The QCDOC Project
Overview and Status

All Hands Meeting

June 1-2, 2005

Norman H. Christ

Outline

- Project goals
- Architecture
- Software
- Construction and packaging
- Construction status
- Final bring-up issues
- Application performance

Project Goals

- Massively parallel machine capable of *strong scaling*: use many nodes on a small problem.
  - Large inter-node bandwidth.
  - Small communications latency.
- $1/sustained Mflops cost/performance.
- Low power, easily maintained modular design.

QCDOC Collaboration

- Columbia (DOE)
  - Norman Christ
  - Saul Cohen*
  - Calin Cristian*
  - Zhihua Dong
  - Changhoan Kim*
  - Ludmila Levkova*
  - Sam Li*
  - Xiaodong Liao*
  - HueyWen Lin
  - Guofeng Liu*
  - Meifeng Lin*
  - Robert Mawhinney
  - Azusa Yamaguchi
- BNL (SciDAC)
  - Chulwoo Jung
  - Konstantin Petrov => Enno Scholz 10/05
  - Stratos Efstathiadis
- UKQCD (PPARC)
  - Peter Boyle
  - Mike Clark
  - Balint Joo
- RBRC (RIKEN)
  - Shigemi Ohta
  - Tilo Wettig
- IBM
  - Dong Chen
  - Alan Gara
  - Design groups:
    - Yorktown Heights, NY
    - Rochester, MN
    - Raleigh, NC

