A Few Thoughts on more than 40 Years in HPC

Arthur S. Bland

September 25, 2020

ORNL is managed by UT-Battelle LLC for the US Department of Energy
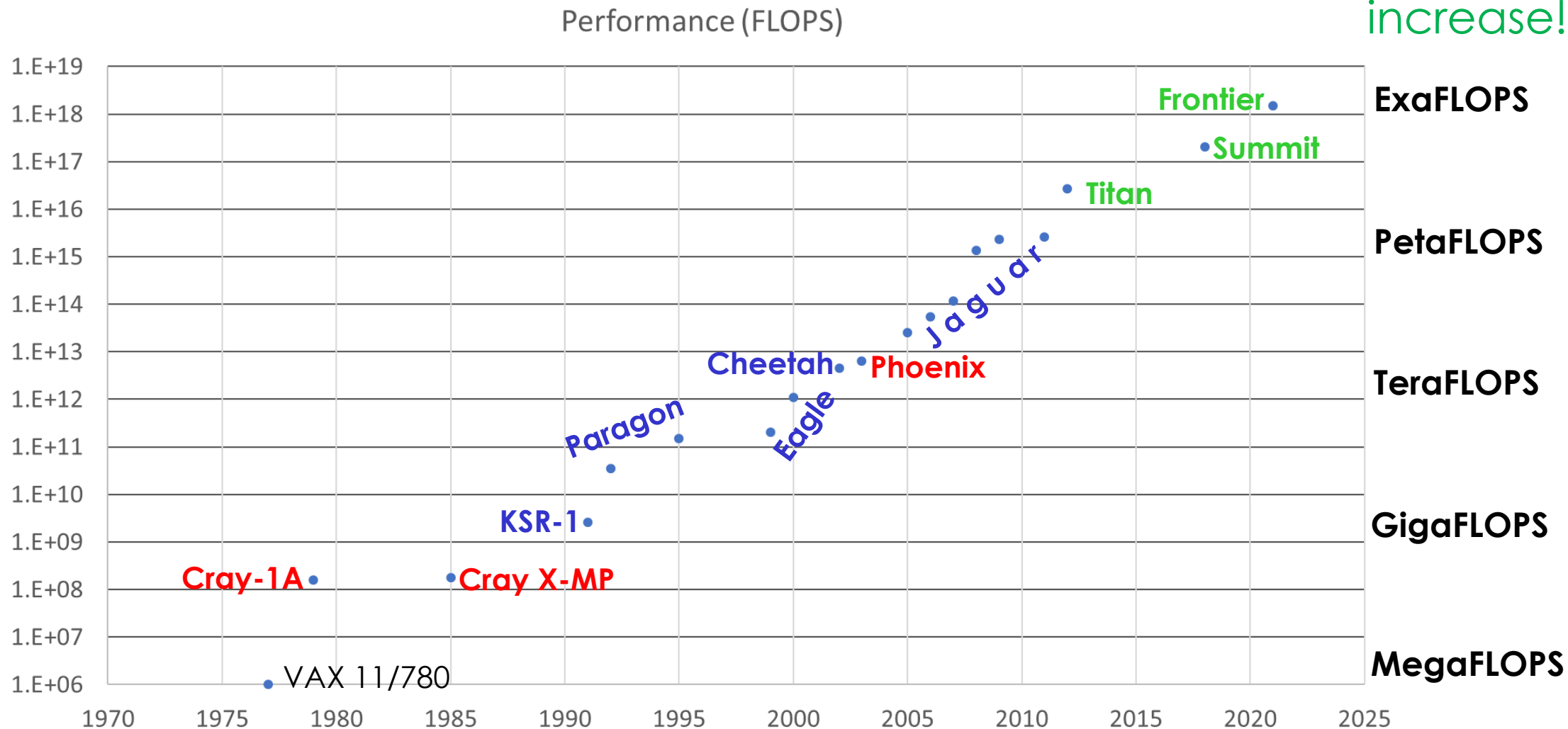
The fine print

- Predictions are difficult and often wrong. This is especially true with technology. I have a "Magic 8 Ball™" that often says:

- The statements in this presentation do not necessarily represent Oak Ridge National Laboratory or the Department of Energy.

