The Structure of a Cluster Supercomputer

Hussein Al-Azzawi
UNM Center for Advanced Research Computing
06/04/2019

CENTER FOR ADVANCED RESEARCH COMPUTING

What’s on the agenda

Intro

What is a Supercomputer

An example

But, Why?

HPC Challenges..

Simulation

Paralasim

Cluster Components

Design, briefly

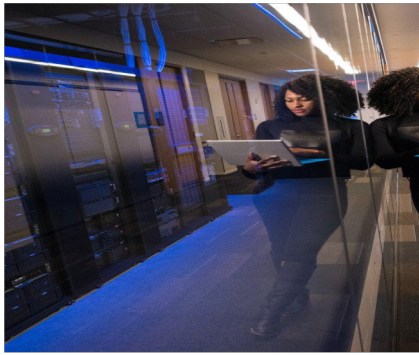
Performance

New Technologies

XSEDE

CENTER FOR ADVANCED RESEARCH COMPUTING

- Biggest public HPC Center in the State of New Mexico
- In 2000, UNM, in collaboration with IBM, built the first-ever Linux beowulf cluster in production
- Main research fields:
  - Nuclear
  - Materials science
  - Astrophysics
  - Genome sequencing



What is a Supercomputer

A computer with very high-levels computational capacities compared to general-purpose computers such as a personal desktop, laptop, cellphone.

Summit in its natural habitat - America's Fastest and Smartest HPC Cluster

image courtesy of Oak Ridge National Laboratory website

But, Why?

1. To resolve tomorrow’s problems, which can’t be studied using today’s infrastructure, by using computer simulations
2. Cluster HPC helps bypassing the limitations of Moore's law
3. Increase the research outcomes by an order of magnitude by manipulating excessive data points and machine learning
4. Solve expensive and dangerous problems in simulations

HPC Challenges..

Staffing

$$$

Power, space, and cooling

image courtesy of medium.com

Computer Simulation

A computer simulation is a program that attempts to reproduce an **abstract model** of a particular system in the **virtual space** before building it in **reality**.

