

Data Intensive Distributed Computing

Ioan Raicu

**Center for Ultra-scale Computing and Information Security
Department of Electrical Engineering & Computer Science
Northwestern University**

EECS 395 / EECS 495

**Hot Topics in Distributed Systems: Data-Intensive Computing
January 7th, 2010**

Famous Quotes

The users should be able to focus their attention on the information content of the data, rather than how to discover, access, and use it.

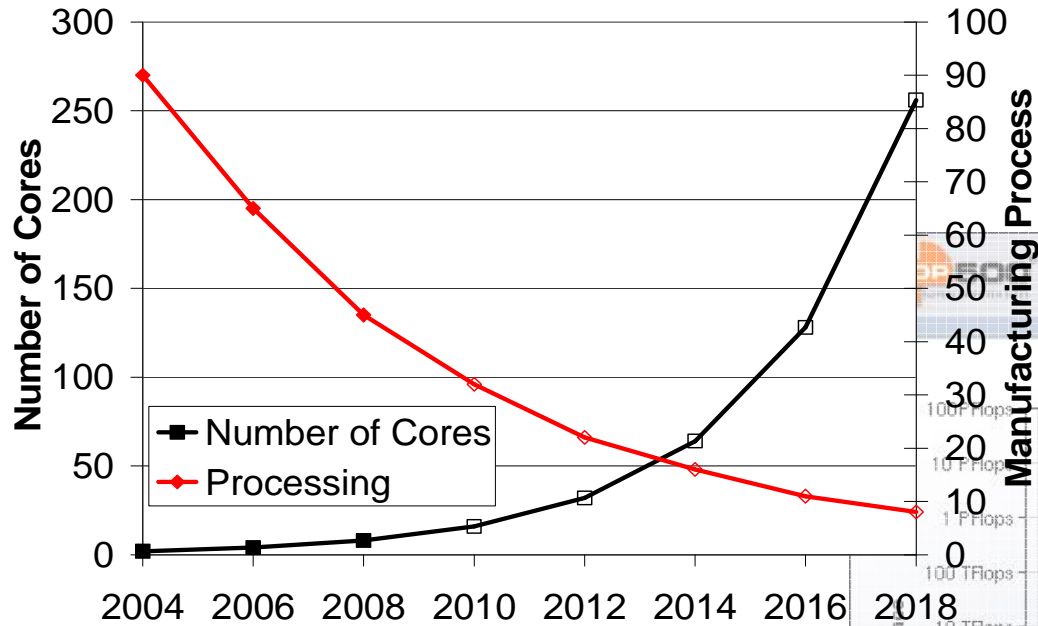
Climate Change Science Program report, 2003

