

March 13, 1989



Jess Martak
Westinghouse Corporation
7301 Sykesville Road
Sykesville, MD
21784

Re: SQQ'89(I) Meeting Review

Dear Jess:

We were sorry other engagements kept you from the Honeywell, Intel, Westinghouse discussion last week. In your absence, your team kept us busy with many interesting questions and discussions. We trust the Intel and Honeywell answers were sufficient and worth the effort to join the Hypercube IV events.

Marty and Mary asked that we send written responses to their prepared questions. I have enclosed answers to those and more. I have also taken the liberty of drawing a few sketches of current configurations (commercially available) and proposed customizations that could be useful. The drawings and their accompanying notes attempt to show the flexibility and robustness of iPSC/2 Parallel Supercomputers.

For Mary, I have enclosed copies of a few current manuals. Updated manuals (that include the I/O system) will be available in two weeks. Additional copies are also available.

For Don, I have enclosed mechanical drawings of the 19" I/O cabinet and a site preparation guide. Copies of the University of Michigan papers on fault-tolerance are also included.

If you have any questions, please do not hesitate to call. I am also available to help relay this information and further explain the enclosures should Westinghouse require additional support in completing the proposal.

Sincerely,

A handwritten signature in black ink, appearing to read "Peter I. Wolochow". The signature is fluid and cursive, with several loops and a long horizontal stroke at the end.

Peter I. Wolochow.
Defense Systems Marketing

CC: Don Evans - Westinghouse, Sykesville
Mary Jones - Westinghouse, Sykesville
Marty Yencha - Westinghouse, Sykesville
Jim Masten - Honeywell, Everett
Fred Marshall - Honeywell, Everett
Scott Glasgow - Honeywell, Everett
Norm Tinkelpaugh - Honeywell, Everett

Harry McGehee - Intel, Greenbelt
Justin Rattner - Intel, Beaverton
Jon Vahle - Intel, Beaverton

SQQ89(I) 3/7/89 Tri-Company Review

Questions and Answers / Meeting Summary

Hypercube Hardware

Configuration of nodes:

1. If we repackage hardware, such that sub-cubes were physically separated substantially, is there a technique for connecting remote cubes, maintaining the fully connected Hypercube topology?
 - Currently 3 meters is maximum separation
 - Store-and-forward S/W needed for longer hops
 - All signals are single-ended; Different transmission technologies (fibre, differential, etc.) could be used – with customized H/W interfaces.

 - Is sub-systems (each a cube) are connected using S/W message routing between system, a more robust system can be implemented. See attached slides for possible configurations.

 - => See attached slides for discussion on options (using today's commercial product and possible modifications)

2. What is the impact (pros/cons) of truncating the hypercube - i.e, few nodes than a power of 2, say 20 or 40?
 - Fewer nodes can be accommodated using provisions for "I/O nodes". For example, an 8-node hypercube can support from 1 to 7 I/O nodes – 8 to 15 total nodes. Transparent communication between these "I/O Nodes" (not between the I/O nodes and the hypercube) would require store-and-forward extensions to the operating system or explicit division of the application S/W.

- 2A. What is the impact (pros/cons) of connecting some nodes with fewer than the 8 lines?
 - In hypercubes with less than 128 nodes, each node already uses less than 8 lines (DCM channels). The number used is the power of 2 dimension (eg: 32-node system uses 5 channels, plus I/O and SRM connection).
 - If fewer lines are needed (ie: for between sub-systems – each a hypercube – store-and-forward additions to the operating system would be required.

3. At runtime, can a 32 node logical system be mapped to a 20 node physical system?
 - Processes can be mapped arbitrarily assuming sufficient memory. iPSC/2 cube sharing does this function automatically whenever a "cube" is assigned

to a loaded application.

- Available applications (eg: Interwork II, from Block Island Technologies) can dynamically move processes between processors in a globally-managed system.

4. How can the hypercube be made redundant (fault tolerant) overall including the SRM? (i.e., are separately powered subcubes feasible? Dual SRM's?)

- Full fault tolerance is not a design feature of the iPSC/2. However, there are many features built in that can help reduce the risk of individual module failure. Some are listed here. Some are described in the attached slides.
- Independent (from the ac line) power is supplied to each subcube (16-node d4) and I/O cardcage
- Dual SRM's are possible, but S/W would change. As we anticipate having no SRM in deployed SQQ89 systems, this issue should become moot. Initialization and system diagnostic functions can be built into the display processor and/or a modified USM board.
- S/W techniques can offer substantial tolerance – especially if reconfiguration is allowed during the power-up sequence.

5. Describe the PM/FD/FL (performance monitoring, fault detection, fault localization) features/characteristics of the hypercube system and its components (the 386, 387, Weitek unit, direct connect circuitry)

- Performance-monitoring signals on backplane (See the attached diagrams) are available for each processor
- A modified backplane (see attached slides) could route PMFL signals to a specially configured microprocessor
- Memory parity errors are detectable on each node
- S/W logging via operating system (NX/2). More comprehensive analysis / logging is possible.
- DCM connections employ checksums. Recovery from checksum errors would require specific operating system modifications

5A. For the system are there features for collecting PM data (connections?)