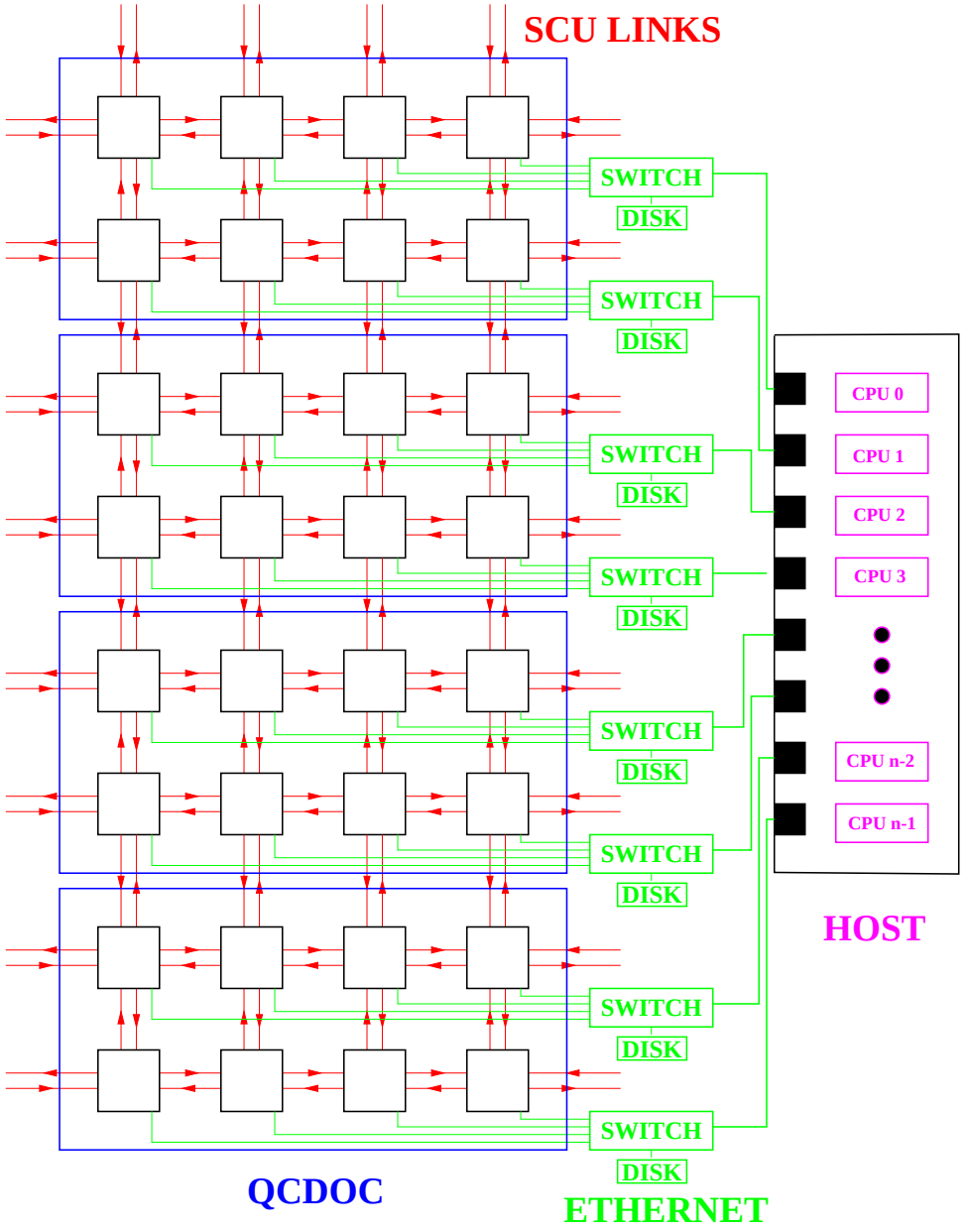
*CU graduate student

QCDOC Architecture

- IBM-fabricated, single-chip node.
  [50 million transistors, 5 Watt, 1.3cm x 1.3cm]
- Processor:
  - PowerPC 32-bit RISC.
  - 64-bit, 1 Gflops floating point unit.
- Memory/node: 4 Mbyte (on-chip) & o2 Gbyte DIMM.
- Communications network:
  - 6-dim, supporting lower dimensional partitions.
  - Global sum/broadcast functionality.
  - Multiple DMA engines/minimal processor overhead.
- Ethernet connection to each node: booting, I/O, host control.
- ~7-8 Watt/node, 15 in^3 per node.

Network Architecture

- Red boxes are nodes.
- Blue boxes mother boards.
- Red lines are communications links.
- Green lines are Ethernet connections.
- Green boxes are Ethernet switches.
- Pink boxes are host CPU processors.



Software Environment

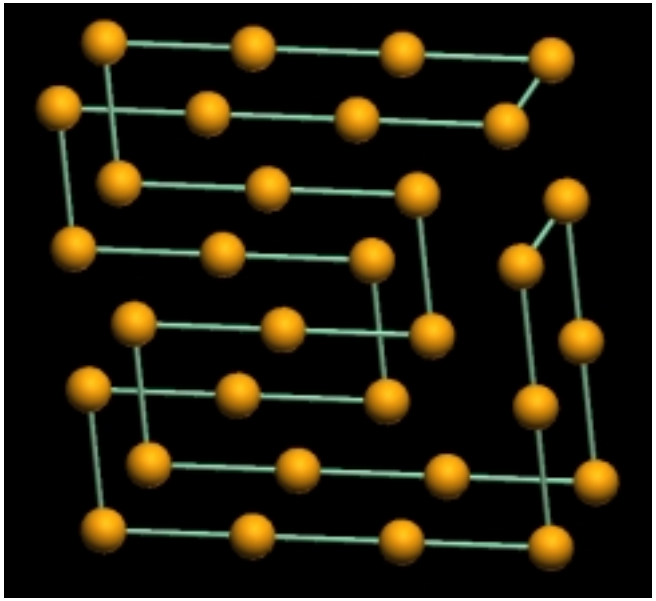
- Lean kernel on each node
  - Protected kernel mode and address space.
  - RPC support for host access.
  - NFS access to NAS disks (/pfs).
  - Normal Unix services including stdout and stderr.
- Threaded host daemon
  - Efficient performance on 8-processor SMP host.
  - User shell (qsh) with extended commands.
  - Host file system (/host).
  - Simple remapping of 6-D machine to (6-n)-D torus.
- Programming environment
  - POSIX compatible, open-source libc.
  - gcc and xlc compilers
- SciDAC standards
  - Level-1, QMP protocol
  - Level-2 parallelized linear algebra, QDP & QDP++.
  - Efficient level-3 inverters
    - Wilson/clover
    - Domain wall fermions
    - ASQTAD
    - p4 (underway)