Famous Quotes

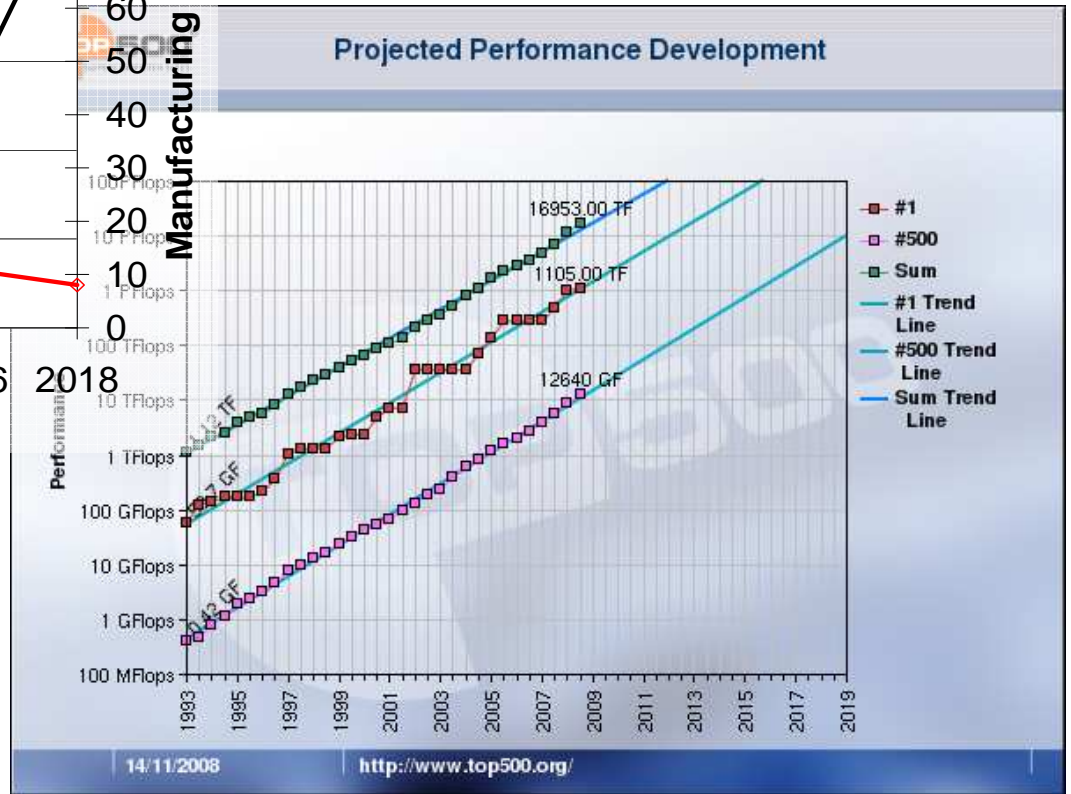
A supercomputer is a device for turning compute-bound problems into I/O-bound problems.

Seymour Cray

Projected Growth Trends



Pat Helland, Microsoft, The Irresistible Forces Meet the Movable Objects, November 9th, 2007



