USC Viterbi
School of Engineering

Exascale Computing Study

Computational Sciences

07 Oct 25
Robert F. Lucas
{rflucas}@isi.edu

Caveat

The study is still being written up

This is only my perspective

This material is based on work sponsored by DARPA, AFRL, and GTRI

Introduction

HPCS Program targets Petascale systems circa 2010

What research is needed for Exascale, circa 2018?

I know of two efforts to address this question:

- **NSA Advanced Computing Systems**
  - **Gary Hughes**
- **DARPA ExaScale Computing Study**
  - **Bill Harrod**

Goal of the Study

Determine what research the government needs to fund to enable its computer vendors to credibly decide, circa 2011, to initiate product development for Petascale systems that would be available later in the next decade.

Participants

Name	Organization	Name	Organization
Shekhar Borkar	Intel	Dean Klein	Micron
Dan Campbell	GTRI	Peter Kogge	Notre Dame
Bill Carlson	IDA	Bob Lucas	USC/ISI
Bill Dally	Stanford	Mark Richards	Georgia Tech
Monty Denneau	IBM	Al Scarpelli	AFRL
Paul Franzon	NC State	Steve Scott	Cray
Bill Harrod	DARPA	Allan Snavely	SDSC
Kerry Hill	AFRL	Thomas Sterling	LSU
Jon Hiller	STA	Stan Williams	HP
Sherman Karp	STA	Kathy Yelick	LBNL & UCB
Steve Keckler	University of Texas		

Bill Harrod is the DARPA Program Manager
Peter Kogge is the Principle Investigator

Participants
Countless Other Contributors

David {Bailey, Koester}	LBNL and MITRE
Keren Bergman	Columbia
Loring Craymer	NSA ACS

Lots of people from each host institution.