Memory Usage

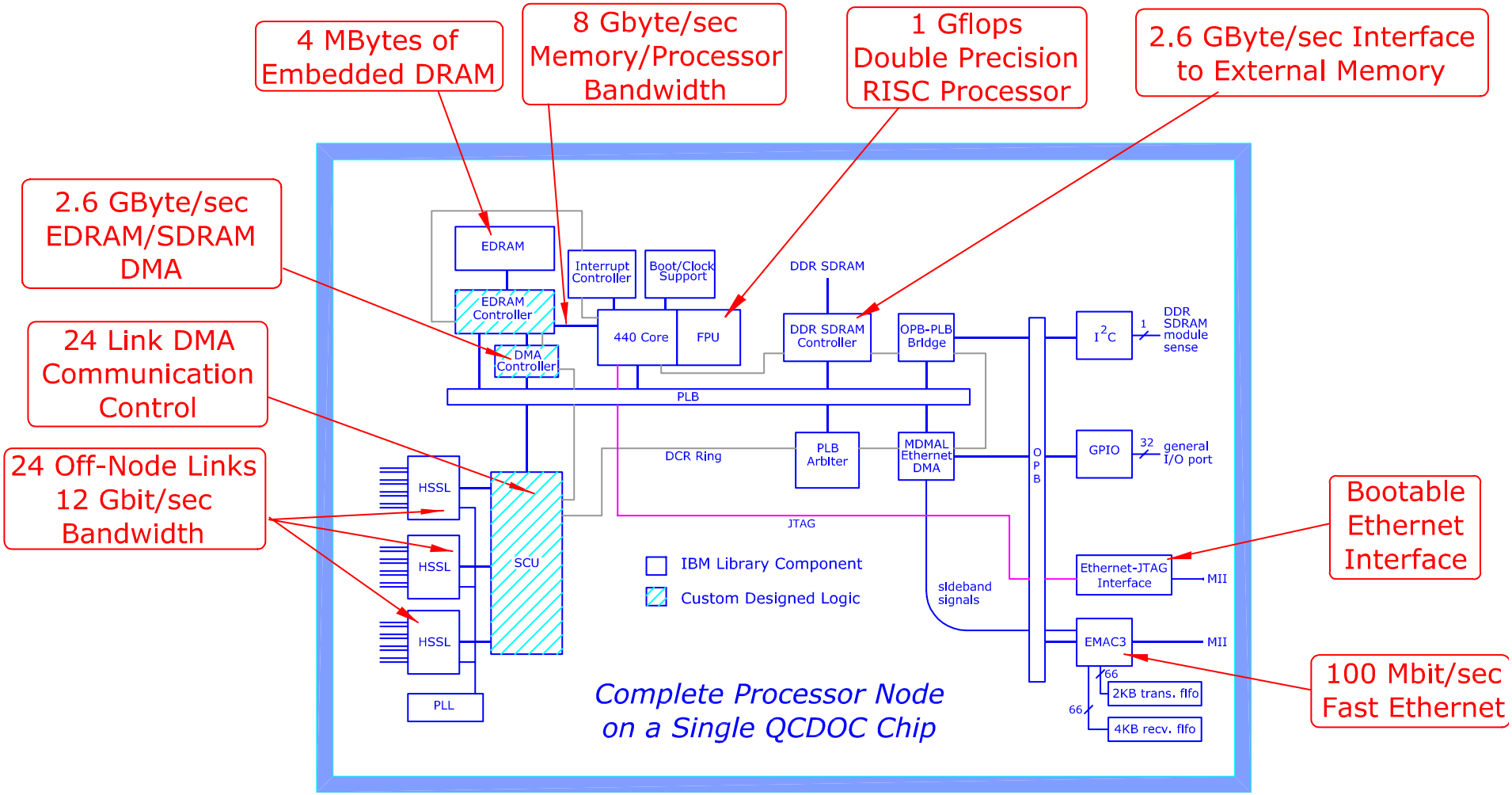
- qos uses the PPC440's MMU to assign specific pages to specific tasks
- Peter has reorganized this in qos-2-6-0 to make more efficient use of the 128 Mbyte, external DIMM memory:
  - **4 Mbytes** (kernel, print buffers, network buffers, kernel thread stacks etc.)
  - **5 Mbytes** (code)
  - **3 Mbytes** (stack)
  - **4 Mbytes** (static data)
  - **112 Mbytes** (heap)
- qalloc routine generalizes malloc and permits a choice of data location EDRAM/external DIMM (and caching options if fancy communications strategies are being implemented).

Mesh geometry

- $N_0 * N_1 * N_2$ mother boards are wired as a N_0 x N_1 x N_2 torus.
- With 2^6 nodes on a mother board, the resulting machine is a $2*N_0$ x $2*N_1$ x $2*N_2$ x 2 x 2 x 2, six-dimensional torus.
- Always refer to machine axes 0-5 and label physics directions with letters x, y, z, t, s, w.
- qpartition_remap -X01 -Y23 -Z4 -T5

