

Network Based Computing: Trends, Issues and Challenges



Dhableswar K. (DK) Panda
Department of Computer Science and Engineering
The Ohio State University

E-mail: panda@cse.ohio-state.edu
<http://nowlab.cse.ohio-state.edu>
<http://www.cse.ohio-state.edu/~panda>

Presentation Outline

- Motivation
 - Technological Trends
 - Applications (Past, Current, and Emerging)
 - Defining Network-Based Computing Systems
 - Challenges and Research Issues
 - Research Challenges and Projects
 - Available Experimental Testbed
 - Related 888 Course
- Conclusions

Dominant Computing System and Paradigm

- Continues to change over the years
- Objective: design cost-effective systems
- Depends on two primary factors
 - technological trends (processor, memory, disk, interconnection)
 - past
 - current and future
 - computational demand of applications
 - past
 - current and emerging

Past and Current Systems

- Single Board Computer, Mainframe, Personal Computer, Workstation, Multiprocessors, Client-Server, Massively Parallel Processor
- Have been **dominant computing systems** during some years
- Where are we headed next?

Basic Resources for Designing Computing Systems

- Computing cycles
 - capacity (total amount of CPU cycles)
 - rate (MHz)
- Memory and Disk
 - capacity (MBytes)
 - bandwidth (MBytes/sec)
 - latency (nano/micro/milli sec)
- Interconnection/network between processor-memory-disk
 - bandwidth (MBytes/sec)
 - latency (microsec)

Presentation Outline

- Motivation
- Technological Trends
- Applications (Past, Current, and Emerging)
- Defining Network-Based Computing Systems
- Challenges and Research Issues
- Research Challenges and Projects
- Available Experimental Testbed
- Related 888 Course
- Conclusions

Technological Trends in the Past

- **Processor**
 - Speed: 1.35x per year (before mid 1980s), 1.58x per year (afterwards)
 - Cost (\$ per MIPS): reduces by a factor of 0.5 per year (83-90)
- **Memory (DRAM)**
 - Capacity: 4x-every-three-years (1.6x per year)
 - Cycle time: 1/3 in 10 years
 - Cost (\$ per MByte): reduces by a factor of 2.7 per year
- **Magnetic Disk**
 - Capacity: 1.25x per year (prior to 1990), 1.50x per year (afterwards)
 - Access time: 1/2 in 10 years (80-90)
 - Cost (\$ per MByte): reduces by a factor of 1.3/year (86-92), 2.0/year (92-95)

Technological Trends in the Past (Contd.)

- **Interconnection schemes**
 - System Bus: 10.0 MBytes/sec (Multibus-I) to 1024 MBytes/sec (Sun SMP system)
 - 100 times improvement
 - Interconnection Network: 1.0 MByte/sec (CM-2) to 300 MBytes/sec (Cray T3D/T3E)
 - 300 times improvement
 - Local Area Network (LAN):
 - Ethernet (1974-): 10 Megabits/sec (1.25 MBytes/sec)
 - Token Ring (1980-): 4/16 Megabits/sec (0.5/2.0MBytes/sec)
 - FDDI (1987-): 100 Megabits/sec (12.5 MBytes/sec)
 - Fast Ethernet (1993-): 100 Megabits/sec (12.5 MBytes/sec)
 - 10 times improvement in 20 years

Impact on Designing Past and Current Computing Systems

- **Computing systems have been designed by ignoring LANs**
 - Mainframes
 - Symmetric MultiProcessors (SMPs)
 - Massively Parallel Systems (MPPs)
 - Personal Computers
 - Desktop Workstations
- **Objective: Design cost-effective systems by taking advantage of on-going advancements in processor, memory, disk, bus, and interconnection network technologies.**
- **LANs/WANs have been primarily used to**
 - provide connectivity among computing systems
 - support file transfers between computing systems (client-server) and resources (disks, printers, scanners, ...)
 - support transactions between system

Technological Trends - Current and Future

- Networking Technology
- Powerful Multi-core Processors
- Accelerators
- Solid State Devices (SSDs)
- Emergence of the Web Technology

