



# An Effective Standard for Developing Performance Portable Applications for Future Hybrid Systems

Supercomputing 2012  
November 12, 2012

**John Levesque**  
Director/CTO Office



## First a confession

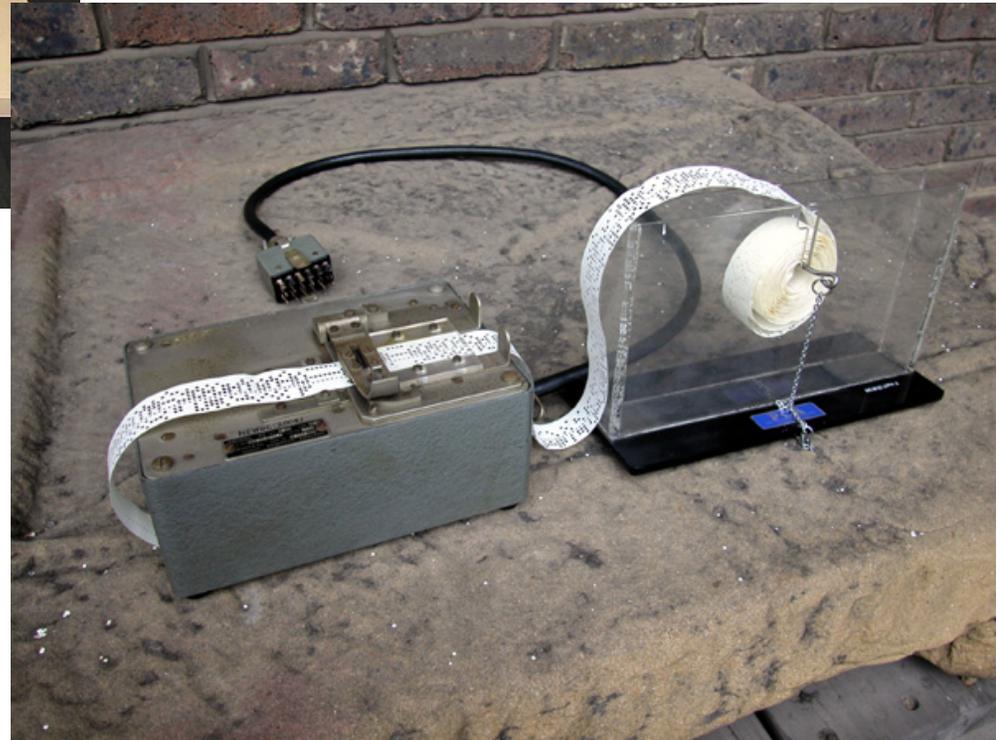
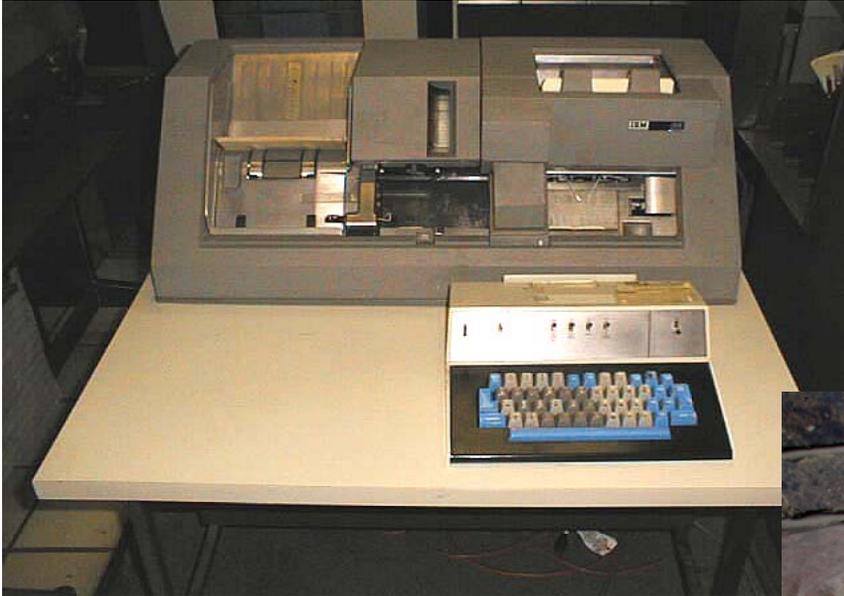
I have never written an application from scratch, everything I have done starts with existing applications which I restructure to run faster.

