determine if your stripe size is set for optimal performance (see page 6-28).

CHANNEL [0-6] DESIGNATION: Each disk on the LUN must be connected to a separate disk channel (or module). The CRD-5000 will support as many disk channels as it has disk modules (up to a maximum of seven). You may designate how each disk on the LUN will be treated by the RAID set. Use the arrow keys to scroll among the disk channel numbers.

EMPTY: This designates that no disk drive exists on this particular LUN and disk channel. If you later attach a drive to this disk channel, the CRD-5000 will detect its presence at power up and report "Unknown" rather than "Empty."

OFFLINE: This designates that the disk will not be used by the RAID set. The drive will spin up with the other drives in the rank, but no data will be written to it.

ONLINE: This designates that the drive will be a fully functioning member of the RAID set. Data will be written to and read from the disk.

WARM SPARE: This drive will be connected to the CRD-5000, but it will not be permitted to spin up until another drive in the array fails and the system *automatically* brings the warm spare online as a substitute.

HOT SPARE: This drive will be connected to the CRD-5000 and will be kept spinning whenever the RAID set is operating. The CRD-5000 will *automatically* substitute the hot spare for a failed drive. Because the hot spare is kept spinning at all times, it is subject to failure even before it is brought on-line.

SYSTEM PARAMS
▸ PASSWORD CHECKING [OFF/ON]
▸ DATE
▸ TIME

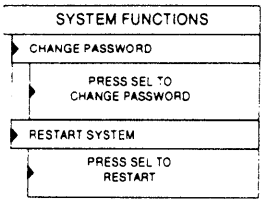
SYSTEM PARAMS

This menu provides access to the security and "housekeeping" functions of the system.

PASSWORD CHECKING: You can control access to the front-panel or monitor-utility menus by selecting this menu item and enabling password checking. To change your password, go to "SYSTEM FUNCTIONS" and select the "CHANGE PASSWORD" option (see page 5-10).

DATE: Select this menu item to change the date of the system's internal clock. Whenever a RAID set is created or a disk fails, the system records the time and date of the event.

TIME: Select this menu item to update the hour and minute of the system's internal clock.



SYSTEM FUNCTIONS

The system functions menu contains just two options, RESTART SYSTEM and CHANGE PASSWORD.

RESTART SYSTEM: The Restart System option enables you to "reboot" the CRD-5000 without having to turn off power and then turn it back on. If you have made any changes to host or disk parameters, you must restart the CRD-5000 before the changes will take effect. The CRD-5000 will not restart until you confirm your selection by pressing "SEL." Press "EXIT" if you do not want to restart the system.

CHANGE PASSWORD: The CRD-5000 offers password protection to prevent unauthorized users from accessing the front-panel or monitor-utility menus. The password must be six digits in length and each digit must be 1, 2, 3 or 4. On the front panel the four available digits are represented by the front panel buttons as the following table shows:

Password Translation Table

Front Panel	↑	↓	SEL	EXIT
Keyboard	1	2	3	4

Use the "CHANGE PASSWORD" menu option to select a new password. When you press "SEL" to change your password, the system will ask for the current password. If the wrong password is entered the display will jump back to the top level of the "CHANGE PASSWORD" menu option. (If you are changing the password for the first time, the default password is 333333, or the "SEL" button pressed six times in succession.) Once you have entered the current password correctly, the system will ask for the new password. The next six front-panel buttons you push will be your new password. The system will ask you to verify your password by reentering it. If you do not enter the same password, the system will jump back to the top level of the "CHANGE PASSWORD" menu option, without changing the password. To enable password checking, you must go to the "SYSTEM PARAMS" menu (see page 5-9).

HOST PARAMS

- CHANNEL [0-2]
  - SCSI ID [0-7]
  - TAG QUEUING [OFF/ON]
  - SYNC MODE [ON/OFF]
  - SYNC RATE [5-10]

HOST PARAMS

This menu presents options for controlling the way the CRD-5000 interacts with the host computer. Any changes to these parameters will be treated as saved parameters and will not become current parameters until the CRD-5000 undergoes a power cycle, either by shutting off and restoring power or selecting the "Restart System" menu option. Make sure that no I/O operations are pending when you shut off power to the CRD-5000, or you may lose data.

CHANNEL [0-2]: The CRD-5000 can work with up to three different host channels, and each host interface can be configured differently. Select the number of the host channel before continuing on to the parameter menu levels.

SCSI ID: All SCSI devices must be identified by a unique SCSI ID number between 0 and 7. The CRD-5000 is considered a target SCSI device by the host. Select a SCSI ID that does not conflict with the host adapter or any other devices on the same SCSI bus.

TAG QUEUING: Tag queuing can greatly increase I/O performance if both the initiator and the target support it. The CRD-5000, which acts as a target for the host, supports tag queuing. If the host adapter also supports tag queuing, enable this option by turning tag queuing on.

SYNC MODE: To enable synchronous communication between the host and the CRD-5000, turn sync mode on. When sync mode is turned off, the CRD-5000 will engage in asynchronous communication with the host.

SYNC RATE: Select the maximum sync rate supported by the host adapter. The CRD-5000 will support maximum sync rates of five, six, seven, eight, nine and 10 megabytes per second. This value applies only when the Sync Mode parameter is turned on.

NOTE: The CRD-5000 uses active SCSI terminators. When the CRD-5000 is at one end of the SCSI bus and a device with passive terminator resistors is at the other end of the bus, you may need to replace the passive resistors with higher quality resistors to achieve successful transfer rates of greater than seven megabytes per second.

DISK PARAMS
- LUN [0-3]
  - TAG QUEUING [OFF/ON]
  - SYNC MODE [ON/OFF]
  - SYNC RATE [5-10]

DISK PARAMS

This menu enables you to configure the interface between the CRD-5000 and each disk array connected to it. Any changes to these parameters will be treated as saved parameters and will not become current parameters until the CRD-5000 undergoes a power cycle, either by shutting off and restoring power or selecting the "Restart System" menu option. Make sure that no I/O operations are pending when you shut off power to the CRD-5000, or you may lose data.

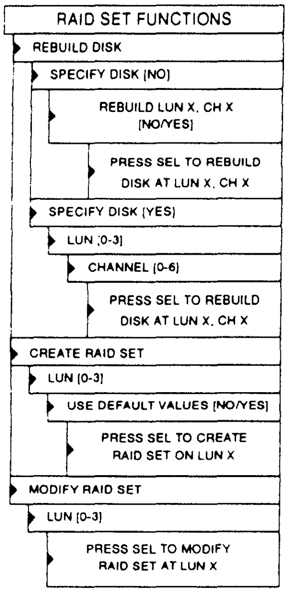
LUN [0-3]: Select the logical unit number of the RAID set to be configured.

TAG QUEUING: Tag queuing can greatly increase I/O performance if both the initiator and the target support it. The CRD-5000, which acts as the initiator to the logical unit, supports tag queuing, provided *all* drives in the logical unit are identical and support tag queuing.

SYNC MODE: To enable synchronous communication between the CRD-5000 and the logical unit, turn sync mode on. When sync mode is turned off, the CRD-5000 will engage in asynchronous communication with the logical unit.

SYNC RATE: Select the maximum sync rate supported by the fastest drive in the logical unit. The CRD-5000 will use this rate as its "opening offer" and work its way down when negotiating transfer rates with each drive. The CRD-5000 will support maximum sync rates of five, six, seven, eight, nine and 10 megabytes per second.

NOTE: The CRD-5000 uses active SCSI terminators. When the CRD-5000 is at one end of the SCSI bus and a device with passive terminator resistors is at the other end of the bus, you may need to replace the passive resistors with higher quality resistors to achieve successful transfer rates of greater than seven megabytes per second.



RAID Set Functions

The RAID set functions menu is where the CRD-5000 acts upon the configuration information input through the other menus. It also offers a rebuild disk option for rebuilding a replacement disk drive.

CREATE RAID SET: Creating a RAID set is the last step in configuring the system. Once given the go ahead, the CRD-5000 will perform a number of tasks to prepare the system and the disks for operation. The process takes several minutes (at a rate of roughly six megabytes per second). To begin the process, select the logical unit number of the array, then press "SEL" again to confirm. Be careful! Creating a RAID set destroys any data on the drives.

MODIFY RAID SET: Only one RAID set parameter can be changed without forcing the RAID set to be recreated from scratch. This is the Channel Designation parameter under the RAID Set Params menu. After modifying this parameter, the change must be recorded by selecting the Modify RAID set menu option. The system will prompt for the logical unit number of the RAID set and ask for confirmation before proceeding.

REBUILD DISK: When a disk fails in a system not equipped with a spare, you must remove the bad drive, insert a new drive and rebuild the disk. The rebuilding process is initiated by this menu option, which prompts for the logical unit number and the disk channel for the drive. Performing a rebuild destroys any data on the disk, so the system will ask for confirmation before proceeding.

Chapter 6

Monitor Utility

While the front-panel display and keypad are sufficient to manipulate all the parameters and functions of the CRD-5000, many users will find the monitor utility more convenient and informative. The basic self-test, dynamic and configuration information displayed via the monitor utility is identical to that displayed on the front panel. Thanks to a monitor's larger screen area, however, the monitor utility is able to display all information pertinent to a particular function on one screen, in an easy-to-read format. Furthermore, the monitor utility can display additional, helpful information that would be impractical to display on the front panel.

Connecting a Monitor

The CRD-5000 is shipped with a three-foot serial cable for making the connection to an external monitor. At one end of the cable is a female connector compatible with the CRD-5000's three-pin RS-232 port. At the other end is a standard male DB-9 connector for attaching to a terminal serial port.

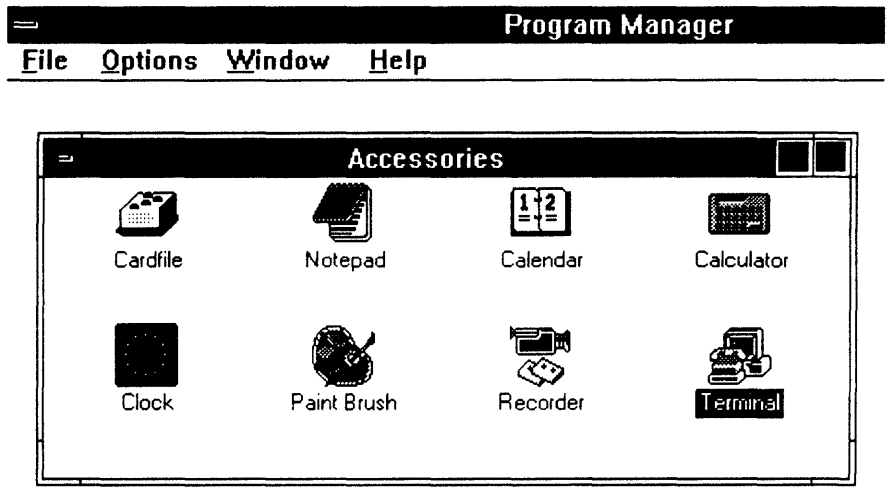
To use the monitor utility, run the cable from the CRD-5000 to your monitor. The terminal should be configured as follows:

Parameter	Value
Baud Rate	9600
Data Bits	8
Stop Bits	1
Parity	None

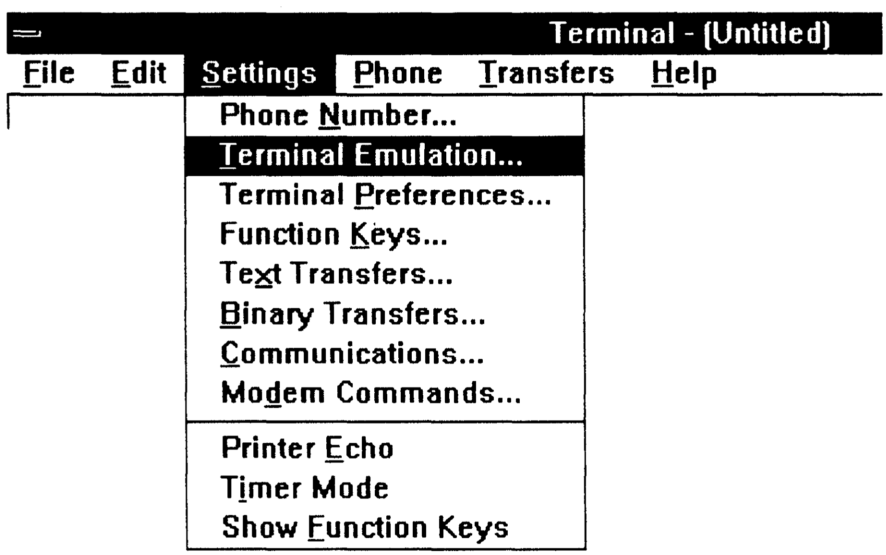
Emulating a Monitor

If you do not have a dedicated monitor at your disposal, you can still access the CRD-5000's monitor utility by using a PC and communications software to emulate a monitor. In this example, we will use the Microsoft Windows Terminal program. Any PC equipped with Windows will have this communications application.

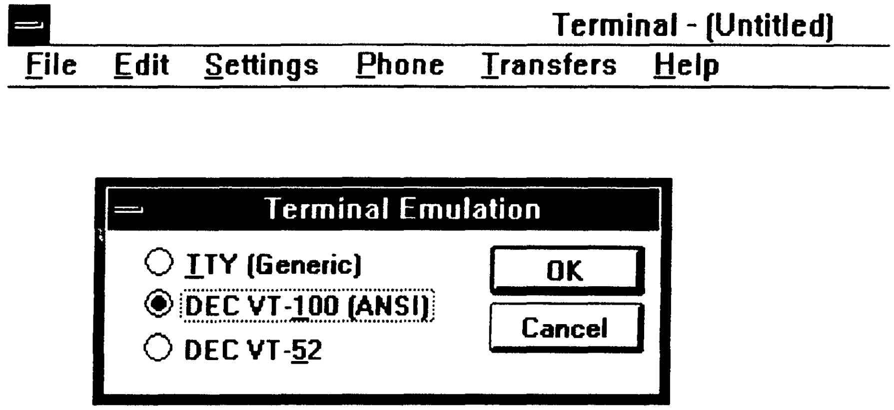
First connect the CRD-5000 to the serial port of your PC with the cable supplied. Then load Windows, open the Accessories window and double-click on the Terminal icon.



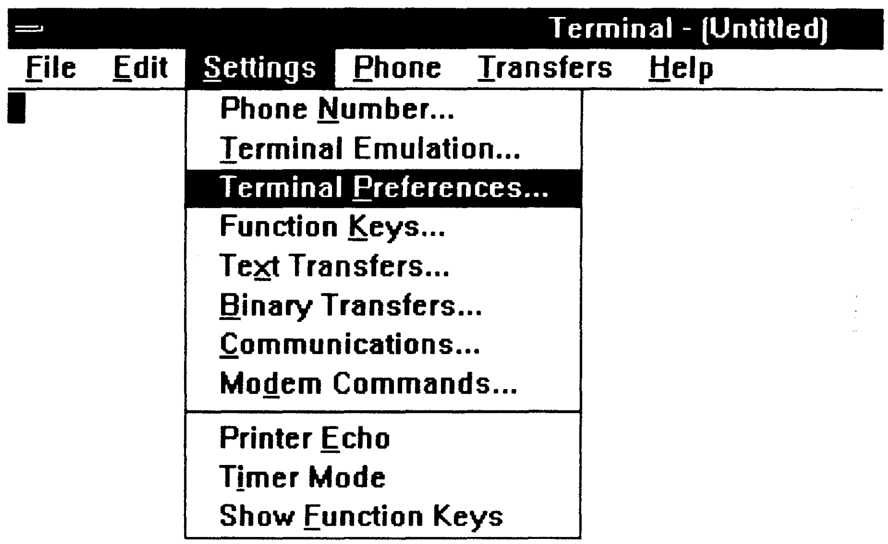
Before you can view the monitor utility, you will have to configure the software to emulate a monitor. Start by clicking on Settings, then Terminal Emulation....