Growth in Commodity Networking Technology

Ethernet (1979 -)	10 Mbit/sec
Fast Ethernet (1993 -)	100 Mbit/sec
Gigabit Ethernet (1995 -)	1000 Mbit /sec
ATM (1995 -)	155/622/1024 Mbit/sec
Myrinet (1993 -)	1 Gbit/sec
Fibre Channel (1994 -)	1 Gbit/sec
InfiniBand (2001 -)	2 Gbit/sec (1X SDR)
10-Gigabit Ethernet (2001 -)	10 Gbit/sec
InfiniBand (2003 -)	8 Gbit/sec (4X SDR)
InfiniBand (2005 -)	16 Gbit/sec (4X DDR)
	24 Gbit/sec (12X SDR)
InfiniBand (2007 -)	32 Gbit/sec (4X QDR)
40-Gigabit Ethernet (2010 -)	40 Gbit/sec
InfiniBand (2011 -)	54.6 Gbit/sec (4X FDR)
InfiniBand (2012)	100 Gbit/sec (4X EDR)

27 times in the last 10 years

Powerful Multicore Processors

- Processor speed was doubling every 18 months (since early 90s to 2002/2003)
- Design and development of powerful and cost-effective uni-processor/multi-processor systems
- Speed could not be enhanced any more due to power dissipation
- Led to multi-core era
- 8/16-core processors are common today
- 64/128-core processors will be common in another 2-3 years
- System on Chip (SOC) like Tilerla already have 128 cores in a single chip
- Up to 512/1024-core commodity processors in the next 5 years

Accelerators

- Accelerators are emerging in full-speed
- Started with nVIDIA, IBM Cell, etc.
- Intel Many-Integrated Cores (MIC) is upcoming
- CUDA and GPGPU framework are providing software integration with the accelerators

Solid State Devices (SSDs)

- Magnetic disks have been the main component for permanent storage
- Capacity has been growing, but not speed
- Performance difference between memory and disk speed steadily increasing
- New Solid State Devices (SSDs) are built with flash memory technology
- Very fast compared to magnetic disk
- Helping to reduce the gap between memory and storage
- Some SSDs (Fusion I/O) can operate on full networking speed

Emergence of the Web Technology

- Web technology to share information and data
- CGI scripts and Java support
 - to integrate computation with communication
 - to blur distinction between heterogeneous computing systems (architectures, operating systems, instruction sets)
- Rapid growth in the last 15 years

Summary of Technology Trends

- Network has been ignored as an integral component of a computing box in the past
- Five important developments in the last decade
 - commodity and affordable networking technologies
 - powerful multi-core processors
 - Accelerators
 - Solid State Devices (SSDs)
 - emergence of the Web technology

Presentation Outline

- Motivation
- Technological Trends
- Applications (Past, Current, and Emerging)
- Defining Network-Based Computing Systems
- Challenges and Research Issues
- Research Challenges and Projects
- Available Experimental Testbed
- Related 888 Course
- Conclusions

Past Applications

- Sequential Batch Jobs
- Grand Challenge (Parallel) Applications
- Transactions-based Applications
 - airline/hotel/car reservation
 - databases

Trend for Computational Demand

- Continuous increase in demand
 - multiple design choices
 - larger data set
 - finer granularity of computation
 - simulation with finer time step
 - low-latency/high-throughput transaction,
- Expectation changes with the availability of better computing systems

Emerging Applications

- High Performance and High Throughput Computing Applications
 - Weather forecasting, physical modeling and simulations (aircraft, engine), drug designs, ...
- Database applications
 - data-mining, data ware-housing, enterprise computing
- Financial
 - e-commerce, on-line banking, on-line stock trading
- Digital Library
 - library of audio/video, global library
- Collaborative computing and visualization
 - shared virtual environment
- Telemedicine
 - content-based image retrieval, collaborative visualization/diagnosis
- Virtual Reality, Education, and Entertainment

High Performance Computing

- Large number of parallel applications in scientific and engineering disciplines
- Grand Challenge Applications
 - weather forecasting
 - manufacturing (computer-aided design, modeling, and simulation)
 - designing and analyzing VLSI circuits
 - analyzing chemical formulae for new medicines
 - simulation and visualization of physical phenomena (astronomy, deforestation, global warming, ...)
 - simulating nuclear reactions (US Department of Energy's ASCI program)

High Throughput Computing

- Simulation experiments for large number of design choices
- Common requirement for scientists, engineers, physicists, chemists, biomedical engineers, ...
- Need results in a short period of time
- Example: design of an aircraft wing
 - 5 degrees of freedom (major parameters)
 - 5 different values for each parameter
 - 25 different simulation experiments (independent)
 - each simulation experiment takes 2 hours
- Availability of 25 computers over a network will allow design decisions to be made in 2 hours (instead of 50 hours using a single computer)

Grid Computing/Cloud Computing

- Applications require computing systems with enormous computing power at affordable cost
- Some applications require sharing of computational cycles of a set of systems over network
 - to provide meta-computing/grid computing environment to users
 - to distribute parallel jobs over a set of systems
- Some applications require high-speed connectivity between systems to support cooperative parallel computation
- Filed is emerging into the area of Cloud Computing