- Yes. See the attached diagrams for specific connections available for performance monitoring, status information, etc.

6. What external interfaces will be available? Are you basically planning that the VME and MULTIBUS II adaptor technique will be the approach with 3rd parties supplying the boards?

- MBII, VME, and SCSI are viewed as primary interfaces for plug-ins.
- The LBX II interface (subset of MBII specification) is current connection to VX (Vector processor extension) and Honeywell's MX module.

Hypercube Software

Ada:

1. (Reference Ada data sheet) What are the "Chap. 13 I/O Facilities"? and Where are supported at each node? Which are supported at SRM?
 - Call standard Ada I/O capabilities are supported through available I/O (at the time of the data sheet printing, only the SRM) services. Each node makes standard I/O calls that are either supported through the SRM, or the Concurrent I/O system.
2. Will the '89 Ada support the VX extension?
 - Unfortunately, the VX option requires advanced (unstudied) modifications to the Ada compiler if full Ada support is required. As an asynchronous processor, the VX exception handling becomes very complex if Ada's predictability/repeatability requirements are supported.
 - If the VX processor is called by subroutines able to manage their own exceptions and not worry about order or results in a multi-processor configuration, no special compiler or development tools are required.
 - It is our understanding that Honeywell anticipated supporting attached processors as part of their Distributed Ada plan.
3. Will the Ada runtime package run on the base machine or will it run under Unix?
 - Runtime support for **both** is part of the iPSC/2 Ada package. A runtime environment for each node support multiple tasks and I/O connections to the SRM and CFS. Message-passing subroutines provide the interface between processors. A second runtime environment builds off the Unix environment of the SRM. This environment supports the host-iPSC/2 interface and access to all files/utilities of the Unix environment.
4. Will the Ada validation be done on basis of a single node or will it be done for multiple nodes on the hypercube?
 - Verdix is managing the Validation process for iSC (they have much more experience, and will significantly reduce the cost by including iPSC/2 Ada as part of quarterly run) during the summer.
 - We anticipate a very simple validation process – Verdix Ada 386 is already validated on VMS and several Unix environments. Only a single node (and SRM) environment is required for validation.
5. Are there any plans for a distributed Ada - i.e., a development tool that would enable easy repartitioning of an Ada application to nodes through automated

node I/O code generation? What is schedule?

- Intel is participating in a joint (Honeywell, Verdix, Intel, and others) development effort. We expect the iPSC/2 to be the first commercial platform for this advanced Ada technology.
6. Will there be cross - compilation capability - i.e., do parts of development on a VAX system, then move code to hypercube for test?
- Verdix Ada Development System (VADS) is available for other Unix and VMS systems. Intel 386 cross compilers are also available for these environments.
 - VADS compiler speeds are best on non-VMS environments. iPSC/2 Ada is expected to compile between 1500 and 1800 Lines Per Minute on the SRM.
7. For Ada, what support tools are planned and what is the planned release schedule for the various tools and features?
- iPSC/2 Ada, scheduled for August, 1989, will be released with the standard VADS development environment supported on the SRM
 - There are no immediate plans to host the compiler on iPSC/2 nodes.
 - Enhancements to the development tools (eg: full support of multi-processor debugging environments from the remote host) are planned for a subsequent (1990) release.

Software Development:

8. How many workstations can be used, practically, connected to SRM to do software development? Can tasks be run on the hypercube from the workstation via SRM?
- Each SRM supports from 4-10 users. Users can be connected directly to the SRM through an expansion RS-232 module, or on workstations connected to the Ethernet network.
 - About 10 sub-cubes can be allocated to multiple users of a single iPSC/2 (assuming at least 10 nodes) system. Ada programs can be written and debugged on either the workstation, SRM, sub-cube, or entire iPSC/2.
 - Multiple SRM's can be configured on the network to support additional users. Each would have exactly the same S/W environment. Only one can be connected to the iPSC/2 at any time (unless special S/W is developed).

Other Design Issues

PMFL:

1. Is there software support for performance monitoring and fault localization?

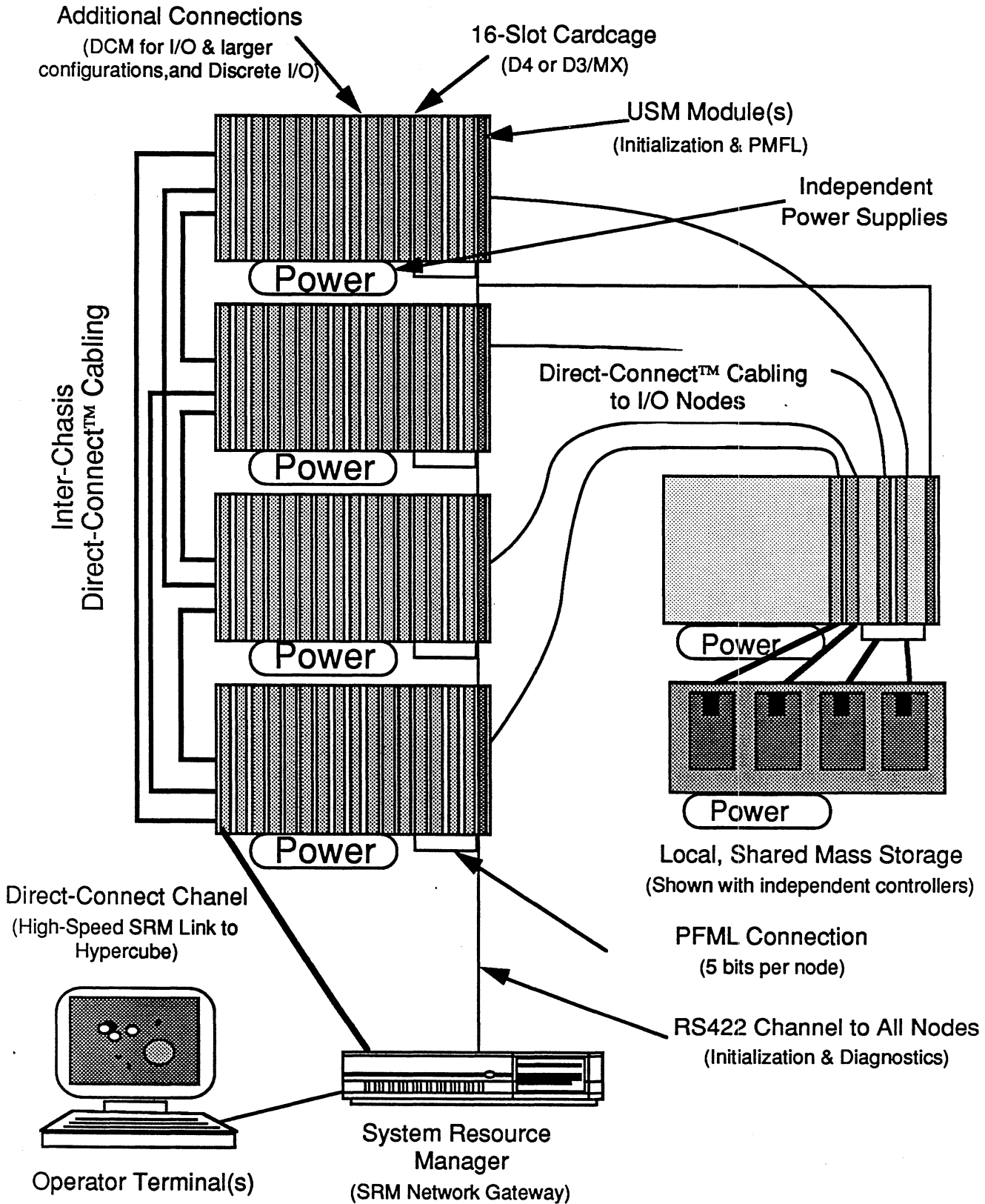
- Some hooks are embedded in the operating system. Application S/W and/or modifications to the operating system would be required to take advantage of these existing hooks.
2. Can we purchase programmer's reference manuals?
- iPSC/2 manual are available. Copies of key manuals are enclosed with this summary. Larger quantities can be purchased through our offices.
3. Is there any work in progress directed toward process control applications (often called real time or event driven - i.e., controlled response time applications)
- Lowell University has used an iPSC/1 to control a robotic work cell
 - iSC is active in a number of near real-time applications requiring similar response times. Some involve 100's of thousands of database operations for 1000's of on-line users.
 - Real-time specifics for the iPSC/2 (and NX/2) are available / measurable ---- under specific conditions. With certain working assumptions (eg: message lengths), message latency limits could be set.
 - Real-time responsiveness of the Ada runtime is an issue for the compiler vendor. iPSC/2 Ada – using a light-weight model for tasking - provides the best working environment.

iPSC/2 future development:

4. Do you have plans for hypercube enhancements - like higher performance CPU's, vector extensions, direct connect circuits, etc.?
- During our discussions, we alluded to some future iPSC implementations. Additional information could be shared under non-disclosure. Intel would be pleased to visit Westinghouse to discuss this issue if required.
 - In general, Intel assumes a 2-4 year product cycle for each major system architecture. this is the very nature of the parallel supercomputer business.
- 4A. Or do you see major advances in the architecture itself?
- As for 12, Intel would be pleased to elaborate, should the need arise.
 - From the I/O perspective (support for MX, VX, disks, etc.) the current architecture will be deployed for several years.
5. Can the SRM be replaced or buried in some other system component?
- As shown on attached slides, SRM functions could be integrated with one intelligent display

