

Measurement, Visualization, and Improvement of Linux Cluster Performance

Paul Howard, Bruce Schulman, and Stephen Fried
Microway, Inc.
12 Richards Road
Plymouth, MA 02360
{paulhoward,bschulman,steve}@microway.com

High Performance Embedded Computing Workshop
23-25 September 2008

Approved for public release; distribution is unlimited.

Improvement of Linux InfiniBand Clusters

Measurement

- MPI latency, bandwidth, and contention
- InfiniBand performance and error registers
- Any or all HCAs, any switch port, whole switch

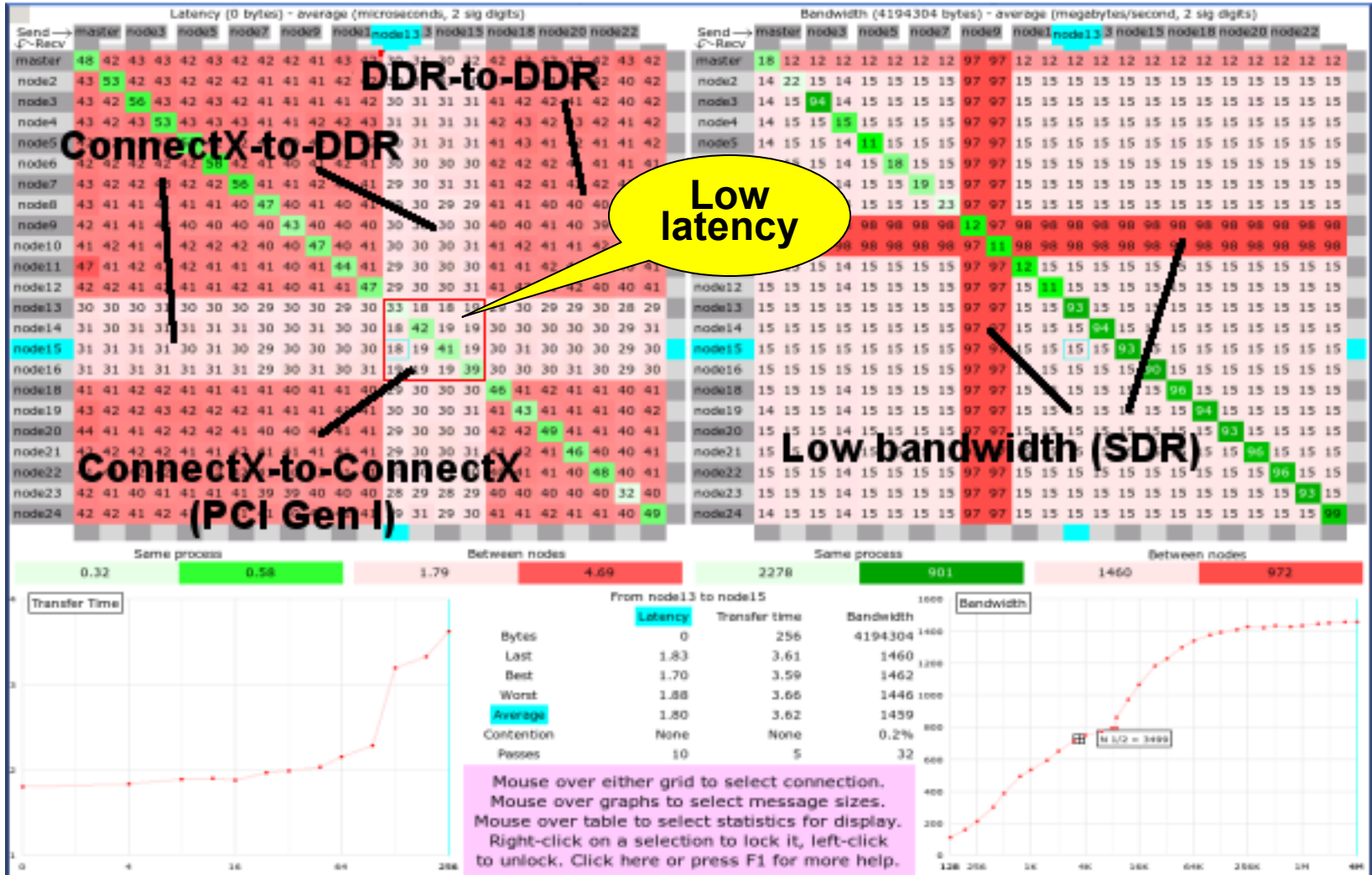
Visualization

- MPI Link-Checker 3D graphic for fabric capability
- InfiniScope for application performance

