Resource Management **Issues** in Large-Scale Cluster

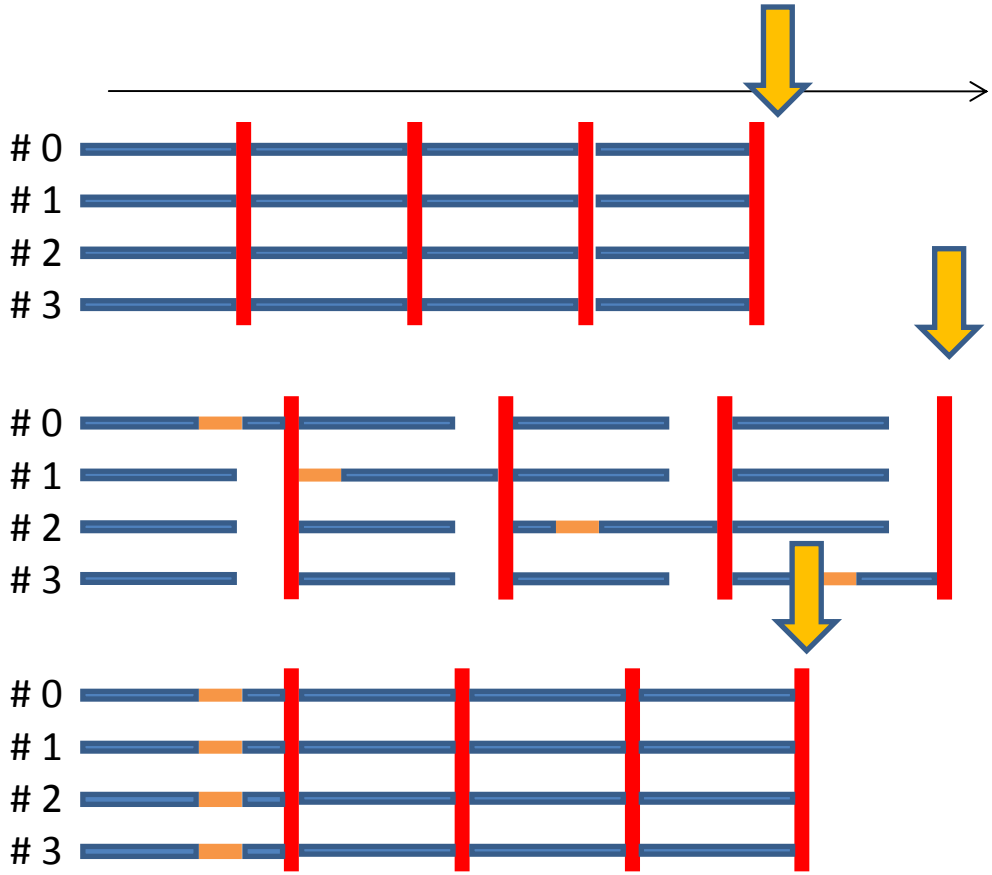
Yutaka Ishikawa
ishikawa@is.s.u-tokyo.ac.jp
Computer Science Department/Information Technology Center
The University of Tokyo
http://www.il.is.s.u-tokyo.ac.jp/
http://www.itc.u-tokyo.ac.jp

Outline

- Jittering
- Memory Affinity
- Power Management
- Bottleneck Resource Management