Top500 Projected Development,

http://www.top500.org/lists/2008/11/performance_development

Growing Storage/Compute Gap

- Local Disk:

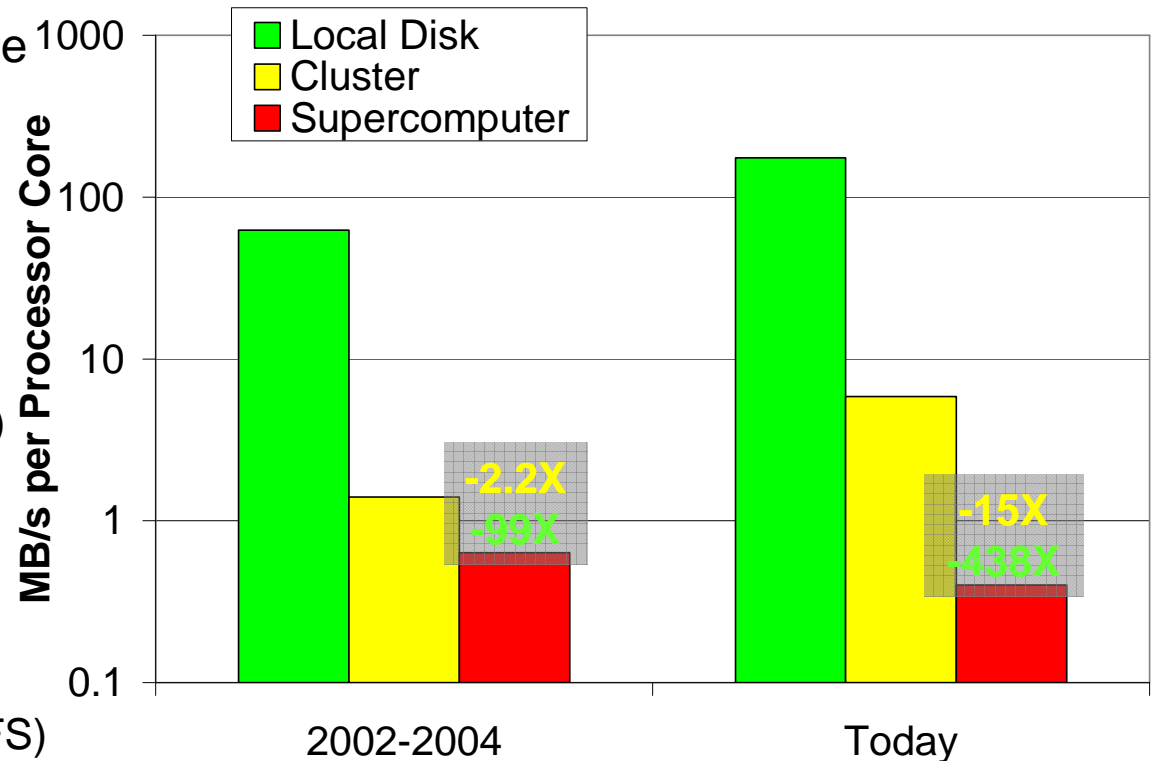
- 2002-2004: ANL/UC TG Site (70GB SCSI)
- Today: PADS (RAID-0, 6 drives 750GB SATA)

