A new 7.5 TFLOP cluster in NOTUR

Cyril Banino-Rokkones

Norwegian University of Science and Technology

Outline

- System presentation
  - IBM p575+ system
  - NTNU configuration
- Early performance tests
- System usage and availibility

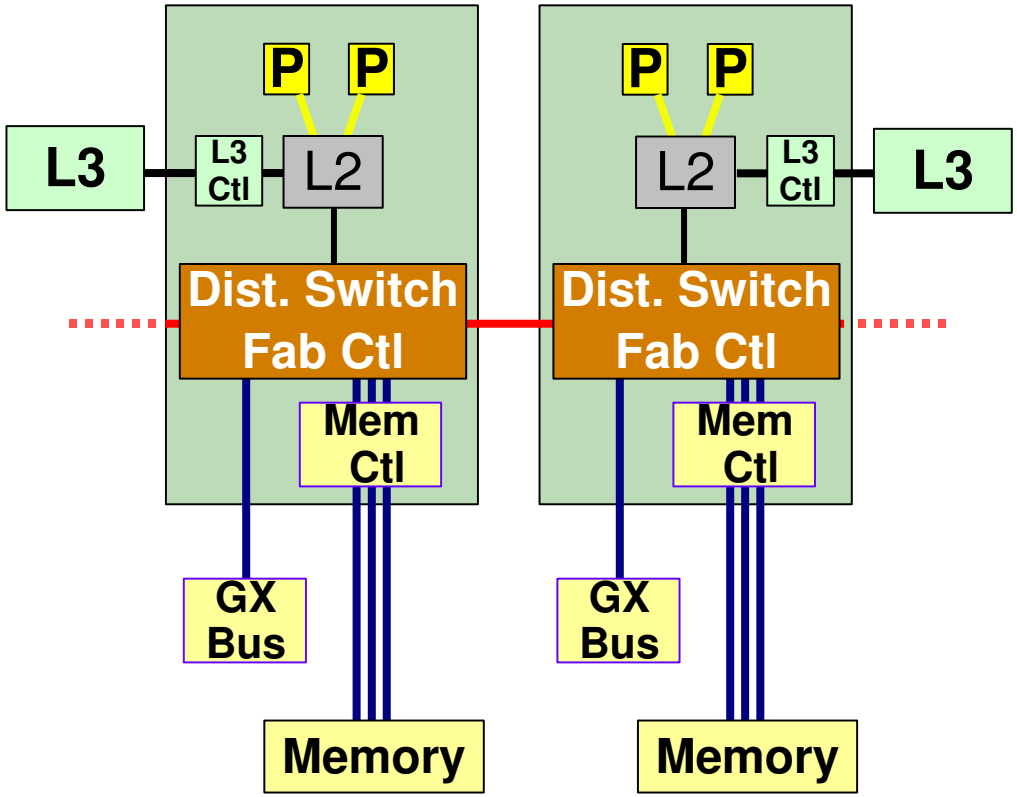
njord.hpc.ntnu.no

- System IBM p575+
- Type: distributed/SMP
- 992 CPUs
- CPU type: power5+
  - theo peak: 7.6 Gflop
- Total theo peak:
  - 7.5 Tflop
- LINPACK interpolation
  - 5.796 Tflop (928)
- Total memory:
  - 2016 GB
- Total disk: 72 TB (formatted raid5)

njord.hpc.ntnu.no (back view)

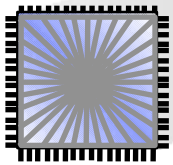
- 64 nodes:
  - 55 compute (32 GB)
  - 1 compute (128 GB)
  - 4 I/O (16 GB)
  - 2 login (32 GB)
  - 2 test
- Each node:
  - shared memory
  - 8 dual-cores p5+ chips
- System well-suited for
  - large-scale MPI
  - OpenMP
  - mix of both