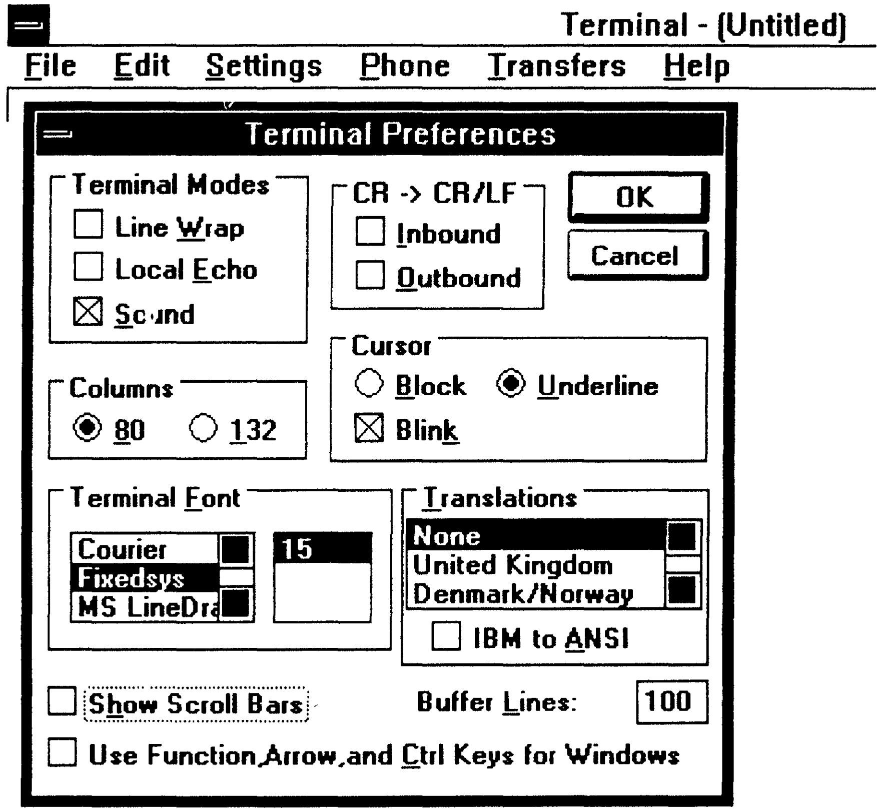
Select the DEC VT-100 (ANSI) emulation.



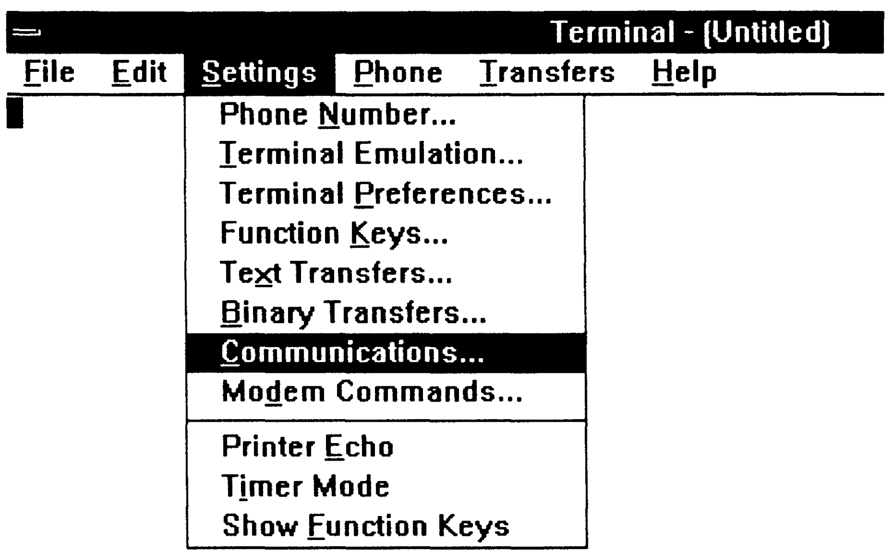
Then click on Settings, Terminal Preferences....



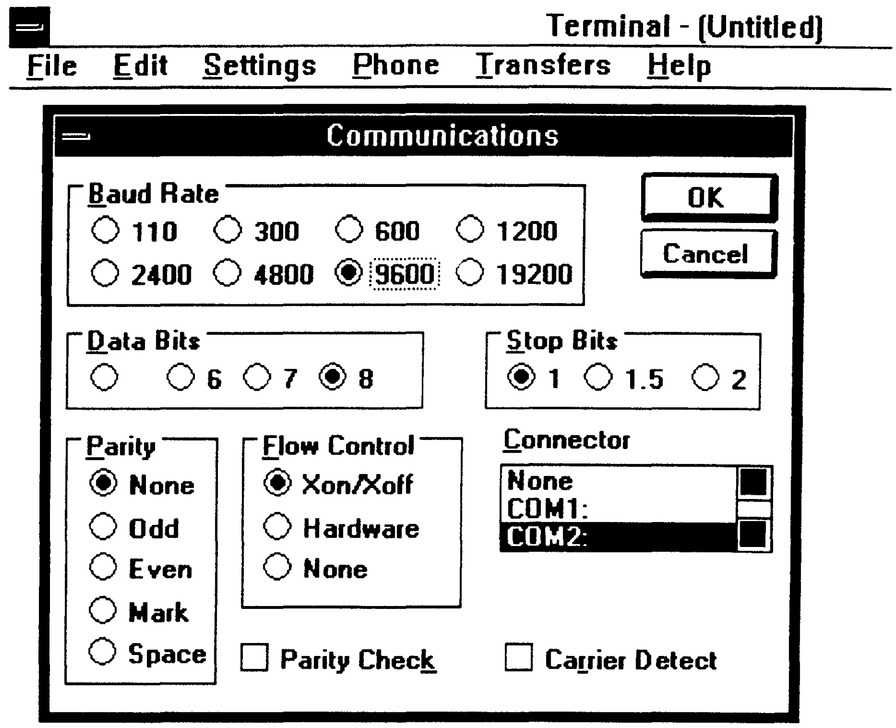
The monitor utility requires the use of the arrow and Ctrl keys, so it is critical that you make sure the box next to "Use Function, Arrow, and Ctrl Keys for Windows" is *not* selected.



Finally click on Settings, Communications....



Select an available COM port. Set the baud rate to 9600, data bits to 8, stop bits to 1, parity to none and flow control to Xon/Xoff.



Your PC should now be set up to run the CRD-5000 monitor utility.

Running the Monitor Utility

At power-on, the CRD-5000 will send self-test information, including error messages, to the monitor automatically. Unlike the front-panel display, however, the monitor does not automatically switch to the dynamic display after the self-test. To see this information or to configure a RAID set from the monitor, you must use the monitor utility. To access the monitor utility, press any key on your monitor or PC keyboard when the monitor utility title box appears.

The Main Menu

The first screen in the monitor utility following the title box displays a menu of options. Use the arrow keys on your keyboard to highlight a menu option and select it by pressing the Enter (Return) key.

CRD-5000 Monitor Utility
MAIN MENU

04/24/93 07:58:45

Raid Set Statistics
Setup Parameters
Raid Set Functions
System Functions
Disk Utilities
I/O Statistics
Rebuild Status

UP ARROW: CURSOR UP	DOWN ARROW: CURSOR DOWN	ENTER: SELECT	CTRL-Z: EXIT

Select RAID Set Statistics to check on the status of your RAID sets. The Setup Parameters item allows you to set RAID set, disk, host and system parameters. RAID Set Functions include rebuilding a disk, creating a RAID set and modifying a RAID set. Select System Functions to change your password, load or save system configuration settings, load new system code, or restart the system. The Disk Utilities option allows you to test a particular disk in a RAID set. I/O Statistics can give you a reading on how well your RAID set is performing. And Rebuild Status displays information about how a disk rebuild is progressing.

RAID Set Statistics

The RAID Set Statistics screen resembles the dynamic display on the front panel. It shows the status of each LUN at any given moment. Unlike the front-panel display, the monitor utility does not cycle through each LUN automatically. You can manually move through the LUNs by pressing the arrow keys.

This screen puts important information about the operation of each RAID set at your fingertips. The RAID Status field lets you know how the overall RAID set is performing. If a disk channel in the RAID set goes bad for whatever reason, this field will indicate that the LUN is operating in degraded mode and identify the defective logical channel.

The Physical Channel (Phy. Ch.), Logical Channel (Log. Ch) and Channel Status columns will help you pinpoint where each drive is located in the array and what its status is. This information is especially important when a drive goes bad. If you misidentify the drives in your array and mistakenly replace a good drive instead of the bad drive, the result is the same as having two drives fail at the same time. No RAID system can survive a second drive's failing before the data from the first bad drive is rebuilt. If a drive goes bad on the CRD-5000, its Channel Status field will switch to OFFLINE and its number will disappear from the Logical Channel field. The associated Physical Channel number will lead you to the drive that needs to be replaced. For more information about replacing and rebuilding data from bad drives, see Chapter 7.

CRD-5000 Monitor Utility
RAID SET STATISTICS
ON LUN 0
04/24/93 07:58:45

Raid Status	NON-DEGRADED	Raid Level	5
Date Stamp	04-08-93 14:38:46	Stripe Size	128 BLKS
Capacity	323 MB	Capacity	663040 BLKS

Phy. Ch.	Log. Ch.	Channel Status	Disk Size	Disk Vendor	Disk Model	Disk Rev.
0		EMPTY				
1	0	ONLINE	100 MB	Acme	LL4000-100	4.07
2	1	ONLINE	150 MB	Elihu Labs	EH482150	3.84
3	2	ONLINE	125 MB	Scion	SO-125E66	6.21
4	3	ONLINE	150 MB	Gigastore	GB100-150	3.32
5	4	ONLINE	125 MB	Ace	LH200B125	2.50
6	5	ONLINE	150 MB	Giant	GS6401250	5.41

UP ARROW: PAGE UP	DOWN ARROW: PAGE DOWN	CTRL-Z: EXIT

Setup Parameters

When you select Setup Parameters from the main menu, the following sub-menu appears.

CRD-5000 Monitor Utility
SETUP PARAMETERS

Raid Set Parameters
Disk Parameters
Host Parameters
System Parameters

UP ARROW: CURSOR UP	DOWN ARROW: CURSOR DOWN	ENTER: SELECT	CTRL-Z: EXIT

Raid Set Parameters

The RAID Set Parameters screen provides the system with the basic information it needs to create a RAID set. This menu is typically the first step in configuring a RAID system.

CRD-5000 Monitor Utility
RAID SET PARAMETERS

04/24/93 07:58:45

Parameter	LUN 0	LUN 1	LUN 2	LUN 3
Raid Level	5	5	5	5
Stripe Size	128 BLKS	128 BLKS	128 BLKS	128 BLKS
Ch. 0 Designation	EMPTY	EMPTY	EMPTY	EMPTY
Ch. 1 Designation	ONLINE	EMPTY	EMPTY	EMPTY
Ch. 2 Designation	ONLINE	EMPTY	EMPTY	EMPTY
Ch. 3 Designation	ONLINE	EMPTY	EMPTY	EMPTY
Ch. 4 Designation	ONLINE	EMPTY	EMPTY	EMPTY
Ch. 5 Designation	ONLINE	EMPTY	EMPTY	EMPTY
Ch. 6 Designation	ONLINE	EMPTY	EMPTY	EMPTY

ARROW KEYS: MOVE CURSOR	ENTER: SELECT	CTRL-Z: EXIT

Press Enter to select a parameter for modification. The status bar at the bottom of the screen will now display a new set of instructions:

UP ARROW: INC	DOWN ARROW: DEC	ENTER: SAVE	CTRL-Z: NO SAVE

Use the arrow keys to change a parameter value. The "↑" key causes the field to step up, one by one, through the legal values. The "↓" key steps down, one by one, through the legal values. If you hold down either arrow key, the value will increase or decrease quickly. Save your change by pressing Enter (Return). Cancel your change by pressing Ctrl-Z.

LUN [0-3]: The CRD-5000 creates a separate RAID set for each logical unit, and each logical unit can have different RAID parameters.

RAID LEVEL: The CRD-5000 supports three RAID levels:

LEVEL 0: Data is striped logical block by logical block, with each block being written to a different drive in parallel. Offers very high throughput but no redundancy, leaving the data vulnerable to disk failures.

LEVEL 3: Data is striped logical block by logical block and written to the drives in the array in parallel. One disk in the array is set aside for storing parity information. Level 3 features high data throughput, but only one write operation can be processed at a time, because the parity drive is involved in every write. This level is best suited for applications such as imaging and graphics, which call for sequential reads of large files.

LEVEL 5: Data is striped in multiple blocks and written to the drives in parallel. Parity information is spread evenly across the drives, rather than stored on a single, dedicated drive. For this reason, Level 5 permits multiple write operations to occur in parallel. This level is ideal for applications requiring many simultaneous random reads and writes, such as databases and transaction processing systems.

STRIPE SIZE: RAID systems "stripe" data when writing to the array. This means that data from the host is divided into smaller units, with each unit then being written simultaneously to a disk in the array. The size of these units (measured in logical blocks) is determined by the stripe size parameter. This parameter does not apply to RAID Levels 0 and 3, which always use a stripe size of one logical block.

The size of the stripe is particularly important in RAID Level 5, where the goal is to confine a read or write task to a single disk and enable multiple, parallel I/O operations. A stripe size that is too small will cause more data to be broken up across two or more drives in the array. A stripe size that is too large will create a "hot spot" on a disk, which will receive more than its share of I/O activity, while the other disks in the array are underused. In general, the stripe size should be three times the average QIO size. If you do not know your average QIO size, the I/O Statistics screen of the CRD-5000 monitor utility displays information that will help you determine if your stripe size is set for optimal performance (see page 6-28).

CHANNEL [0-6] DESIGNATION: Each disk on the LUN must be connected to a separate disk channel (or module). The CRD-5000 will support as many disk channels as it has disk modules (up to a maximum of seven). You may designate how each disk on the LUN will be treated by the RAID set. Use the arrow keys to scroll among the disk channel numbers.

EMPTY: This designates that no disk drive exists on this particular LUN and disk channel. If you later attach a drive to this disk channel, the CRD-5000 will detect its presence at power up and report "Unknown" rather than "Empty."

OFFLINE: This designates that the disk will not be used by the RAID set. No data will be written to the disk.

ONLINE: This designates that the disk will be a fully functioning member of the RAID set. Data will be written to and read from the disk.

WARM SPARE: This disk will be connected to the CRD-5000, but it will not be permitted to spin up until another disk in the array fails and the system *automatically* brings the warm spare online as a substitute.

HOT SPARE: This disk will be connected to the CRD-5000 and will be kept spinning whenever the RAID set is operating. The CRD-5000 will *automatically* substitute the hot spare for a failed drive. Because the hot spare is kept spinning at all times, it is subject to failure even before it is brought online.

Disk Parameters

Use this screen to modify the parameters for the disk channels. The CRD-5000 will support tag queuing and synchronous data transfer rates of up to 10 megabytes per second with the disks on each logical unit.