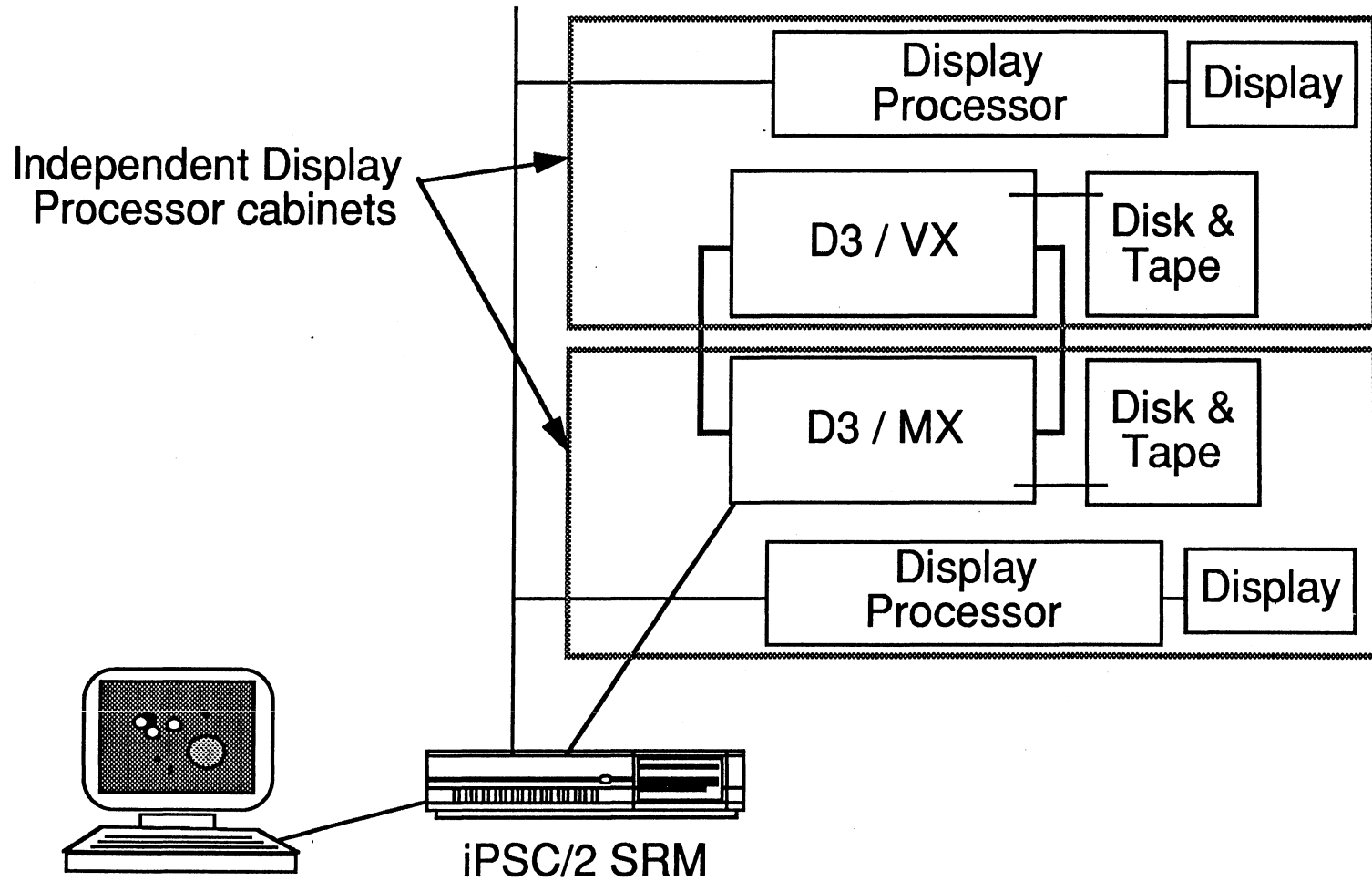
- The SRM replacement would be responsible for down-loading S/W and cube allocation (get-cube)
 - In replacing the SRM, location of disks would need to be evaluated
6. Connectivity effects and effects of degradation
- Current paper on effects at Hypercube IV conference (see attached copy)
 - See attached slides for possible impact of single failure in a hypercube
7. List of diagnostics available.
- iSC uses a comprehensive set of diagnostic packages during manufacturing and test and during power-up / initialization sequences. Diagnostics are done at the module-level. Many of these routines can be made available to Westinghouse / Honeywell.
8. How to take best advantage of existing Mil-spec cabinets (with displays)?
- The iPSC/2 uses standard 19" components suitable for mounting in almost any cabinet.
 - The attached slides show a possible configuration employing standard (commercial) modules and/or a customized 27-slot chassis.

Commercial Hypercube Configuration



SQQ89(I)

Suggested Demonstration Configuration



SQQ89(I)

Inter-Cube Communication Options (H/W)

- I. DCM network (un-modified)
 - Possible driver modifications (simple) to extend distance
=> Probably limited to 5 meters.
 - No operating S/W notifications
 - Initialization sequence changes to support independent power-up and diagnostic sequence.