Power5+ Design

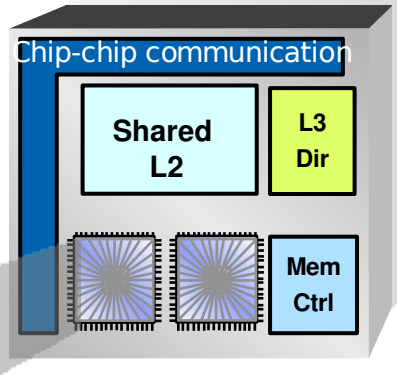


POWER5+ Chip / MCM structure buildup

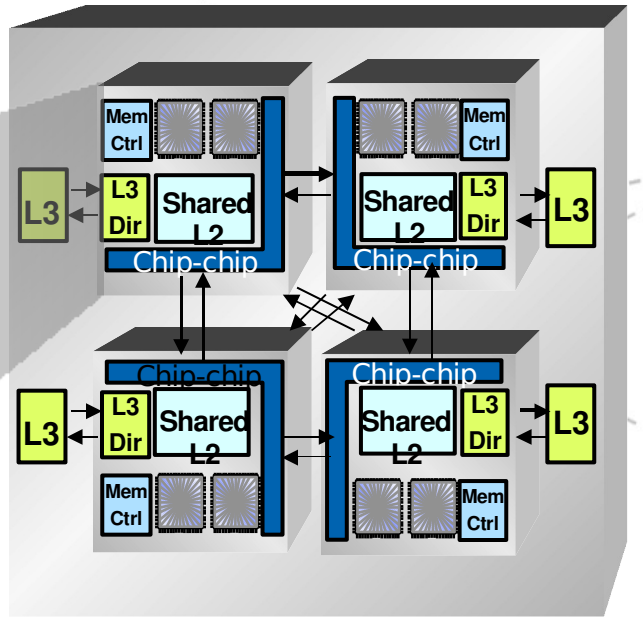
- Memory controller on chip for better performance
- L3 cache resides on the Multi-Chip Module
- L3 cache not in path of memory transfers



POWER5+ 1.9 GHz microprocessors



2-way POWER5+ SMP system dual core on a single chip



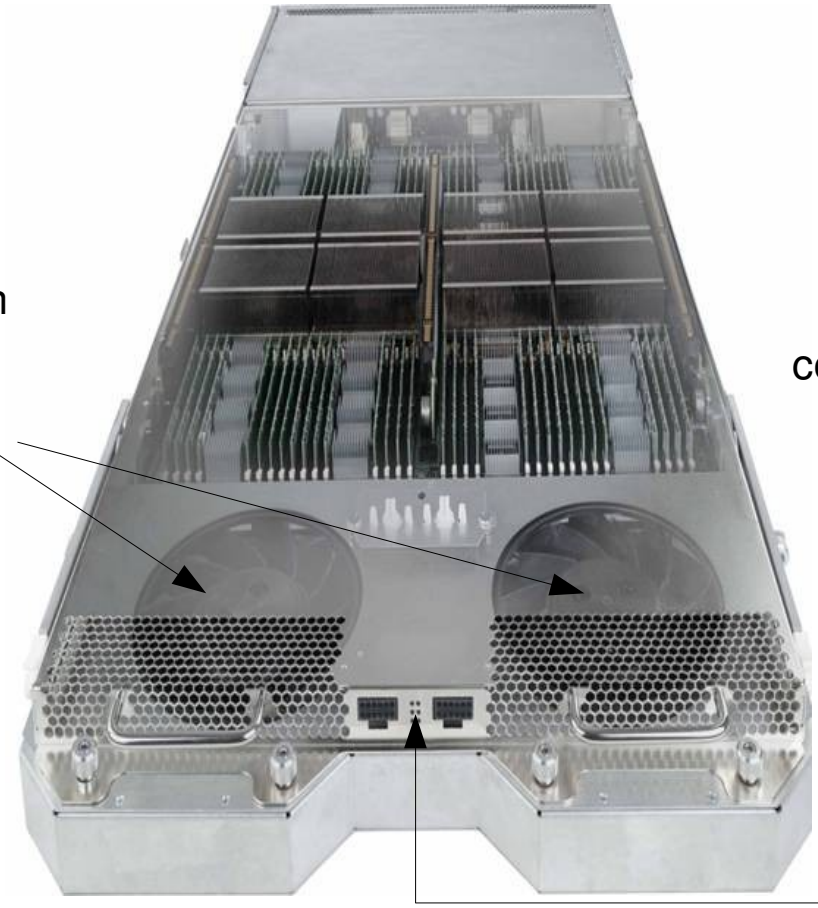
8-way (four chip) POWER5+ SMP system on a Multi-Chip Module (MCM)

p5-575 node

- 110 kgs
- 8 procs in 2U

Front view of p5-575 node

Air intake ventilation grids for two custom designed blowers in node.