Issues

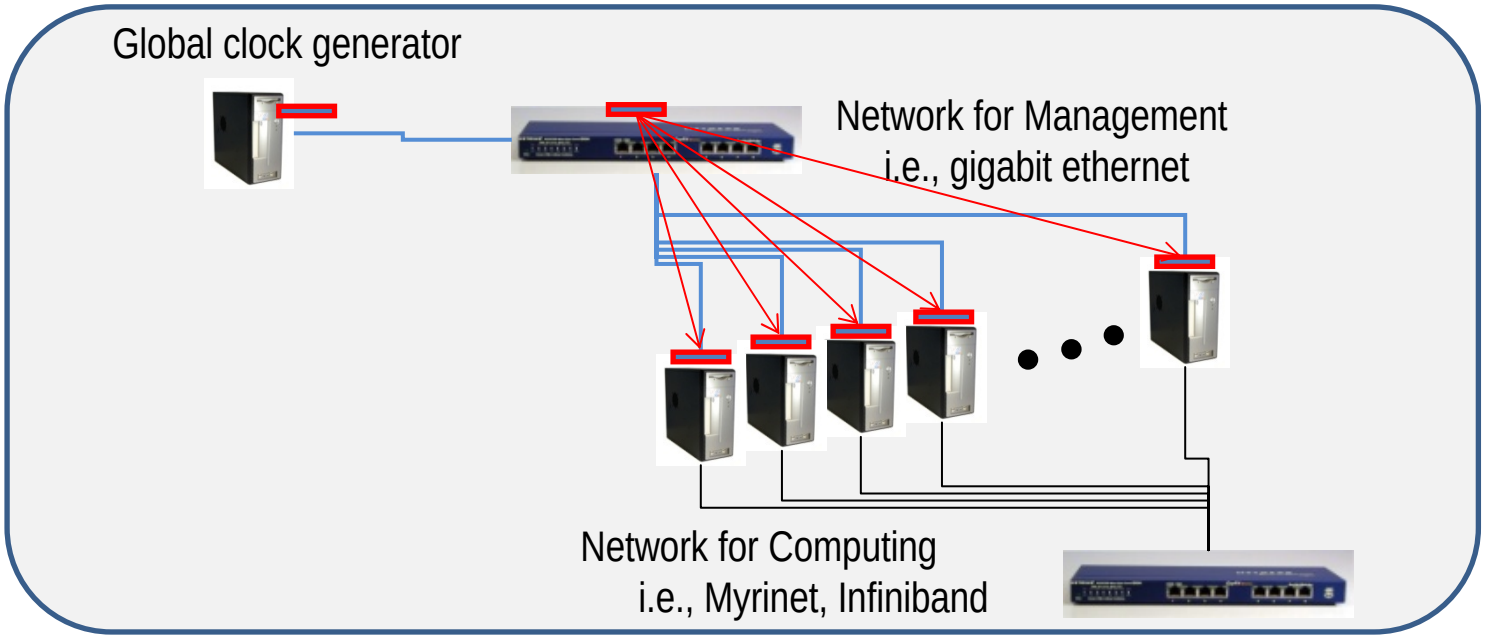
- Jittering Problem
  - The execution of a parallel application is disturbed by system processes in each node independently. This causes the delay of global operations such as allreduce



References:
•Terry Jones, William Tuel, Brain Maskell, “Improving the Scalability of Parallel Jobs by adding Parallel Awareness to the Operating System,” SC2003.
•Fabrizio Petrini, Darren J. Kerbyson, Scott Pakin, “The Case of the Missing Supercomputer Performance: Achieving Optimal Performance on the 8,1928 Processors of ASCI Q,” SC2003.

Jittering Problem

- Our Approach
  - Clusters usually have two types network
    - Network for Computing
    - Network for Management
  - The Management network is used to deliver the global clock
    - Interval Timer is turned off
    - Broadcast packet is sent from the global clock generator
  - Gang scheduling is employed for all system and application processes