Reply Hazy, Try Again

OAK RIDGE | LEADERSHIP COMPUTING FACILITY
National Laboratory

40+ years of exponential performance growth in HPC hardware

1 Trillion times increase!

Performance (FLOPS)

ExaFLOPS
PetaFLOPS
TeraFLOPS
GigaFLOPS
MegaFLOPS

Frontier
Summit
Titan
Jaguar
Cheetah   Phoenix
Paragon
Eagle
KSR-1
Cray-1A   Cray X-MP
VAX 11/780

Minicomputer   **Vector**   **MPP CPU**   **MPP CPU-GPU**

Software improvements account for a large part of the performance improvements of applications

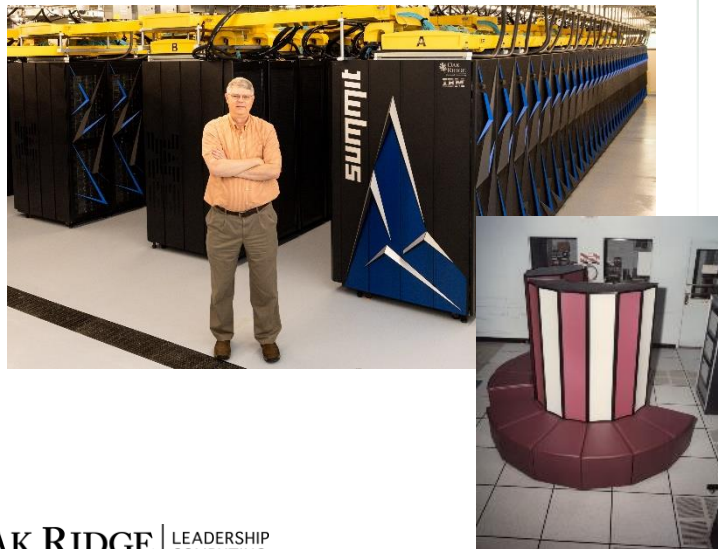| Mathematical Methods	Solver Libraries	Solver Libraries and Tools

• Differential Equations
  – Time discretization methods
• Partial Differential Equations
  – Hyperbolic conservation laws
  – Incompressible and low-Mach number fluid dynamics
  – Adaptive mesh refinement
  – Complex geometries and fronts
  – Fast integral-equation-based methods
  – Compatible discretization methods for PDEs
	• GEAR
	• DASSL
	• LINPACK
	• EISPACK
	• BLAS
	• LAPACK & ScaLAPACK
	• SuperLU
	• PETSc
	• HYPER & ML
		• ARPACK
		• MINPACK
		• TAO
		• NEOS
		• DAKOTA
		• Trilinos
		• DDT
		• FORGE
		• Compilers

Investments in software tools and applications through SciDAC and ECP have been critical

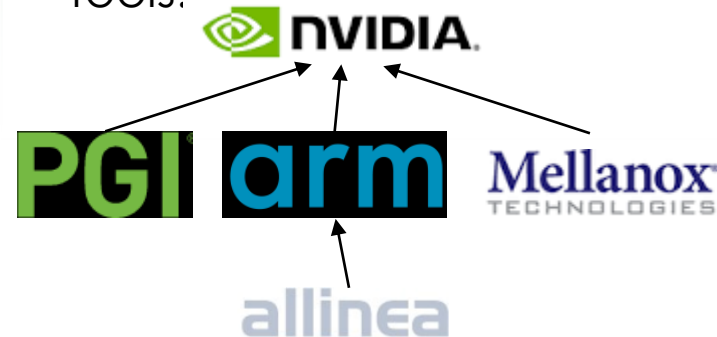# System Complexity has grown tremendously

Hardware	Software	Facilities
• The Cray-1 had approx. **200,000 gates**. It used 70 square feet of floor space. • On Summit, each POWER9 has **8 billion** transistors and each V100 GPU has **21 billion**. Summit uses 8,000 square feet.	• Hybrid systems are much harder to program than traditional CPU only systems. • Larger companies are purchasing independent software and hardware companies, stranding us without important technology such as compilers, debuggers, interconnects and other tools.	• Leadership systems require more power and cooling than systems from the past. Preparing a facility for an exascale system takes time and money. In just the last 17 years we have gone from 2 MW of power and cooling to 100 MW of power and cooling.