- Cluster:

- 2002-2004: ANL/UC TG Site (GPFS, 8 servers, 1Gb/s each)
- Today: PADS (GPFS, SAN)

- Supercomputer:

- 2002-2004: IBM Blue Gene/L (GPFS)
- Today: IBM Blue Gene/P (GPFS)



State of the Art: Storage Systems

- Segregated storage and compute
 - NFS, GPFS, PVFS, Lustre
 - Batch-scheduled systems: Clusters, Grids, and Supercomputers
 - Programming paradigm: HPC, MTC, and HTC
- Co-located storage and compute
 - HDFS, GFS
 - Data centers at Google, Yahoo, and others
 - Programming paradigm: MapReduce
 - Others from academia: Sector, MosaStore, Chirp

State of the Art: Storage Systems

- Segregated storage and compute

- NFS, GPFS, PVFS, Lustre

- Batch-scheduled
Supercomputers

- Programming pa

- Located storage

- Data centers at

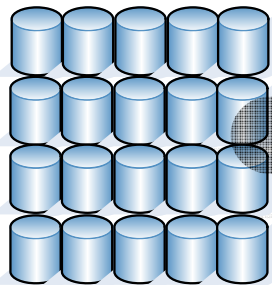
- Programming pa

- Others from aca

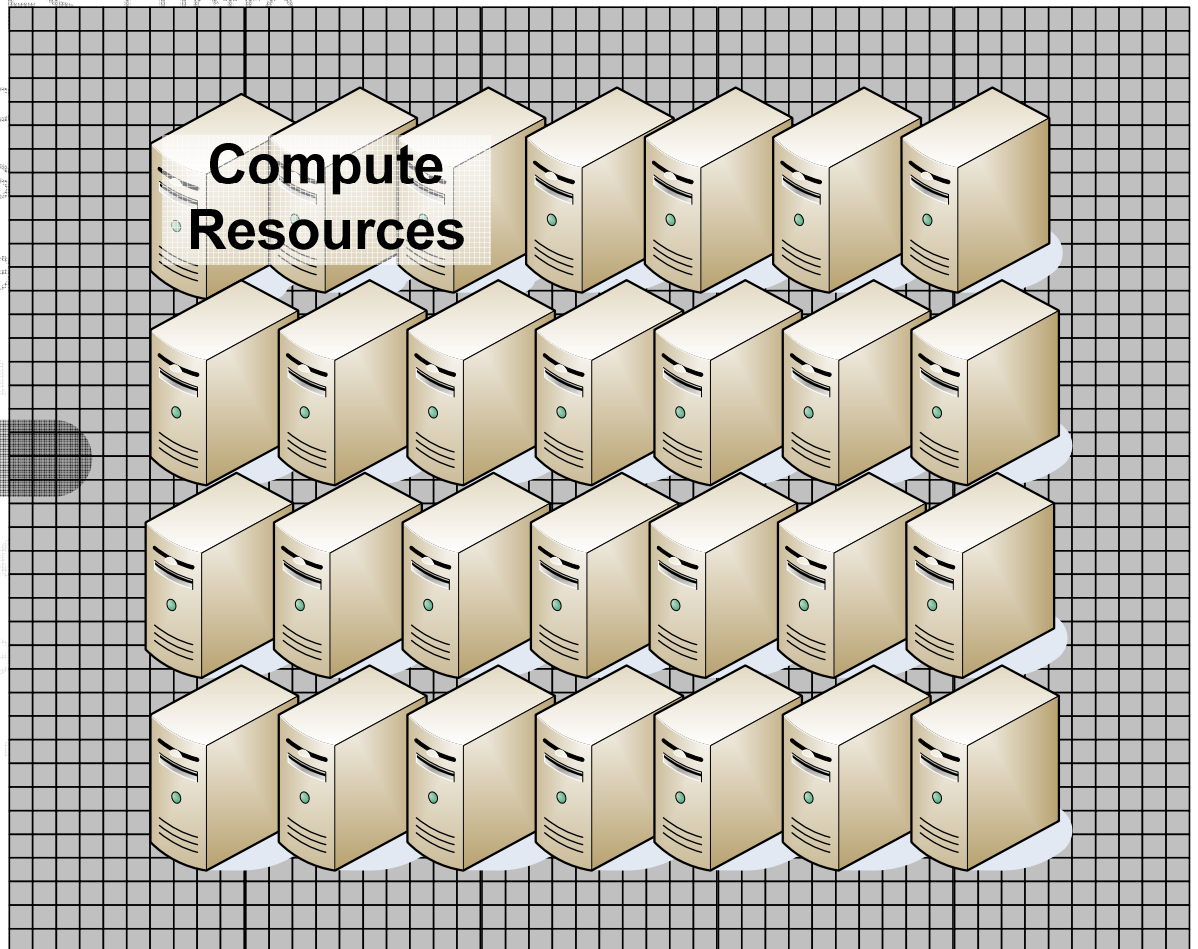
**Network
Fabric**

**Compute
Resources**

NAS



Network Link(s)