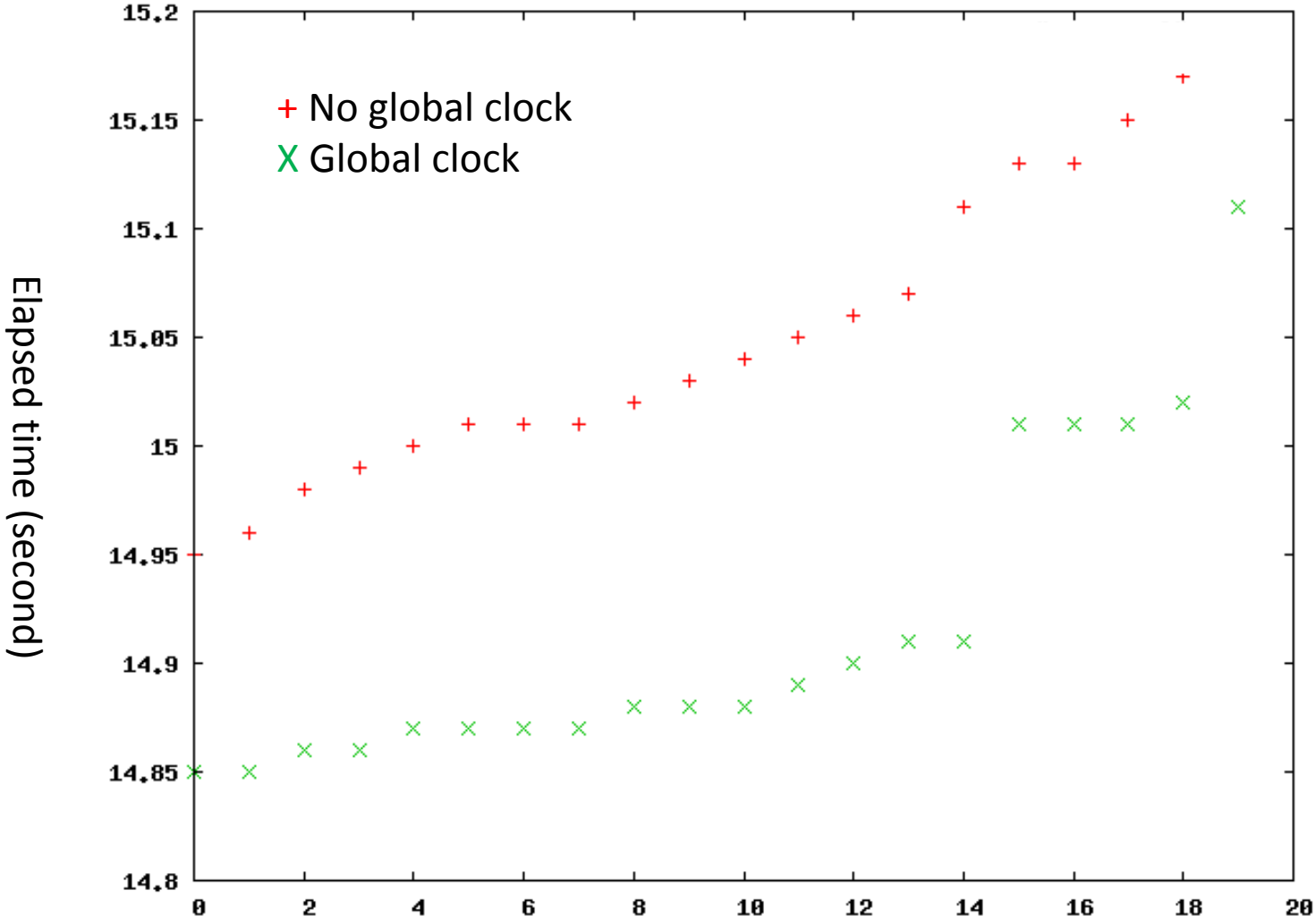
Jittering Problem

- Preliminary Experience
  - The Management network is used to deliver the global clock
  - The Interval Timer is turned off
  - Each arrival of the special broadcast packet, the tick counter is updated (The kernel code has been modified)
  - No cluster daemons, such as batch scheduler nor information daemon, are running, but system daemons are running

CPU : AMD Opteron 275 2.2GHz
Memory : 2GHz
Network : Myri-10G
: BCM5721 Gigabit Ethernet
# of Host : 16
Kernel : Linux 2.6.18 x86_64 modified
MPI : mpich-mx 1.2.6
MX : MX Version: 1.2.0
Daemons: syslog, portmap, sshd, sysstat, netfs, nfslock, autofs, acpid, mx, ypbind, rpcgssd, rpcidmapd, network