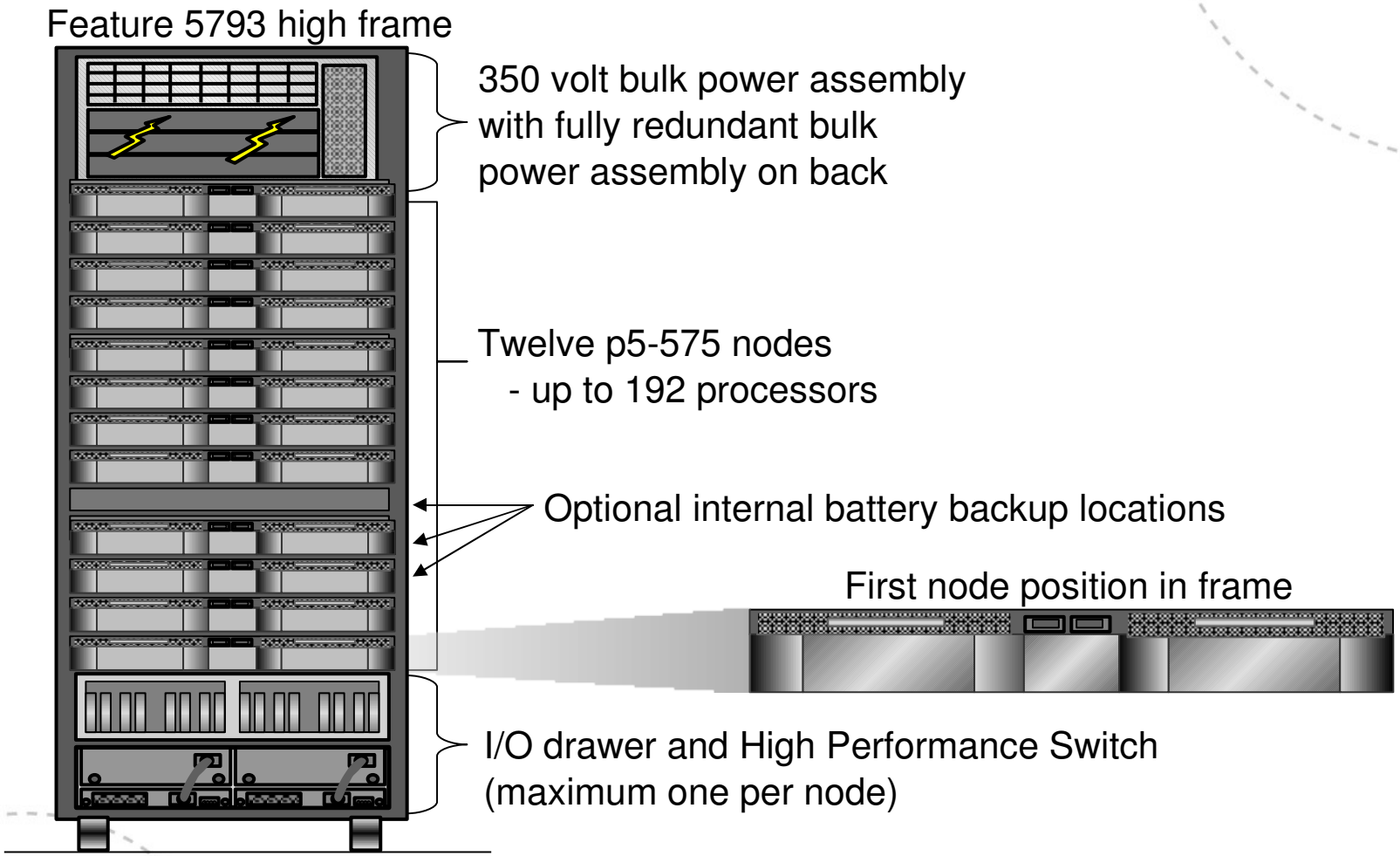
Two 350v power connectors, one redundant for connection to second bulk power assembly in rear of primary rack

p5-575 node may be powered down by the Hardware Management Console (HMC) without affecting other nodes, and slid out on the rails for in-place maintenance actions

IBM High Performance Switch (HPS)

- Peak link Bandwith 2GB/s - Latency < 5µs
- 16 ports for server-to-switch and 16 ports for switch-to-switch connections
- Supports up to 16 p5-575 servers and 32 links