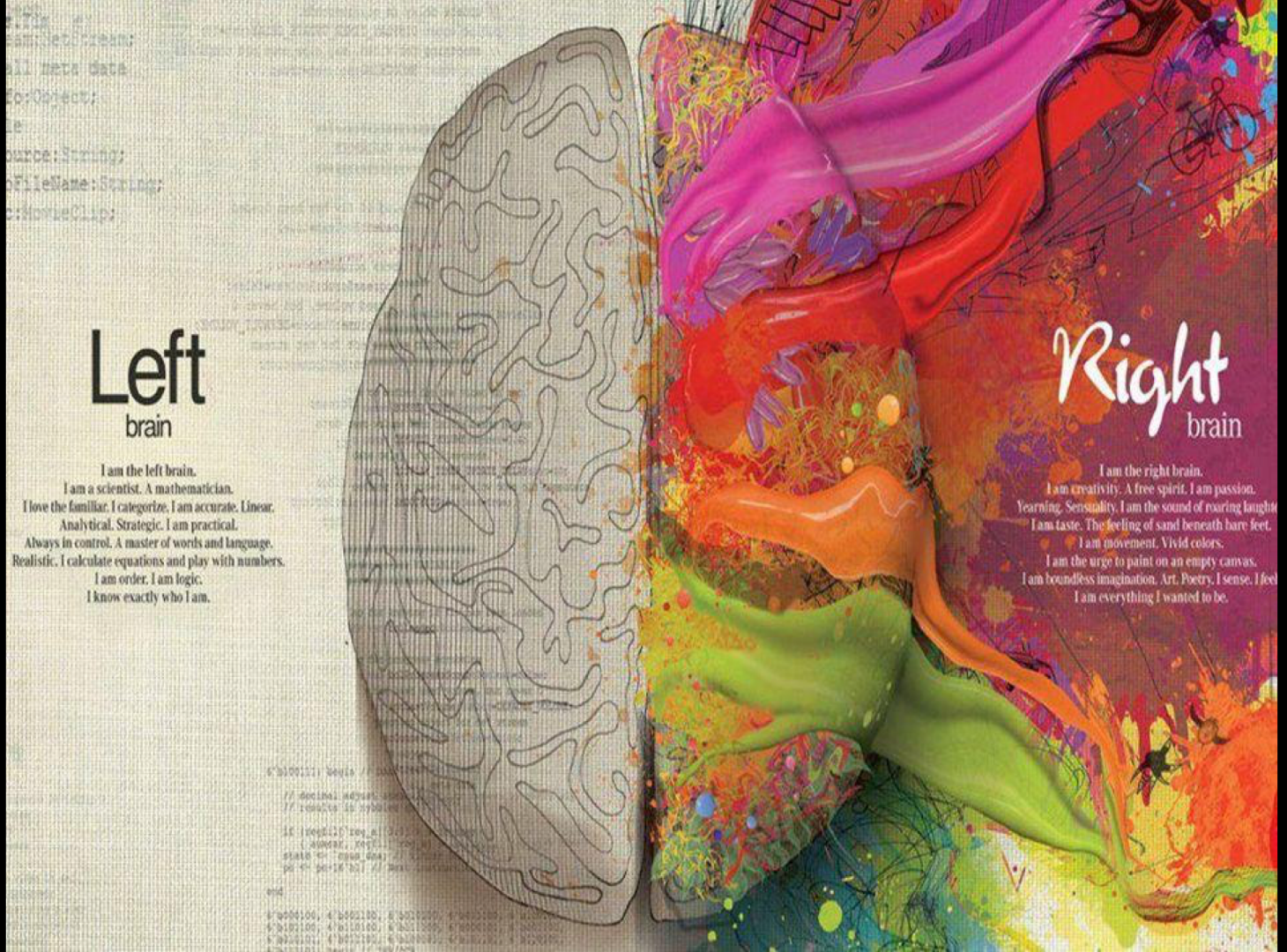
Parallelism

Performing computations in **parallel** by carrying out many calculations **simultaneously**. Utilizing thousands of general-purpose computers all working together to solve the same **problem** at the **same time**. This in fact is an excellent case for how **modern** supercomputers work.