CRD-5000 Monitor Utility
DISK PARAMETERS

07/24/93 07:58:45

Parameter	LUN 0	LUN 1	LUN 2	LUN 3
Tag Queuing	OFF	OFF	OFF	OFF
Sync Mode	ON	ON	ON	ON
Sync Rate	8 MB/SEC	8 MB/SEC	8 MB/SEC	8 MB/SEC

ARROW KEYS: MOVE CURSOR	ENTER: SELECT	CTRL-Z: EXIT

Press Enter to select a parameter for modification. The status bar at the bottom of the screen will now display a new set of instructions:

UP ARROW: INC	DOWN ARROW: DEC	ENTER: SAVE	CTRL-Z: NO SAVE

Use the arrow keys to change a parameter value. The "↑" key causes the field to step up, one by one, through the legal values. The "↓" key steps down, one by one, through the legal values. Save your change by pressing Enter (Return). Cancel your change by pressing Ctrl-Z.

LUN [0-3]: The CRD-5000 supports up to four logical units.

TAG QUEUING: Tag queuing can greatly increase I/O performance if both the initiator and the target support it. The CRD-5000, which acts as the initiator to the logical unit, supports tag queuing, provided *all* drives in the logical unit are identical and support tag queuing.

SYNC MODE: To enable synchronous communication between the CRD-5000 and the logical unit, turn sync mode on. When sync mode is turned off, the CRD-5000 will engage in asynchronous communication with the logical unit.

SYNC RATE: Select the maximum sync rate supported by the fastest drive in the logical unit. The CRD-5000 will use this rate as its "opening offer" and work its way down when negotiating transfer rates with each drive. The CRD-5000 will support sync rates of five, six, seven, eight, nine and 10 megabytes per second.

NOTE: The CRD-5000 uses active SCSI terminators on each of its disk modules. When the CRD-5000 is at one end of the SCSI bus and a device with passive terminator resistors is at the other end of the bus, you may need to replace the passive resistors with higher quality resistors to achieve successful transfer rates of greater than seven megabytes per second.

Host Parameters

This screen is where you configure the parameters for each host channel. The CRD-5000 will support tag queuing and synchronous data transfer rates of up to 10 megabytes per second with the host.

CRD-5000 Monitor Utility
HOST PARAMETERS

04/24/93 07:58:45

Parameter	Channel 0	Channel 1	Channel 2
Status	INSTALLED	EMPTY	EMPTY
SCSI ID	0	0	1
Tag Queuing	ON	ON	ON
Sync Mode	ON	ON	ON
Sync Rate	8 MB/SEC	10 MB/SEC	8 MB/SEC

ARROW KEYS: MOVE CURSOR	ENTER: SELECT	CTRL-Z: EXIT

Press Enter to select a parameter for modification. The status bar at the bottom of the screen will now display a new set of instructions:

UP ARROW: INC	DOWN ARROW: DEC	ENTER: SAVE	CTRL-Z: NO SAVE

Use the arrow keys to change a parameter value. The "↑" key causes the field to step up, one by one, through the legal values. The "↓" key steps down, one by one, through the legal values. Save your change by pressing Enter (Return). Cancel your change by pressing Ctrl-Z.

CHANNEL [0-2]: The CRD-5000 can work with up to three different host channels, and each host interface can be configured differently.

SCSI ID: All SCSI devices must be identified by a unique SCSI ID number between 0 and 7. The CRD-5000 is considered a target SCSI device by the host. Select a SCSI ID that does not conflict with the host adapter or any other devices on the same SCSI bus.

TAG QUEUING: Tag queuing can greatly increase I/O performance if both the initiator and the target support it. The CRD-5000, which acts as a target for the host, supports tag queuing. If the host adapter also supports tag queuing, enable this option by turning tag queuing on.

SYNC MODE: To enable synchronous communication between the host and the CRD-5000, turn sync mode on. When sync mode is turned off, the CRD-5000 will engage in asynchronous communication with the host.

SYNC RATE: Select the maximum sync rate supported by the host adapter. The CRD-5000 will support sync rates of five, six, seven, eight, nine and 10 megabytes per second. This value applies only when the SYNC MODE parameter is turned on.

NOTE: The CRD-5000 uses active SCSI terminators. When the CRD-5000 is at one end of the SCSI bus and a device with passive terminator resistors is at the other end of the bus, you may need to replace the passive resistors with higher quality resistors to achieve successful transfer rates of greater than seven megabytes per second.

System Parameters

This screen provides access to the security and "housekeeping" functions of the system.

CRD-5000 Monitor Utility
SYSTEM PARAMETERS

04/24/93 07:58:45

Parameter	Value
Password Checking	OFF
Date	07-19-93
Time	14:23:59

ARROW KEYS: MOVE CURSOR	ENTER: SELECT	CTRL-Z: EXIT

The password checking field is where you inform the system whether you wish to enable password protection. If you have entered "On" in this field and the user does not provide the correct password, the monitor utility will bar access to all but the status pages (RAID Set Statistics, I/O Statistics and Rebuild Status). To change a password, you must select the Change Password option on the System Functions submenu (see page 6-18).

The date and time fields allow you to set the CRD-5000's internal clock.

Raid Set Functions

When you select RAID Set Functions the monitor utility will display a secondary menu, asking you whether you want to rebuild a disk, create a RAID set, or modify a RAID set.

CRD-5000 Monitor Utility
RAID SET FUNCTIONS

Rebuild Disk
Create Raid Set
Modify Raid Set

UP ARROW: CURSOR UP	DOWN ARROW: CURSOR DOWN	ENTER: SELECT	CTRL-Z: EXIT

The Rebuild Disk option enables you to replace and rebuild the data from a failed member of a RAID set. For more information about this option, see page 7-5.

Creating a RAID set refers to the process by which the CRD-5000 configures its internal logic to conform to the parameters you selected under the RAID Set Parameters menu. During the creation, the unit also writes zeros to every bit in the array to "clean the slate" for parity calculations and stores identifying information on each drive in the array. The CRD-5000 will not operate until you perform a RAID set creation on each rank of drives connected to the unit. You may create a RAID set with the CRD-5000's default settings (RAID level five, 128 block stripe size, all eligible drives included in the RAID set), or with your own parameters. If you want to use your own parameters, you must specify them before initiating the creation process by selecting Setup Parameters from the main menu and then RAID Set Parameters from the sub-menu (see page 6-7).

CAUTION: Creating a RAID set will destroy any data on the disk drives. If you don't want to lose this data, be sure to back it up before creating the RAID set.

The Modify RAID Set function permits you to make certain changes to an existing RAID set. Unlike the Create RAID Set function, the Modify RAID Set function does not erase the data on the disks in the RAID set. Only one parameter—the Channel Designation parameter—may be modified using the Modify RAID Set function.

Create RAID Set

When you select Create RAID Set from the submenu, the following warning message appears:

```
                 CRD-5000 Monitor Utility
                     CREATE RAID SET

                 04/24/93       07:58:45

          +-------------------------------------+
          |              WARNING                |
          |                                     |
          |  Creating A RAID Set Will Erase All |
          |  The Information On The Selected LUN|
          +-------------------------------------+

| ENTER: SELECT       | CTRL-Z: EXIT        |                     |
```

Press Enter to proceed with the operation. If the disk drives in the RAID set have data that you want to keep, press Ctrl-Z to exit. Return to the Create RAID Set function, after you have backed up the data or found different disk drives to use in the RAID set.

Select The LUN of the RAID Set To Be Created
LUN 0

UP ARROW: INC	DOWN ARROW: DEC	ENTER: SELECT	CTRL-Z: EXIT

Use the arrow keys to scroll through the LUN display until you reach the desired LUN. Press Enter to proceed to the next step in the RAID set creation.

Create RAID Set With Default Values?
YES

UP ARROW: INC	DOWN ARROW: DEC	ENTER: SELECT	CTRL-Z: EXIT

If you select "YES" at this point, the CRD-5000 will create a RAID set with the default settings: RAID level five, 128 block stripe size and all eligible drives included as online members of the RAID set. Select "NO" to have the CRD-5000 create the RAID set with parameters you selected in the RAID Set Parameters screen.

CRD-5000 Monitor Utility
CREATE RAID SET

04/24/93 07:58:45

Press Enter To Create RAID Set On LUN 0

ENTER: SELECT	CTRL-Z: EXIT	

Press Enter to begin the creation of the RAID set. The following screen will appear, showing the status of the process:

CRD-5000 Monitor Utility
CREATE RAID SET

04/24/93 07:58:45

Creating RAID Set On LUN 0

#####

10% Complete

Estimated Time: 00:09:37 Elapsed Time: 00:00:51

PLEASE WAIT...

The RAID set creation will take a few minutes, with larger capacity arrays requiring more time than smaller capacity arrays. Creating a RAID set on an array with five 340MB disk drives takes a little over four and a half minutes.

When the RAID set creation has been completed, the monitor utility will display the following message:

CRD-5000 Monitor Utility
CREATE RAID SET

04/24/93 07:58:45

The RAID Set Was Successfully Created
On LUN 0

CTRL-Z: EXIT

If the creation operation detects any problems with any of the drives on the logical unit, it will abort and display an error message:

CRD-5000 Monitor Utility
CREATE RAID SET

04/24/93 07:58:45

CREATION ERROR
Unable To Create RAID Set on LUN 0
____

CTRL-Z: EXIT	

RAID Set Creation Error Messages

The third line inside the box will contain one of four possible error messages:

"No Drive Detected on Channel _": This indicates the number of the physical disk channel that according to the configuration parameters was supposed to have a drive connected to it but none was detected, or the drive would not spin up.

"Bad Drive Detected on Channel _": This is the number of the physical channel that contains a drive found to be unreliable or inoperable by the CRD-5000.

"Too Few Drives for a Valid RAID Set": There were not enough drives to make up a valid RAID set. At least two drives are required to create a RAID set.

"The System Is Busy": The CRD-5000 cannot create a RAID set while it is processing I/O requests from the host.

Modify RAID Set

Most changes to RAID set parameters require that the RAID be recreated, which destroys any data on the disk drives. A channel designation, however, may be changed without affecting your data by using the Modify RAID Set function rather than a RAID creation. You may wish to use the Modify RAID Set function in conjunction with the RAID Set Parameters screen to add or delete spare drives, or change one type of spare to another. You also change a active drive from "Online" to "Offline" to make a good RAID set degraded if you think that drive is showing early signs of failing or if you want to upgrade the drive.

After you have changed a channel designation from the RAID Set Parameters screen (see page 6-7), select Modify RAID Set from the System Functions submenu of the monitor utility. The following screen will appear:

Select The LUN of the RAID Set To Be Modified
LUN 0

UP ARROW: INC	DOWN ARROW: DEC	ENTER: SELECT	CTRL-Z: EXIT

Use the arrow keys to scroll among the available LUNs and press Enter when you have reached the LUN you want to modify.

CRD-5000 Monitor Utility
MODIFY RAID SET

04/24/93 07:58:45

Press Enter To Modify RAID Set On LUN 0

ENTER: SELECT	CTRL-Z: EXIT	

Select the logical unit number of the RAID set and press Enter. This operation will bring the RAID set in line with any new channel designations you have specified. It does not affect the data on the RAID set.

CRD-5000 Monitor Utility
MODIFY RAID SET

04/24/93 07:58:45

MODIFICATION ERROR Unable To Modify RAID Set on LUN 0 ________________

CTRL-Z: EXIT	

RAID Set Modification Error Messages

The third line inside the box will contain one of four possible error messages:

No Drive Detected on Channel _: The CRD-5000 could not find a drive on channel number _.

Drive Too Small On Channel _: The drive on channel number _ is too small to be used with this RAID set.

Raid Set Too Small: The drive could not be deleted because the RAID set is already degraded.

Illegal Choice: The CRD-5000 will not permit you to change a drive's status to "Online" from any other designation, or to redesignate an "Online" drive as a spare.

System Functions

The System Functions sub-menu is where to go to change your password, load or save configuration parameters, load updated system code when it becomes available or restart your system to activate new host or disk parameters.

Change Password

The CRD-5000 offers password protection to prevent unauthorized users from accessing the front-panel or monitor-utility menus. The password must be six digits in length and each digit must be 1, 2, 3 or 4. The same password will apply to the monitor utility as well as the front panel. With the monitor utility, you would use the keyboard to enter the password digits. With the front panel display, you would use the front panel buttons according to the following translation:

Password Translation Table

Front Panel	↑	↓	SEL	EXIT
Keyboard	1	2	3	4

When you select Change Password from the System Functions submenu, the following screen will appear:

CRD-5000 Monitor Utility
CHANGE PASSWORD

04/24/93 07:58:45

Enter Current Password

ENTER: SELECT	CTRL-Z: EXIT	

The system asks for the current password to prevent unauthorized password changes. (If you are changing the password for the first time, the default password is 333333.) This screen will continue to appear until either the current password is entered correctly or the user exits by pressing Ctrl-Z.

Once you have entered the correct current password, the system will ask you to enter a new password.

CRD-5000 Monitor Utility
CHANGE PASSWORD

04/24/93 07:58:45

Enter New Password

ENTER: SELECT	CTRL-Z: EXIT	

Enter a new password. Remember, the password must be six digits in length, and the only valid digits are 1, 2, 3, or 4. The password you enter here will also apply to the front panel. If you enter an invalid password, the system will disregard your selection and ask you to enter a new password again. When you have entered a valid password, the following screen will appear:

CRD-5000 Monitor Utility
CHANGE PASSWORD

04/24/93 07:58:45

Reenter New Password

ENTER: SELECT	CTRL-Z: EXIT	

The system will verify your password selection by asking you to reenter it. If the password you reenter does not match the password recorded in the previous screen, the system will force you to start over again at the Enter New Password screen.

To enable password checking, you must go to the System Parameters menu (see page 6-12).

Load Configuration

By using the Load Configuration option, you may load a set of configuration parameters from storage. See the next section "Save Configuration" for instructions on how to save a set of configuration parameters.

CRD-5000 Monitor Utility
LOAD CONFIGURATION

04/24/93 07:58:45

Send Configuration Data

CTRL-Z: EXIT	

When the "Send Configuration Data" screen appears, issue the necessary commands to your communications software to send a text file. (For the Windows Terminal application, click on Transfers, Send Text File....) Select the filename of the configuration you want to load and initiate the file transfer. The CRD-5000 monitor utility will respond with the following screen.

CRD-5000 Monitor Utility
LOAD CONFIGURATION

04/24/93 07:58:45

Receiving Configuration Data

CTRL-Z: EXIT	

If the configuration is loaded successfully the monitor utility will display a confirmation screen.

CRD-5000 Monitor Utility
LOAD CONFIGURATION

04/24/93 07:58:45

Valid Configuration Data Received

CTRL-Z: EXIT	

If the transfer was not successful or the configuration file is not in the proper format, the following message is displayed.

CRD-5000 Monitor Utility
LOAD CONFIGURATION

04/24/93 07:58:45

ERROR

Invalid Configuration Data Received

CTRL-Z: EXIT	

You must restart the system after loading a new set of configuration parameters The Load Configuration option will not change your password.

Upon the successful completion of this menu option, the setup parameters contained in the configuration file will overwrite whatever parameters were set before the transfer. You must restart the system for the new host, disk and system parameters to take effect. Your password will be unaffected by this operation.

Save Configuration

Use this menu option to save your current parameters to a file. You may then use the "Load Configuration" option to recall these parameters at a later time.