p5-575 24-inch frame node layout

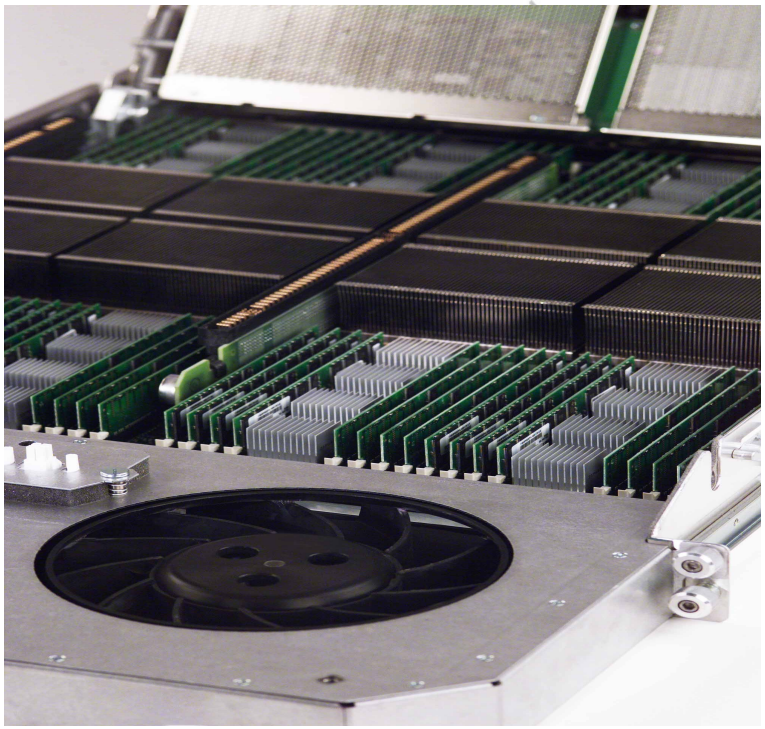


Second p5-575 cluster frame added

- p5-575 servers may be added
- Scaling to 16 servers

Software stack

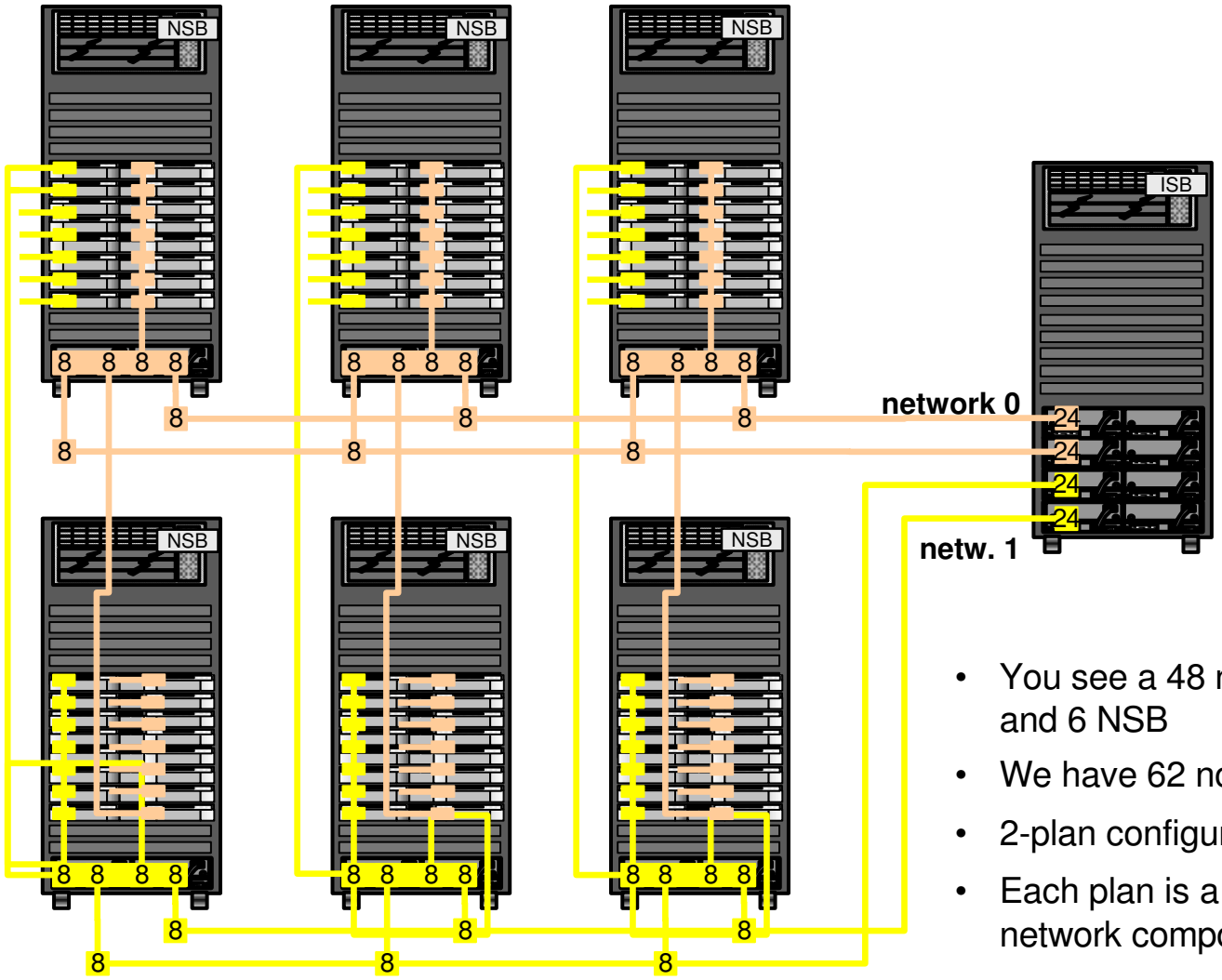
- AIX 5.3
- System administration tools (CSM)
- IBM compilers XLF, XLC, XLC++
- Parallel Operating Environment (POE)
- Libraries, ESSL, pESSL, libmass
- Scheduling software LoadLeveler
- Global Parallel File System GPFS
- HPC toolkit