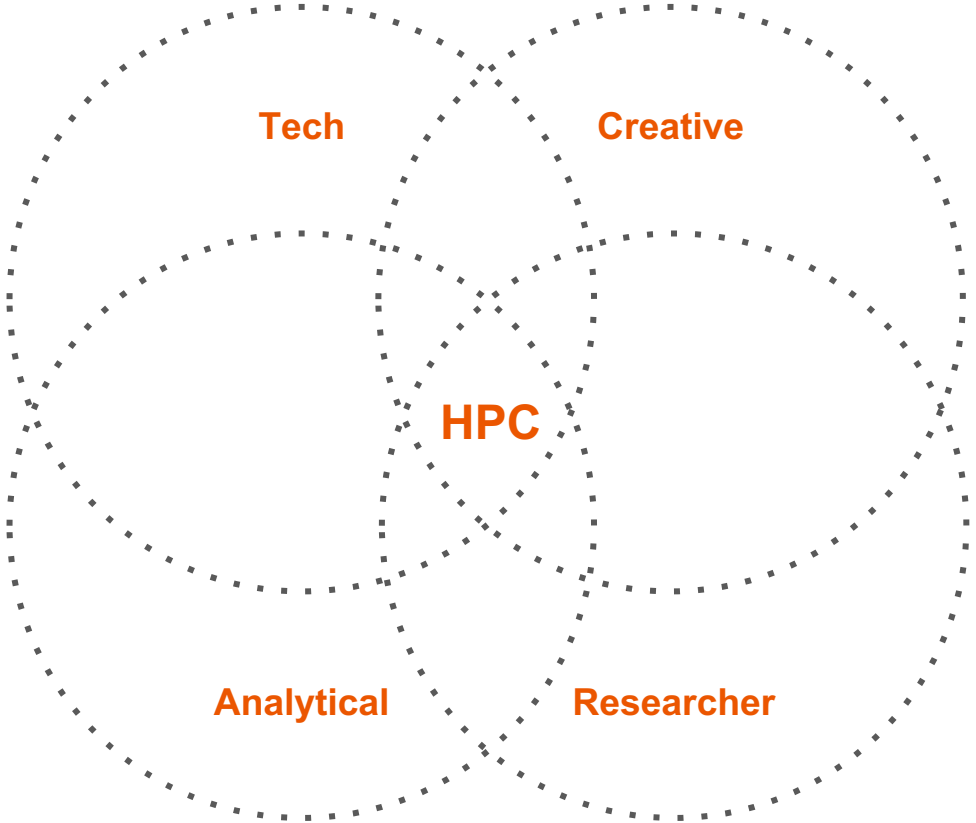
Parallelism

Multithreading vs Multiprocessing

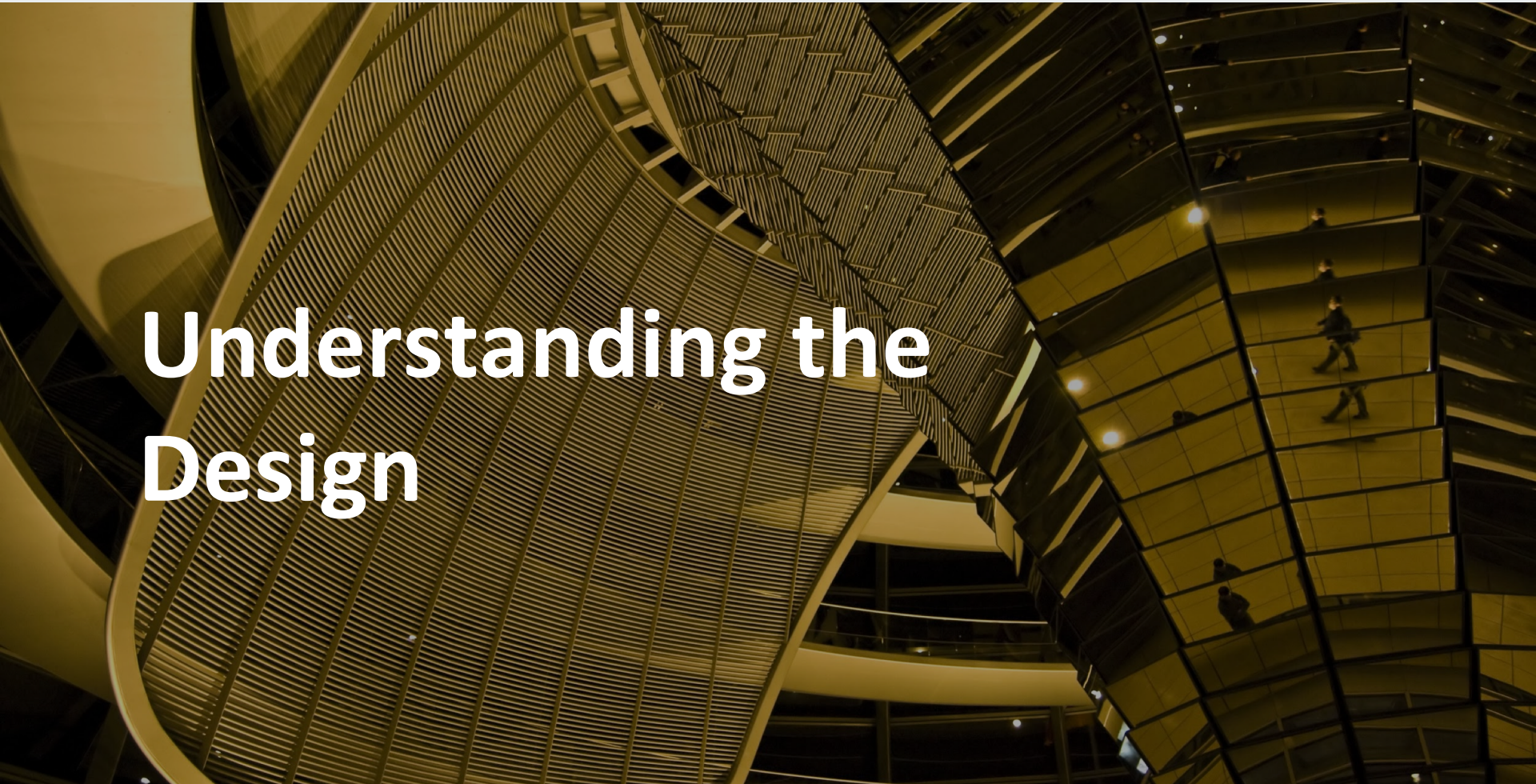
- Shared Memory Parallelism is concerned with threads
- Distributed Parallelism is concerned with processes.


Left
brain
I am the left brain.
I am a scientist. A mathematician.
I love the familiar. I categorize. I am accurate. Linear.
Analytical. Strategic. I am practical.
Always in control. A master of words and language.
Realistic. I calculate equations and play with numbers.
I am order. I am logic.
I know exactly who I am.
Right
brain
I am the right brain.
I am creativity. A free spirit. I am passion.
I am taste. The feeling of sand beneath bare feet.
I am movement. Vivid colors.
I am the urge to paint on an empty canvas.
I am everything I wanted to be.