- II. DCM network (modified)
 - Possible fibre-optic links
 - Same S/W issues as I.

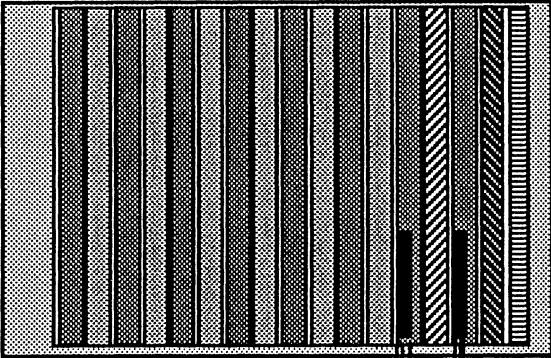
- III. Network (Ethernet or FDDI)
 - No DCM initialization S/W to develop
=> operate as independent sub-system
 - Store-and-forward S/W needed to route messages to remote inter (if bandwidth tollerable)
 - Relies on future product; I/O node network interface
 - Adaptive routing is simpler (S/W - based between subsystems)

SQQ89(I)

Sub-System Implementation (Available Commercial)

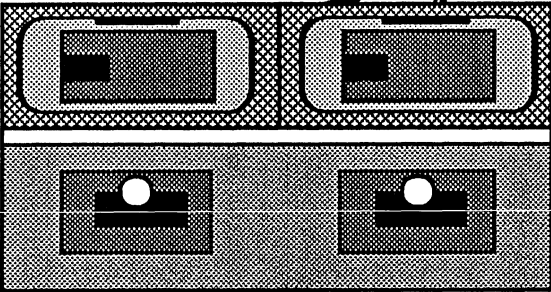
8 Nodes & 6 VX or MX Accelerators
(32 MIPs, 120 or 600 FLOPS)

2 I/O Nodes & Option Slots
(8 MIP & 8 MBytes/Sec I/O)



Standard 19-inch Rack
(16 nodes, I/O, & PMFL support)

SCSI Cabline



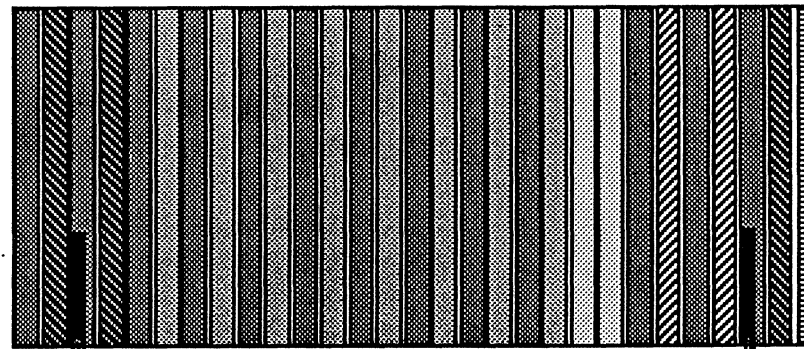
Standard, Dual Disk & Tape
(Removable 760 MByte Drives)

SQQ89(I)

Possible Sub-System Implementation

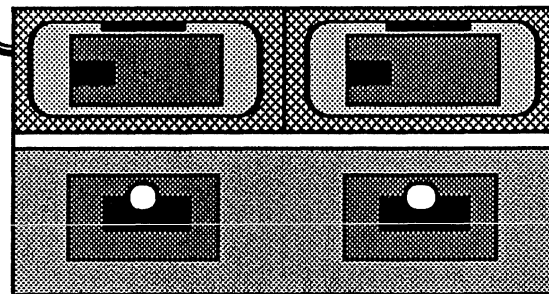
8 Nodes & 8 VX or MX Accelerators
(32 MIPs, 160 or 800 MFLOPS)

5 I/O Nodes & Option Slots
(8 MIP & 20 MBytes/Sec bandwidth for NTDS, SCSI, other I/O)



Custom 32-inch Rack
(29 Slots, I/O, & PMFL support)