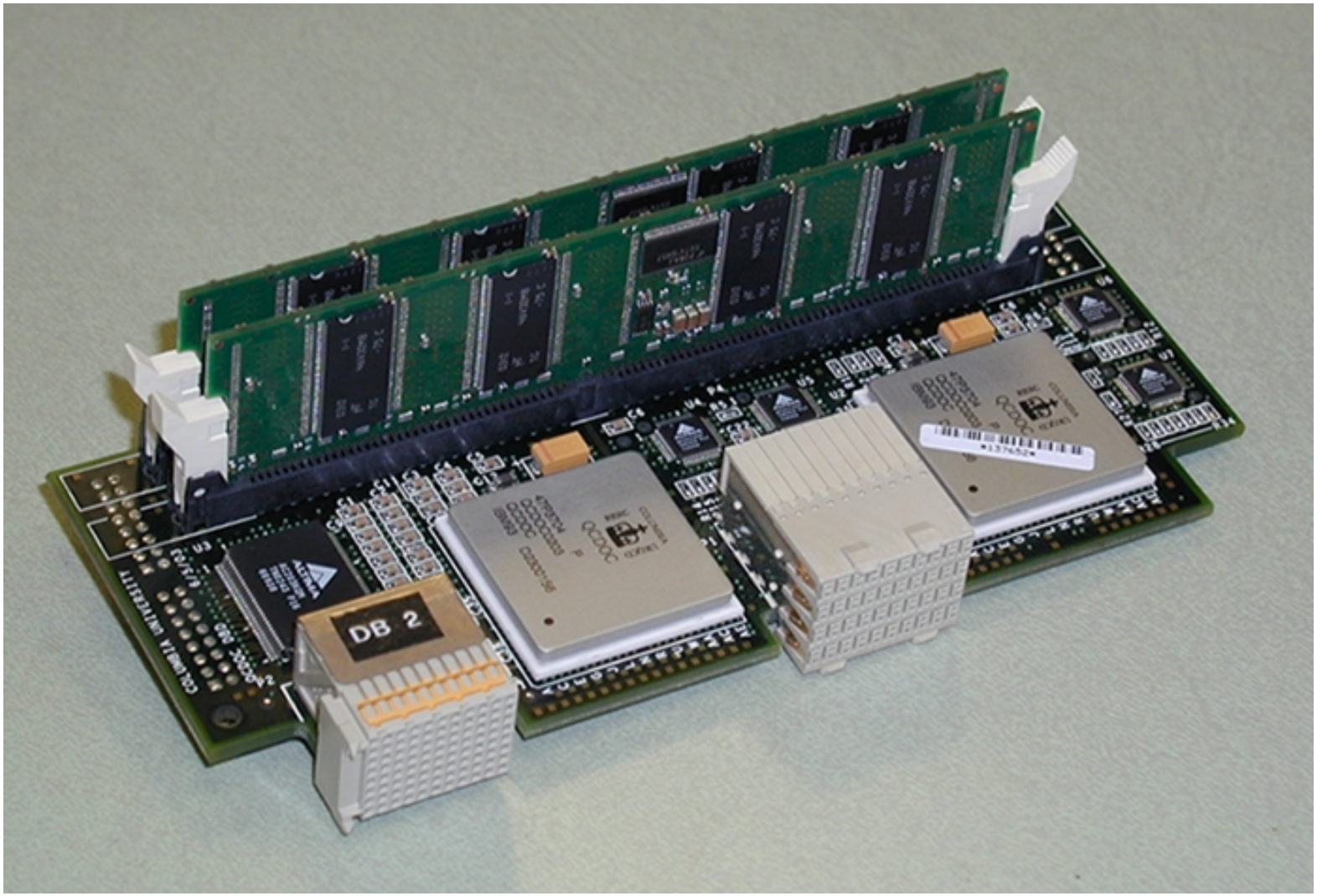


QCDOC Chip

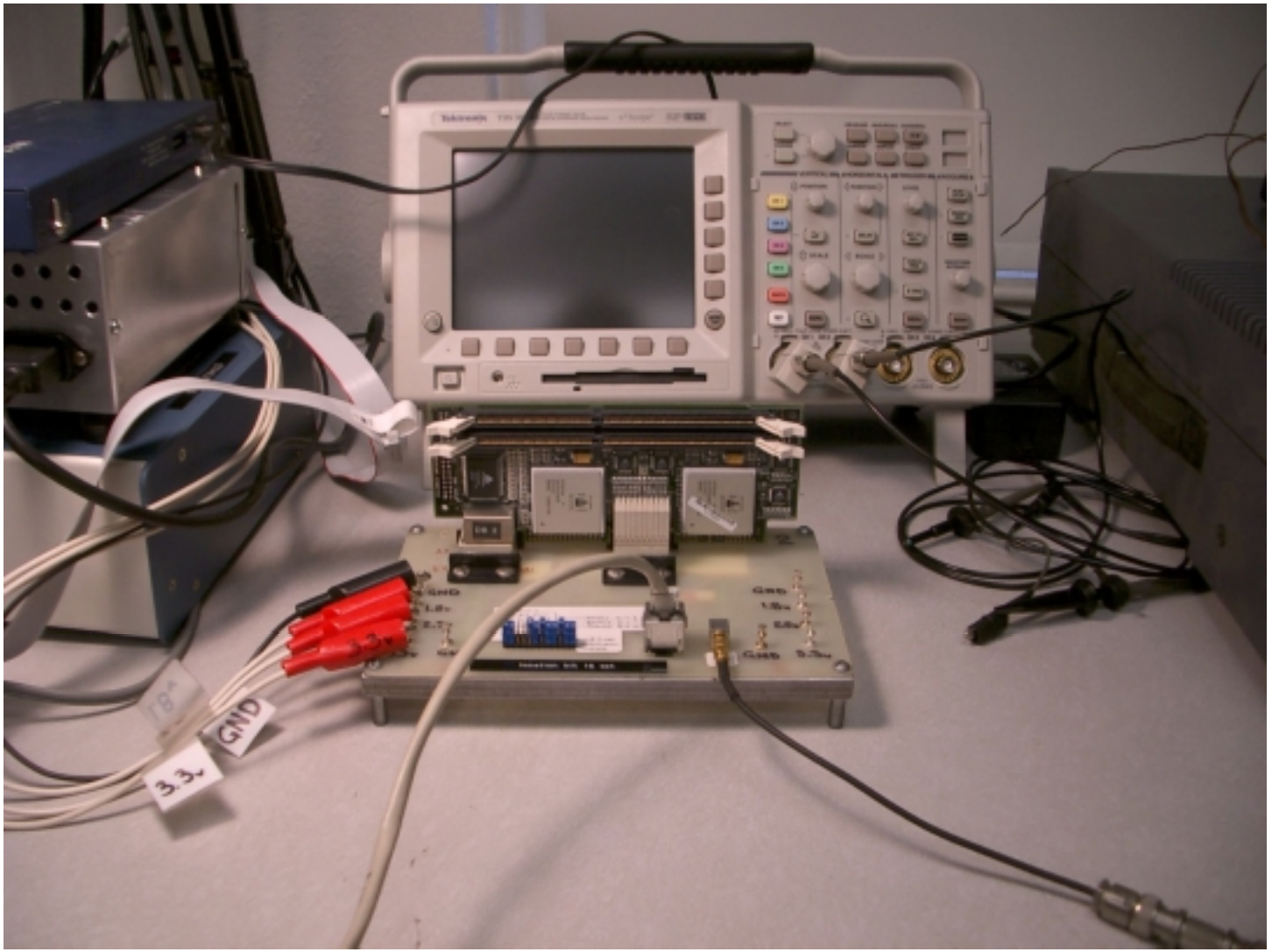