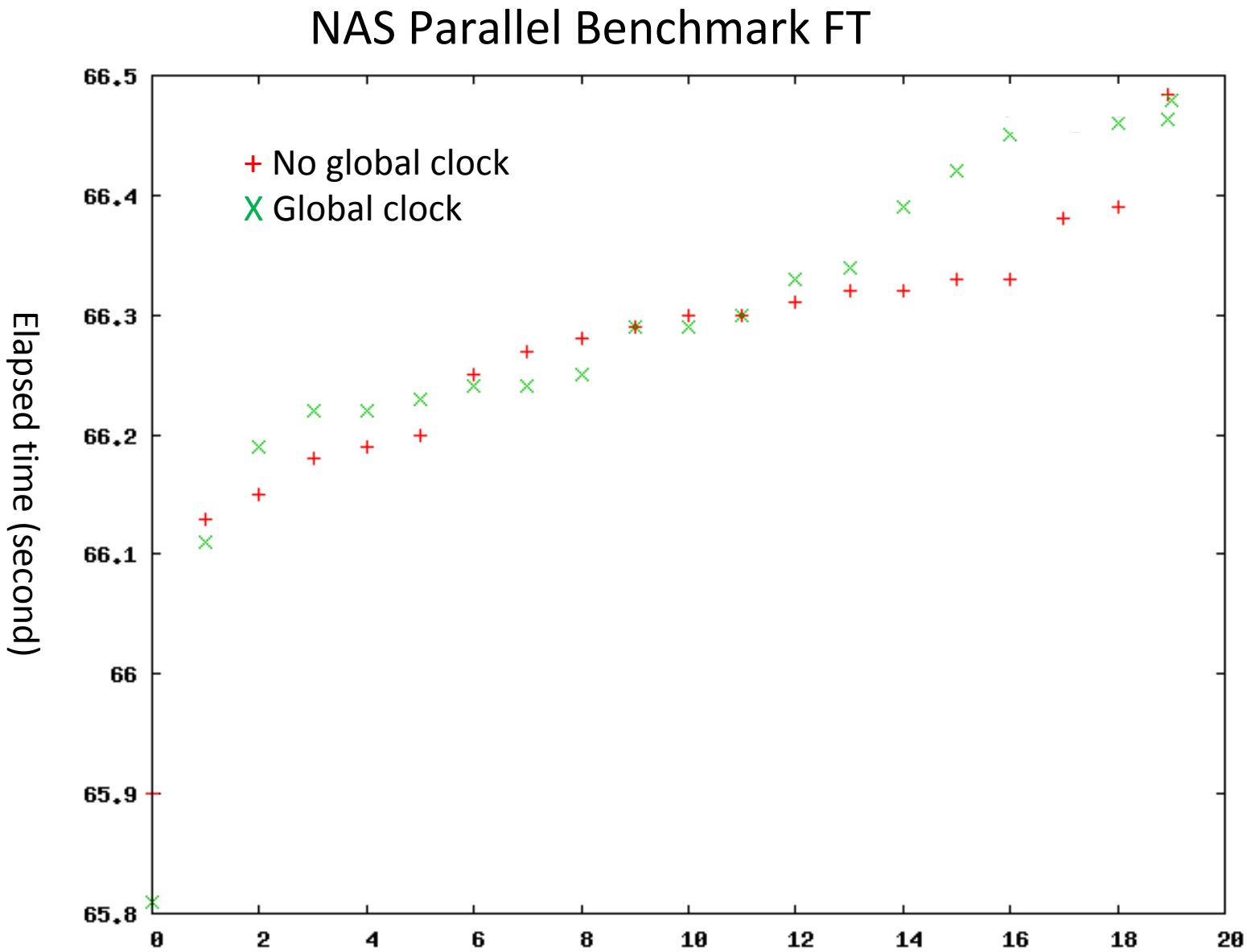
Preliminary Global Clock Experience

NAS Parallel Benchmark MG

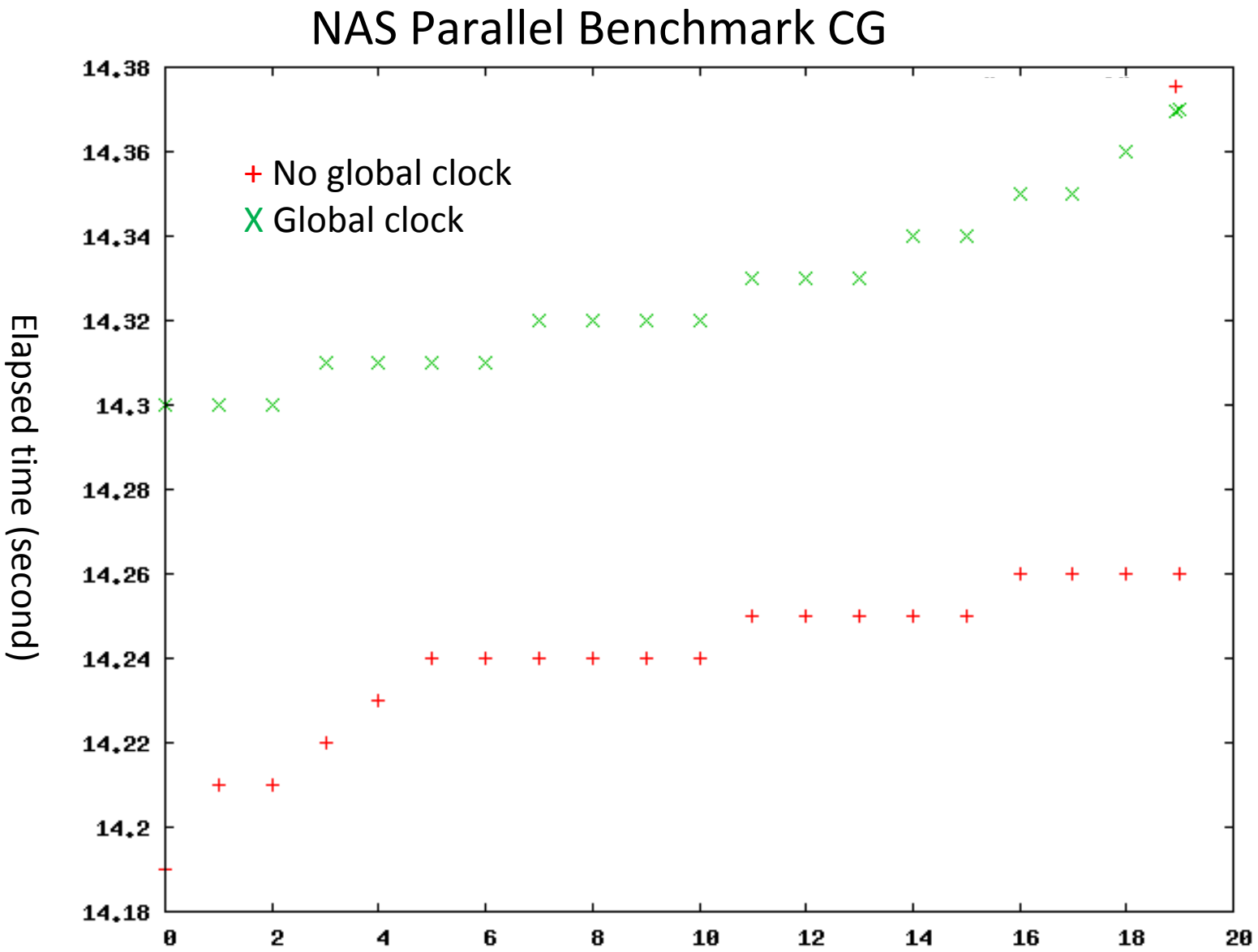


20 times executions are sorted

Preliminary Global Clock Experience

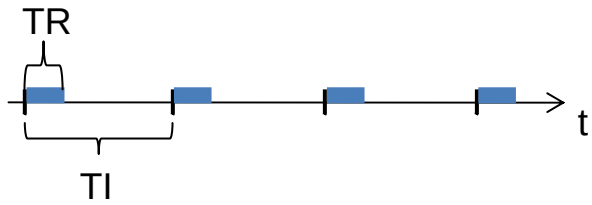


Preliminary Global Clock Experience



What kind of heavy daemon running in cluster

- Batch Job System
  - In case of Torque