CRD-5000 Monitor Utility
SAVE CONFIGURATION

04/24/93 07:58:45

Press Enter To Receive Configuration Data

ENTER: SELECT	CTRL-Z: EXIT

Before you press Enter, issue the necessary commands to your communications software to receive a text file. (In the Windows Terminal application, click on Transfers, Receive Text File....) Once your software is ready to receive, press Enter to begin the transfer. The Monitor Utility will display the following message and handle the rest of the transfer.

CRD-5000 Monitor Utility
SAVE CONFIGURATION

04/24/93 07:58:45

Sending Configuration Data

PLEASE WAIT...	

When the transfer is completed, the following message will appear.

CRD-5000 Monitor Utility
SAVE CONFIGURATION

04/24/93 07:58:45

Data Transmission Completed

CTRL-Z: EXIT	

At this point, instruct your communications software to stop receiving text. This is important, because if you fail to do this, your software will continue to append monitor utility screens to the configuration file, which could render it invalid.

Restart System

To activate any host or disk parameter changes you must restart the system, either by recycling power or using the Restart System menu option. When you select Restart System from the System Functions submenu, the following screen appears.

CRD-5000 Monitor Utility
WARM RESTART

04/24/93 07:58:45

WARNING
Pressing Enter Will Restart System

ENTER: SELECT	CTRL-Z: EXIT	

Load System Code

The Load System Code option enables you to load the most recent firmware version into the CRD-5000's EPROM (erasable programmable read-only memory). The version number of the firmware currently in use by the system is displayed in the title box of the screen that invokes the monitor utility.

CMD Technology makes the latest version of the CRD-5000 firmware available on its BBS. Set your communications software to 8 data bits, 1 stop bit and no parity. The BBS will support up to 9600 baud. The primary BBS line is (714) 454-1134. Use (714) 454-0795 as an alternate line. You may also receive the firmware code via Internet. Just send your request for code to TECH-SUPPORT@CMD.COM. A member of the technical support staff will respond by sending the code to your address. In either case, the code will be an ASCII-format file, with a .HEX extension.

To transfer the update firmware to your CRD-5000, select Load System Code from the System Functions submenu. The monitor utility will respond with the following warning:

CRD-5000 Monitor Utility
LOAD SYSTEM CODE

04/24/93 07:58:45

WARNING
Pressing Enter Will Erase The Current System Code Before It Loads The New System Code

ENTER: SELECT	CTRL-Z: EXIT

If you press Enter to proceed, the system will respond with the following message:

*** Ready to Load System Code ***

Use your communication application's file transfer function to send the firmware update file. If you are using the Windows Terminal application, click on "Transfer," then "Send Text File...." Specify the location and name of the firmware update file and click on OK. The monitor utility will display the following information:

*** Loading Hex Data File ***

*** Flash Eprom Erase Completed ***

During the course of the transfer, the monitor utility will show how many bytes of code have been loaded and how many have been programmed in the EPROM. You may also monitor the progress of the transfer by watching the gauge at the bottom of the Terminal screen.

When the transfer has been completed successfully, the screen will display the following message:

*** Flash Eprom Load Completed ***

*** Start Self-Test (Y) ***

Press "Y" to initiate the CRD-5000's self-test. Once it completes the self-test, your unit will be ready for operation. The new firmware will be fully compatible with any existing RAID sets.

If for any reason the file transfer is aborted prematurely, the CRD-5000 will be inoperable. Since the EPROM must be erased before the new code can be loaded, the old code will not be available to control the system if the new code is not fully loaded. In such an event, the monitor utility will display the following message:

*** Eprom is not Programmed ***

*** Do You Wish to Download System Code ***

Press "Y" to restart the download process (see above for instructions).

You may update your firmware without entering the monitor utility by pressing Ctrl-C during the self-test. Once the self-test is finished, the system will display the following messages:

SELF TEST COMPLETE
*** 00 Mbytes of Cache installed ***
*** Loop in Self Test (L)
*** Burn in Test (B)
*** Download Hex Data to Flash Eprom (D)

Press "D" to initiate the update process, which will be identical to process described above. The Loop in Self Test and Burn in Test are diagnostic tests intended for trained technicians.

Disk Utilities

The Disk Utilities option takes you to the CRD-5000's Test Disk utility, which performs either a read-only or a write/read test on a single disk drive or on an entire LUN.

Test Disk

The Test Disk utility will help you verify that you can write to and read from your disk drives, but it can wreak havoc with your data if you are not careful. To avoid endangering your data, please heed the following guidelines.

- Never use Test Disk if there is the possibility of I/O activity coming from the host to the LUN being tested.
- The write/read portion of Test Disk will overwrite any data on a drive. Performing this test on an online member of a non-degraded RAID set will cause the RAID set to become degraded. You will have to use the CRD-5000's rebuild function to bring the RAID set back to non-degraded operation. This applies only to online disk drives. You may perform a write/read test on a hot or warm spare, with no impact on the RAID set.
- If you perform a write/read test on an online member of a *degraded* RAID set, the RAID set will go offline, cutting off access to your data. Similarly, if you perform a write/read test on *all* drives in a degraded or a non-degraded RAID set, you will loose access to your data.

The Test Disk utility will flash warning messages if you initiate a test that could damage your data.

CRD-5000 Monitor Utility
TEST DISK

04/24/93 07:58:45

Select The Type Of Test

Read Only

UP ARROW: INC	DOWN ARROW: DEC	ENTER: SELECT	CTRL-Z: EXIT

Use the arrow keys to toggle between the read-only test and the write/read test. Press Enter to move to the next screen.

CRD-5000 Monitor Utility
TEST DISK

04/24/93 07:58:45

Select The LUN To Be Tested

LUN 0

UP ARROW: INC	DOWN ARROW: DEC	ENTER: SELECT	CTRL-Z: EXIT

Use the arrow keys to toggle among the LUNs. Press Enter when the LUN you want to test is displayed.

CRD-5000 Monitor Utility
TEST DISK

04/24/93 07:58:45

Select The Channel To Be Tested

All Channels

UP ARROW: INC	DOWN ARROW: DEC	ENTER: SELECT	CTRL-Z: EXIT

You may test just one channel or all channels on the LUN. Use the arrow keys to toggle between the two choices and press Enter.

CRD-5000 Monitor Utility
TEST DISK

04/24/93 07:58:45

Press Enter To Begin Read Test
On LUN 0, All Channels

ENTER: SELECT	CTRL-Z: EXIT	

The test will begin selecting blocks at random. If you selected the read-only test, the utility will try to read from these random blocks. If you selected the write/read test, the utility will write data to the blocks and then attempt to read the data. While the test is underway, the utility will display a screen with a progress report.

CRD-5000 Monitor Utility
TEST DISK

04/24/93 07:58:45

Performing Random Read Test On LUN 0, Channel 2

Testing Block Number: 212370

Blocks Tested: 1038 Elapsed Time: 00:01:21

CTRL-Z: EXIT

Test Disk will run indefinitely until you exit by pressing Ctrl-Z. If the utility discovers an error, it will display a message pinpointing the location of the error and describing the error with either its SCSI Status or the Request Sense Data (or both), depending on the type of error.

CRD-5000 Monitor Utility
TEST DISK

04/24/93 07:58:45

TEST ERROR

Test Failed During A Write Operation
On LUN 0, Channel 4, Block 4294967295

SCSI Status: 02

Request Sense Data:

70	00	05	00	00	00	0A	00
00	00	00	00	00	00	00	00

CTRL-Z: EXIT

Consult a SCSI manual for the meaning of the SCSI Status or Request Sense Data values.

I/O Statistics

As the CRD-5000 carries out its RAID functions, it keeps track of how many reads and writes take place on single and multiple disk channels. The results are posted on the I/O Statistics screen, which is particularly useful when you are fine-tuning the stripe size in RAID level five.

CRD-5000 Monitor Utility
I/O STATISTICS (PRELIMINARY PAGE)
ON LUN 0
04/24/93 07:58:45

Channel Statistics	Ch. 0	Ch. 1	Ch. 2	Ch. 3	Ch. 4	Ch. 5	Ch. 6
Single-Ch. reads	0	0	0	0	0	0	0
Multi-Ch. reads	0	0	0	0	0	0	0
Single-Ch. writes	0	0	0	0	0	0	0
Multi-Ch. writes	0	0	0	0	0	0	0

ARROW UP: PAGE UP	ARROW DOWN: PAGE DOWN	CTRL-Z: EXIT

If you find that multi-channel reads or writes outnumber their single-channel counterparts in a RAID five rank, you may wish to experiment with a larger stripe size. If you find that one particular disk channel is performing more reads and writes than the other channels, your stripe size probably is too large. For more information about setting the stripe size, see page 6-9.

To view the I/O statistics for another logical unit, press either the up or down arrow key. Ctrl-Z takes you back to the main menu.

Rebuild Status

The CRD-5000 will continue to process I/O requests from the host while rebuilding a disk on one or more logical units. You may monitor the progress of each rebuild operation by selecting the Rebuild Status option from the monitor utility main menu. The Rebuild Status screen looks like the following diagram. The screen you see on your monitor will depend on the status of your system.

CRD-5000 Monitor Utility
REBUILD STATUS
ON LUN 0
04/24/93 07:58:45

LUN 0 Rebuilding Channel 2 ############### 44% Complete	TIME Elapsed: 00:01:42 Estimated: 00:03:50
LUN 1 No Rebuild In Progress	
LUN 2 No Rebuild In Progress	
LUN 3 Rebuild Completed Successfully On Channel 2	Elapsed: 00:05:32 Estimated: 00:05:31

CTRL-Z: EXIT	

Chapter 7

Drive Failures

The failure of a disk drive controlled by the CRD-5000 is not the catastrophic event that a similar failure would occasion on a system served by a single, large disk drive. Even if one of the drives on a RAID 3 or 5 array grinds to a halt, the CRD-5000 will continue to handle all I/O operations, with little or no deterioration in performance. This should not be taken as a sign that CRD-5000 is operating normally, however. To assure continued data availability, the operator must take prompt steps to either replace the failed drive or switch to a spare and then instruct the CRD-5000 to rebuild the data from the failed drive. For while the CRD-5000 can withstand the loss of one drive on a logical unit, it cannot withstand the loss of a second drive before the first has been rebuilt.

Of course, if you have selected RAID level 0 and a drive fails on the array, your data will be lost. RAID 0 does not employ a parity drive, so there is no way to rebuild data if a drive fails.

When a Drive Goes Bad

The moment that the CRD-5000 detects that an online drive channel is no longer functioning it will flash the "Failure" LED on the front panel and indicate which disk channel has failed on the front-panel display.

When the CRD-5000 is operating normally, the front-panel display for a logical unit with six disk channels would look like this:

```
LUN 0        STAT NNNNNN_-N
          CACHE HITS 0%
```

A failure on disk channel four would appear on the front-panel display as follows:

```
LUN 0        STAT NNNONN_-D
          CACHE HITS 0%
```

When an "O" replaces a "N," the disk channel represented by that character has failed. The last character on the first line will change to a "D," indicating that the RAID set is running in degraded mode. In degraded mode, the CRD-5000 will respond to I/O requests in one of three ways:

- ↳ If the operation can be done without accessing the failed drive, it takes place as a normal RAID I/O operation.
- ↳ RAID 3: If the write request is targeted for all drives in a stripe, the system proceeds as usual, except it skips the degraded channel. If the write is targeted for a partial stripe that includes the degraded channel, the system proceeds with a normal partial-stripe write (i.e. the excluded channels are read for parity information) and skips the degraded channel. If the partial-stripe write does not include the degraded channel, the system reads the old data from the parity drive and the remaining data drives and performs an exclusive-or operation with the new data being written to derive the new parity data.
- ↳ RAID 5: If a write request must access a failed data drive, the CRD-5000 reads the data from all the other data drives, generates parity by performing an exclusive-or operation with that data and the new data coming from the host, and writes the parity to the parity drive. In a read request from a failed data drive, the CRD-5000 reads the data from the other data drives and parity from the parity drive, and performs an exclusive-or operation to derive the data from the failed drive.
- ↳ If the parity drive has failed, the CRD-5000 simply performs the I/O operation without calculating parity.

If you have a monitor connected to the CRD-5000, the "RAID Set Statistics" screen of the monitor utility will indicate when a drive channel fails and identify the offending channel. A typical monitor display following a disk channel failure might look like this:

CRD-5000 Monitor Utility
RAID SET STATISTICS
ON LUN 0
04/24/93 07:58:45

Raid Status	DEGRADED (Log. Ch. 3)	Raid Level	5
Date Stamp	04-08-93 14:38:46	Stripe Size	128 BLKS
Capacity	323 MB	Capacity	663040 BLKS

Phy. Ch.	Log. Ch.	Channel Status	Disk Size	Disk Vendor	Disk Model	Disk Rev.
0		EMPTY				
1	0	ONLINE	100 MB	Acme	LL4000-100	4.07
2	1	ONLINE	150 MB	Elihu Labs	EH482150	3.84
3	2	ONLINE	125 MB	Scion	SO-125E66	6.21
4	3	OFFLINE	150 MB	Gigastore	GB100-150	3.32
5	4	ONLINE	125 MB	Ace	LH200B125	2.50
6	5	ONLINE	150 MB	Giant	GS6401250	5.41

UP ARROW: PAGE UP	DOWN ARROW: PAGE DOWN	CTRL-Z: EXIT

In this example, the "Raid Status" field shows that the logical unit is running in degraded mode. Additionally, the field identifies logical channel 3 as the failed channel. Do not confuse *logical* disk channels with *physical* disk channels. Logical channels are the CRD-5000's internal representations of the paths to each disk drive in the array. They do not necessarily coincide with physical channels, which are the actual hardware slots on the back of the CRD-5000 to which the drive is cabled. In the example above, the channel that the CRD-5000 views as logical channel three is actually located on slot four.

The ability to distinguish between logical and physical channels is crucial to identifying and replacing a failed disk drive, as the next section will explain.

Replacing a Failed Drive

The proper way to replace a failed drive will depend on a number of factors, including the number of logical units (ranks) connected to the CRD-5000, the type of enclosure that houses the CRD-5000 and its drives, and whether or not there is a designated spare on the logical unit. In any event, it is important that you promptly identify the failed drive and either replace it or switch to a spare as soon as possible, because the data on the logical unit is vulnerable until the failed drive's data is rebuilt.

Designating one drive on the logical unit as a either a hot or warm spare eases the pressure to replace a failed drive. The CRD-5000 will automatically bring the spare online to take the place of a failed drive. You can then replace the failed drive—and designate the replacement as the new spare—the next time the system is powered down for regular maintenance.

At some point, you will have to replace the failed drive. If you do not have a spare drive on the logical unit, you will need to replace the drive right away. A spare drive will give you a little breathing room, but you still need to replace the failed drive eventually. The first step in replacing a bad drive is to identify which one it is. The preceding section explained how the CRD-5000's front-panel display and monitor utility indicate a failed drive's logical unit and physical disk channel. If your RAID system enclosure does not have some means of indicating a drive's LUN and physical disk channel at the drive itself, you will have to trace the SCSI cable for that disk channel from the CRD-5000 to the drive.

Properly identifying the failed drive is important, because if you replace the wrong drive, you effectively will have created a double failure in the RAID set, and your data will be lost. A RAID set cannot survive the failure of a second drive before the first failed drive is rebuilt.

Once you have located the failed drive, you can replace it with any manufacturer's drive. The capacity of the replacement drive must be equal or greater than the capacity of the smallest drive remaining on the logical unit.

When the CRD-5000 is housed in a specially-designed enclosure, it will support hot-plugging—the ability to replace a failed drive without shutting off power (see Chapter 3 for more information). Otherwise, you are better off powering down the RAID unit before replacing a drive.