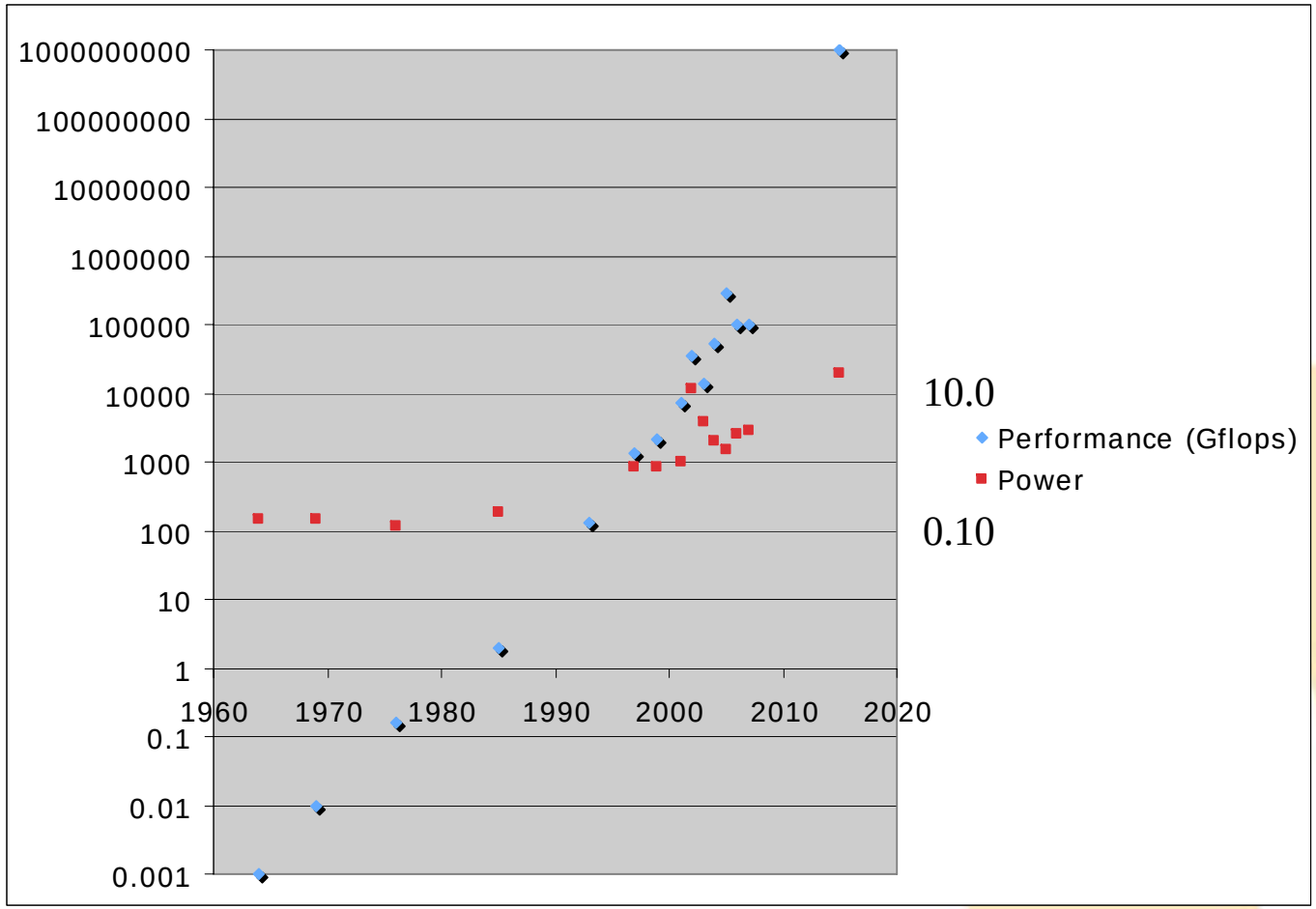
Four Meetings

Meeting #1:	May 30, STA
Meeting #2:	June 26-27, HP
Topic #1: Packaging:	July 17-18, Georgia Tech
Meeting #3	July 24-25, Intel
Meeting #4 Memory Roadmap and issues	August 16-17, Micron
Topic #2: Architectures and Programming	August 30, Stanford University
Topic #3: Applications, Storage, and I/O	September 6-7, UC Berkeley
Topic #4: Optical interconnects	Sept. 25-26, Stanford University
Meeting #5:	October 10-11, USC/ISI
Meeting #6:	November 15, SC|07