  - Every 1 second, the daemon takes 50 microseconds
  - Every 45 seconds, the daemon takes about 8 milliseconds
- Monitoring System
  - Not yet majored
- Simple Formulation



$$\text{Worst Case Overhead} = \sum \frac{\text{MIN}(TIi,\ TRi \times N)}{TIi}$$

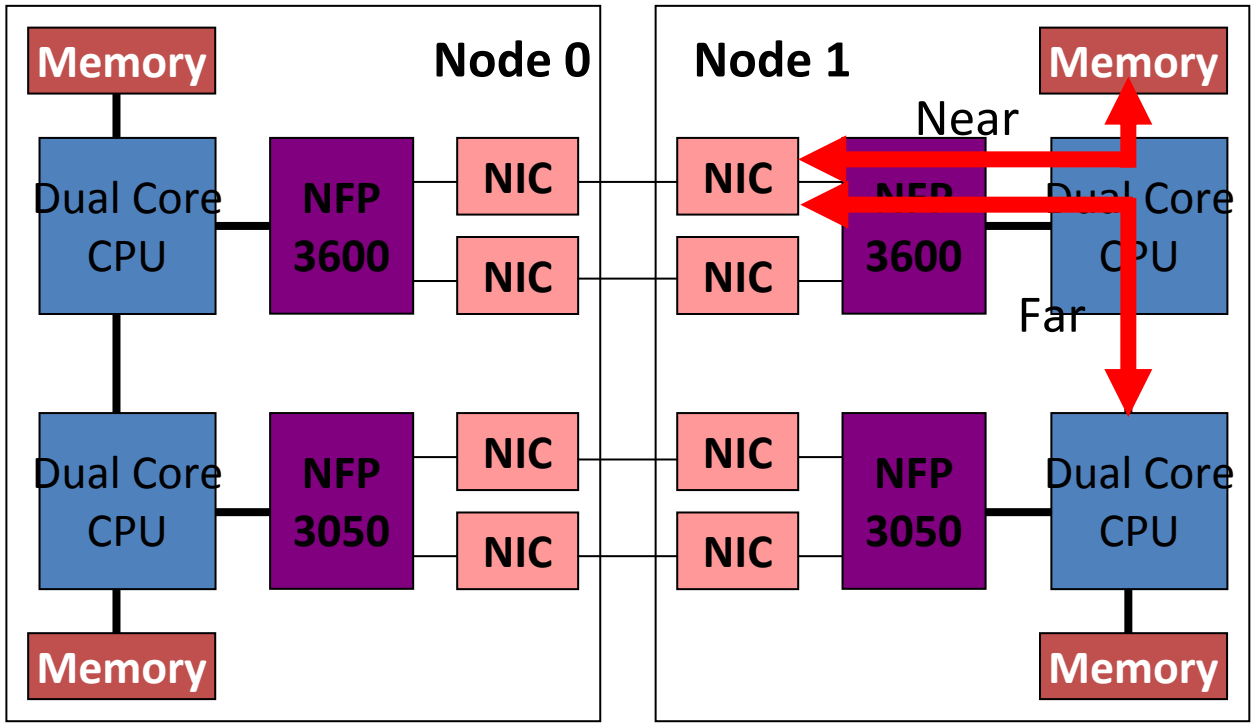
N: Number of nodes
TIi: Interval time in daemon i
TRi: Running time in daemon i