Outline

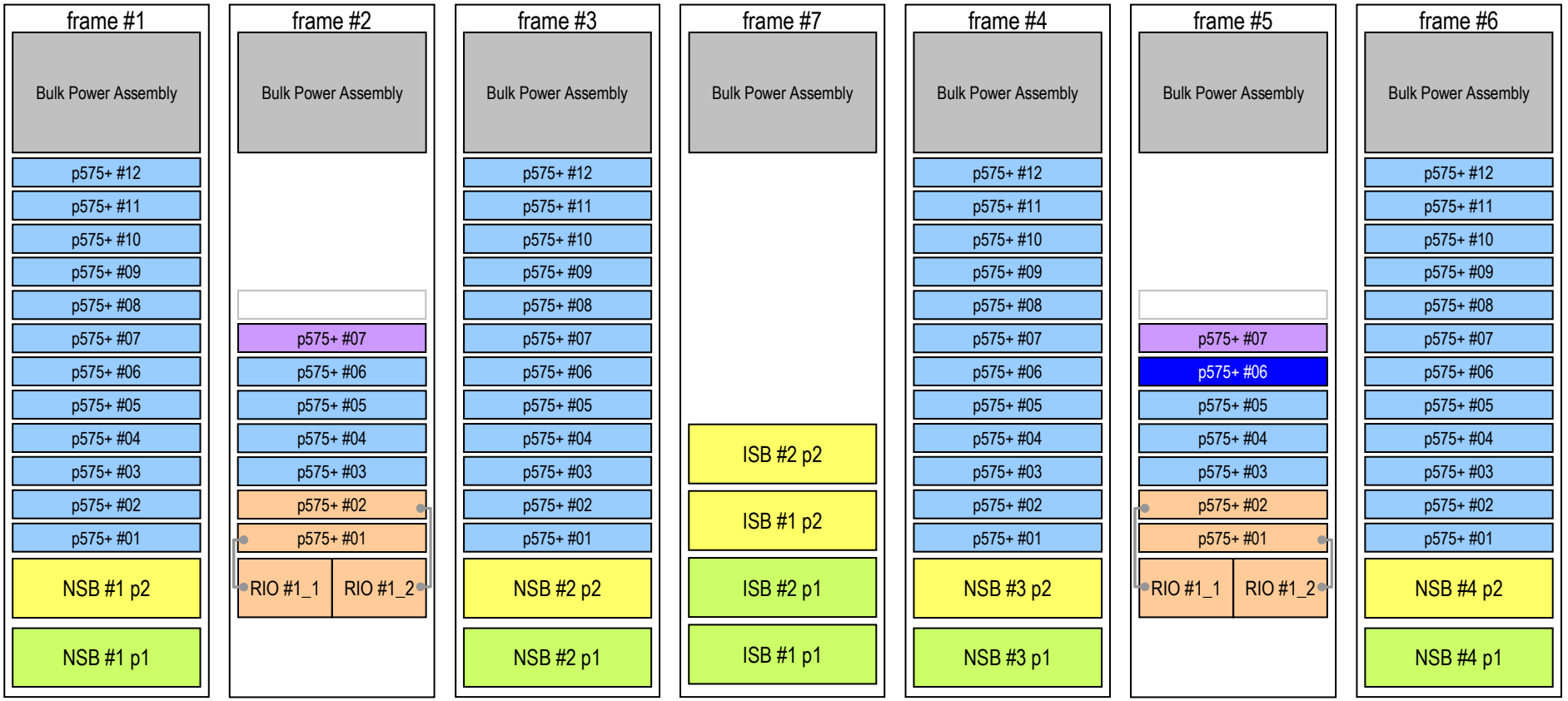
- System presentation
  - IBM p575+ system
  - NTNU configuration
- Early performance tests
- System usage and availibility