Connect the drive to a SCSI cable, either the failed drive's cable or another available cable. (If you select another cable, it must not be connected to a drive already on the same logical unit.) Be sure to make the drive's SCSI ID the same as the other drives on the logical unit and properly terminate the drive's SCSI bus (see Chapter 3 for information about termination). If power is off, turn it back on now and proceed to the next section, which describes how to rebuild the data on the replacement drive.

Rebuilding a Drive

Following a drive failure, a RAID set will remain in degraded mode until you rebuild the data from the failed drive onto a replacement drive. In the case of a hot or warm spare, the CRD-5000 will automatically initiate the rebuilding process. If you are manually replacing a failed disk, you must use the CRD-5000's Rebuild Disk function, available through the front panel configuration menu or the monitor utility.

There is no need to designate a replacement disk drive as online under the front panel's RAID SET PARAMS menu or the monitor utility's RAID Set Parameters menu. In fact, if you try to do this, the CRD-5000 will issue an error message. The Rebuild Disk function takes care of bringing the replacement disk online and making it a part of the RAID set.

The CRD-5000 will perform concurrent rebuild operations on multiple logical units.

Front Panel Rebuild

To initiate the Rebuild Disk function from the front panel, press the "SEL" button after the CRD-5000 has been powered on and completed its self-test. Use the "↑" or "↓" button to cycle through the menu choices until you reach "RAID SET FUNCTIONS." Press the "SEL" button to enter this menu. Press the "↑" or "↓" button until "REBUILD DISK" appears on the display. Press "SEL."

At this point, you have a choice between entering the LUN and channel number of the replacement drive or having the CRD-5000 locate it for you. If you answer "NO" to the prompt "SPECIFY DISK," the CRD-5000 will search for the first disk drive that is cabled up but not online and display its LUN and channel number. If this is the correct disk drive, select "YES" and the CRD-5000 will begin rebuilding the disk drive after you confirm. If this is not the correct disk drive, select "NO" and the CRD-5000 will search for the next eligible disk drive.

If you respond with "YES" to the prompt "SPECIFY DISK," You will be presented with two prompts, one for the LUN of the replacement disk drive and the other for the channel number of the drive. Once you have provided this information, the CRD-5000 will proceed with the rebuild operation after you pres "SEL" a final time to confirm the action.

CAUTION: The rebuild function destroys any data on the disk. Please make sure the replacement drive does not contain any data that you wish to keep.

Monitor Utility Rebuild

You also may use the monitor utility to rebuild a disk. For help in accessing the monitor utility, see chapter 6. Once you are in the monitor utility, select "RAID Set Functions" from the main menu. Then select "Rebuild Disk." The following screen will appear:

CRD-5000 Monitor Utility
REBUILD DISK

04/24/93 07:58:45

Specify Disk To Be Rebuilt?

(Selecting "NO" Will Cause The Disk To Be Selected For You)

NO

UP ARROW: INC	DOWN ARROW: DEC	ENTER: SELECT	CTRL-Z: EXIT

Use the arrow keys to toggle between "YES" and "NO." If you select "YES," the monitor utility will ask you for the LUN and channel number of the disk drive to be rebuilt. Use the arrow keys to scroll among the possible choices and press Enter to select. Press Enter a final time to proceed with the rebuild operation.

If you select "NO," the monitor utility will take it upon itself to search for a likely candidate disk drive for rebuilding. When it finds a disk drive that is cabled up but not online, it displays the following message. (In this example, the first "Unknown" disk drive is on LUN 0, channel 6.)

CRD-5000 Monitor Utility
REBUILD DISK

04/24/93 07:58:45

Rebuild Disk on LUN 0, Channel 6?

YES

UP ARROW: INC	DOWN ARROW: DEC	ENTER: SELECT	CTRL-Z: EXIT

Use the arrow keys to toggle between "YES" and "NO." If you select "NO," the monitor utility will search for the next eligible disk drive and repeat the above display with that drive's LUN and channel. A "YES" selection will begin the rebuilding process and the following screen will appear.

CRD-5000 Monitor Utility
REBUILD DISK

04/24/93 07:58:45

To View The Progress Of The Rebuild(s), Select
The "Rebuild Status" Option From The Main Menu

YES

CTRL-Z: EXIT			

See page 6-28 for an explanation of the Rebuild Status option.

The Rebuild Disk function reconstructs the data on the failed drive by performing a bit-by-bit exclusive-or operation using the data from the functioning data drives and the parity drive. During the rebuild process, the CRD-5000 will continue to handle all I/O requests, either in degraded mode or normal mode, depending on how far the rebuild has progressed and where the data is located. You may monitor the progress of the rebuild operation on all logical units by accessing the monitor utility's Rebuild Status screen (see page 6-28).

The rebuild utility will issue an error message under the following conditions:

- If you attempt to initiate a rebuild on a RAID set that is already undergoing a rebuild.
- If the drive you are attempting to rebuild is already online.
- If the RAID set is not degraded.

Appendix A

Glossary

ASYNCHRONOUS TRANSMISSION: Transmission in which each byte of information is acknowledged individually, through the use of Request (REQ) and Acknowledge (ACK) signals.

CACHE: A large bank of random access memory used for temporary storage of information read from the disk array. When the host requests data from the disk array, the CRD-5000 will first look for the data in the cache. If the data previously was stored in the cache, the CRD-5000 can retrieve the data much faster than it would if it had to go to the disk array.

CONNECT: The function that occurs when an initiator selects a target to start an operation, or a target reselects an initiator to continue an operation.

CONTROL SIGNALS: The set of nine lines used to put the SCSI bus into its different phases. The combinations of asserted and negated control signals define the phases.

DISCONNECT: The function that occurs when a target releases control of the SCSI bus, allowing the bus to go to the Bus Free phase.

FREE: In the context of Bus Free phases, "free" means that no SCSI device is actively using the SCSI bus and, therefore, the bus is available for use.

GIGABYTE: One billion bytes; equal to one thousand megabytes.

HOST: A processor, usually consisting of the central processing unit and main memory. Typically, a host communicates with other devices, such as peripherals and other hosts. On the SCSI bus, a host has a SCSI address.

HOST CHANNEL: A SCSI bus connecting the CRD-5000 with one or more hosts.

INITIATOR: A SCSI device that requests another SCSI device (a target) to perform an operation. Usually, a host acts as an intiator and a peripheral device acts as a target.

I/O: Input/Output. The transmission of data from a peripheral device to the host computer or from the host computer to a peripheral device.

LOGICAL DISK CHANNEL: A *logical* disk channel is the CRD-5000's logical representation of a path to a particular disk drive in the array. A logical disk channel does not necessarily have to correspond with the physical disk channel with the same number.

LOGICAL BLOCK: The logical representation of a physical area on a SCSI hard disk where data is stored. On most disks, a logical block can hold up to 512 bytes of data.

LOGICAL UNIT: The logical representation of a physical or virtual device, addressable through a target. A logical unit can have more than one physical device.

LUN: Logical Unit Number. Used to identify a logical unit.

MTBDL: Mean Time Between Data Loss. The length of time before a hardware malfunction will cause availability of data to be lost.

MTBF: Mean Time Before Failure. The estimated length of time a device will operate before a failure occurs.

MTDA: Mean Time of Data Availability. The probability of two drives in an array failing simultaneously divided by the mean time to replace and regenerate data.

MTTF: Mean Time To Failure. The length of time from the first disk drive failure to the second disk drive failure.

MTTR: Mean Time To Replace. The length of time needed to replace a failed disk drive.

PHASE: One of the eight states to which the SCSI bus can be set. During each phase, different communication tasks can be performed.

PHYSICAL DISK CHANNEL: A *physical* disk channel is a SCSI bus that connects one or more disk drives to the CRD-5000. Each disk drive on a physical disk channel must be part of a different logical unit. Each disk drive in a logical unit must be connected to a different physical disk channel.

PORT: A connection into a bus. The SCSI bus allows eight ports.

PRIORITY: The ranking of the devices on the bus during arbitration.

RAID: Redundant Array of Independent Disks

RAID 0: Striping is implemented at the logical block level. All disks in array are dedicated to data storage. Since there is no redundancy, RAID 0 cannot recover from a disk failure.

RAID 3: Striping is implemented at the logical block level, with error-correction (or parity) information stored on a single, dedicated disk. I/Os generally involve all drives in the array, working in parallel. RAID 3 is best suited for sequential reads of large files, such as in imaging or graphics applications.

RAID 5: Data is striped in multiple logical block units. Parity data is spread evenly across the drives in the array. There is no dedicated parity drive. Handles multiple simultaneous random I/O operations more quickly than RAID 3, because a properly configured system can perform operations by accessing only one disk at the same time as other operations are carried out on other disks in the array. RAID 5 is best suited for database and transaction processing applications.

RAID Set: A group of disk drives that have been configured and initialized to be active members of a disk array.

RANK: A group of disk drives, each on a separate disk channel, that are either part of a RAID set or designated as spares for that RAID set. All the drives on a rank, whether they are active members of a RAID set or spares, must have the same SCSI ID. Each rank connected to the CRD-5000 is assigned a unique logical unit number (LUN).

RECONNECT: The function that occurs when a target reselects an initiator to continue an operation after a disconnect.

RESELECT: A target can disconnect from an initiator in order to perform a time-consuming function, such as a disk seek. After performing the operation, the target can "reselect" the initiator.

SCSI: Small Computer System Interface

SCSI Address: The octal representation of the unique address (0-7) assigned to a SCSI device. This address is normally assigned and set in the SCSI device during system installation.

SCSI ID: The bit-significant representation of the SCSI address referring to one of the signal lines DB0 through DB7.

SLED: Single Large Expensive Disk

SPINDLE SYNCHRONIZATION: The coordination of the spindles of all drives in the array to facilitate parallel transfers of data.

STRIPING: The division of data into smaller sections, which are written in parallel to the disks in the array.

SYNCHRONOUS TRANSMISSION: Transmission in which the sending and receiving devices operate continously at the same frequency and are held in a desired phase relationship by correction devices. For buses, synchronous transmission is a timing protocol that uses a master clock and has a clock period.

TAG QUEUING: A featured introduced in the SCSI-2 specification that permits each initiator to issue commands accompanied by instructions for how the target should handle the command. The initiator can specify whether each command should be executed by the target at the first available opportunity, in the order in which it was received, or at a time deemed appropriate by the target. Tag queuing permits each initiator to have multiple commands outstanding and each target to optimize the commands it receives.

TARGET: A SCSI device that performs an operation requested by an initiator.

TERMINATION: The electrical connection at each end of the SCSI bus, composed of a set of resistors.

Appendix B

Disk Drive Geometry

Although a RAID set on the CRD-5000 appears as one drive to the host, it is in fact a logical representation of several disk drives. This means that the traditional vocabulary for describing a disk drive's geometry—sectors, heads and cylinders—has no meaning in the RAID environment. However, for the benefit of operating systems that require disk geometry information, the CRD-5000 reports the following values for sectors, heads and cylinders:

Sectors	64
Heads	16
Cylinders	Divide the total number of blocks in the RAID set by 1024 and drop the remainder.

The sector and head values are constant, while the cylinder value varies according to the size of the RAID set.

Example

Let's continue with the RAID set example presented in Chapter 6. To determine the number of cylinders go to the monitor utility's RAID Set Statistics screen, which displays the total number of blocks in the RAID set (663040 blocks in this example).

CRD-5000 Monitor Utility
RAID SET STATISTICS
ON LUN 0
04/24/93 07:58:45

Raid Status	NON-DEGRADED	Raid Level	5
Date Stamp	04-08-93 14:38:46	Stripe Size	128 BLKS
Capacity	323 MB	Capacity	663040 BLKS

Phy. Ch.	Log. Ch.	Channel Status	Disk Size	Disk Vendor	Disk Model	Disk Rev.
0		EMPTY				
1	0	ONLINE	100 MB	Acme	LL4000-100	4.07
2	1	ONLINE	150 MB	Elihu Labs	EH482150	3.84
3	2	ONLINE	125 MB	Scion	SO-125E66	6.21
4	3	ONLINE	150 MB	Gigastore	GB100-150	3.32
5	4	ONLINE	125 MB	Ace	LH200B125	2.50
6	5	ONLINE	150 MB	Giant	GS6401250	5.41

UP ARROW: PAGE UP	DOWN ARROW: PAGE DOWN	CTRL-Z: EXIT

The geometry for this RAID set therefore would be as follows:

Sectors	64
Heads	16
Cylinders	647 (663040÷1024, rounded *down* to nearest integer)

Appendix C

Troubleshooting

Symptom	Solution
CRD-5000 does not recognize disk drives	1. Check cabling 2. Check SCSI termination 3. Are drives supplied with power?
Alarm sounds and fault light is on	1. If alarm sounds in slow bursts, the warm temperature sensor has reached its first threshold. Increase air circulation or ventilation space around unit. 2. If alarm sounds in quick bursts, the warm temperature sensor has reached its second threshold. System will shut down. Before turning system back on 3. A disk drive channel has failed.
CRD-5000 does not pass self-test	Note error code displayed on front panel or monitor and call CMD Technology technical support at (714) 454-0800 or (800) 426-3832.
There is no display on the front panel when unit is powered on	Your power supply may be inadequate. Power supply must deliver 12V ±5%, 0.25 Amp and 5V ±5%, 6 Amp.
I/Os occur very slowly	The CRD-5000 will appear to operate very slowly while it is rebuilding a disk drive.
Operating system reports error regarding size of disk drive	Some operating systems cannot handle extremely large disk drives. You may need to remove a disk drive from the array or use smaller disk drives to reduce the total capacity of the array.

Appendix D

Part Numbers

Use the following part numbers when ordering spare or replacement parts for your CRD-5000.

Part No.	Description
SBA-RD5010-000	System Board
SBA-RD5020-000	Data Board
SBA-RD5030-000	SCSI Controller Board
SBA-RD5050-000	Back Panel Board

Index

F

H

M

P

R

S

T

V

Quick Reference

Making the Connections

The CRD-5000 can accommodate up to four ranks (RAID sets), with as many as seven SCSI disk drives in each rank. You may connect as many as four disk drives to a disk channel, but each of these drives must be part of a different RAID set. The CRD-5000 can be equipped with as many as three host channel modules for multi-hosting applications.

Cabling

Use 50-pin SCSI cables to connect the CRD-5000 to each disk drive and to the host. Run one cable from the CRD-5000's host channel port to the SCSI adapter on the host. If you are installing just one rank of drives, run cables from each disk drive to a disk channel port on the CRD-5000. If you are installing more than one rank, attach the additional drives to the disk channel cables. Each member of a rank must be on a separate disk channel.

Termination

The host and disk channel modules in the CRD-5000 are shipped with active termination installed. If the CRD-5000 is located at one end of the SCSI cable to the host and each SCSI cable to the disk drives, then all you have to do is make sure that the SCSI devices or adapters at the *other* end of each cable are properly terminated. If the CRD-5000 does not lie at the end of a host or disk channel SCSI cable, you must open up the CRD-5000 and remove the terminator resistors for that particular module. The terminator resistors are labeled RP1 and RP2.

SCSI ID

Each disk drive in a rank must have the same SCSI ID. Follow the disk drive manufacturer's instructions for setting the ID of each drive. Set the ID for all the drives in the first rank to 0. If you are installing additional ranks, set the ID for all the drives in the second rank to 1, and so on for ranks three and four. The common ID for each rank will become the logical unit number (LUN) that the CRD-5000 and the host will use to identify that rank.