50 million transistors, 0.18 micron, 1.3 x 1.3 cm die, 5 Watt

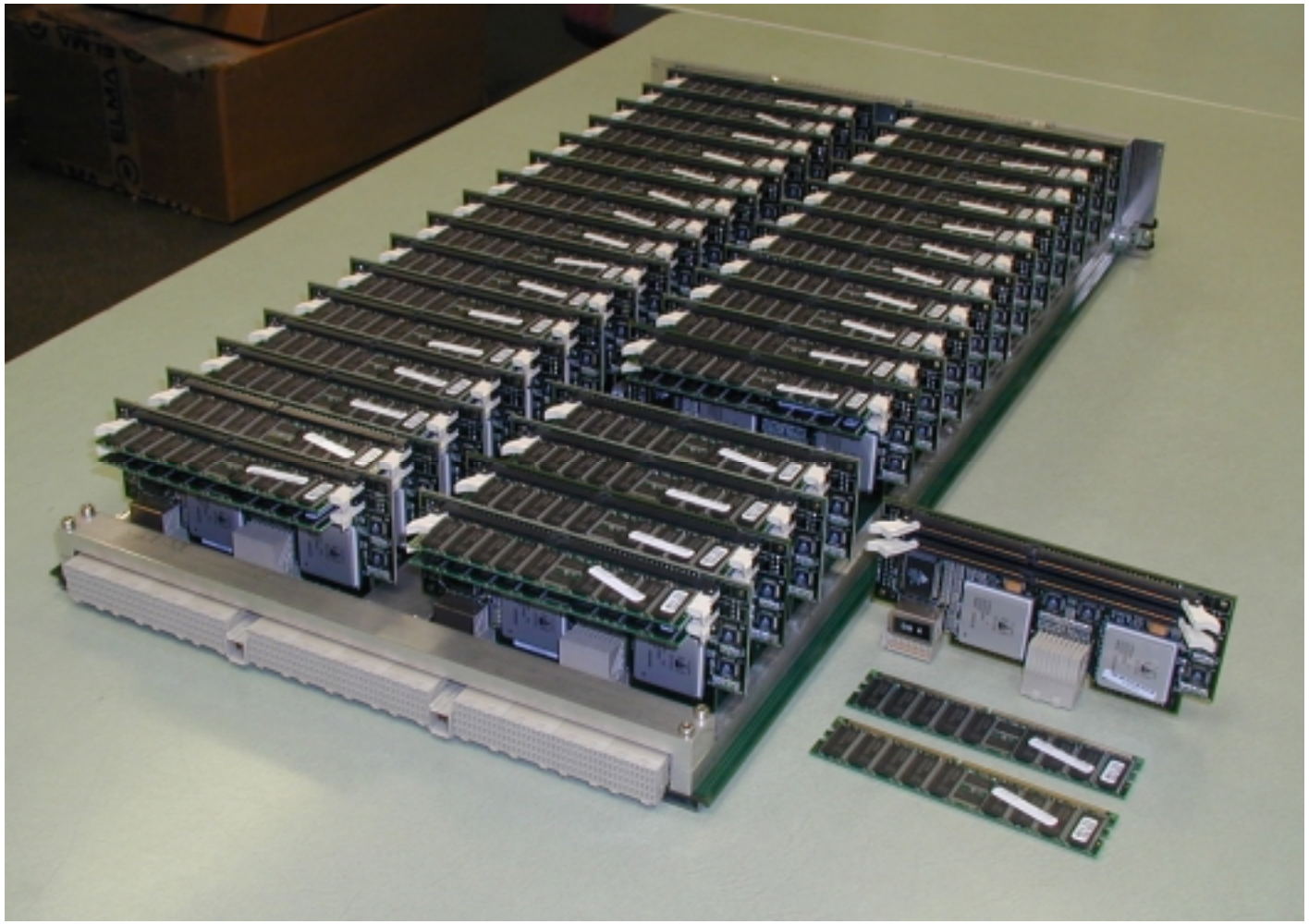
Daughter board (2 nodes)