Point of View

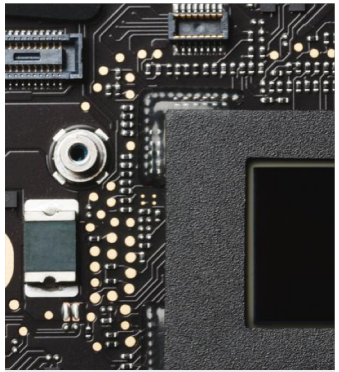


Computational Science vs Computer Science

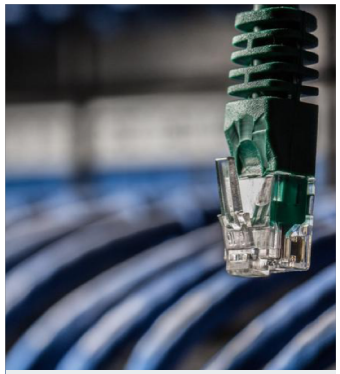


Understanding the Design

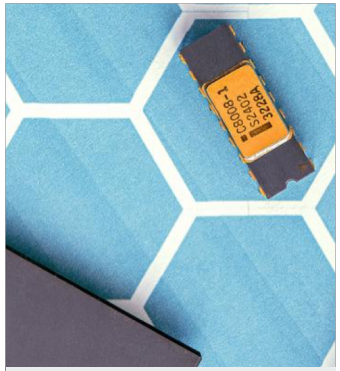
Cluster Components



Compute

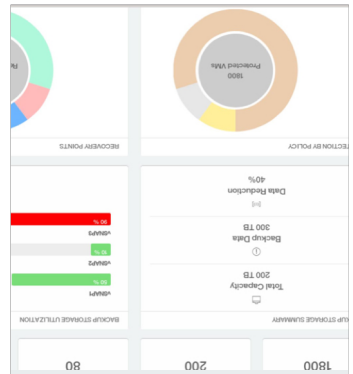


Network



Storage

Batch Scheduler



HPC SW

Compute

- CPU
- GPU
- Memory
- Interconnects
- Diskful/Diskless

Network Interconnects

Main factors:

- Latency
- Bandwidth

Types

- InfiniBand
  - Low latency (~1.5 µsec)
  - Cluster interconnect (100G, 56G, 40G)
- Ethernet
  - High Latency (~12 µsec)
  - Data management (1G)
  - Communications (10G)

image courtesy of Mellanox Store website

Storage

Storage Usage

- Home directories
- Scratch space
- Project space

Storage Types

- Traditional file systems
- Parallel file systems

Resource Manager/ Scheduler

Resource Manager oversees the available resources of the cluster

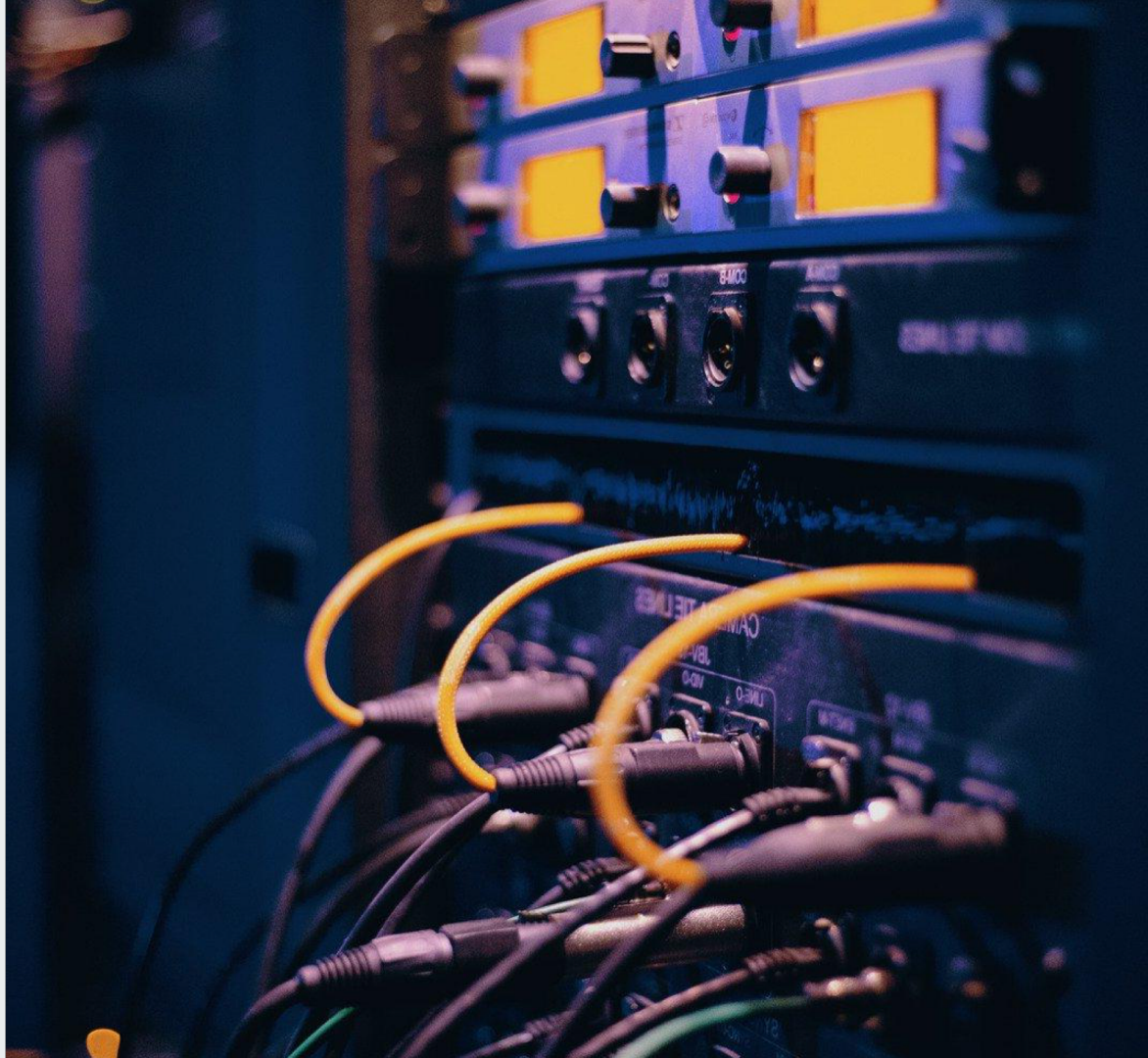
Scheduler determines what runs where..