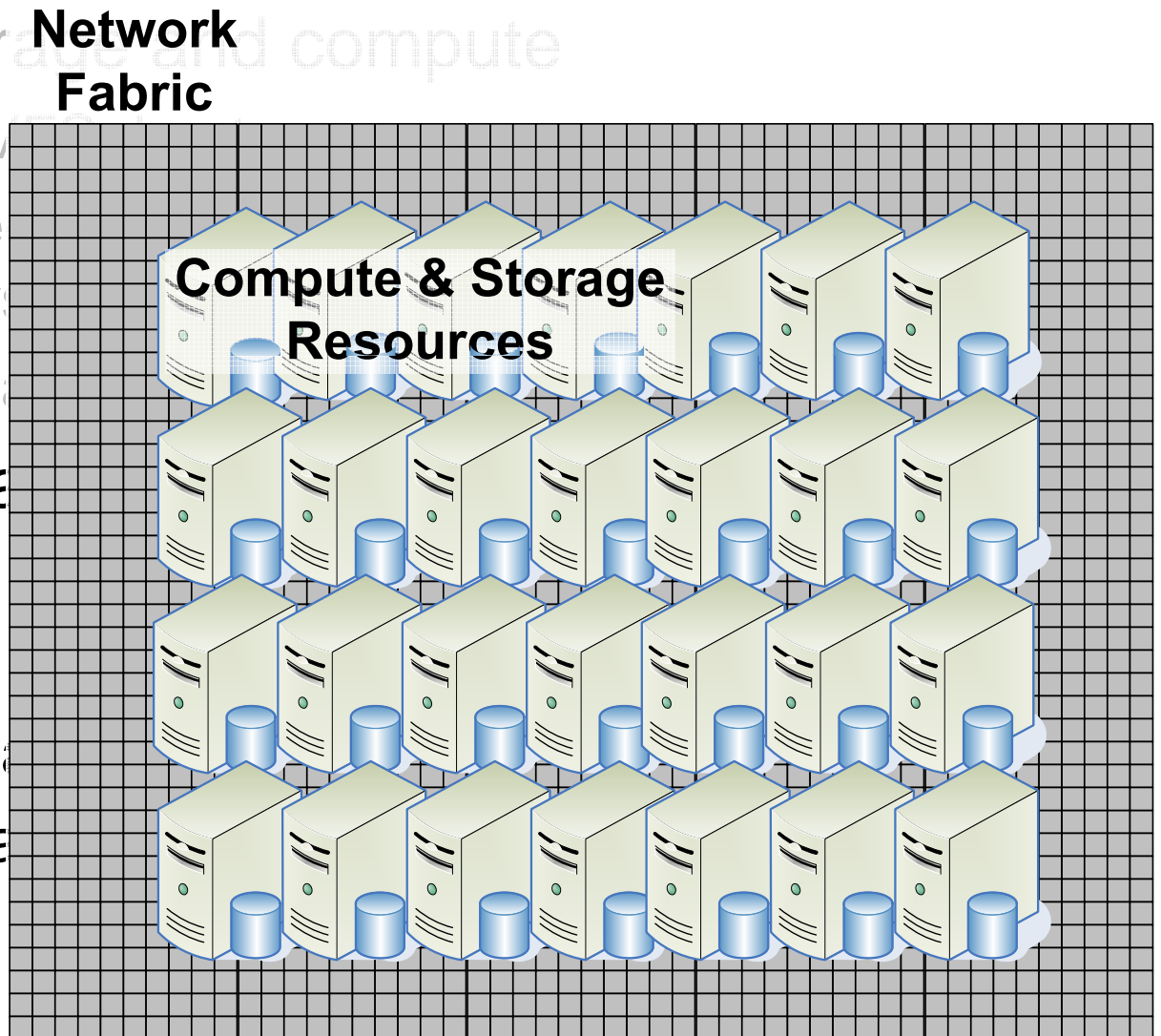


State of the Art: Storage Systems

- Segregated storage and compute
 - NFS, GPFS, PVFS, Lustre
 - Batch-scheduled systems: Clusters, Grids, and Supercomputers
 - Programming paradigm: HPC, MTC, and HTC
- Co-located storage and compute
 - HDFS, GFS
 - Data centers at Google, Yahoo, and others
 - Programming paradigm: MapReduce
 - Others from academia: Sector, MosaStore, Chirp

State of the Art: Storage Systems

- Segregated storage and compute
 - NFS, GPFS, PV
 - Batch-schedule Supercomputers
 - Programming p
- Co-located storage
 - HDFS, GFS
 - Data centers at
 - Programming p
 - Others from acc