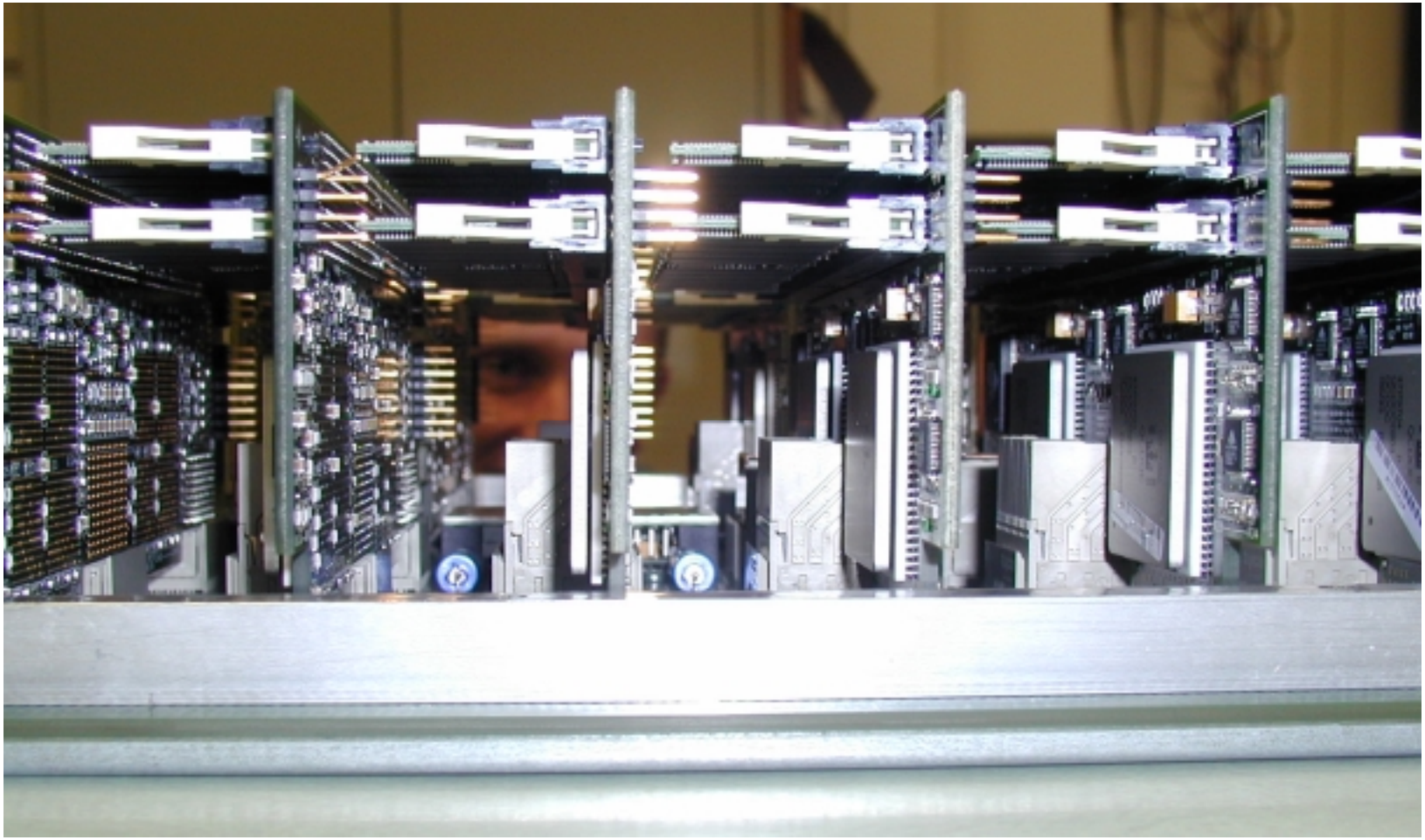
BNL – constructed test jig



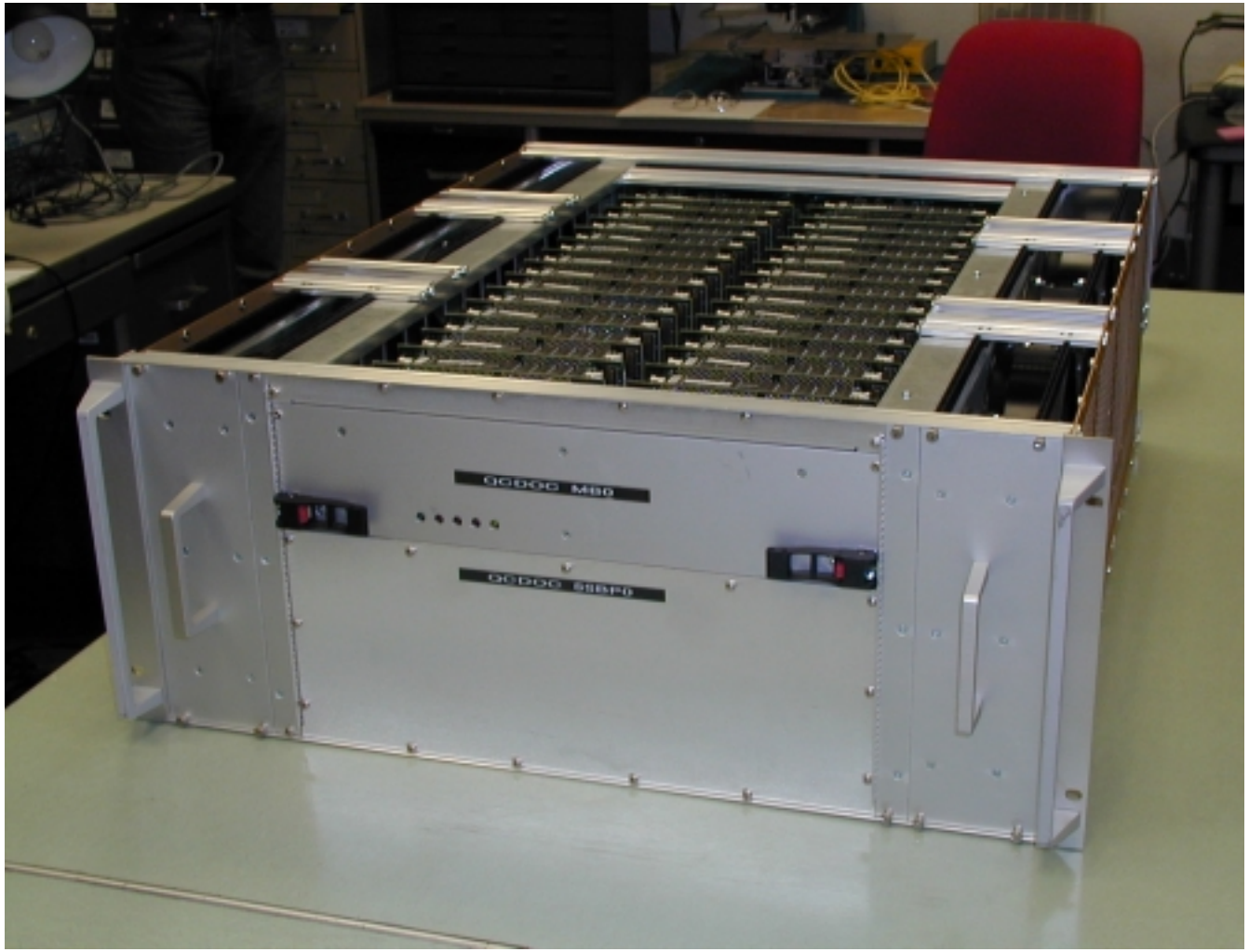
Mother board (64 nodes)



Edge view of mother board



Single mother board test jig



512-Node Machine

First 4 racks installed at Columbia

UKQCD Machine (12,288 nodes/10 Tflops)

Brookhaven Installation

RBRC (right) and DOE (left) 12K-node QCDOC machines

Project Status

- UKQCD – 13,312 nodes --$5.2M – 3-5 Tflops sustained.
  - Installed in Edinburgh 12/04.
  - Running production at 400 MHz/10% reprod.
- RBRC – 12,288 nodes -- $5M – 3-5 Tflops sustained.
  - Installed at BNL 2/05.
  - 1/4 in production/100% reprod.
  - 3/4 performing physics tests.
- DOE – 12,288 nodes -- $5.1M – 3-5 Tflops sustained
  - Installed at BNL 4/05.
  - 5/6 performing physics tests.
  - 1/6 being debugged.