Examples:

Slurm, Torque, Moab..

HPC Software

- General usage
- Field dependant



Other HPC Components

- Benchmarks
- Security and logs
- Monitoring utilities
- Backup systems
- Cooling systems
- Backup power
- Licensing

Traditional HPC Cluster

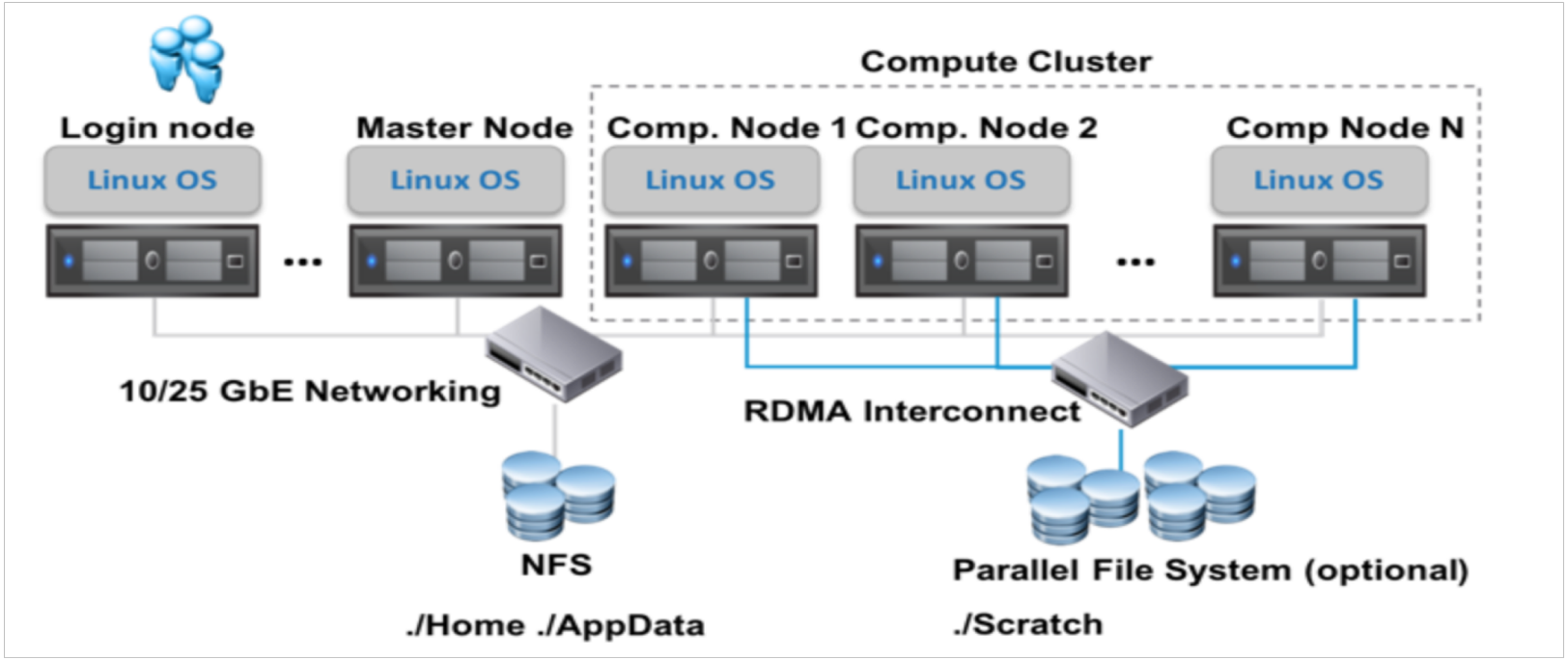


image courtesy of VMware Blogs website

Moab HPC Cluster

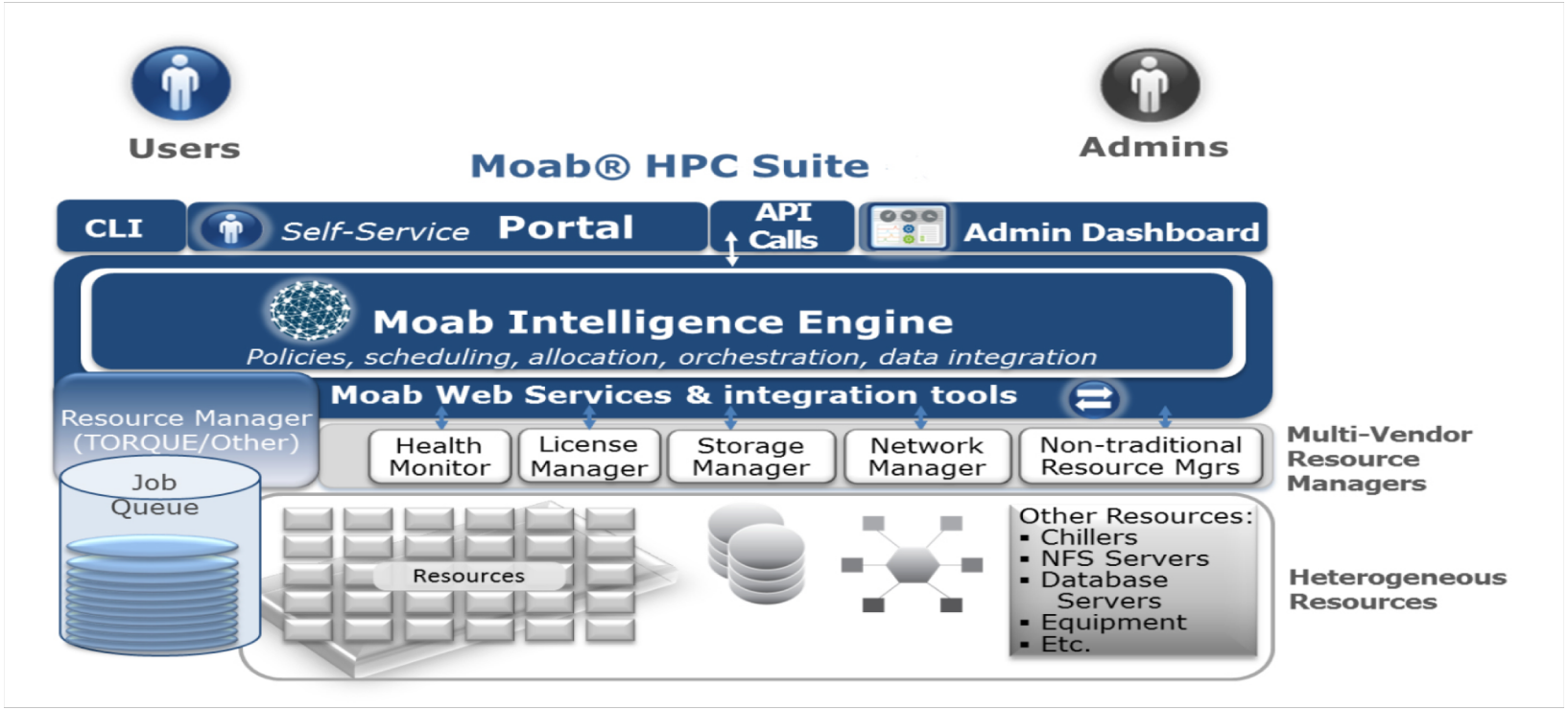


image courtesy of Adaptive Computing website

Virtualized HPC Cluster

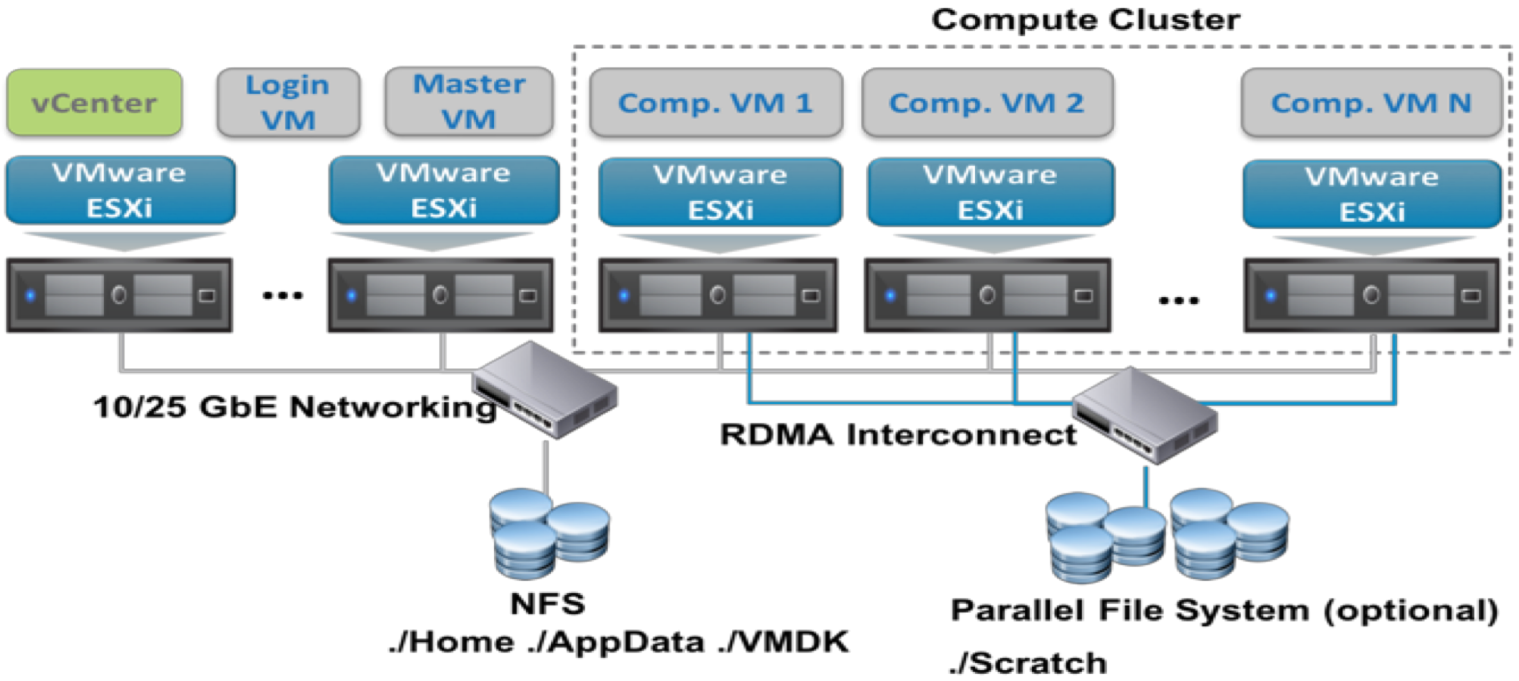


image courtesy of VMware Blogs website

How long is a trillion seconds?

Computers store numbers in floating-point format, so they are called floating-point numbers.