One might say that I am one of the reasons some legacy applications are still around.

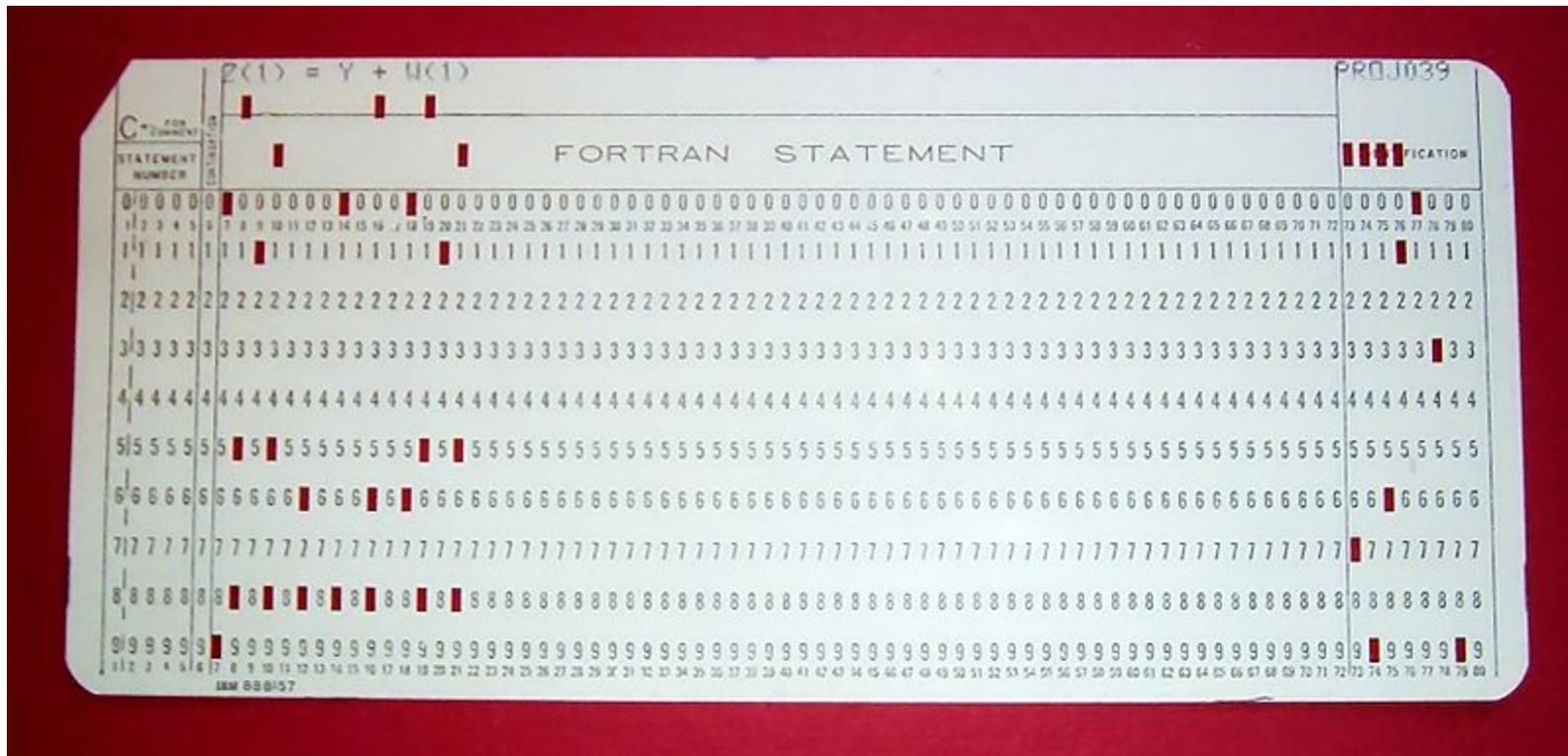
# What Supercomputer Was this?