eHPS network



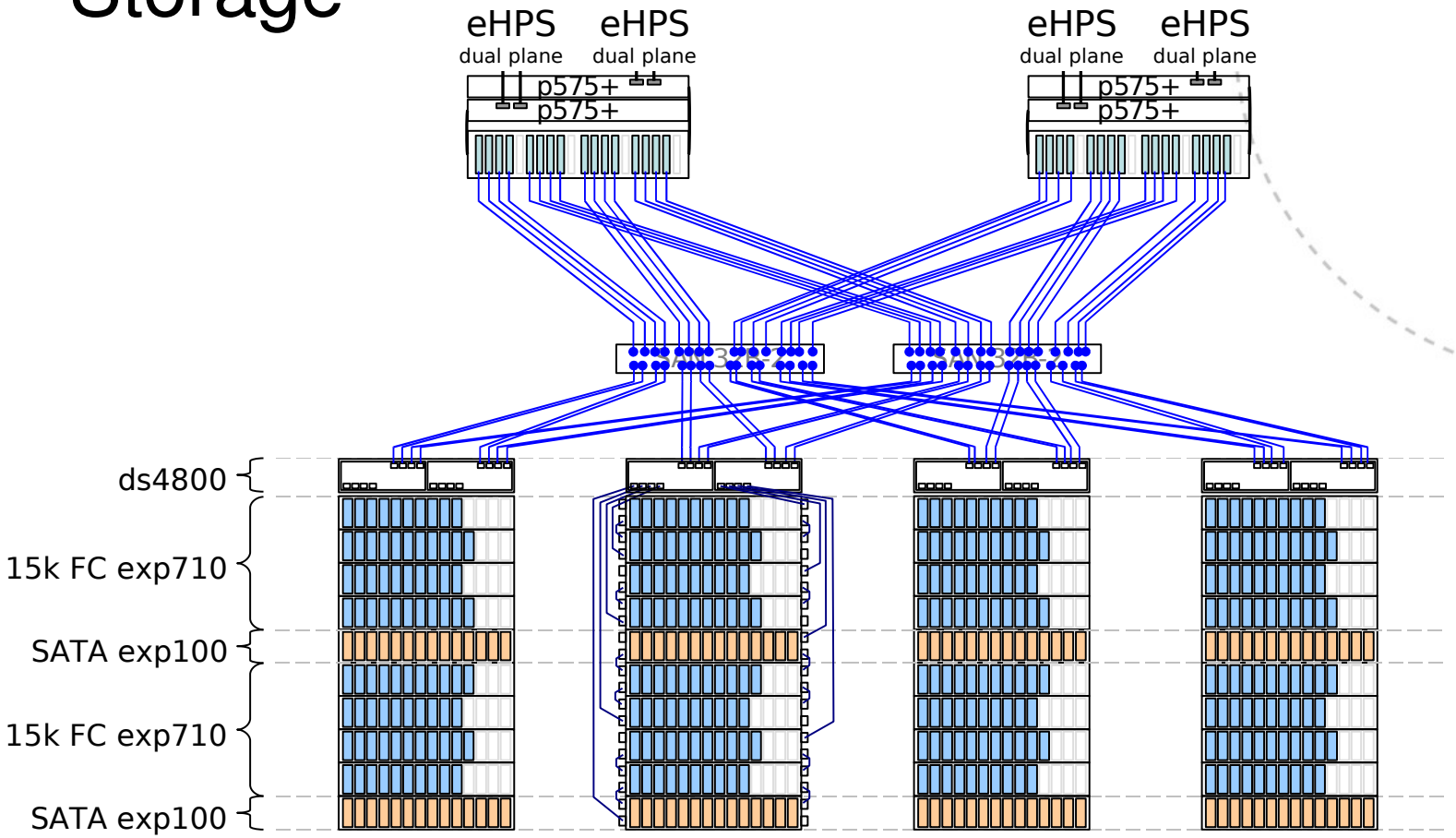
- You see a 48 nodes configuration with 4 ISB and 6 NSB
- We have 62 nodes with 4 ISB and 8 NSB
- 2-plan configuration
- Each plan is a stand-alone, redundant network composed of 2 ISB and 4NSB

Frame layout – node placement



p575+ #01	IO node (GPFS/VSD server); 16-way 1.9 GHz, 16 GB memory [4]
p575+ #07	Login node (FEN); 16-way 1.9 GHz, 32 GB memory [2]
p575+ #03	Compute node; 16-way 1.9 GHz, 32 GB memory [55]
p575+ #06	Compute node; 16-way 1.9 GHz, 128 GB memory [1]

Storage



4 storage controller pairs connected to 4 GPFS IO-nodes via redundant 4Gbps SAN fabrics

Outline

- System presentation
  - IBM p575+ system
  - NTNU configuration
- Early performance tests
- System usage and availibility

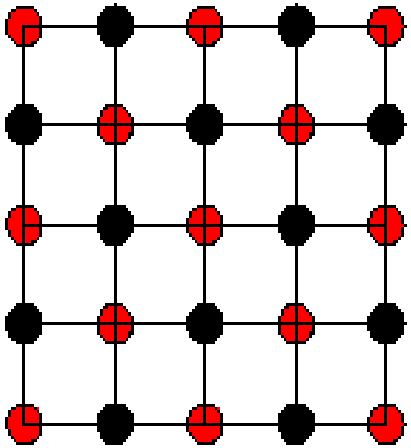
Uniprocessor tests

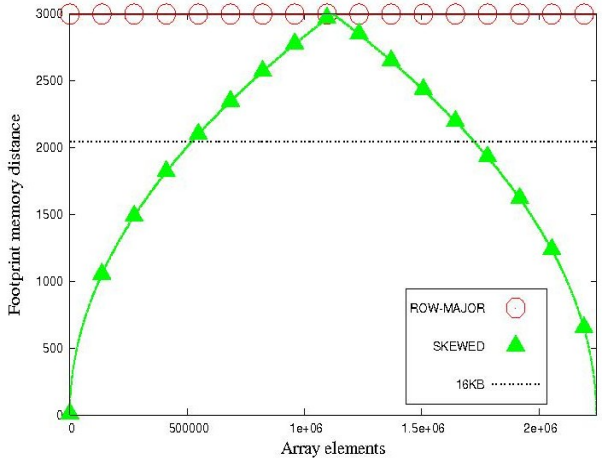
(Iterative method Gauss-Seidel: 5-point and 7-point stencils)