A single instruction like addition or multiplication is called an operation,

HPC speed is measured in terms of floating-point operations per second or Flop/s (Flops)

Rmax - Maximal LINPACK performance achieved

Rpeak - Theoretical peak performance

Ops per second	Scientific Notation	Metric Prefix	Unit
1 000	10^3	Kilo	Kflops
1 000 000	10^6	Mega	Mflops
1 000 000 000	10^9	Giga	Gflops
1 000 000 000 000	10^{12}	Tera	Tflops
1 000 000 000 000 000	10^{15}	Peta	Pflops
1 000 000 000 000 000 000	10^{18}	Exa	Eflops

Performance

CPU clock frequency in GHz
Clock speed in cycles per second (Hertz)
# of floating-point operations per cycle

Example:
2.6 GHz processor
2.6 billion cycles per second
16 floating-point operations per cycle
16 x 2.6 billion = 41.6 Billion Flops
41.6 GFlops

Top 500 History - Peak Speed (Rmax)

1993
Loremujitsu Numerical Wind Tunnel 124.50 GFLOPS
National Aerospace Laboratory, Tokyo, Japan

1997
Intel ASCI Red/9152 1.338 TFLOPS
DoE-Sandia National Laboratories, New Mexico, USA

2008
IBM Roadrunner 1.026 PFLOPS
DoE-Los Alamos National Laboratories, New Mexico, USA

2018
IBM Summit 122.3 PFLOPS
DoE-Oak Ridge National Laboratory, Tennessee, USA

20XX
TBD - First EFLOPS
I'm certain folks attending this conference are working on this right now!