Specific DOE Machine Status

Rack #	Use	Person
16	p4 thermo	Kostya Petrov/Chulwoo
17	RHMC Dyn. DWF traj. length study	Chulwoo Jung/Meifeng Lin
18	Dynamical DWF, chroma	Robert Edwards
19	2+1 flavor ASQTAD	Dru Renner
20	Ready for ASQTAD production	Dru/Carleton/Steve
21		
22		
23		
24	p4, T=0 scaling study	Michael Cheng
25	Testing of RBC 3-point code	Saul Cohen
26	Being debugged	BNL staff
27	Being debugged	BNL staff

Final Bring-up issues

- FPU errors
  - Lowest two bits infrequently incorrect (not seen at 400MHz).
  - Remove slow nodes at 432MHz and run at 400MHz.
- Serial communication errors.
  - Induced by Ethernet activity.
  - 0.25/month at 400 MHz/1K nodes.
  - Further reduced by PLL tuning.
  - Protected by hardware checksums with no performance loss.

Final Bring-up issues

- PLB bus errors
  - PLB arbiter detects an invalid bus transfer when accessing external DDR memory—a fatal error.
  - Corrected by configuring arbiter to use a less aggressive protocol.
  - Greatest impact for ASQTAD CG performance: 40% → 38% for 4^4 local volume, worse for larger local volumes—improvement may still be possible.
- Lost Ethernet contact
  - Happens infrequently during booting, less frequently during I/O.
  - Often can be recovered by “qreset_sys”, which can be scripted.
  - Less frequently requires a power cycle, not yet callable from software.
  - May be improved by more OS work.

Final Bring-up issues

- Parallel disk system
  - 24 2-Tbyte RAID servers (Anacapa).
  - 512-nodes achieve ≥12 Mbytes/sec.
  - Disks shipped last week.
- Larger machine partitions
  - Four 4096-node partitions assembled, world-wide.
  - Expect to run as DOE computer as 4096 + 8192 node machines.
- Manufacturing failures
  - 1% non-functioning daughter boards
  - 1.5% non-functioning mother boards
- Single mother board units
  - (30 spares) – (4 not repairable) = 26?
  - Put in two 8-slot and two 1-slot crates

ASQTAD Performance

	Sites/ node	Optimized					MILC		
Precision		Single	Double				Single		Double
Machine size		1024	1024	sim	128	128	128	16	16
		CG		Dslash	Dslash	CG	CG	CG/QDP	CG
Asqtad.CG	2^4			43%	22%	19%		2%	
	4^4	44%	38%		42%	40%	9.5%	15%	7.5%
	6^4	44%	27%		36%	36%	13.6%	19%	8.6%
	8^4	34%	18%		28%	28%	12.3%		8.1%
Asqtad force	2^4				19%				
	4^4				35%		10.0%		6.1%
	6^4				29%		10.8%		6.1%
	8^4				20%		8.2%		6.0%
Asqtad/ Symanzik link products	2^4				23%				
	4^4				50%		7.2%		7.1%
	6^4				49%		8.1%		6.4%
	8^4				27%		8.5%		6.8%

Application Performance

(double precision)

1024-node machine:

Fermion action	Local volume	Dirac performance	CG performance
Wilson	2^4	44%	32%
Wilson	4^4	44%	38%
Clover	4^4	54%	47.5%
DWF	4^5	47%	42%
ASQTAD	4^4	40%	38%

4096-node machine (UKQCD):

DWF/24^3 x 64/RHMC (Local vol: 6x6x6x2x8)	CG: Complete code:	1.1 Tflops (34%) 0.97 Tflops (29%)

QCDOC Summary

- Present DOE QCDOC machine use:
  - Alpha users developing code on 1 mother board machines.
  - MILC (staggered 2+1 flavor) using one 1K-node machine.
  - JLAB (DWF 2+1 flavor) using one 1K-node machine.
  - RBC/BNL (QCD thermo) using two 1K-node machines.
  - RBC (DWF) using two 1K-node machine.
  - Two racks being tested.
- 4K node machine ready for MILC use.
- Most of machine in production by early June?