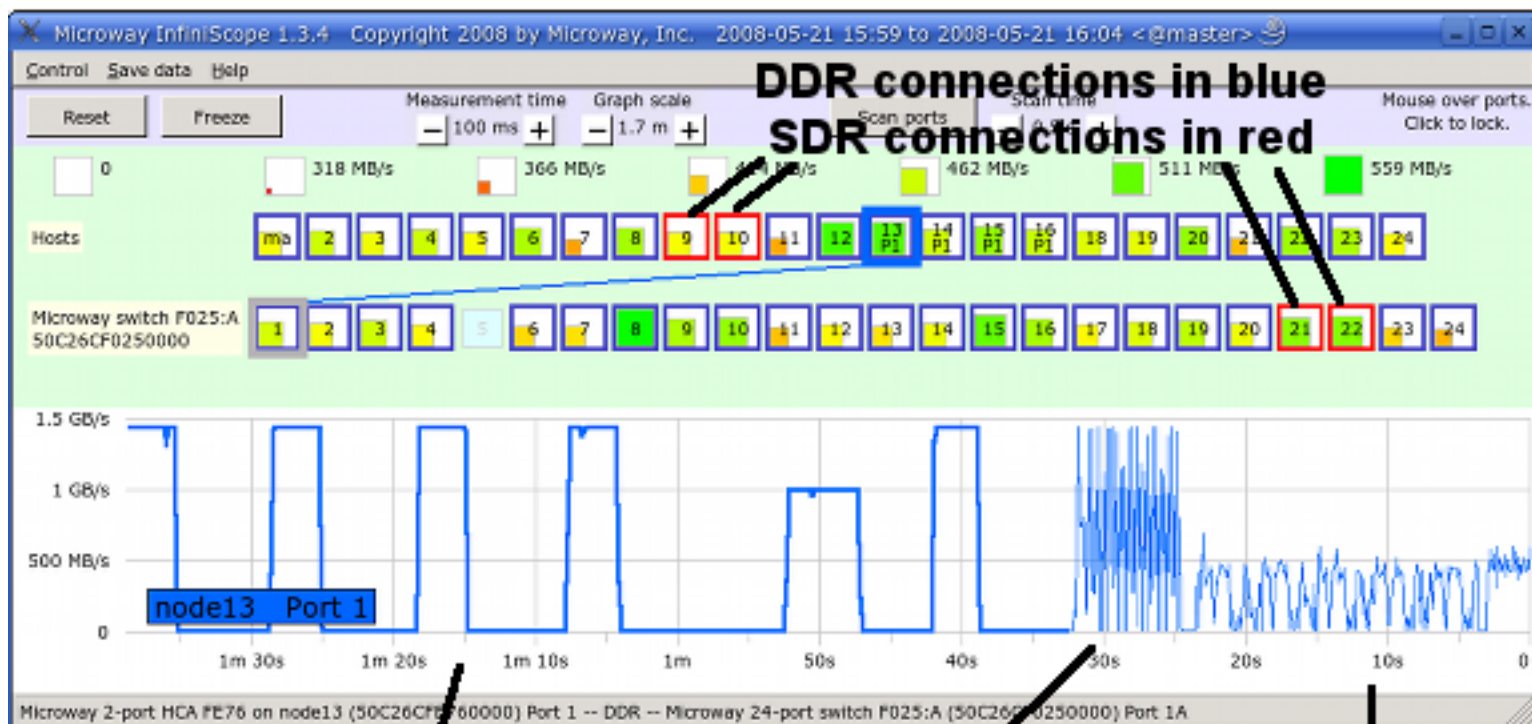
Improvement

- Verify fabric, find underperforming components
- Compare new hardware
- Find application data flow hot spots in fabric

MPI Link-Checker™



InfiniScope™



Fabric Loading Program

**Contention from
both at once**

MPI Link-Checker

Interconnect Technologies Compared

Interconnect	Bandwidth (MB/s)	Latency (μ s)	Cable length (m)
GigE	115	30	100
10GigE	860	9	100
InfiniBand DDR	1460	3.5	100*
TriCom-X™ InfiniBand DDR	1910	1.2	100*

* Emcore (Intel) cables.