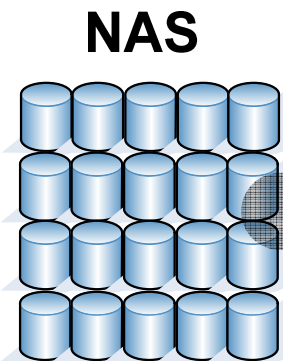


Question

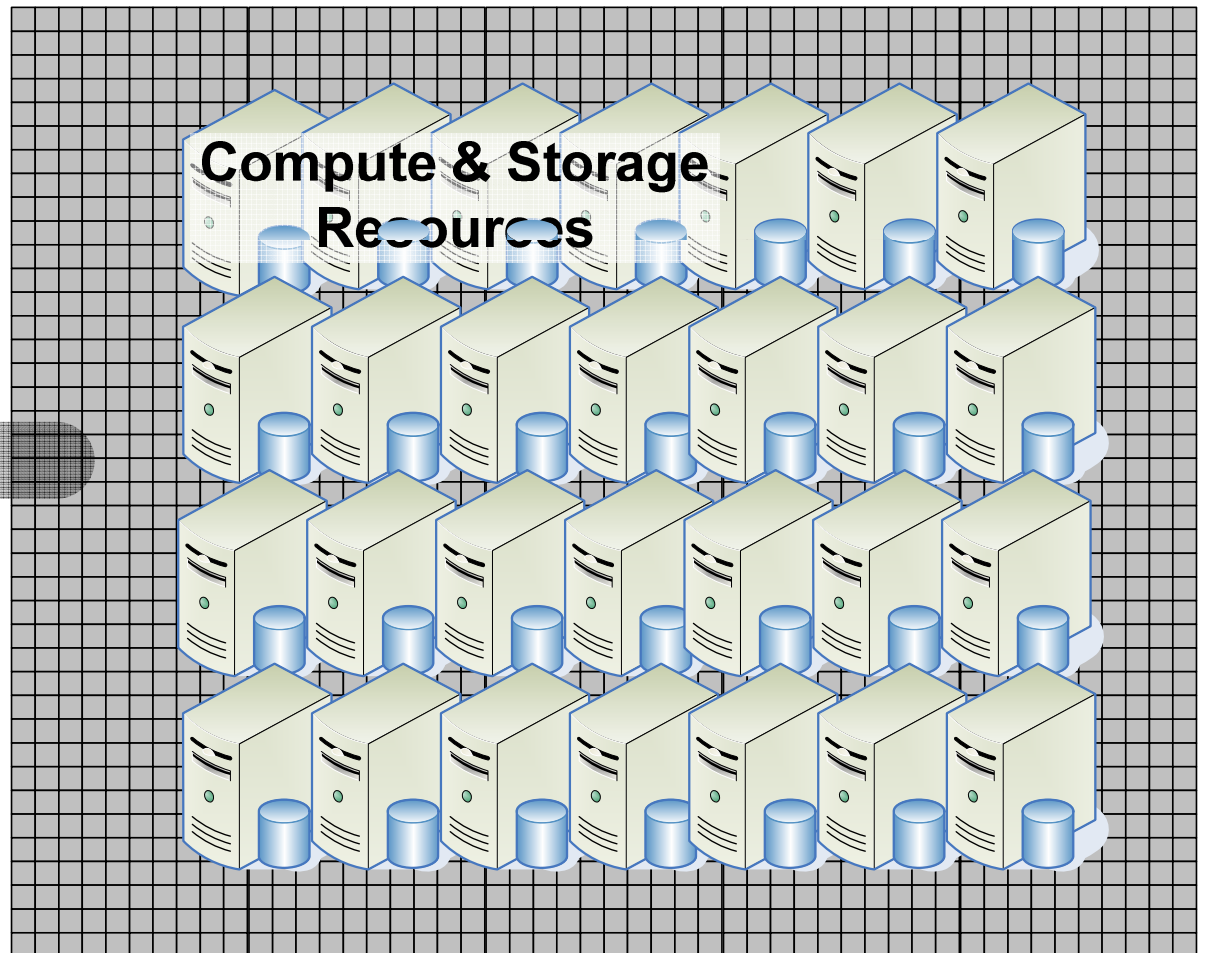
What if we could combine the scientific community's existing programming paradigms, but yet still exploit the data locality that naturally occurs in scientific workloads?

Combine State of the Art Systems

**Network
Fabric**



Network Link(s)

A horizontal grey arrow pointing from the NAS to the network fabric, labeled "Network Link(s)".

