SE ZG522: Big Data Systems

03 Parallel and Distributed Sys

Dr. Kamlesh Tiwari
Assistant Professor, Department of CSIS,
BITS Pilani, Pilani Campus, Rajasthan-333031 INDIA

Feb 12, 2021 ONLINE (WILP @ BITS-Pilani Jan-May 2022)

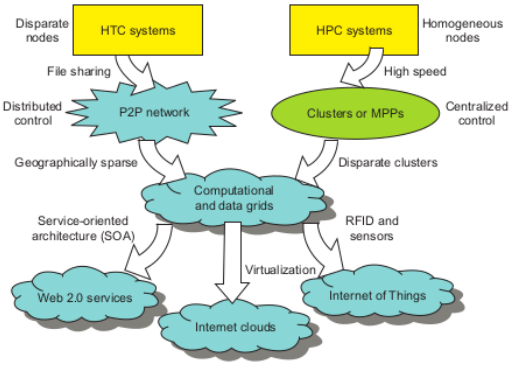
http://ktiwari.in/bds

Computing Paradigm

- **Centralized Computing**
  All computer resources are centralized in one physical system.
- **Parallel Computing**
  All processors are either tightly coupled with centralized shared memory or loosely coupled with distributed memory.
- **Distributed Computing**
  Multiple autonomous computers, each having its own private memory. Communicating is accomplished through message passing.
- **Cloud Computing**
  A form of utility computing or service computing. Can either be centralized or distributed, Parallel/distributed or both. Can be built with physical or virtualized resources.

Moore's Law: Processor speed doubles every 18 month

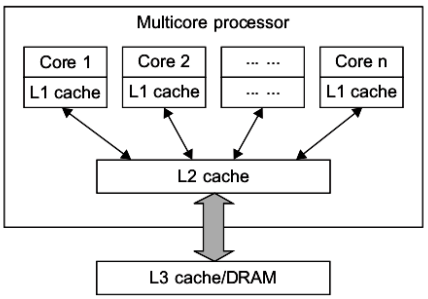
Trends



High Throughput Computing, High Performence Computing, massively parallel processors (MPP)

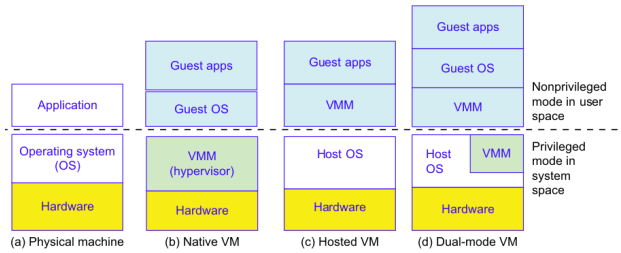
Multicore CPUs

- Clock rate reached its limit on CMOS-based chips due to power limitations



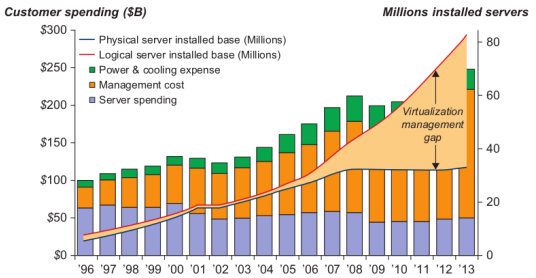
Many-Core GPU, and Multithreading

Virtual Machines



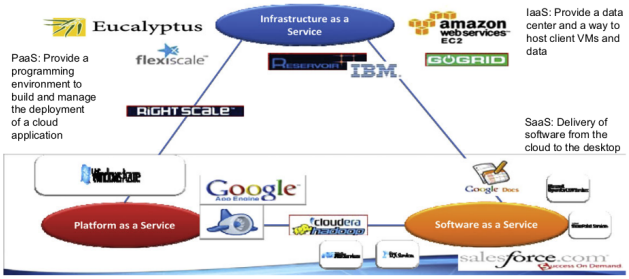
- The VM is built with virtual resources managed by a guest OS to run a specific application.

Growth and cost breakdown of data centers



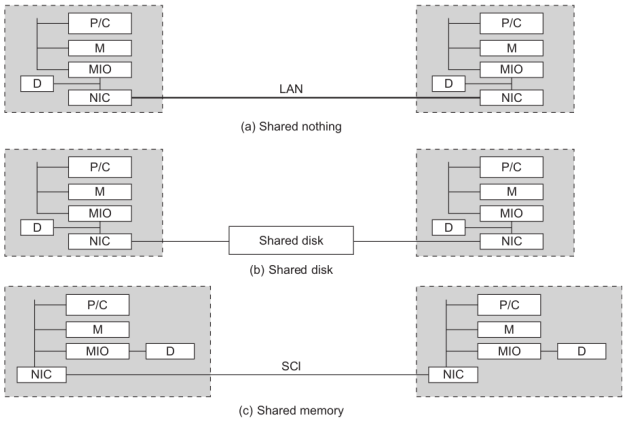
- IDC report 2009: only 30% of data center costs goes to purchase equipment. 33%: chiller, 18%: power supply (UPS), 9%: room air conditioning (CRAC), and 7% power distribution, lighting, and transformer costs.

Cloud Landscape



- **Infrastructure as a Service (IaaS)** servers, storage, networks, data center fabric. User deploy and run on multiple VMs using guest OSes
- **Platform as a Service (PaaS)** User deploy applications. Includes middleware, databases, development tools, Web 2.0 and Java.
- **Software as a Service (SaaS)** Browser-initiated application software over thousands of paid cloud customers.

ways to connect cluster nodes



ACID

- Atomicity: Transactions are guaranteed to either completely occur, or have no effects.
- Consistency: Always transfers from one consistent state to another.
- Isolation: Result not revealed to other transactions until committed.
- Durability: Once committed, the effect of transition persists even if system fails.

Client server and peer to peer network?

Thank You!

Thank you very much for your attention!

Queries ?