The host channel modules on the CRD-5000 are set at the factory for SCSI ID 0. Make sure this ID does not conflict with any other device connected to the same SCSI cable. If necessary, you may change the CRD-5000's host channel ID via the "Host Parameters" menu on the front-panel or monitor utility. See page 5-11 or 6-11 for more information.

Setting RAID Set Parameters

After connecting the cables and supplying power to the CRD-5000 and its drives, you must configure the system in preparation for creating a RAID set on each logical unit. This can be done via the front panel or a monitor.

Front Panel

Enter the front-panel utility by pressing the "SEL" button after the completion of the self-test. (If password protection is enabled, you will have to enter your password to access the utility.) Press either the "↑" or "↓" buttons until "RAID SET PARAMS" appears on the display. Press "SEL" to drop down to the first submenu, which asks you to enter the logical unit number of the RAID set. Use the "↑" or "↓" button to cycle through the possible logical unit numbers until the display reaches the number you wish to configure. Press "SEL."

The first submenu asks you to select a RAID level. Cycle through the RAID levels with the arrow keys and press "SEL" when the level you desire appears in the display. For more information about RAID levels, see Chapter 1 "Introduction."

The next submenu applies only to RAID level five, which requires the user to select a stripe size. The optimal stripe size (which is measured in logical blocks) should be large enough to enable most read operations to be satisfied on one drive and small enough to prevent excessive amounts of data to be concentrated on one drive. Use the arrow buttons to increase or decrease the stripe size and press "SEL" when you reach the desired value.

Next, designate each drive in the rank as empty, offline, online, hot spare and warm spare. Active members of the RAID set should be online. Use the arrow buttons to cycle through the disk channels. Press "SEL" when you arrive at the channel you wish to designate. Then cycle through the possible designations and press "SEL" again to confirm you choice. Repeat these steps for each drive channel in the RAID set.

Return to the top menu level by pressing "EXIT" until "RAID SET PARAMS" appears on the display. Then use the arrow buttons to scroll the display to "SYSTEM FUNCTIONS." This is where you will perform the final step in configuring the CRD-5000: creating a RAID set.

Monitor

See Chapter 6 for instructions on connecting a monitor to the CRD-5000. Enter the monitor utility by pressing any key on your keyboard after the self-test is completed. (If password protection is enabled, you will have to enter your password to access the utility.) Use the arrow keys on your keyboard to highlight "Setup Parameters" on the main menu and press Enter. Select "Raid Set Parameters" from the submenu and press Enter.

From the "Raid Set Parameters" screen, select a parameter for modification by using the arrow keys and Enter. Use the arrow keys to change a parameter value. Save your change by pressing Enter (Return). Cancel your change by pressing Ctrl-Z.

For explanations of the parameters see the previous section or page 6-7. When you are finished, press Ctrl-Z until you reach the main menu. Then select "RAID Set Functions."

Creating a RAID Set

Creating a RAID set is the process the CRD-5000 uses to configure its internal logic to conform to the RAID set parameters you have selected, among other functions.

Front Panel

Use the arrow and "SEL" buttons to select "CREATE RAID SET" from the "RAID SET FUNCTIONS" submenu. Use the arrow buttons to display the logical unit number of the RAID set. If you are satisfied with the default RAID set parameters (RAID level five, 128-block stripe size, all drives on the LUN online with no spares), toggle to "YES" and press "SEL." If you want to use your own RAID set parameters, toggle to "NO" and press "SEL." Press "SEL" once again to begin creating the RAID set.

If your system has multiple logical units, you must run a separate RAID set creation on each LUN.

Monitor

After you have selected RAID Set Functions from the main menu, the monitor utility will display a secondary menu. Highlight "Create Raid Set" and press Enter.

Press Enter again to proceed past the warning message. On the next screen, use the arrow keys to select the logical unit number of the RAID set to be created. Select "YES" or "NO" to the the CRD-5000's default parameters (RAID level five, 128-block stripe size, all drives on the LUN online with no spares). If you answer "NO" the RAID set will be created with the parameters you have set in the RAID Set Parameters menu. Press Enter at the next prompt to begin the creation. When the RAID set has been successfully created, press Ctrl-Z to return to the main menu.

Replacing a Bad Drive

When a disk drive that is an active member of a RAID set fails, the RAID set will go into degraded mode. The CRD-5000 will continue to handle I/O requests from the host, reconstructing data from the failed disk "on the fly" by using parity information.

Spares

If you have designated a disk drive as a hot or warm spare, the CRD-5000 will automatically bring the spare online in the event of a drive failure and rebuild the data from the failed drive.

Identifying the Bad Drive

The first step in replacing a failed drive is identifying which drive has failed. If you mistakenly replace a functioning drive, you have effectively created a double failure on the RAID set, which will result in the loss of data.

When a drive fails, the front panel will indicate that the RAID set is in degraded mode and identify which drive has failed.

```
LUN 0        STAT NNNONN_-D
          CACHE HITS 0%
```

In this example, disk channel three, which used to be online "N" is now offline "O," indicating that it has failed. The "D" at the far right of the first line indicates that the RAID set is in degraded mode.

The monitor utility's "RAID Set Statistics" screen also shows when a RAID set is degraded and which drive has failed. In such an event, the word "DEGRADED" and the logical channel number of the failed drive will appear in the "Raid Status" field. Do not confuse the logical channel number with the physical channel number. When you go to replace the failed drive, you must replace the drive connected to the *physical* channel that corresponds with the failed logical channel reported by the CRD-5000.

Before you replace the failed drive, it is a good idea to power down the CRD-5000 and all the drives, unless your RAID enclosure is designed to permit "hot plugging" or "hot swapping." Remove the failed drive and replace it with a drive of equal or greater capacity. Be sure to set it to the same SCSI ID as the other drives in the RAID set. Power the system back up and enter either the front-panel or monitor utility.

Rebuilding a Drive

Following a drive failure, a RAID set will remain in degraded mode until you rebuild the data from the failed drive onto a replacement drive. In the case of a hot or warm spare, the CRD-5000 will automatically initiate the rebuilding process. If you are manually replacing a failed disk, you must use the CRD-5000's Rebuild Disk function, available through the front panel configuration menu or the monitor utility. During the rebuild process, the CRD-5000 will continue to handle all I/O requests.

Front Panel Rebuild

To initiate the Rebuild Disk function from the front panel, press the "SEL" button after the CRD-5000 has been powered on and completed its self-test. Use the "↑" or "↓" button to cycle through the menu choices until you reach "RAID SET FUNCTIONS." Press the "SEL" button to enter this menu. Press the "↑" or "↓" button until "REBUILD DISK" appears on the display. Press "SEL."

Choose between entering the LUN and channel number of the replacement drive or having the CRD-5000 locate it for you. If you answer "NO" to the prompt "SPECIFY DISK," the CRD-5000 will search for the first "Unknown" disk drive and display its LUN and channel number. If this is the correct disk drive, select "YES" and the CRD-5000 will begin rebuilding the disk drive after you confirm. If this is not the correct disk drive, select "NO" and the CRD-5000 will search for the next "Unknown" disk drive.

If you respond with "YES" to the prompt "SPECIFY DISK," You will be presented with two prompts, one for the LUN of the replacement disk drive and the other for the channel number of the drive. Once you have provided this information, the CRD-5000 will proceed with the rebuild operation after you pres "SEL" a final time to confirm the action.

Monitor Utility Rebuild

You also may use the monitor utility to rebuild a disk. Once you are in the monitor utility, select "RAID Set Functions" from the main menu. Then select "Rebuild Disk." Have the monitor utility find the replacement drive or specify the LUN and channel number yourself. Press Enter to begin the rebuilding operation.

Computing Surface

CS-2 Software Installation

84-CB074

Meiko Internal Use Only

Meiko's address in the US is:

Meiko
130 Baker Avenue
Concord MA01742

508 371 0088
Fax: 508 371 7516

Meiko's address in the UK is:

Meiko Limited
650 Aztec West
Bristol
BS12 4SD

Tel: 0454 616171
Fax: 0454 618188

Issue Status:		
	Draft	
	Preliminary	x
	Release	
	Obsolete	

Circulation Control: *Internal. This documentation contains installation notes for Meiko engineers only.*

Contents

Preface *i*

About this Document

This document describes how to install the Meiko version of the Solaris operating system, the operating system patches and packaged software on to a CS-2. It assumes that the hardware is correctly assembled, configured and tested.

The installation procedure installs the operating system from a CD-Rom and the patches and packages from a release tape.

Upgrading the operating system on an existing installation is also covered.

Unix Installation and Configuration 1

This chapter describes how to install the Meiko Solaris operating system. Patches to the OS are covered in the next chapter.

Hostnames

A generic hostname for the machine is required and should be specified by the customer. The examples in this document use `cs2-` as the generic hostname.

CS-2 nodenames are constructed from the generic name to which is appended each processor's Elan Id. The generic name should be short and must not end in a numeric character. The Elan Id can be padded with leading zeros if you wish. This has the advantage that the clients will always be listed in numerical order but the disadvantage that some of the names will be longer than if they were not padded.

The hostname on the Elan network for each node will be the same as the nodename for that node.

Some nodes may have multiple network interfaces, particularly the main server node, and thus will require unique hostnames for each network. Usually the Ethernet hostname is the nodename appended by `-le0` so that it is easily identified in a list of CS-2 hostnames.

Internet Addresses

An Internet (IP) address for the main server on the local Ethernet and a network number to be used for the internal Elan network created within the machine are also required. The latter should be a class C network or a subnet of a class A or B network and should be supplied by the customer.

Adding the CS-2 to Meiko's Local Network

Allocating a Local IP Address

During the initial installation the CS-2 main server processor may be connected to Meiko's Ethernet network. To do this we need to allocate a temporary local Ethernet IP address. This can be done by our System Administrator, who can also provide a temporary internal Elan IP network address if necessary.

For example:

Hostname	IP Address	Network
`cs2-0-le0`	`192.131.108.190`	`Ethernet`
`cs2-0`	`193.39.185.1`	`Elan-net`
`cs2-1`	`193.39.185.2`	`Elan-net`
`cs2-2`	`193.39.185.3`	`Elan-net`

If the Ethernet interfaces on additional nodes are to be used, Ethernet IP addresses can also be allocated for those.

For example:

Hostname	IP Address	Network
`cs2-4-le0`	`192.131.108.191`	`Ethernet`

Adding the CS-2 to our Local Network

The NIS maps on the NIS server (currently `hub`) need to be updated to incorporate the Elan and Ethernet hostnames and IP addresses of the machine. This should be undertaken by our System Administrator.

Hostids

A unique hostid needs to be allocated for the machine. It is essential that every node in the machine has the same hostid, in order for NFS over Elan to work.

Installing the Operating System

Preliminary

Check that the ethernet address, hostid, serial number and modification level are all correctly set:

```
ok banner
ok .idprom
```

Check that the I/O devices are all accessible over the SCSI buses:

```
ok probe-scsi-all
```

You can use `devalias` to check what device aliases are available.

Booting the Main Server from a local CD-Rom Drive

If the CD-Rom drive is attached to SCSI bus 0, the server can be booted from it with the following command:

```
ok boot cdrom
```

If the CD-Rom drive is attached to SCSI bus 1, and the Open Boot Prom is Rev. 136 or later, use the following command:

```
ok boot cdrom1
```

Otherwise the following command will have to be used:

```
ok boot /iommu/sbus/dma@3,400000/esp@3,800000/sd@6,0:b
```

If the CD-Rom drive is attached to a differential SCSI card in SBus slot 1, the following command should be used:

```
ok boot /iommu/sbus/dma@1,81000/esp@1,80000/sd@6,0:b
```

During the boot process you will need to pre-configure the system and answer some questions from the `suninstall` program.

Pre-configuring the System

See *Configuring a Networked System after Booting Solaris 2.3* in Chapter 5 of the *Solaris 2.3 System Configuration and Installation Guide.*

1. **Specify the terminal type.**
   Enter the number corresponding to the type of terminal you are using. This will usually be `DEC VT100` for a tip or telnet connection in an xterm window. This question only appears if you are *not* using the graphics console.
2. **Specify the hostname of this processor.**
   Give the Ethernet hostname (usually the nodename appended by `-le0`).
3. **Specify that the processor is connected to a network.**
4. **Specify the processor's IP address.**
   Enter the ethernet IP address.
5. **Select a network naming service.**
   Select `none` so that AdminTool will use the local `/etc` files.
6. **Identify sub-nets.**
   You should specify that there are no sub-networks, unless you know that sub-networks are used at the site where the machine is to be installed.

7. **Specify your geographic region and timezone.**

8. **Enter the current date and time.**

If the machine is connected to the Meiko internal network and the ethernet address of the processor is that of a machine which has already been installed on the Meiko network, items 2, 3 and 4 will be assumed to relate to that machine and you will not see the questions.

suninstall

See Chapter 7 of the *Solaris 2.3 System Configuration and Installation Guide.*

`suninstall` is started automatically following your pre-configuration. On-screen messages will guide you through selecting items. Generally, you use <Tab> to move to the next field, <Shift-Tab> or <^R> to move to the previous field and <Return> to select a field; <F1> provides help if you are working on the Sun graphics console, <^W> if not. Sometimes you can scan choices with the up/down arrow keys, in which case you can use <^D> or <^U> to scroll down or up a page. If the display gets corrupted, you can use <^L> to refresh the screen.

1. **Format/label your hard disk .**
   If your hard disk has never been used under Solaris it has to be labeled. You should exit `suninstall`, label the disk and then restart `suninstall`. To exit `suninstall` select `Exit Install...` from the `Solaris Installation` menu, to label a disk use the `format` command and to restart suninstall type `suninstall`.

   If you attempt to use an unlabelled disk with the `suninstall` program the size of this disk will be shown as 0 Mbytes.

2. **Specify the Installation Type.**
   Select `Custom Install...` from the `Solaris Installation` menu.

3. **Specify the System Type.**
   Select `System Type...` from the `Custom Install Configuration` menu. Since the processor you are installing is to be a server for diskless clients, you should select the `Server` system type from the `Choose System Type` menu. Then select `Apply`.

4. **Specify the Server Parameters.**
Select `Server Params...` from the `Custom Install Configuration` menu and check that `sparc.dino1` has an asterisk against it in the `Client Architectures` section if you have a Dino1 board. The Dino2 board has the `sparc.sun4m` architecture. You cannot deselect the architecture of the server but you can add others.

   Specify the number of diskless clients and the amount of swap space required by each. In general, the amount of swap space will be similar to the amount of memory that each client has, unless the clients will swap to local disks, in which case you can specify 1 MB.

5. **Specify the Software Selection.**
Select `Software Selection...` from the `Custom Install Configuration` menu. Then select `Entire Distribution` and use `Edit...` to deselect `SunButtons and Sundials support`. The <space> key toggles selection. Then select `Done`.

6. **Specify the Disk Configuration.**
Select `Disks/File Systems...` and configure each disk that is listed in the menu. You should increase the default size of the `/opt` partition by about 45Mbytes to accommodate the Meiko software packages (installed later) and allow a further 30 Mbytes if you intend to install the Sun compilers. Also increase the default root (`/`) partition size to about 64 Mbytes and increase `/export` by about 10 Mbytes per client, to allow space for installing patches. Use `Space...` to check on the minimum required for `/export/swap` and use that. If your partition sizes are not large enough to continue the installation you will be prompted to find more space. The installation process will not touch any unconfigured disks.

7. **Start the installation.**
Select `Begin Install`. The installation takes about 1 hour. Once the installation process has completed, the processor will automatically reboot from its local disk. This behaviour can be changed in `Props...` if required. You will be prompted to select a root password as the processor reboots.