Scalable Database Servers

- Continuous increase in demand for organizing data, searching them, and making decisions
- Designing database servers which
 - deliver high performance and throughput and scalable
 - support a range of thick and thin clients over network
- Data-mining and knowledge discovery
 - establishing association/correlation between data
 - deriving patterns from data
- Requires enormous amount of computing power
 - to integrate data generated on a daily/weekly/monthly basis
 - to search all possible association/correlation/pattern between data

Web/Multimedia Server

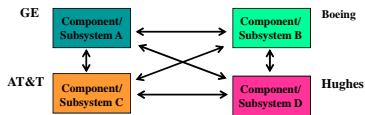
- Web technology is providing a new dimension
 - to share information and data (text, audio, video, image, ..) across geographically distributed systems
 - to search for information
 - to allow collaboration
- Number of URLs keep increasing in an exponential manner
- Requires scalable servers which
 - provide fast search/access
 - allow fast transfer of data between servers and clients
 - support both **push** and **pull** technologies

Electronic Commerce

- On-line banking
- On-line shopping
 - virtual distributors
 - virtual warehouse
 - real-time interaction between
 - consumer
 - merchant
 - consumer's bank
 - merchant's bank
- On-line buying and selling stock
 - real-time interaction between
 - consumer
 - broker
 - stock exchange

Manufacturing

- Component-based simulation of a product
 - parts of an aircraft
 - parts of a car
- Integrated simulation of a product
 - parts of an aircraft
 - parts of a car
- Integrated simulation of a product with components from different vendors
 - different vendors need not disclose the architecture/model of their components/subsystems



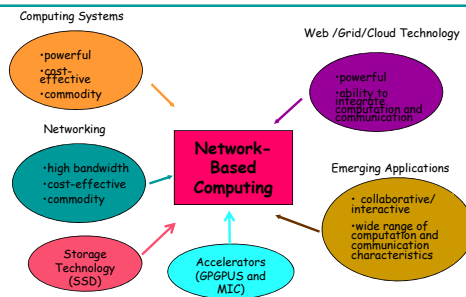
Summary of Applications

- Continuous increase in computational demand
- Wide range of applications with different computation and communication characteristics
- Many applications are collaborative/interactive in nature
- How to build computing systems to support modern applications while taking advantage of advancement in technologies?
 - computing
 - networking
 - web

Presentation Outline

- Motivation
- Technological Trends
- Applications (Past, Current, and Emerging)
- Defining Network-Based Computing Systems
- Challenges and Research Issues
- Research Challenges and Projects
- Available Experimental Testbed
- Related 888 Course
- Conclusions

Network-Based Computing



30

A closer look at the applications

Characteristics

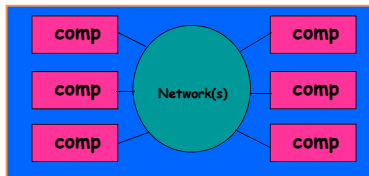
- tasks
 - parallel
 - independent sequential
- computation
 - non-interactive
 - interactive
 - user(s)/client(s) over the LAN
 - user(s)/client(s) over the WAN
- data
 - centralized
 - distributed

Different Types of NBC Systems

- Five major categories
 - Dedicated Cluster over a LAN
 - Clusters with Interactive clients
 - Globally-Interconnected Systems
 - Interconnected Clusters over WAN
 - Multi-tier Data Centers

Dedicated Cluster over a SAN/LAN

- Packaged inside a rack/room interconnected with SAN/LAN

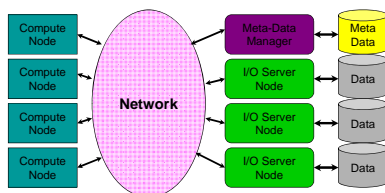


Trends for Commodity Computing Clusters in the Top 500 List (<http://www.top500.org>)