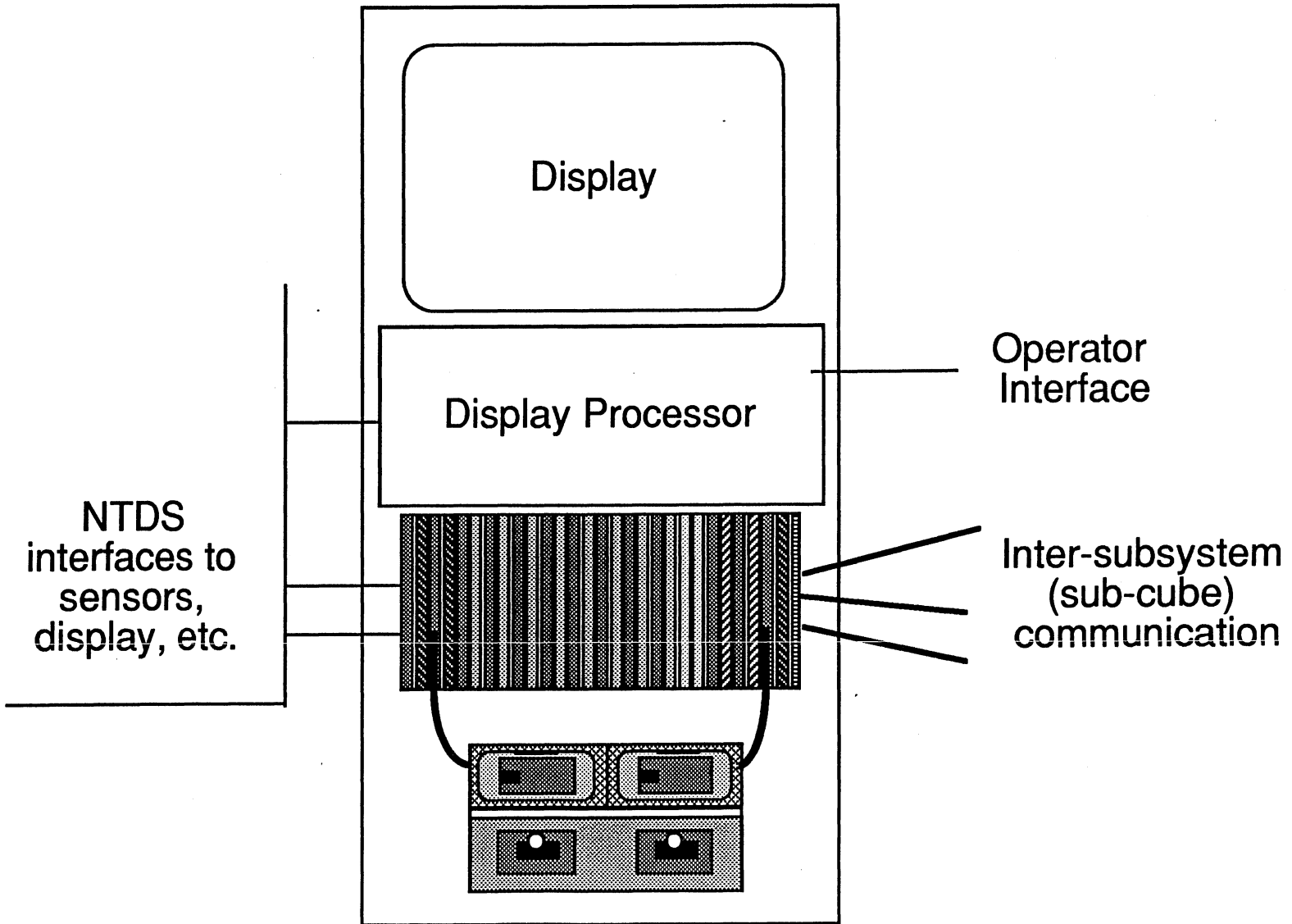
SCSI Cabline



Standard, Dual Disk & Tape
(Removable 760 MByte Drives)

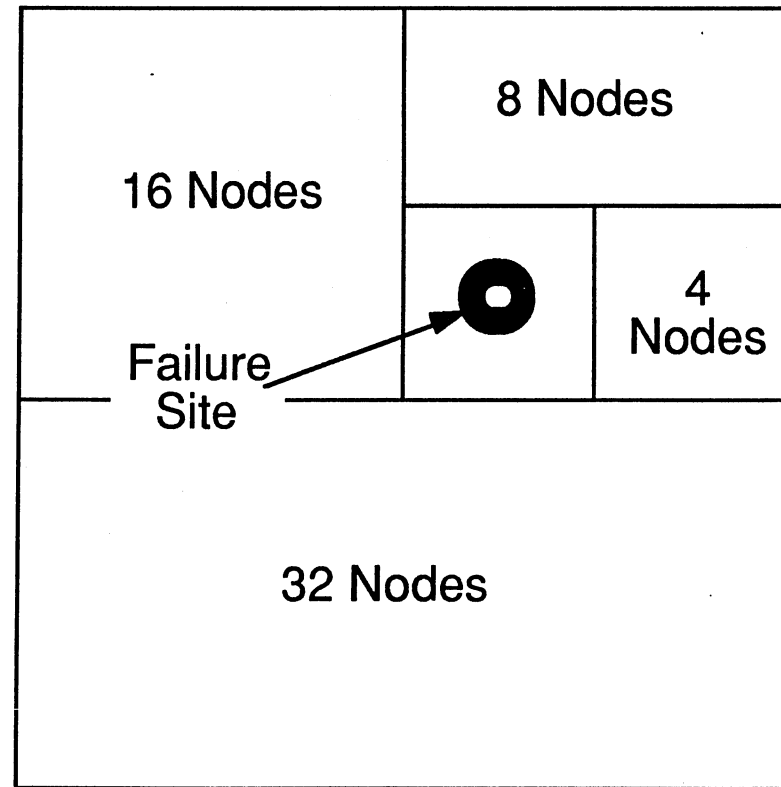
SQQ89(I)