array

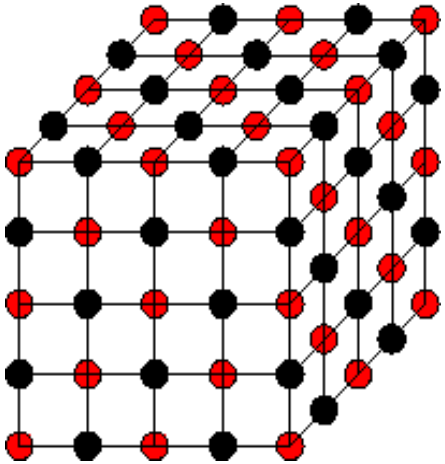
stencil footprint

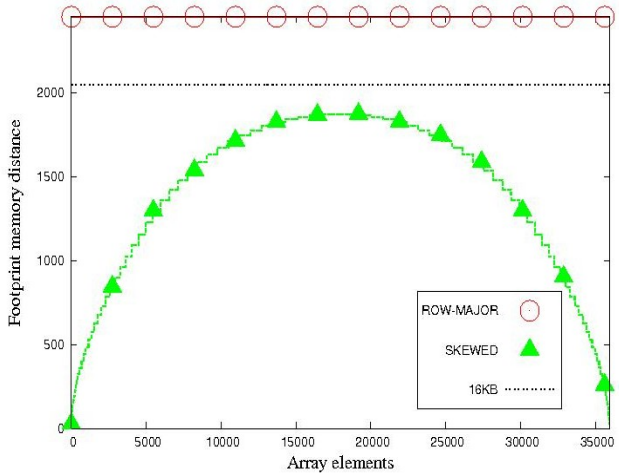
2D



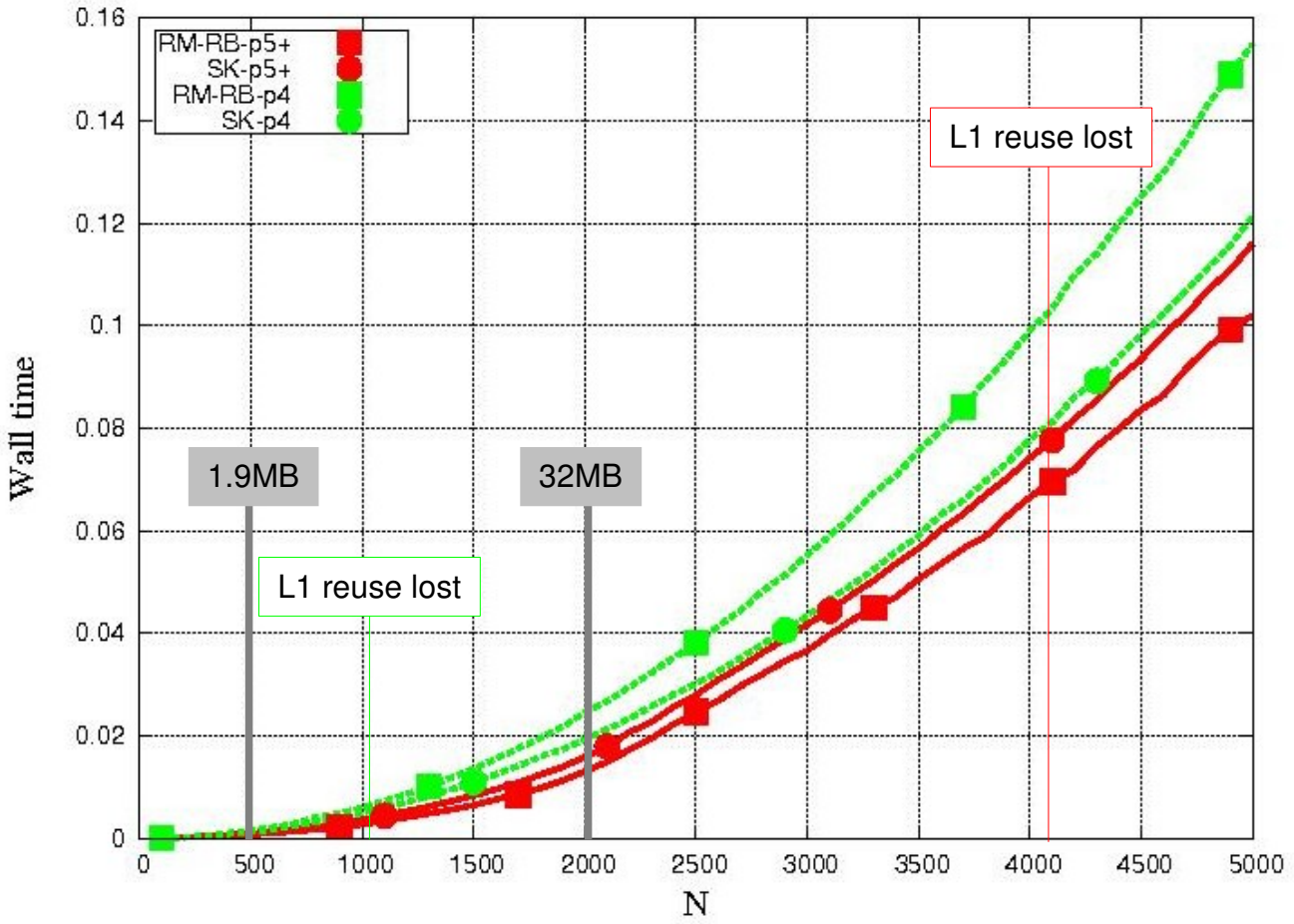


3D



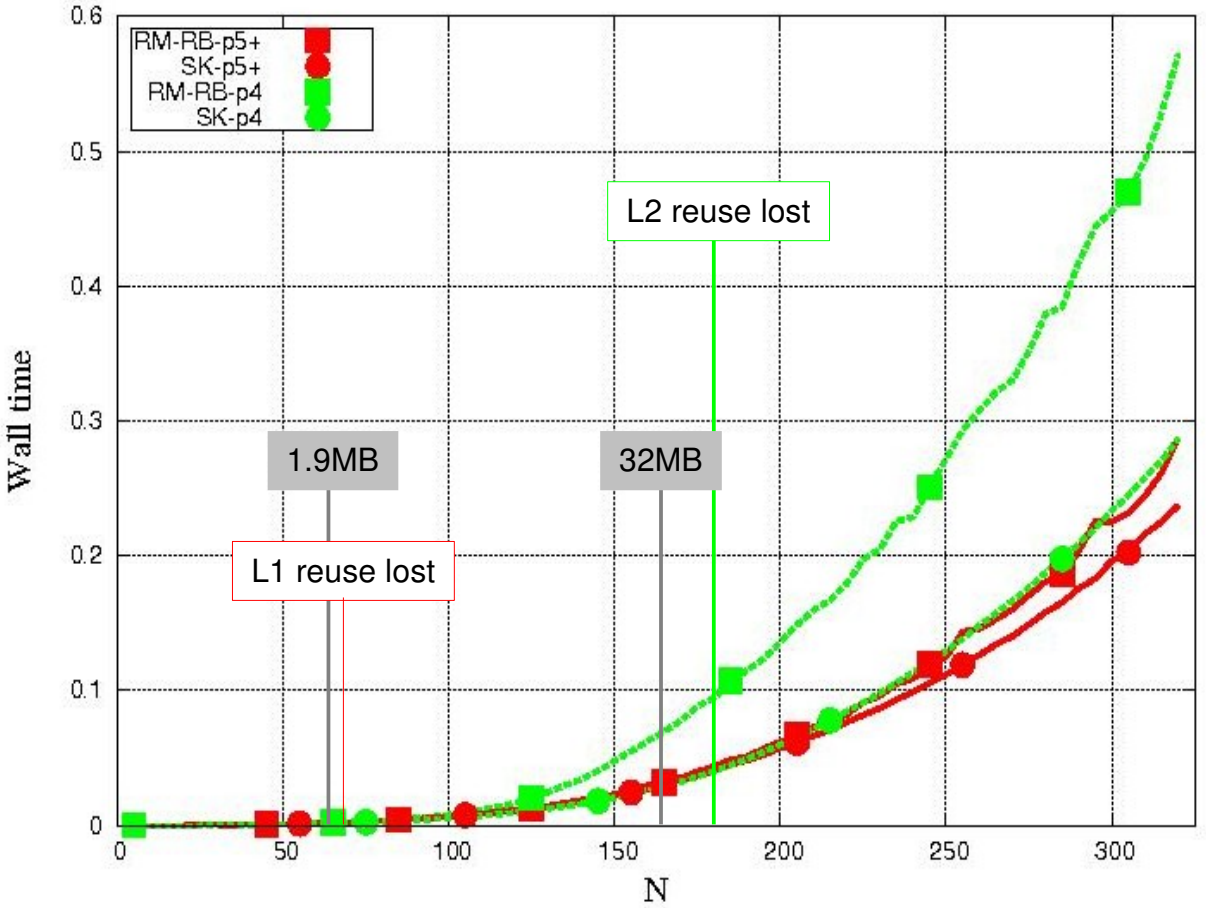


2D Gauss-Seidel



	Clock freq	L1	L2	L3	TLB
power5+	1.9 Ghz	64 KB	1.9 MB	32 MB	2048
pentium4	3.4 Ghz	16 KB	512 KB	-	128

3D Gauss-Seidel



	Clock freq	L1	L2	L3	TLB
power5+	1.9 Ghz	64 KB	1.9 MB	32 MB	2048
pentium4	3.4 Ghz	16 KB	512 KB	-	128

Outline

- System presentation
  - IBM p575+ system
  - NTNU configuration
- Early performance tests
- System usage and availibility

System availability

- Start of acceptance test on 19th October
- The acceptance test will take 4 weeks
- The system should be available around mid-November

Expected use

- Operational weather forecast
- Physics, simulation of superconducting/superfluid behavior
- Computational fluid dynamics
- Applied geophysics
- Structural engineering
- Climate modeling
- Computational chemistry

More information

- www.notur.no
- www.hpc.ntnu.no