Some Stuff We've Discussed

- Power
- Memory volume
- Programming
- Reliability
- Packaging

Power Perspective

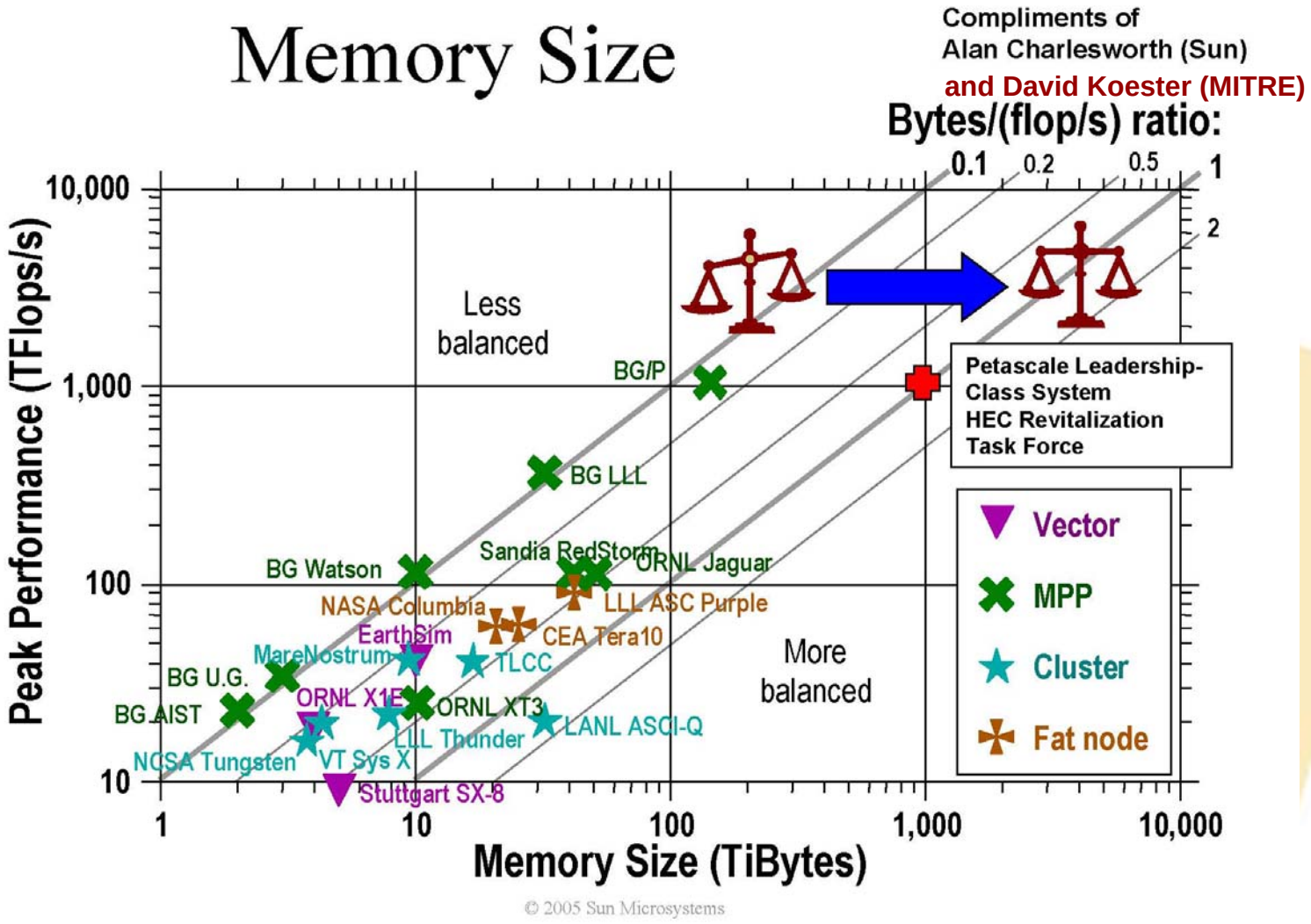


Power Research

Pervasive problem, requires range of solutions!

- **Architecture:**
  - Intel Polaris spent most power issuing instructions
- **Process and circuit technology**
  - Sub threshold devices?
- **Optics:**
  - Interconnect (DARPA MTO program)
  - Clock distribution (save 10 W per socket?)
- **Memory:**
  - ~5W for a GByte DRAM