Possible Deployable System



Degradation of Hypercube Systems

| Single-Node/Router Failure Example | |
|------------------------------------|--|
| Before | After |
| 1 64-node | 1 32-node 1 16-node 1 8-node 1 4-node |
| Case A | |
| 1 32-node | 1 32-node |
| 2 16-node | 1 16-node 1 8-node 1 4-node |
| Case B | |
| | 3 16-node 1 8-node 1 4-node |



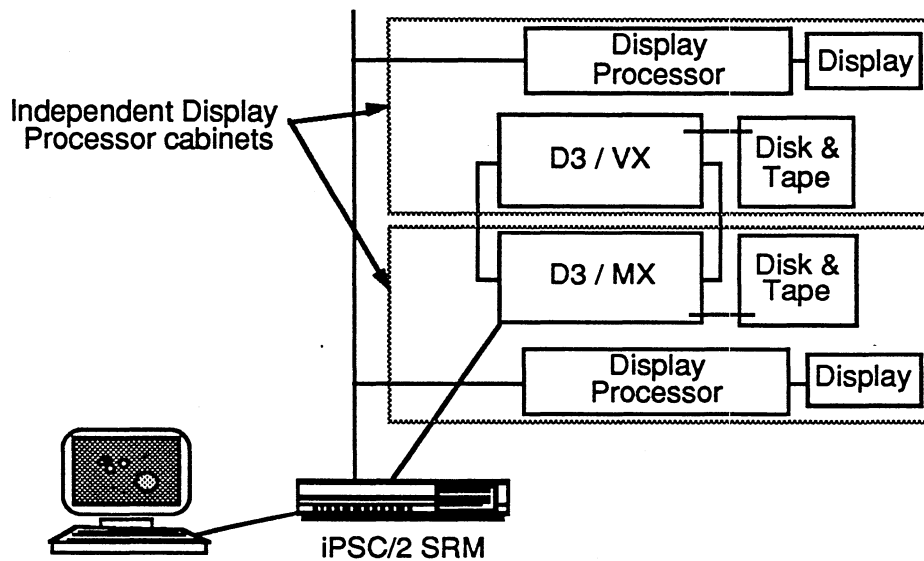
8 X 8 (64 node) Hypercube

CHACTERISTICS

- All commerically available products (Except Honeywell 'MX')
- 2X capacity with equal capacity / cabability in each half
- Physically separate 8-node sub-cubes
- Separate SRM for initialization and network interface

SQQ89(I)

Suggested Demonstration Configuration



SQQ89(I)

2

Inter-Cube Communication Options (H/W)

- I. DCM network (un-modified)
 - Possible driver modifications (simple) to extend distance
=> Probably limited to 5 meters.
 - No operating S/W notifications
 - Initialization sequence changes to support independent power-up and diagnostic sequence.

- II. DCM network (modified)
 - Possible fibre-optic links
 - Same S/W issues as I.

- III. Network (Ethernet or FDDI)
 - No DCM initialization S/W to develop
=> operate as independent sub-system
 - Store-and-forward S/W needed to route messages to remote inter (if bandwidth tollerable)
 - Relies on future product; I/O node network interface
 - Adaptive routing is simpler (S/W - based between subsystems)

Each back plane:

- Is independently powered
- Provides a 'd4' (16-node) hypercube interconnection
- Provides connectors to interconnect up to 7 other backplanes (total 'd7', 128-node hypercube)
- Includes a 'USM' slot for diagnostics and initialization sequences
- Provides an 80-pin (5 / node) interface to external status / info strobe

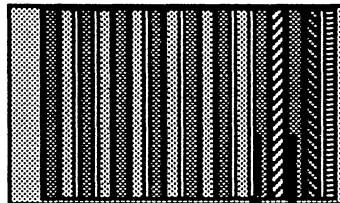
- Provides independent (channel 7) connections to I/O nodes from each node slot=> Could be used for direct MX I/O
- Provides LBX II connections between pairs of slots (8 pair/backplane)
- Buses power, reset, and USM interrupt

SQQ89(I)

Sub-System Implementation (Available Commercial)

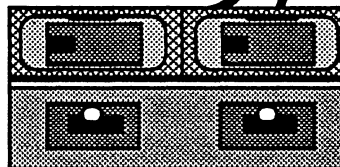
8 Nodes & 6 VX or MX Accelerators
(32 MIPs, 120 or 600 FLOPS)

2 I/O Nodes & Option Slots
(8 MIP & 8 MByte/Sec I/O)



Standard 19-Inch Rack
(16 nodes, I/O, & PMFL support)

SCSI Cabline



Standard, Dual Disk & Tape
(Removable 760 MByte Drives)

POSSIBLE (STAGE 1) H/W MODICIATIONS

Relatively easy ways to extend / adapt functions for SQQ '89 (I)

1) New backplane to:

- More efficiently fill available space.
- Incorporate connections to PMFL processor in USM location.
- Provide spacing for up to 8 nodes and 2 (optionally devoted) I/O nodes.
- Possible backplane connections to NTDS

2) Modified/extended USM funcitons to support

- a) PMFL functions
- b) Disply processor bootstrap

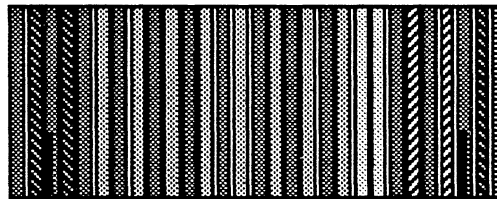
3) Mil-spec connectors and board guide rails

SQQ89(I)

Possible Sub-System Implementation

8 Nodes & 8 VX or MX Accelerators
(32 MIPs, 160 or 800 MFLOPS)

5 I/O Nodes & Option Slots
(8 MIP & 20 MBytes/Sec bandwidth for NTDS, SCSI, other I/O)



Custom 32-Inch Rack
(29 Slots, I/O, & PMFL support)

SCSI Cabline

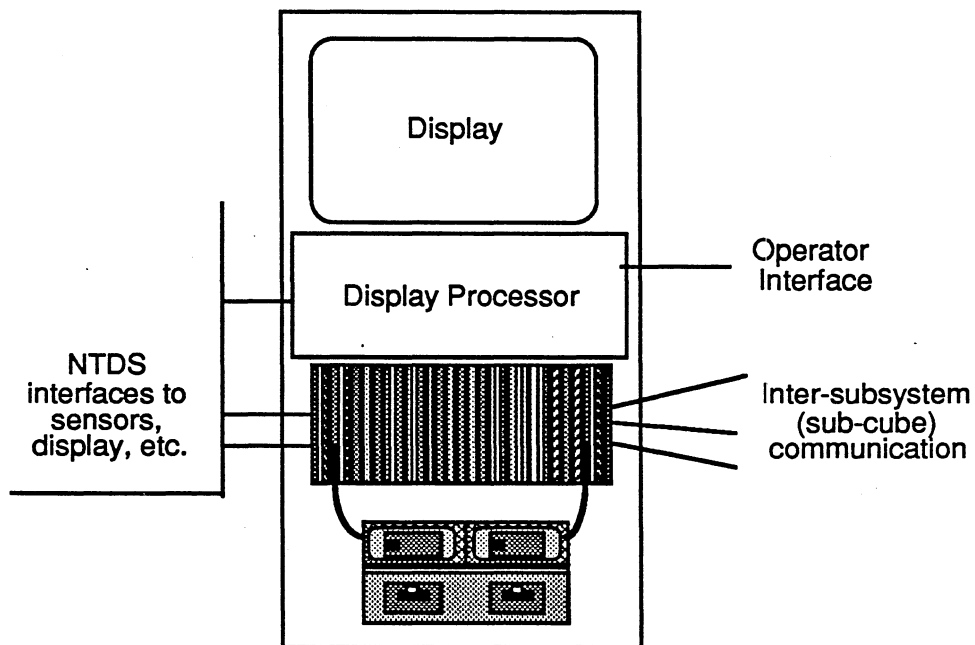
Standard, Dual Disk & Tape
(Removable 760 MByte Drives)

CHARACTERISTICS:

- Stand alone functional sub-system with self-contained disks, displays, processors, and I/O
- Inter-cube communication dependent on required bandwidth and technology
- Redudant I/O communications to NTDS for sensor input and display of the output
- Specific initialization and diagnostic channel (RS-4222) to display (master – taking role of SRM) processor

SQQ89(I)

Possible Deployable System

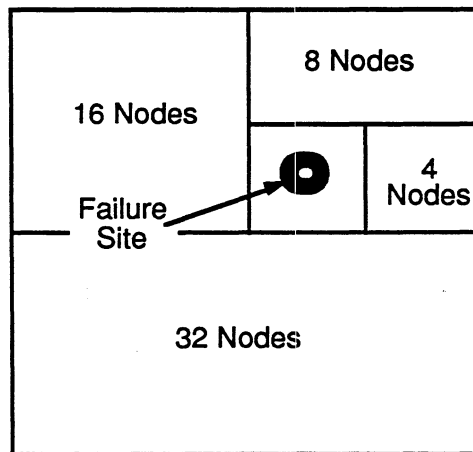


SQQ89(I)

Degradation of Hypercube Systems

6

| Single-Node/Router Failure Example | |
|------------------------------------|-----------|
| Before | After |
| 1 64-node | 1 32-node |
| | 1 16-node |
| | 1 8-node |
| | 1 4-node |
| Case A | |
| 1 32-node | 1 32-node |
| 2 16-node | 1 16-node |
| | 1 8-node |
| | 1 4-node |
| Case B | |
| | 3 16-node |
| | 1 8-node |
| | 1 4-node |



8 X 8 (64 node) Hypercube