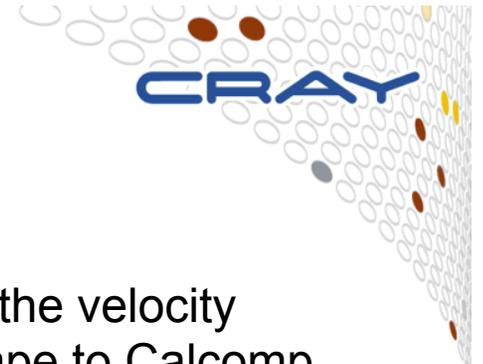


# What are these?



# Why was the Fortran line length 72?





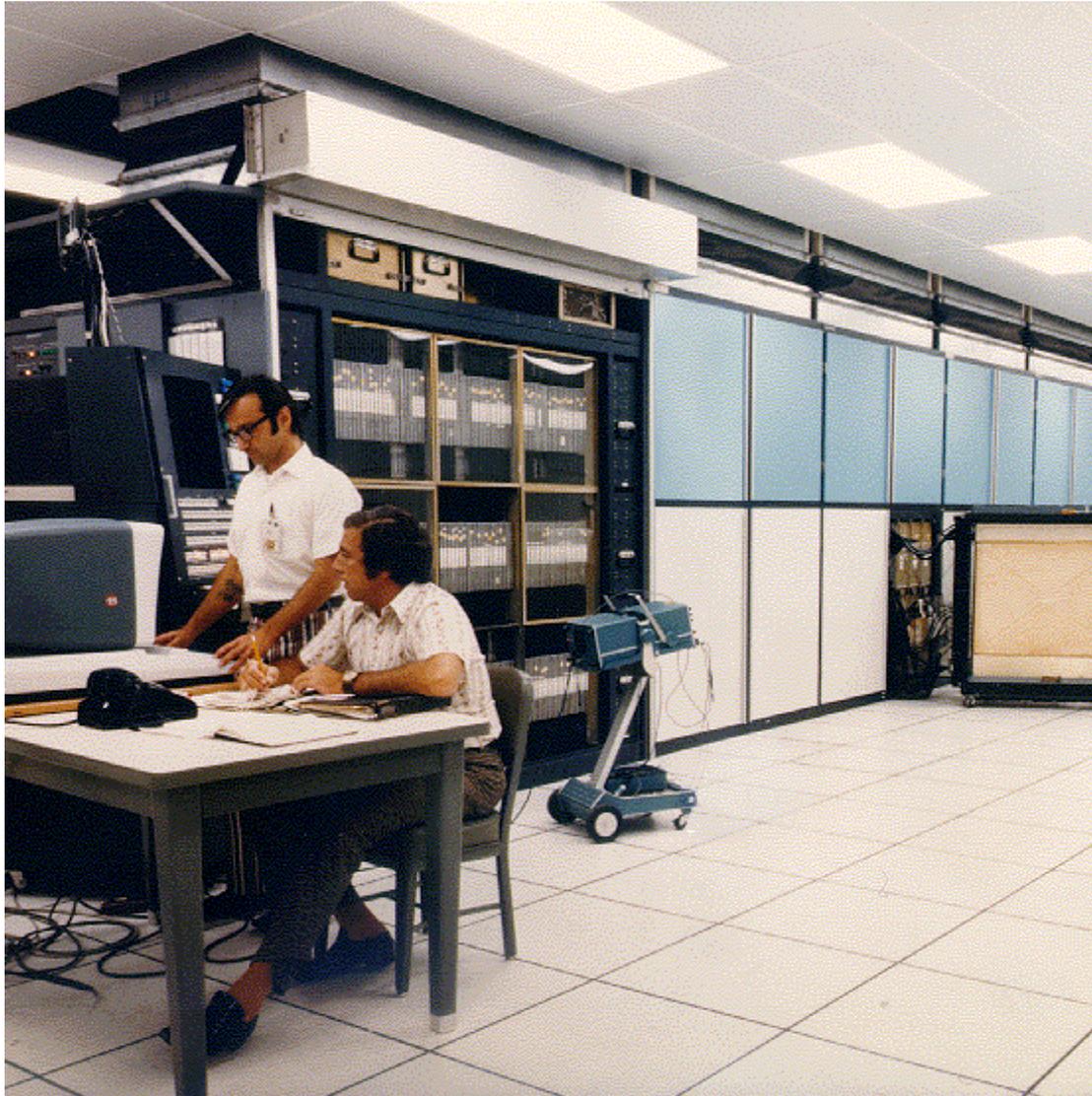
# Computational Steering in the 70s



Set a sense switch, dump the velocity Vector field to tape, take tape to Calcomp Plotter.