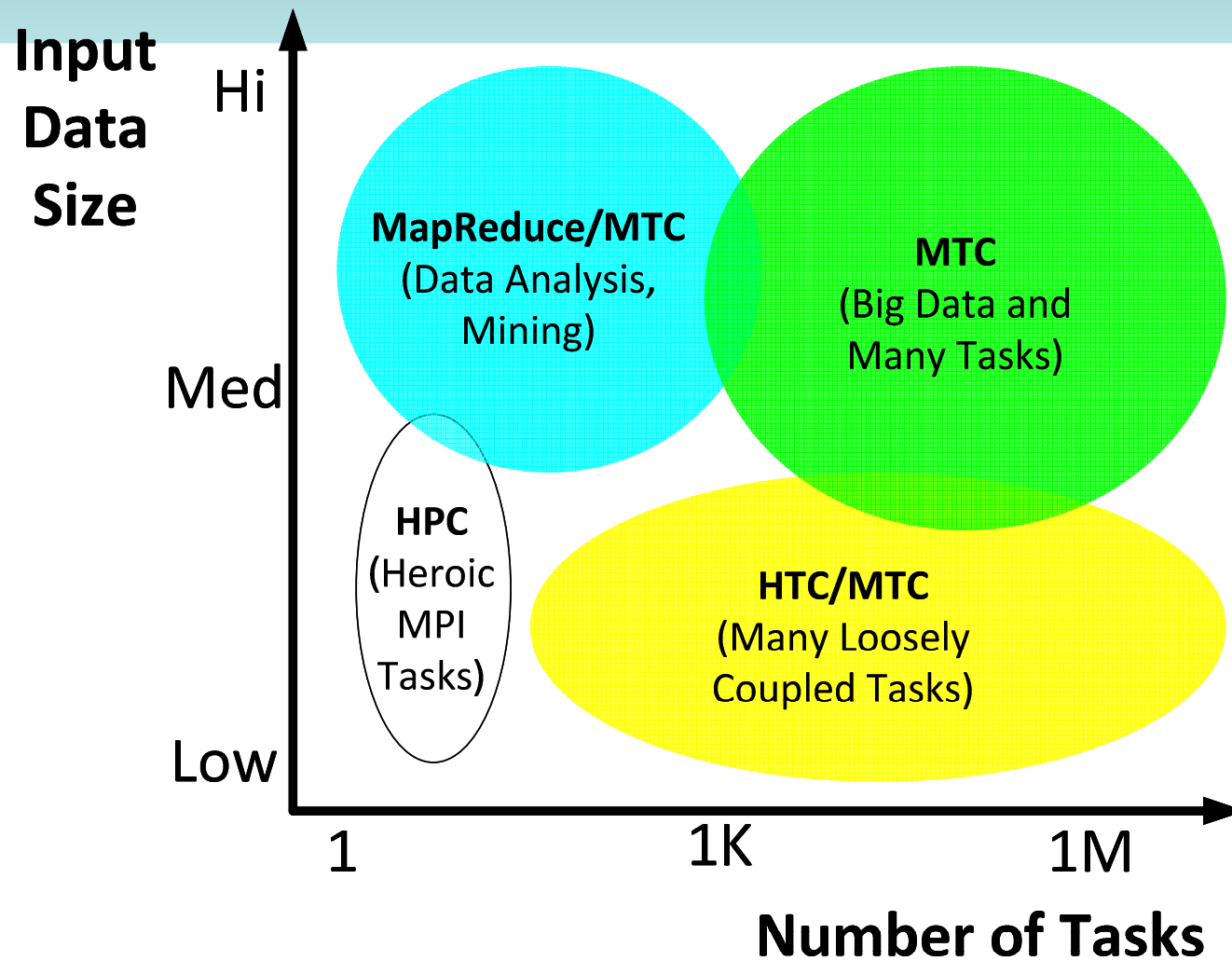
High-Throughput Computing & High-Performance Computing

- **HTC: High-Throughput Computing**
 - Typically applied in clusters and grids
 - Loosely-coupled applications with sequential jobs
 - Large amounts of computing for long periods of times
 - Measured in operations per month or years
- **HPC: High-Performance Computing**
 - Synonymous with supercomputing
 - Tightly-coupled applications
 - Implemented using Message Passing Interface (MPI)
 - Large of amounts of computing for short periods of time
 - Usually requires low latency interconnects
 - Measured in FLOPS

MTC: Many-Task Computing

- Bridge the gap between HPC and HTC
- Applied in clusters, grids, and supercomputers
- Loosely coupled apps with HPC orientations
- Many activities coupled by file system ops
- Many resources over short time periods
 - Large number of tasks, large quantity of computing, and large volumes of data

Problem Space



Questions