Nov. 1996: 0/500 (0%)	Jun. 2002: 80/500 (16%)	Nov. 2007: 406/500 (81.2%)
Jun. 1997: 1/500 (0.2%)	Nov. 2002: 93/500 (18.6%)	Jun. 2008: 400/500 (80.0%)
Nov. 1997: 1/500 (0.2%)	Jun. 2003: 149/500 (29.8%)	Nov. 2008: 410/500 (82.0%)
Jun. 1998: 1/500 (0.2%)	Nov. 2003: 208/500 (41.6%)	Jun. 2009: 410/500 (82.0%)
Nov. 1998: 2/500 (0.4%)	Jun. 2004: 291/500 (58.2%)	Nov. 2009: 417/500 (83.4%)
Jun. 1999: 6/500 (1.2%)	Nov. 2004: 294/500 (58.8%)	Jun. 2010: 424/500 (84.8%)
Nov. 1999: 7/500 (1.4%)	Jun. 2005: 304/500 (60.8%)	Nov. 2010: 415/500 (83%)
Jun. 2000: 11/500 (2.2%)	Nov. 2005: 360/500 (72.0%)	Jun. 2011: 411/500 (82.2%)
Nov. 2000: 28/500 (5.6%)	Jun. 2006: 364/500 (72.8%)	Nov. 2011: 410/500 (82.0%)
Jun. 2001: 33/500 (6.6%)	Nov. 2006: 361/500 (72.2%)	
Nov. 2001: 43/500 (8.6%)	Jun. 2007: 373/500 (74.6%)	

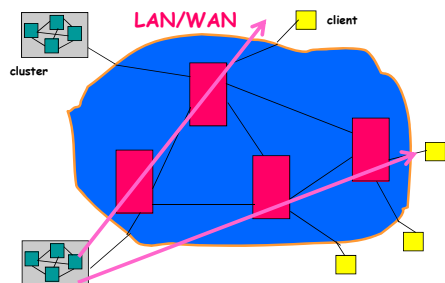
34

Clustered File Systems: Parallel Virtual File System (PVFS)

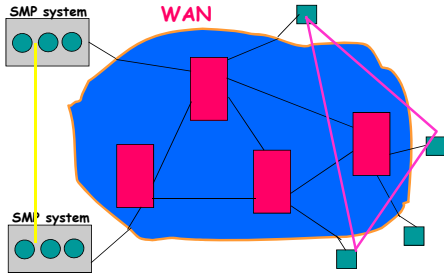


- Relies on Striping of data across different nodes
- Tries to aggregate I/O bandwidth from multiple nodes
- Utilizes the local file system on the I/O Server nodes

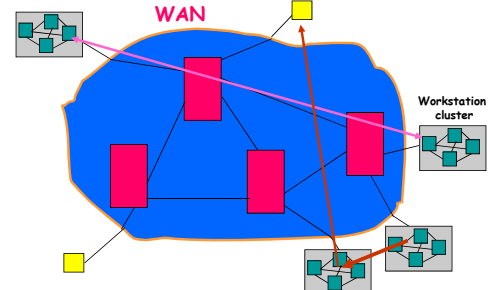
Cluster with Interactive Clients



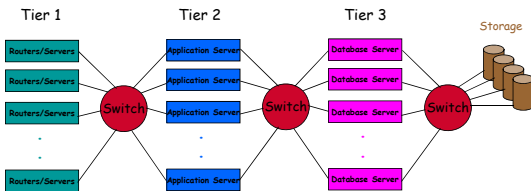
Globally-Interconnected Systems



Interconnected Clusters over WAN

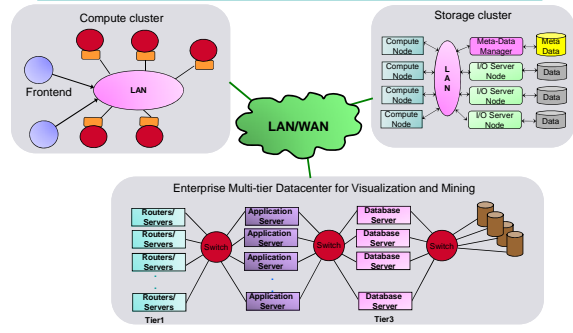


Generic Three-Tier Model for Data Centers



• All major search engines and e-commerce companies are using clusters for multi-tier datacenter
 • Google, Amazon, Financial institutions,

Integrated High-End Computing Environments



Supercomputing '09

40

Presentation Outline

- Motivation
- Technological Trends
- Applications (Past, Current, and Emerging)
- Defining Network-Based Computing Systems
- Challenges and Research Issues
- Research Challenges and Projects
- Available Experimental Testbed
- Related 888 Course
- Conclusions

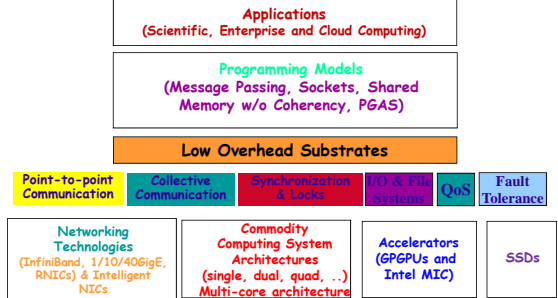
Requirements for Modern Clusters, Cluster-based Servers, and Datacenters