In case of 1000 node cluster
0.000050*1000/1 + 0.008*1000/45 = 22.8 %

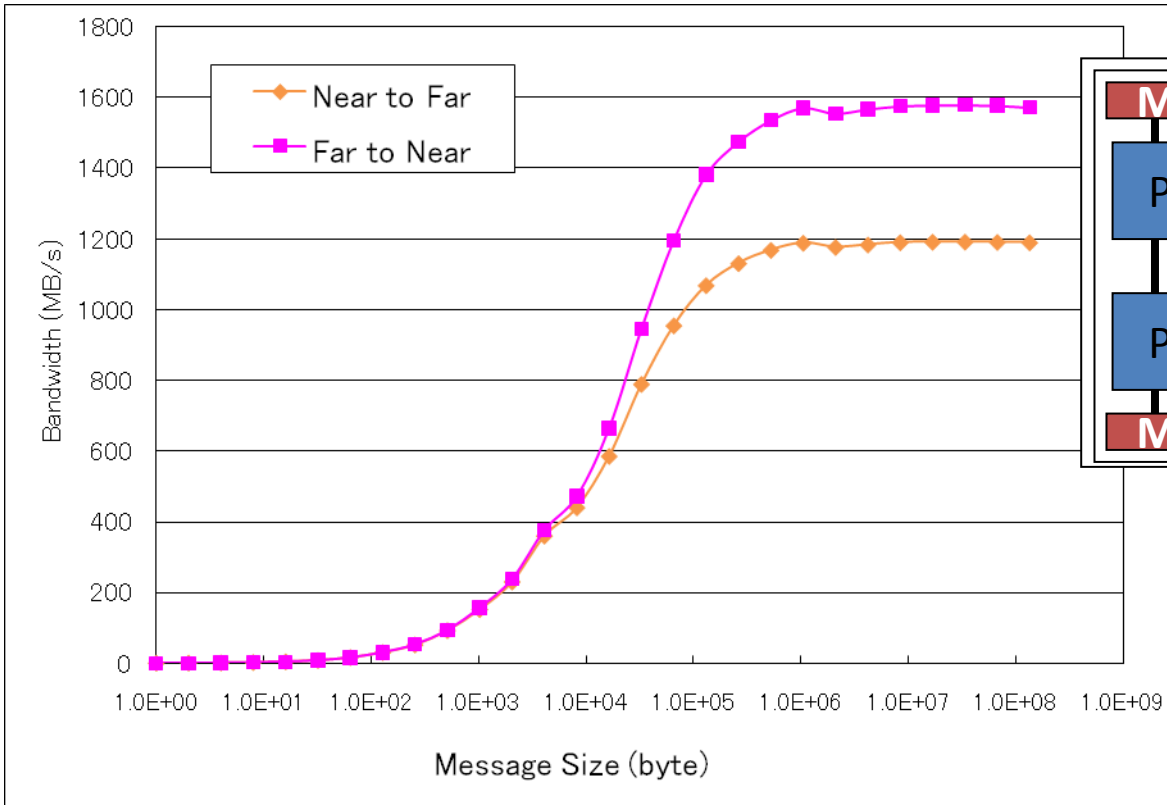
The worst case might never happen !