# What Supercomputer is This?



Anyone know why  
the door is open?

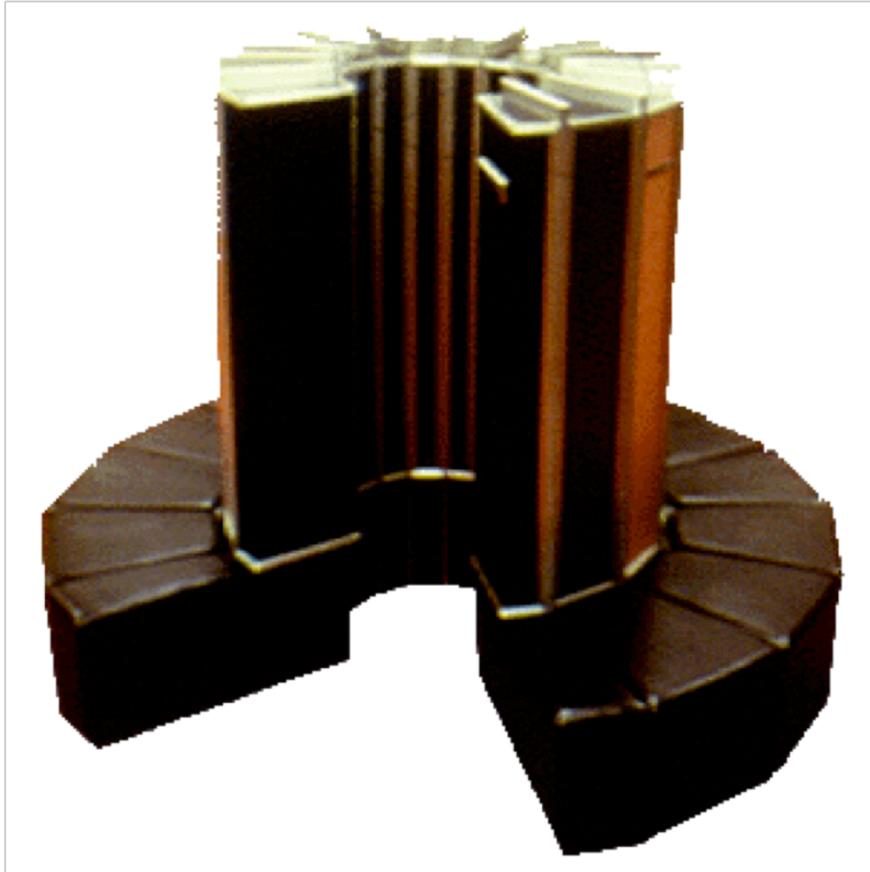
# Who Invented the Internet?



## Another Seymour Masterpiece



How much memory  
did this system have?