- Requires
 - good Systems Area Network with excellent performance (low latency and high bandwidth) for interprocessor communication (IPC) and I/O
 - good Storage Area Networks high performance I/O
 - good WAN connectivity in addition to intra-cluster SAN/LAN connectivity
 - Quality of Service (QoS) for supporting interactive applications
 - RAS (Reliability, Availability, and Serviceability)

Our Vision

- Network-Based Computing Group to take a lead in
 - Proposing **new designs for high performance NBC systems** by taking advantages of modern networking technologies and computing systems
 - Developing **better middleware/API/programming environments** so that modern NBC applications can be developed and implemented in a scalable fashion
- Carry out research in an integrated manner (systems, networking, and applications)

Research Challenges



44

Major Research Directions

- High Performance MPI for InfiniBand, 10GigE/iWARP and RoCE
- Programming Models (MPI+PGAS (UPC)) for Exascale Systems
- High Performance Computing with GPGPUs and Accelerators
- Networking, Storage and Middleware Support for Cloud Computing and Datacenters
- Scalable and High-Performance File Systems and Storage
- Virtualization (High Performance Computing and Cloud Computing)
- WAN-IB support for Grid and Cloud Computing
- QoS-aware Designs
- Power-aware Designs
- Fault-Tolerant Designs
- More details on <http://nowlab.cse.ohio-state.edu/> → Projects

45

External Collaboration

- Industry
 - IBM TJ Watson and IBM
 - Intel
 - Cisco
 - Dell
 - RNet
 - SUN Microsystems/Oracle
 - Apple
 - Linux Networkx
 - NetApp
 - NVIDIA
 - Mellanox, Qlogic and other InfiniBand companies
 - Chelsio, NetEffect and other iWARP companies
- Research Labs
 - Sandia National Lab
 - Los Alamos National Lab
 - Pacific Northwest National Lab
 - Argonne National Lab
 - Lawrence Livermore National Lab
- Other units in campus
 - Ohio Supercomputer Center (OSC)

46

Research Funding

- National Science Foundation
 - Multiple grants
- Department of Energy
 - Multiple grants
- Sandia National Lab
- Los Alamos National Lab
- Pacific Northwest National Lab
- Ohio Board of Regents
- IBM Research
- Intel
- RNet
- SUN
- Mellanox
- NVIDIA
- Cisco
- Linux Networkx
- Network Appliance

47

Presentation Outline

- Motivation
- Technological Trends
- Applications (Past, Current, and Emerging)
- Defining Network-Based Computing Systems
- Challenges and Research Issues
- Research Challenges and Projects
- Available Experimental Testbed
- Related 888 Course
- Conclusions

NBCLab Clusters



Latest InfiniBand Cluster (Deployed in August '10)



- A new research equipment funding (\$900K) has enabled us to install the following cluster:
 - 1,400-core Intel quad-core Westmere cluster with 32TBytes of disks, connected with InfiniBand QDR (40 Gbps)
 - 8 GPU nodes with latest Nvidia Fermi adapters

50

Wright Center for Innovation (WCI)

- A past research equipment funding (\$600K) has enabled us to install the following two clusters:
 - 576-core Intel quad-core cluster with 32TBytes of disk, connected with InfiniBand DDR (20 Gbps)
 - 128-core AMD quad-core cluster, connected with dual-rails of InfiniBand DDR (20 Gbps)

51

Presentation Outline

- Motivation
- Technological Trends
- Applications (Past, Current, and Emerging)
- Defining Network-Based Computing Systems
- Challenges and Research Issues
- Research Challenges and Projects
- Available Experimental Testbed
- [Related 888 Course](#)
- Conclusions

888.08P course


- Research Seminar on Network-Based Computing
- Offered every quarter
- Presentation and Project-based course
- 3-credits
- [Two-quarter rule](#)

Presentation Outline

- Motivation
- Technological Trends
- Applications (Past, Current, and Emerging)
- Defining Network-Based Computing Systems
- Challenges and Research Issues
- Research Challenges and Projects
- Available Experimental Testbed
- Related 888 Course
- [Conclusions](#)



Conclusions

- Network-Based computing is an emerging trend
 - Medium to large-scale network-based computing systems are getting ready to be used
 - Many open research issues need to be solved before building ultra-scale network-based computing systems
 - Will be able to provide affordable, portable, scalable, and ..able computing paradigm for the next 10-15 years
 - Exciting area of research
- 
- 