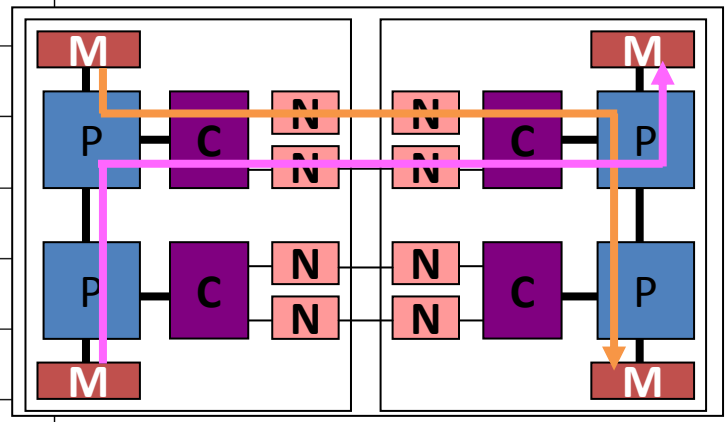
Issues on NUMA

- Memory Affinity in NUMA
  - CPU ←→Memory
  - Network ←→Memory
- An Example of network and memory



Memory Location and Communication





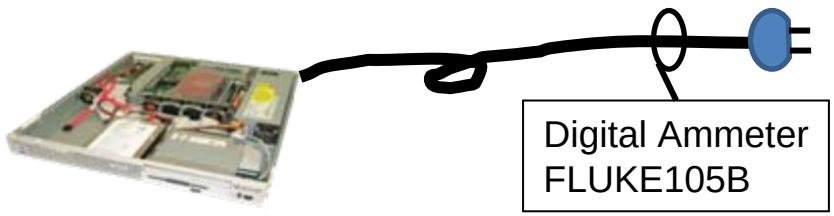
Note: The result depends on the BIOS settings.

- Communication performance depends on data location.
- Data is also accessed by CPU.
- The location of data should be determined based on both CPU and network location.
  - Dynamic data migration mechanism is needed ??

Power Management

Power Consumption Issue

- 100 Tflops cluster machine
  - 1666 Nodes
- If 80 % machine resource utilization (332 nodes are idle)
  - 66 KW power is wasted in case of idle
    - 55K$(660 万円)/year
    - This is under estimation because memory size is small and no network switches are included
  - 10.6KW power is wasted though the power is turned off!!
    - 9K$ (110万円)/year



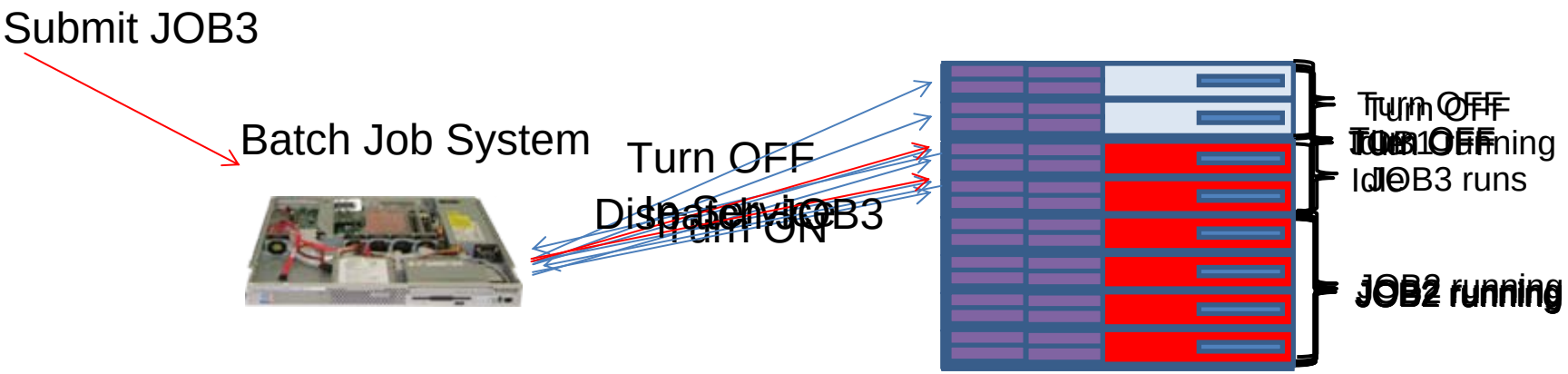
Power Consumption in single node

	Power Consumption (Amp)	
HPL running (Not optimized)	2.92	
Idle (1.9GHz)	2.44	
Idle (1.0GHz)	2.02	
Suspended	1.61	??
No Power but power cable is plugged in (BMC running)	0.32	

Supermicro AS-2021-M-UR+V
Opteron 2347 x 2
(Balcerona 1.9 GHz, 60.8 Gflops)
4 Gbyte Memory
Infiniband HCA x 2
Fedora Core 7