Fine Tuning the Installation

Warning – The CSUTILS product described below is constantly being updated. You should consult the OVD documentation to determine the correct version to install.

---

1. **Disable the automounter.**
   After the reboot has completed, you should log in as root, disable the automounter and then reboot the node again:

```
# cd /etc/rc2.d
# mv S73autofs -S73autofs
# reboot -- -v
```

2. **Install Root Environment Packages.**
   After the reboot has completed, you should log in as root and install the root environment product CS2UTILS from tape. The CS2UTILS product contains two packages, MKrootenv and MKmisc. These are described in appendix A.

   The Meiko software tape will first have to be unloaded in a suitable directory on the local disk, such as `/export/home/packages`. For example:

```
# mkdir -p /export/home/packages
# cd /export/home/packages
# tar xvf /dev/rmt/0
# cd CS2UTILS.n.n.nn
# ./meikoinstall MKmisc MKrootenv
```

   Log out and log in again to pick up the new root environment.

3. **Rename the server.**
   The server currently has the Ethernet hostname as its nodename. This must now be changed so that the Elan hostname is used as the nodename. Add the new name into the `hostname.elanip0` file:

```
cs2-0-le0# uname -S cs2-0
cs2-0-le0# uname -n > /etc/hostname.elanip0
```

4. **Edit** `/etc/inet/hosts`.
   The `/etc/inet/hosts` file must now be edited to include the Elan internet address and hostname of the server processor:

```
192.131.108.190 cs2-0-le0
193.39.185.1 cs2-0
```

5. **Tune the clients' swap filesystem.**
   You can make better use of the `/export/swap` filesystem by tuning it so that it doesn't reserve 10% of its space:

```
cs2-0-le0# umount /export/swap
cs2-0-le0# tunefs -m 0 /export/swap
cs2-0-le0# mount /export/swap
```

Installing Patches on the Main Server 2

The Operating System has a number of patches that need to be installed from the Meiko release tape. The patches are located in the `patches` directory where the release tape was unloaded, e.g. `/export/home/packages/patches`.

A script `meikoinstallpatches` is supplied on the tape, to facilitate the installation of the patches.

1. **Install the patches.**
   Move into the release directory and run the patch installation script:

```
cs2-0-le0# cd /export/home/packages/MSOL23.nn.nn
cs2-0-le0# ./meikoinstallpatches
```

2. **Check the Elan NFS Network Number.**
   The Elan NFS system requires a unique network number to be assigned to each CS2. This is normally read from the machine's 'hostid' ROM parameter. *All nodes in the machine must have the same hostid.*

If the machine has Open Boot Proms earlier than Version 127, it is not possible to set the hostid independently of the board serial number. In this case, the `/etc/system` file should be edited to contain the following:

```
* Set NFS over Elan network number
* (Should really be defined by machine hostid)
set nfs:nfs_elanNet = 0x25000000
```

Replace `25000000` with the hostid of the main server node.

3. **Reboot the machine.**
   Reboot the machine again:

```
cs2-0# init 0
ok boot -v
...
```

4. **Confirm the network is operational.**
   Log in as `root` and check the Elan network:

```
login: root
Password: password

# /sbin/ifconfig -a

elanip0:
 flags=1863<UP,BROADCAST,NOTRAILERS,RUNNING,MULTICAST,MULTI_BCAST> mtu 8232
 inet 193.39.185.1 netmask ffffff00 broadcast 193.39.185.255
 group 60001(nobody)
 ether 0:0:0:0:0:0
```

First Client Installation 3

We build diskless client nodes by first building one client, installing the kernel patches and root packages on it, booting and configuring it, and then we clone it to make the other clients.

This chapter describes how to install the operating system and patches on to the first client processor.

This procedure must be repeated on all of the servers in the CS-2 system.

1. **Install the operating system for the first client.**
   Log in to the server node as root, start OpenWindows (if on the graphics console) and run the AdminTool in a suitable window.

```
cs2-0# /usr/openwin/bin/openwin -nobanner
. . .
cs2-0# admintool &
```

   Use the Admintool's Host Manager to install the client processor and specify `Naming Service: None,Client Type: diskless.` When prompted for an Ethernet address for the client, you should enter its MAC address. This is a 6 field hexadecimal address in which normally only the last field is used, and this will be the same as the processor's Elan Id (for example, processor 12 has a MAC address of `0:0:0:0:0:c`). The processor's Elan Id is stored in the `elan-node-id` variable in the NVRAM.

You should note that `admintool` allows the swap file size to be no more than 100MB. If you want the clients to have more swap than this, specify 1MB of swap at this stage and use `mkfile` later to create the real swapfile:

```
# rm /export/swap/cs2-1
# mkfile 128m /export/swap/cs2-1
```

Confirm that the client has been built correctly by checking the size of the root partition on the boot server disk. The size should be around 20418 blocks.

```
# du -s /export/root/*
```

2. **Add the root environment product to the client.**
   You should now install the root environment product CS2UTILS on the client, by running the `meikoinstall` script on the server:

```
# cd /export/home/packages/CS2UTILS.n.n.nn
# ./meikoinstall MKrootenv
```

   The meikoinstall script will not attempt to reinstall the packages on the server but will perform the installation on the client.

3. **Add the patches to the client.**
   Run the `meikoinstallpatches` script on the server again:

```
# cd /export/home/packages/MSOL23.nn.nn
# ./meikoinstallpatches
```

   This will install the patches on the client. Again the script will see that the patches have already been installed on the server and will not attempt to reinstall them.

4. **Check the Client's Elan NFS Network Number.**
If the client has an Open Boot Prom earlier than Version 127, the file /export/root/cs2-1/etc/system should be edited to contain the following:

```
* Set NFS over Elan network number
* (Should really be defined by machine hostid)
set nfs:nfs_elanNet = 0x25000000
```

Replace 25000000 with the hostid of the main server node.

Package Installation 4

This chapter describes how to install the software packages. The procedures in this chapter should be used on all server processors in your system. Packages of type `root` will be applied to any clients already installed on the current server, as well as to the server itself. Packages of type `usr` will only be installed on the server and packages of any other type (except `root_na` and `usr_na` - see later) will only be installed on the main server.

Warning – The products described below are constantly being updated. You should consult the OVD documentation to determine the correct software to install.

Package Installation

The software products have already been unloaded from tape into a release directory, e.g `/export/home/packages`. Each software product is a collection of packages. The basic products are MEIKOcs2.n.n.n (Meiko CS-2 utilities) and MEIKOab.1.3.00 (the Meiko AnswerBook).

The installation procedure for all products is to move to the release directory and execute the local meikoinstall script. On the main server only (i.e. the one serving /opt), specify the main argument to meikoinstall, as shown below.

```
cs2-0# cd /export/home/packages/MEIKOcs2.n.n.nn
cs2-0# ./meikoinstall main
```

Further information on meikoinstall can be found in Appendix B.

When you install the Meiko AnswerBook product it will ask you to specify the parent of the AnswerBook home directory. It is usual to specify /opt but note that this will use about 15 MB of disk space in this directory.

Special Packages

Certain packages, such as the FDDI device driver, do not need to be installed on all nodes and they have the type root_na or usr_na, where na stands for not automatic. These packages will *not* get installed by meikoinstall - they must be installed by hand on the server responsible for the node on which the device is installed.

For example, if there is an FDDI interface on cs2-9, a diskless node booted from cs2-4, the following sequence of commands should be entered on cs2-4:

```
cs2-4# pkgadd -d /export/home/packages/FDDI.n.n.nn/packages MKfddiu
cs2-4# pkgadd -d /export/home/packages/FDDI.n.n.nn/packages -R \
        /export/root/cs2-9 MKfddir
```

The Setupcs2 Scripts. 5

Various system files now need to be configured to support the CS-2 architecture, for which purpose the scripts `setupcs2files` and `setupcs2server` are provided.

The setupcs2 process can be divided into three parts:-

- Set up the CS-2 description file.
- Create the templates of all the files to be copied, using `setupcs2files`.
- Copy the files on to every server and its clients, using `setupcs2server`.

Set up the CS-2 Description File

The `setupcs2files` script depends on the existence of a data file, named `/opt/MEIKOcs2/etc/cs2.des`, which describes the machine being set up. A template file, `/opt/MEIKOcs2/etc/cs2.des.template`, is included in the release and should be used to create the `cs2.des` file as shown below. This file is well commented. Edit the file and fill in the blanks.

```
cs2-0# cd /opt/MEIKOcs2/etc
cs2-0# cp cs2.des.template cs2.des
cs2-0# vi cs2.des
```

It is important to describe the full machine, not just the server and single client built so far.

Create the CS-2 Distribution Files

All the files which will be distributed by the `setupcs2server` script are created by the `setupcs2files` script into `/opt/MEIKOcs2/dist`. More information on the structure of this directory can be found in Appendix C. The `setupcs2files` script needs to be run by root, on the main server only, as shown:

```
cs2-0# cd /opt/MEIKOcs2/etc/scripts
cs2-0# ./setupcs2files
```

The `-h` option can be used to give some basic information about the script and the `-n` option can be used to show what files will be created without actually creating them.

Copy the CS-2 Files to Server and Clients

Once the `setupcs2files` script has been run, the `setupcs2server` script should be used on every server to copy the files to that server and its clients. It requires two arguments, the name of the server you are running on followed by the name of the main server:

```
cs2-0# ./setupcs2server cs2-0 cs2-0
```

On secondary servers, the `/opt` directory should be mounted from the main server before it is run:

```
cs2-4# mount cs2-0:/opt /opt
cs2-4# cd /opt/MEIKOcs2/etc/scripts
cs2-4# ./setupcs2server cs2-4 cs2-0
```

The files which are updated have been divided into groups and you can choose only to update the files in a particular group by giving a third argument, so the following command:

```
cs2-0# ./setupcs2server cs2-0 cs2-0 basic
```

would only update the files in group `basic`. Currently there are three groups, `basic`, `yp` and `vfstab`, as described in `/opt/MEIKOcs2/etc/setupcs2.notes` (or checkout `cs2parts/admin/setupcs2.notes` for a full description of the groups and how files are updated according to node type). If no group argument is given to `setupcs2server` the files in all groups will be updated.

NB. If you are going to clone clients you should run `setupcs2server` on the server for the client from which you are going to clone the rest of the clients before you do the cloning.

Booting the First Client 6

This chapter describes how to boot and configure the first client node.

The procedure in brief is:

1. Reboot the server
2. Start openwindows
3. Set up the machine description file.
4. Set up the machine manager defaults file.
5. Start the machine manager.
6. Set up the pandora defaults file.
7. Start pandora.
8. Configure the client's eeprom settings.
9. Boot the client.
10. Configure the client.

Setting Up.

Reboot the server.

Before trying to boot the client for the first time, the server must be rebooted in order to start up the daemons necessary to support diskless clients.

Start up openwindows.

If working on the CS-2 graphics console, openwindows should be started up at this stage to provide an environment in which pandora can run. Otherwise, you should work on a suitable graphics display on another machine and set the environment variable `DISPLAY` to the name of that machine.

Set up the machine description and defaults files.

These files are described in the CS-2 system administration documentation. The file `/opt/MEIKOcs2/etc/<name>/machine.des` must describe the hardware configuration of your machine. Some example files are provided which can be used as a basis for the new file, named `template[123].des`.

The `setupcs2` process will have created a machine manager defaults file `/opt/MEIKOcs2/etc/<name>/defaults` which should be examined carefully. Before starting the `mmanager` daemon for the first time, check that the `access-control` status is set to `off`. It may also be necessary to set `halt-on-error` to `off`.

Start the `mmanager` *daemon.*

Start the machine manager daemon on the main server processor in a separate window. The `-irll` flags cause the mmanager daemon to start verbosely and log to `stdout` rather than to a file.

```
cs2-0# /opt/MEIKOcs2/bin/mmanager -irll
```

Set up the pandora defaults file.

The file `/opt/MEIKOcs2/etc/<name>/pandora/defaults` should be examined carefully and edited if necessary.

In particular, you might want to check the elan nfs packet size is 64k rather than the default 16k. This involves setting the `ep_big_msg_size` line to 69568. This should match the `ep_bigmsgsize` value in the eeprom on each node.

Warning – This size must be consistent across the whole machine. It may be necessary to set this value on the server nodes and then reboot them.

Start pandora

Start the pandora utility.

```
cs2-0# /opt/MEIKOcs2/bin/pandora &
```

You should perform the following actions with pandora.

1. Configure in the module controllers, switches and nodes you know to be functional.
2. Set the configuration of all the Elite network switches.
3. Set the eeprom values of all the client nodes.

Boot the client

Using pandora, obtain a console on the first client and then check that the ethernet address, hostid, serial number and modification level are all correctly set:

```
ok banner
ok .idprom
```

Test the Elan interface:

```
ok reset
ok test /elan
```

If this fails, check the machine description file and the configuration of the network switches and eeprom values.

Then boot the node.

```
ok boot -v
```

Configuration

In response to the questions from the client, specify that you are using a VT100 (3), Sun Command Tool (7) or Sun Workstation (8) as appropriate, that you do not want to configure this system as a client of a name service and that there are no sub-networks. Set the timezone, time and root password when prompted.

This will have the effect of overwriting `/etc/nsswitch.conf` on the client, so you need to run `setupcs2files` and `setupcs2server` again on the server:

```
cs2-0# cd /opt/MEIKOcs2/etc/scripts
cs2-0# ./setupcs2files
cs2-0# ./setupcs2server cs2-0 cs2-0
```

Log in to the client as root, disable the automounter and then reboot the node:

```
cs2-1# cd /etc/rc2.d
cs2-1# mv S73autofs -S73autofs
cs2-1# reboot -- -v
```

Local Disks

See the Solaris documentation *SunOS 5.3 Routine System Administration* for more information about the following procedures.

If the client has a local disk, it will need to be formatted and labeled using the `format` command and then have a unix file system built on it using the `newfs` command. It is customary to configure the disk with one single large partition in slice 7 and mount it on `/var/tmp`.

For disk c0t0d0 the following entry should be added to the /etc/vfstab file.

```
/dev/dsk/c0t0d0s7 /dev/rdsk/c0t0d0s7 /var/tmp ufs 1 yes -
```

Once the local disk is mounted, it can be used for local swapping instead of swapping back to the server node. In order to do this a swapfile must be made on the local disk using the mkfile command.

```
cs2-1# mkfile -v 160m /var/tmp/swapfile
```

The /etc/vfstab file should be edited to disable remote swapping and enable local swapping:

```
#cs2-0:/export/swap/cs2-1 - /dev/swap nfs - - -
#/dev/swap - - swap - - -

/var/tmp/swapfile - - swap - - -
```

If you disable swapping back to the server node then the swapfile can be removed from the list of exported files in /etc/dfs/dfstab on the server node and the swapfile in /export/swap deleted.

If you reuse the file space freed by deleting swap files then it is a good idea to use the tunefs command to reserve some space that can be used only by root, to avoid excessive fragmentation of the filesystem.

Other Configuration

If you have edited the files /etc/passwd, /etc/shadow, /etc/groups, /.cshrc or /.profile on the server, you might like to copy them on to the client node.

Cloning 7

This chapter describes how to create the remaining clients on a server.

The procedure is as follows, and should be carried out on each server:

1. Create the clone datafiles,
2. Create new clients from the clone datafiles,
3. Boot each client.

All the scripts necessary for cloning the client nodes are stored in the directory `/opt/MEIKOcs2/etc/scripts/clone`.

Create the clone datafiles.