OAK RIDGE
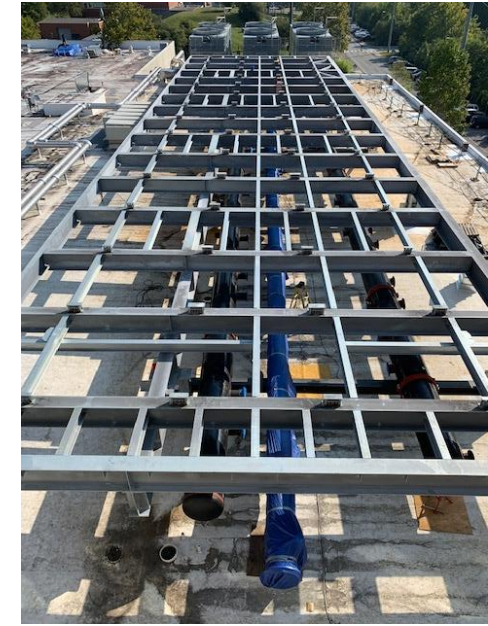National Laboratory | LEADERSHIP COMPUTING FACILITY

Facilities

Cooling Water Pipes

Crane to lift cooling towers to the roof

Cooling tower platform

Piping with cooling towers

Rebuilding the 5600 computer room

Disk cabinet dolly

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

Leadership computing requirements are outpacing the commercial market - What are the implications?

- The number of leadership systems is very small. If defined as the top 5% on Rmax there are only 10 leadership systems worldwide on the June 2020 Top500 list.

- Traditional HPC applications require high memory and interconnect bandwidth in addition to FLOPS, while many commercial applications can get by with less bandwidth.

- It is possible to build commercial cloud systems that can meet leadership requirements, but the cost of bandwidth is prohibitive unless it is needed for most applications.

- Few vendors in the business due to the high cost of delivering very large systems. R&D funding through government channels is required.

- Bandwidth is the most expensive thing to deliver. If most systems sold don't require it, the vendors are not going to build systems with very high bandwidth without NRE funds to pay for the R&D. The systems will be more expensive.

- For Leadership systems, the cost of using commercial cloud systems is at least 2x to 3x more than having the computers in-house. There are instances of cloud systems that have supercomputer architectures, but none that are leadership class. It's too expensive if you don't run them at nearly 100% utilization.

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

What's next?

- The transition to accelerator-based nodes using CPUs and GPUs gave us a big increase in performance and efficiency. What will we do next to get similar gains?
  - Quantum has promise but with more complexity than today's systems.
  - Artificial Intelligence may let us get the answers to some problems without having to "compute" the answer, but certainly not for all types of problems.
  - The cost of current leadership systems is approaching $1 Billion per system and it has a nominal 5-year lifetime. This is a capital-intensive business!

- We will need an "all of the above" strategy to successfully build next-gen systems.
  - We need R&D into new devices that are much more efficient.
  - We must leverage commercial technology. There isn't enough money to build these systems if we must develop custom technology.
  - Every calculation must be done in the most efficient way possible. Be very aggressive in turning off circuits that aren't active. We can't afford to waste power and cooling.
  - We must continue to improve the algorithms to be more efficient. Reduced precision?

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

What might an this look like?

Instead of a leadership computer, it may be an exascale ecosystem:

- – More heterogeneity to improve efficiency but use COTS technology wherever possible to reduce costs
- – Distributed, powerful computers and storage so we don't have to move all the data back to the central exascale system
- – Each of the components in the environment is tailored to the tasks at that location
- – Only move data and calculations to the central computer if it is needed to solve a problem

Reply Hazy, Try Again

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

Federated Instruments:
Computing on the Edge

Manufacturing Demonstration Facility

Supercomputing Facility

Center for Nanophase Materials Sciences

Biological Sciences

Spallation Neutron Source

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

Truth is, predictions in technology space are dangerous

The internet 1995:

'The truth is no online database will replace your daily newspaper, no CD-ROM can take the place of a competent teacher and no computer network will change the way government works … How about electronic publishing? Try reading a book on disc. Yet Nicholas Negroponte, director of the MIT Media Lab, predicts that we'll soon buy books and newspapers straight over the Internet. Uh, sure.'

- Clifford Stull

Many print publications went exclusively online within 15 years

The iPhone in 2006

'Everyone's always asking me when Apple will come out with a cellphone. My answer is, 'Probably never.' ...It just ain't gonna happen'

– *David Pogue, tech columnist for The New York Times, writing less than 9 months before the first iPhone was released*

OAK RIDGE | LEADERSHIP
National Laboratory | COMPUTING FACILITY

Slide courtesy of Jim Hack

What's next for me?

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

Questions?

I can't think of a better place or group of people to have spent the last 40 years with than where I have been. I will miss you all.

OAK RIDGE | LEADERSHIP COMPUTING FACILITY
National Laboratory