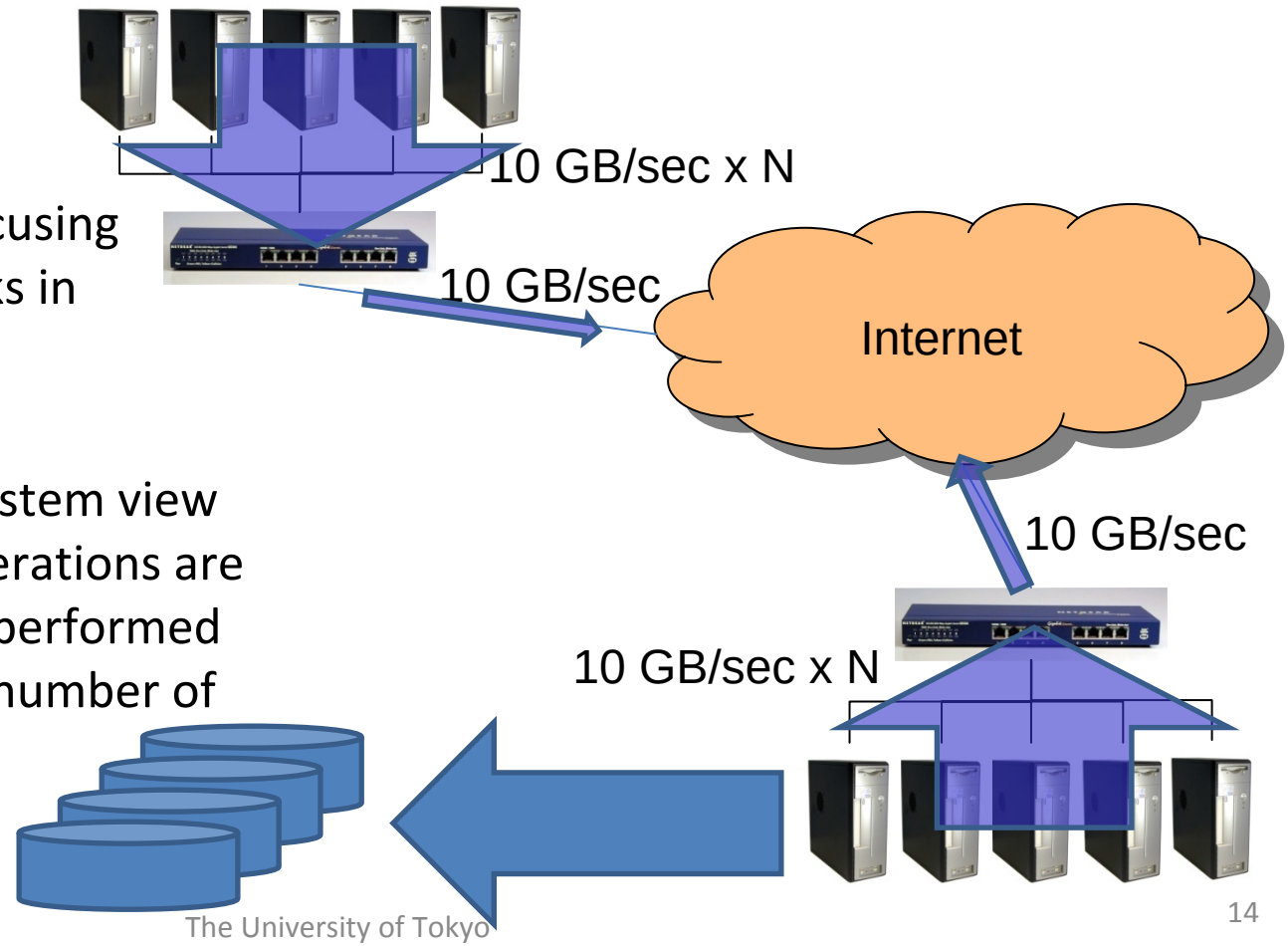
Power Management

- Cooperating with Batch Job system
  - Idle machines are turned off
  - When those machines are needed, they are turned on using the IPMI (Intelligent Platform Management Interface) protocol (BMC).
  - However, still we lose 300 mA for each idle machine
- Quick shutdown/restart and synchronization mechanism



Bottleneck Resource Management

- What are bottleneck resources
  - A cluster machine has many resources while other resources are limited.
  - When the cluster accesses such a resource, overloading or congestion happens
- Examples
  - Internet
    - We have been focusing on bottleneck links in GridMPI
  - Global File System
    - From the file system view point, N file operations are independently performed where N is the number of node



Summary

- We have presented issues on large-scale clusters
  - Jittering
  - Memory affinity
  - Power management
  - Bottleneck resource management