HPC In The Cloud

Pros

- Cost
- Easy to build
- Cloud bursting
- Huge business investments

Cons

- Cost
- Compliance
- Vendor specific
- Harder to track on the institutional level

Quantum Computers

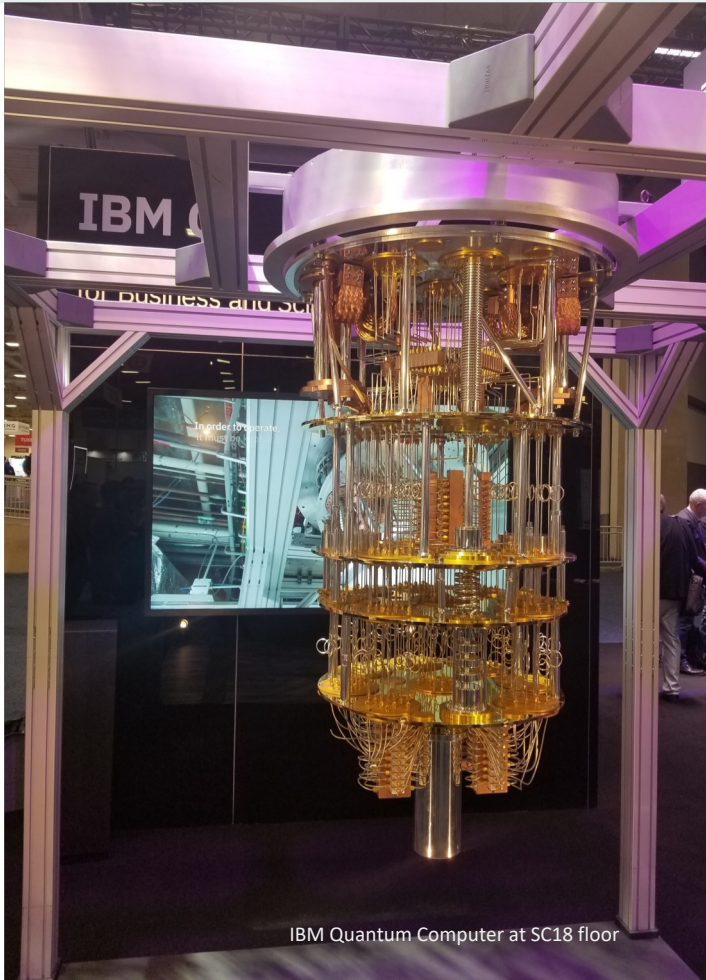
Transistors are already at the atom size,, what's next!

The use of quantum-mechanical phenomena such as superposition and entanglement to perform computations.

Qubits vs Binary bits

- 1 or 0 quantum state
- 1 and 0 superposition states

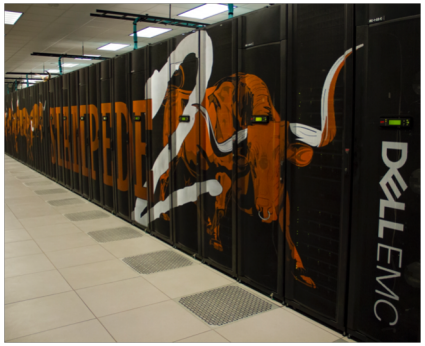
Using quantum algorithms to solve specific complex problems beyond the capabilities of a classical HPC



IBM Quantum Computer at SC18 floor

XSEDE HPC Systems

01 | Stampede2 | TACC
02 | Comet | SDSC
03 | Bridges | PSC
04 | Jetstream | Indian U
05 | Open Science Grid | OSG
06 | Ranch | UT Austin



XSEDE

Extreme Science and Engineering Discovery Environment

Thank You!

Hussein Al-Azzawi
UNM Center for Advanced Research Computing
linkedin.com/in/hazzawi
06/04/2019

CENTER FOR ADVANCED RESEARCH COMPUTING