Execute the `createClone` script, giving the name of the first client as a parameter.

```
cs2-0# cd /opt/MEIKOcs2/etc/scripts/clone
cs2-0# createClone cs2-1
Creating /tmp/cloneMaster.cpio from cs2-1
Creating /tmp/cloneVfstab ... done
Packing client clone ... 24016 blocks
Finished creating a cloneMaster from cs2-1.
```

This will create two datafiles, named /tmp/cloneMaster.cpio and /tmp/cloneVfstab.

Cloning more nodes.

A script should be created to build all of the remaining clients. The sample file clones.example is provided to copy and edit as appropriate. A separate script is required for each server. The Ethernet addresses should be given in hexadecimal. When creating your script, ensure that it does not cause the original client to be overwritten. Remember to give the script execute permission.

It takes about 3 minutes to create each new client. Assuming you have created a script called clones which clones a single node named cs2-2, you should see something like:

```
cs2-0# clones
Creating client cs2-2
Warning: Entry already exists in /etc/hosts.
Creating 128m swapfile ... done
Unpacking client clone ... 24016 blocks
Finished creating client cs2-2.
```

The warning is because the setupcs2server script will already have made an entry for each client in the /etc/hosts file, and can be ignored.

Boot the new clients.

Once all the clients have been created they can be booted by using the pandora utility. It is not necessary to boot each node from its console since the requisite configuration has already been performed on the first client. Instead the boot procs option in the reset menu can be used. However, you must ensure that the idprom and nvram are set correctly on each client, as described under *Booting the First Client* above.

It may be necessary to reboot the server before the clients will boot.

Secondary Server Installation 8

This chapter describes how to build additional server nodes. Building secondary servers is very similar to building the main server, except that it is not necessary to run the `setupcs2files` script, start up mmanager or use the `main` argument to `meikoinstall`.

Secondary servers can be built in one of two ways. One option is to attach the CD-Rom drive to each server in turn and install the software directly from CD as described above for the main server. This has the disadvantages that only one server can be built at a time and that it may involve moving SBus cards around if the CD-Rom drive has a differential SCSI interface.

The alternative is to set up the main server node as an installation server for the secondary servers and to install the software over the Elan network. This method is described below.

Setting up the Main Server Node

Log in to the main server and start the Administration Tool in a suitable windows environment, with the command `admintool`. You need to use admintool's Host Manager, select the local file option (`Naming Service: None`) and specify the secondary server processors as standalone clients. Enable remote install. This will allow the processors to download the operating system from the CD-Rom on the main server and create local boot disks. Note that this will fail if the new nodes are already in the `hosts` file!

Building Secondary Server Nodes

Start pandora on the main server and open a console for each server. Set the nvram parameters as for the client nodes. The `elan-boot-id` parameter must be set to the elan id of the main server rather than the value set by pandora.

Check that the ethernet address, hostid, serial number and modification level are all correctly set:

```
ok banner
ok .idprom
```

Check that the I/O devices are all accessible over the SCSI buses:

```
ok probe-scsi-all
```

Test the Elan interface:

```
ok reset
ok test /elan
```

If this fails, check the machine description file and the configuration of the network switches and eeprom values.

Then boot the node over Elan:

```
ok boot elan -v
```

During the boot process you will need to pre-configure the system and answer some questions from the `suninstall` program, as documented for the main server.

Unlike on the main server, the `/opt` partition on secondary servers will be over-mounted by the master `/opt` directory from the main server. Thus the `/opt` directory need not be increased in size to hold the Meiko software. In fact, it can be reduced to the minimum suggested by `Space...`

When booted, the automounter should be disabled and the name and IP address of the main server should be added to the local `/etc/hosts` file (this will later be updated by `setupcs2server`).

The packages directory should now be mounted by hand from the main server, making sure that you do *not* specify the `elan` option, and the CSUTILS product should be added and the relevant patches installed as described above.

Reboot, after which NFS over elan should be working and you should be able to mount filesystems from the main server using the `elan` option.

Build the first diskless client using AdminTool, as described above.

Install the software packages, making sure that you do *not* use the `main` argument to `meikoinstall`.

The `/opt` partition of the main server should now be mounted by hand and the `setupcs2server` script run to configure the local node. This will use the configuration file (`cs2.des`) that was previously created for the main node.

E.g. For node `cs2-4` you should execute:

```
cs2-4# mount cs2-0:/opt /opt
cs2-4# cd /opt/MEIKOcs2/etc/scripts
cs2-4# ./setupcs2server cs2-4 cs2-0
```

Among other things this will add the `/opt` mount command to the files in the `/etc/vfstab.dir` directory.

Reboot the node again to create the final configuration.

Adding Clients

Client node root partitions can be cloned in exactly the same way as on the main server.

In order to save time it is possible to copy the `/tmp/cloneMaster.cpio` and `/tmp/cloneVfstab` files from the main server, instead of using AdminTool to build the first client. However you must then run the `setupBootServer` script to create a `/tftpboot` directory and perform various other setup actions normally performed by admintool.

Upgrading from CD-Rom 9

This chapter describes how to upgrade the Meiko Solaris operating system from a local CD-Rom drive.

The procedure in brief is as follows:

1. Boot the main server from CD-Rom
2. Pre-configure the system
3. Select upgrade from suninstall
4. Install the patches
5. Reboot
6. Install the packages
7. Run the setupcs2 scripts
8. Upgrade any additional server nodes.

Booting the Main Server

If the CD-Rom drive is attached to SCSI bus 0, the server can be booted from it with the following command:

```
ok boot cdrom
```

If the CD-Rom drive is attached to SCSI bus 1, and the Open Boot Prom is Rev. 136 or later, use the following command:

```
ok boot cdrom1
```

Otherwise the following command will have to be used:

```
ok boot /iommu/sbus/dma@3,400000/esp@3,800000/sd@6,0:b
```

If the CD-Rom drive is attached to a differential SCSI card in SBus slot 1, the following command should be used:

```
ok boot /iommu/sbus/dma@1,81000/esp@1,80000/sd@6,0:b
```

During this boot process you will need to pre-configure the system as described above.

When `suninstall` starts up, select `Upgrade...` from the `Solaris Installation` menu. This will then check to ensure that there is sufficient space available in each partition for the upgrade to take place. If the space check fails, you will be told how much free space is needed in each partition and you will have to find some additional space. This information is also stored in the file `/var/sadm/install_data/upgrade_space_required`.

Once the space check succeeds, the upgrade will proceed, taking up to an hour.

Reboot the node and start OpenWindows, if you are working on the CS-2 graphics console.

The Meiko software tape will have to be unloaded in a suitable directory on the local disk. For example:

```
# mkdir -p /export/home/packages
# cd /export/home/packages
# tar xvf /dev/rmt/0
```

Then the Meiko patches and packages should be installed and the setupcs2 scripts run, all as described above.

Once this has been done, pandora can be started up to give access to the consoles of the secondary server nodes.

Booting the Secondary Servers

The secondary servers can now be booted over elan from the CD-Rom on the main server, as described in the section entitled *Secondary Server Installation* above. They can then be upgraded as described earlier in this section.

Additional Notes 10

Backing Out of Patches

First back out of each patch on the clients and then back out on the servers.

In the following examples the MEIKO-1-023-00-000 patch is removed.

For the clients, log in to each server and type:

```
cs2-0# foreach r (/export/root/cs2-*)
> cd $r/var/sadm/patch/MEIKO-1-023-00-000
> ./backoutpatch -R $r MEIKO-1-023-00-000
> end
```

For the servers:

```
cs2-0# cd /var/sadm/patch/MEIKO-1-023-00-000
cs2-0# ./backoutpatch MEIKO-1-023-00-000
```

Installing SPARC Compilers

The SunSoft compilers are installed from a CD-ROM. If there is not one there already, connect a CD-ROM drive to the SCSI bus on the boot server processor. To install the Sun C compiler and SparcWorks, type:

```
cs2-0# mkdir -p /cdrom
cs2-0# mount -r /dev/sr0 /cdrom
cs2-0# cd /cdrom/Solaris_2.0
cs2-0# ./spro_install cc
cs2-0# ./spro_install sw
```

You will also have to install compiler licences.

Using rdist to Distribute Files

One of the files that the `setupcs2files` script creates on the main server is `/opt/MEIKOcs2/dist/distfile`. In this file a macro ALLNODES is defined to represent all the nodes' names. Another macro ALLNODE-FILES is also defined — currently this represents `/etc/passwd` and `/etc/group`.

Files can be added to the list of files and `rdist` used to distribute them to all nodes of the machine:

```
cs2-0# rdist -f /opt/MEIKOcs2/dist/distfile
```

There is scope for more subtle use of this file — see Robin if you have any suggestions.

The CS2UTILS Product — A

The CS2UTILS product is an internal Meiko product to facilitate the building and configuration of CS2 machines. It contains two packages, MKrootenv and MKmisc.

This product is intended to be a complete replacement for the uncontrolled tar archives that have been used in the past. If there is something missing from this product please add it to the packages rather than creating a new tar archive of missing bits.

The MKrootenv Package.

This is a root package which will be installed on all servers and clients, and contains the following files:

1. Environment configuration files for root: `.cshrc`, `.profile`, `.openwin-init`, `.openwin-menu`, `.openwin-menu-cs2`, `.Xdefaults` and `.xintrc`.
2. New versions of the system configurations files: `/etc/nsswitch.conf` and `/etc/default/login`.
3. The `/etc/updatepatches` script. This is used for updating the Solaris OS patches on the machine.
4. The `/etc/cs2build` script. This script is run by the MKrootenv package postinstall script and performs the following actions:

a. Stops the `in.rdisc` daemon if running and disables the executable so it will not restart at boot time. The `setupcs2` scripts will install a new version of the `/etc/rc2.d/S69inet` script later in the installation procedure.

b. Stops the automounter daemon if running and disables the `/etc/rc2.d/autofs` script that starts it at boot time.

c. Disables NIS+.

5. `/.rhosts` containing only a `+`. Although this is insecure it allows the `setupcs2server` scripts to be run remotely and `/.rhosts` is immediately overwritten by the `setupcs2server` script to contain the names of all the nodes in the machine.

The MKmisc Package.

This is a usr package which will be installed on all servers and contains the following binary files:

`/usr/bin/bash`, `/usr/bin/tcsh`, `/usr/bin/top`, `/usr/bin/ue` and `/usr/openwin/bin/olvwm`.

These files are simply held in the CVS repository in their binary form and are solely for use by Meiko personnel when building and configuring a CS-2 machine.

Warning – This package should be removed before a machine is shipped to a customer site.

meikoinstall Manual Page **B**

meikoinstall — Install Meiko products

Synopsis

`meikoinstall [-i][-r][-x][-f][-h][-n][-R root | `*client*`]`
`[main]` [*pkg*...]

Description

`meikoinstall` is used to install Meiko Products. Meiko products are collections of packages which can be installed using the Solaris `pkgadd` command. Using `meikoinstall` ensures that all required packages are added and removes the need for interactive installation. Before adding a package `meikoinstall` will remove (using the Solaris `pkgrm` command) any package of the same name already installed.

`meikoinstall` is released with all Meiko CS-2 products and is invoked from the top level. When a Meiko product is untarred there will be a subdirectory `packages` and `meikoinstall` will process all the packages in this subdirectory.

`meikoinstall` needs to be run on every server on a CS-2 machine. It will install the relevent packages on the server and its clients. There are three types of packages which can be released, application packages, root packages and usr packages. Application packages will only be installed on the main server (`meikoinstall` needs to be invoked with the `main` argument when running on the main server). `usr` packages will be installed on servers but not clients and `root` packages will be installed on servers and clients (i..e. everywhere).

Arguments

`-i` Interactive mode. If you use this flag the packages will be installed in interactive mode and will require user input. The default mode is non-interactive but this is impossible for some packages (e.g. MKab001 — the Meiko Answerbook).

`-r` Remove mode. This option will remove all the packages which are present in the `packages` subdirectory. It will NOT remove the packages from the subdirectory but will look to see if the packages are currently installed on the server and its clients and will remove any which it finds.

`-x` Remove all mode. This option will remove ALL Meiko packages which are installed on the server and its clients. Use with care.

`-f`	Force reinstallation mode. Normally `meikoinstall` will not install a package if the same version of the package is already installed. This option forces the installed package to be removed and the one in the packages subdirectory to be installed. This is of most use to Meiko Software Engineers when developing packages. They may want to reinstall the same version of a package.
`-n`	No execution mode. Use this option to see what `meikoinstall` will do without causing any real installation to occur.
`-R root` | *client*	You can use the `-R` option to install the appropriate packages on a particular node only. If you want to install the packages on the server node only use `-R root`. If you want to install them on a particular client only use e.g. `-R cs2-4`.
`main`	When running on the main server (the `/opt/MEIKOcs2` server) you must give this argument to `meikoinstall`. If you don't, the application packages will not be installed.
pkg	If *pkg* arguments are given `meikoinstall` will only install (or remove) the given packages.

setupcs2files/server Manual Pages **C**

setupcs2files

Create CS-2 template files.

Synopsis

/opt/MEIKOcs2/etc/scripts/setupcs2files [-n] [-h]

Description

setupcs2files creates templates of all the files which need to be distributed over the nodes on a CS-2 machine at installation time. It depends on the existence of the /opt/MEIKOcs2/etc/cs2.des file which needs to be set up to describe the machine. To create this file copy /opt/MEIKOcs2/etc/cs2.des.template to /opt/MEIKOcs2/etc/cs2.des and edit it; the template is well commented so you just need to fill in the blanks.

setupcs2files then uses the cs2.des file to create templates of all the files which will get distributed by the setupcs2server script. It creates the files in the /opt/MEIKOcs2/dist directory. The files which get created can be described in terms of which setup group they are in (the files in a group can be selectively installed by setupcs2server) and in terms of which type of node they will get installed on. For example the file /etc/hosts will be the same on all nodes and is in the setup group basic. The template of it will therefore be created in /opt/MEIKOcs2/dist/basic/allnodes/etc/hosts. The /etc/defaultrouter file is different on the node which is the network gateway and all others. Two copies are therefore made, one in /opt/MEIKOcs2/dist/basic/netgateways/etc/defaultrouter and one in /opt/MEIKOcs2/dist/basic/notnetgateways/etc/defaultrouter. For a full description of the possible setup groups, node types and the files which are created see the /opt/MEIKOcs2/etc/setupcs2.notes file.

Arguments

-n Will just display the file names of the files which it would create.

-h Help

See Also

setupcs2server

setupcs2server **Distribute CS-2 files.**

Synopsis

`/opt/MEIKOcs2/etc/scripts/setupcs2server [-r] [-n] [-h]` *thisnode mainnode* `[`*grp* `... ]`

Description

`setupcs2server` is used to distribute the files created by the `setupcs2-files` script. It needs to be run on all the server nodes of a CS-2 machine and will update files on a server and its clients.

thisnode is the node name of the server you are running on `mainnode` is the node name of the main (`/opt/MEIKOcs2`) server

Full information on the files which get updated is given in `/opt/MEIKOcs2/etc/setupcs2.notes`.

Arguments

`-r` This flag will reverse the effects of running `setupcs2server` and restore the original versions of the files.

`-n` No execute mode. Just display the names of the files which will get updated.

`-h` Help

grp If no groups are entered the files in all setup groups will be updated. To selectively update files in particular groups only give the groups as arguments. Currently there are three groups, `basic`, `yp` and `vfstab`. See `/opt/MEIKOcs2/etc/setupcs2.notes` for full information on which files are in which groups.

See Also

`setupcs2files`, `/opt/MEIKOcs2/etc/setupcs2.notes`