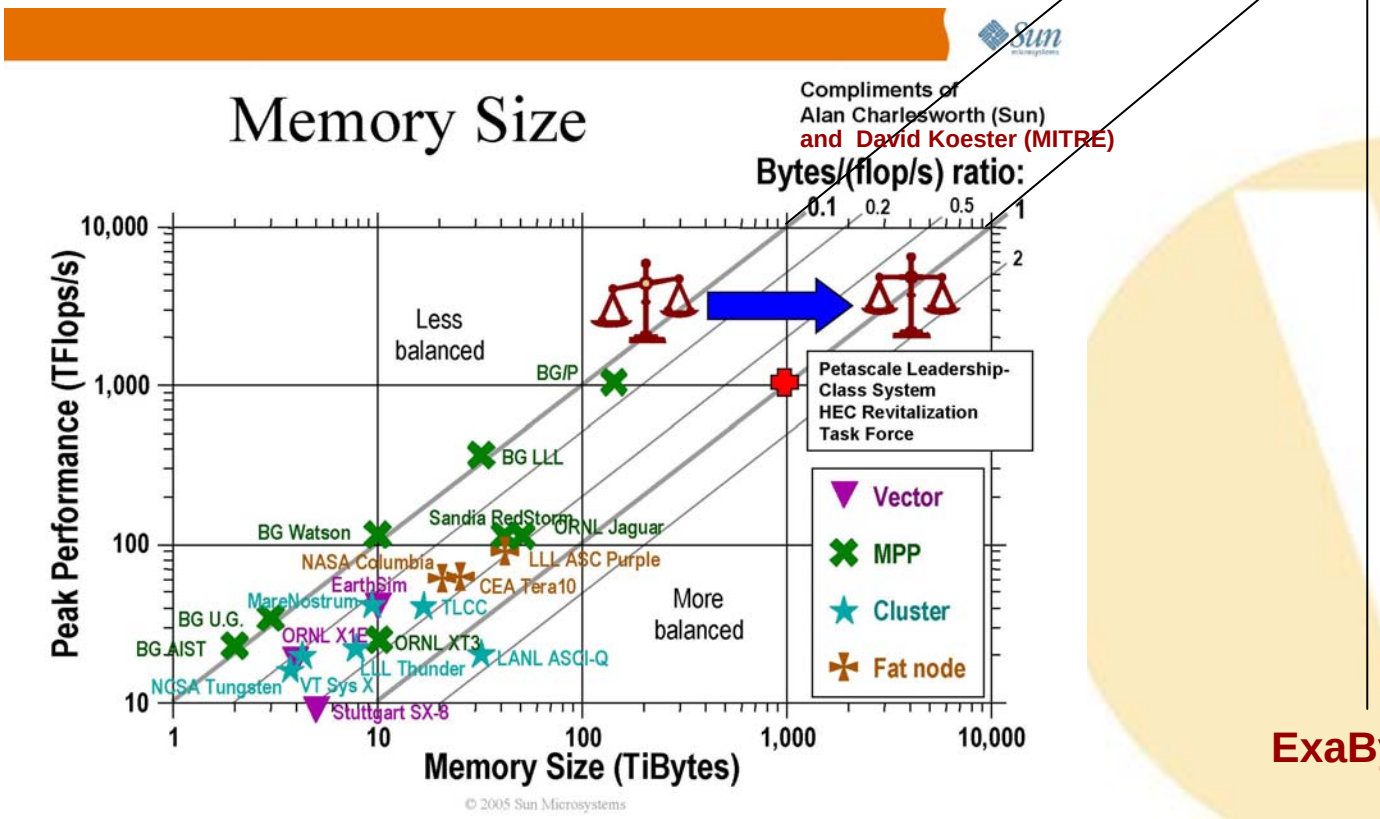
Ops vs. Memory Balance



Memory Capacity Challenge

1,000,000 DRAM chips, circa 2014

ExaFlop/s



ExaBytes

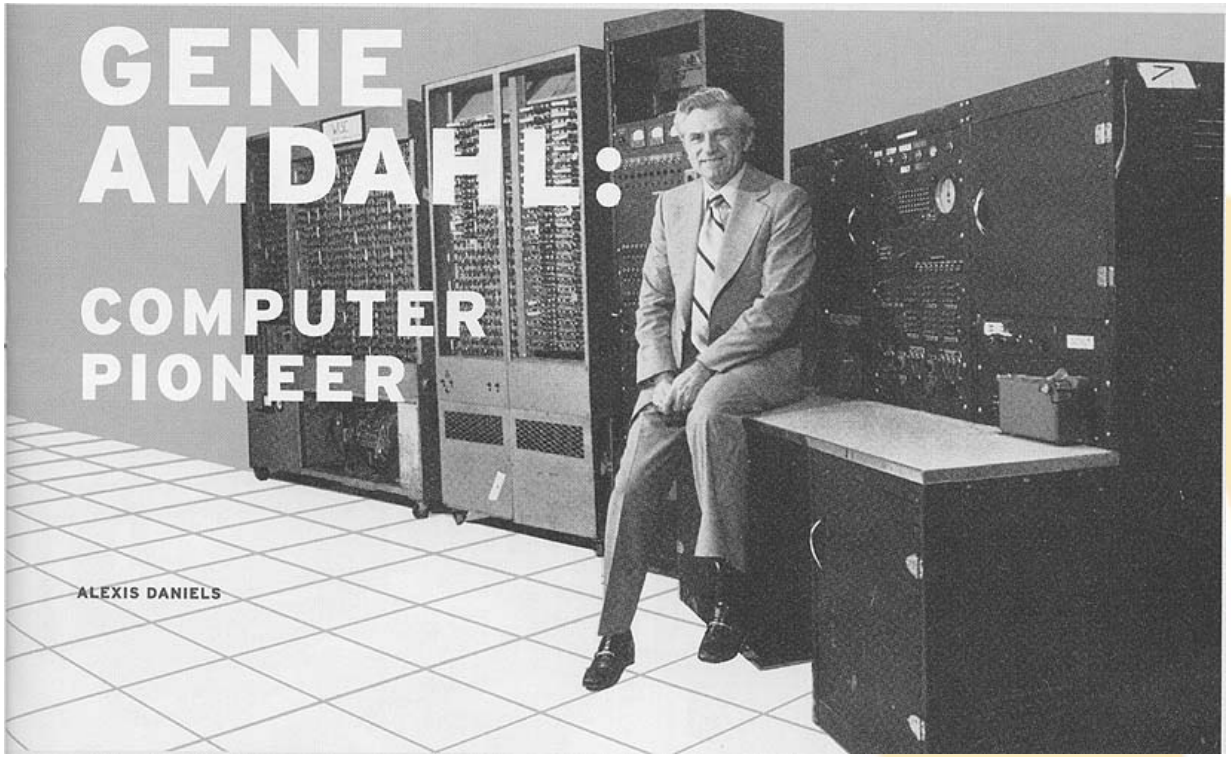
Memory Research

- A PetaByte main memory won't be useless
  - There are applications that look like Linpack
  - Others like sPPM have small footprints
- For scaled speedup, we'll need more main memory
  - Novel technology for main memory
  - EDRAM for L2
  - DRAM as L3
- What about scratch and the file system?
- What about archives?

Programming Challenge

ExaSt

Perhaps ten billion threads!
What would Mr. Amdahl think?



Programming Research

MPI will suffice for a few stunts.

- **MPI + OpenMP per socket isn't much better.**

I'm hoping for something like UPC.

- **UPC has impact because applications adopted it**

God forbid, but is CUDA the model for the future?

- **Multi-core**
- **Multi-threaded**
- **SIMD extensions**
- **Explicit memory hierarchy**

Fault Tolerance

Some say its “only a factor of two away”.

(i.e., build two systems and compare results)

Actually, for much of the system, its better than that today.

Memories and transmission lines already protected

Effectively protecting logic still an issue.

Packaging

Information Sciences Institute

ExaSt

Don’t want to measure computers by the acre.

- **In the best case, distance equals latency.**
- **Minimize power.**

Its not clear that this is a key bottleneck to achieving Exascale.

Personal Observation

This was been a really fun, enlightening exercise ☺

Remarkably conservative! Exotic technology may not be required.

- **E.g., SiGe or SFQ**
- **Nor cooling the system to 70K, much less 4K**

The space of applications is getting smaller

- **How many will run effectively at O(1B) threads?**
- **Five orders-of-magnitude from today's extreme.**

Impact on DOE SC

Good news:

- Growth in raw computing power will continue unabated
- Enables scientific discoveries beyond imagination today

Bad news:

- Even after twenty years, we're still not done porting codes to parallel systems.
- Concurrency will increase 4-5 orders-of-magnitude.
- System balance will change dramatically.
- Number of successful codes (even whole fields) will decline.
- Facilities will have to transform (again!) to adapt (e.g., power).

Summary

An Exascale computing system within a decade is plausible.

There are a number of significant problems that will need to be overcome. The DOE should look towards addressing them now, while there's time. DOE should continue its partnership with DOD (DARPA & NSA).

Bonus Slides