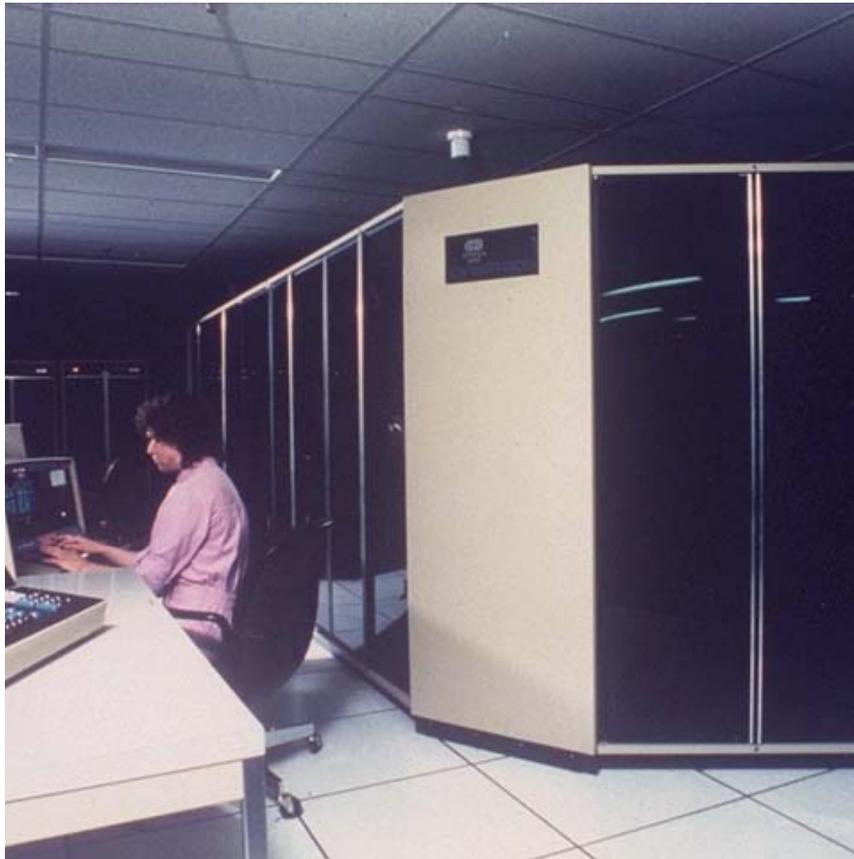


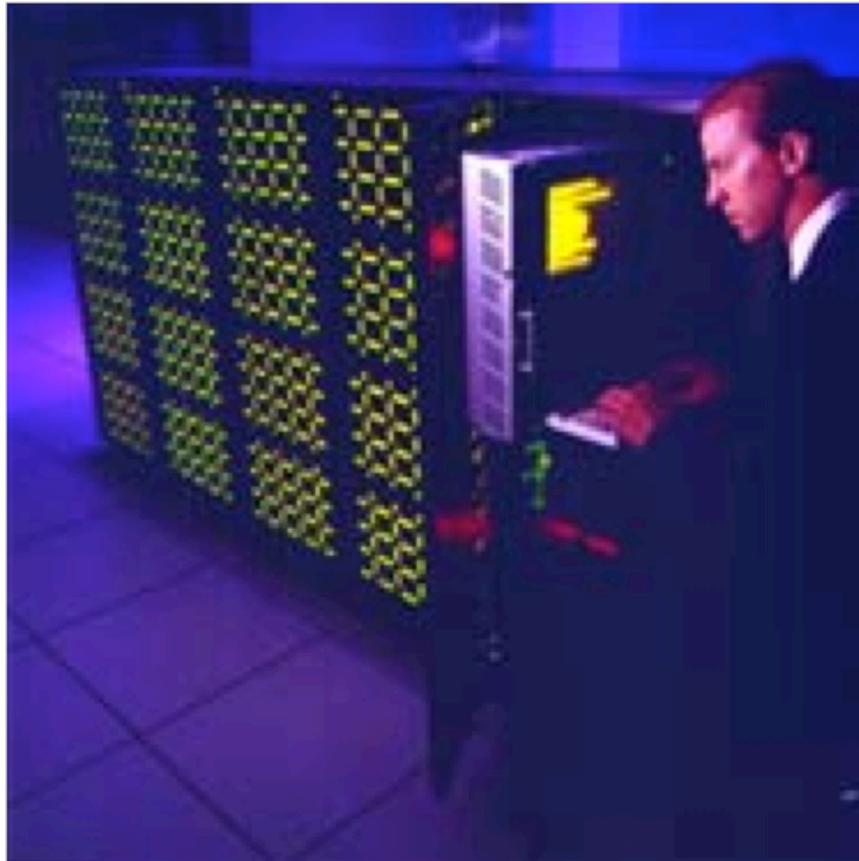
What Supercomputer is this?

# My first laptop



**Door Prize for anyone guessing what this is.**





## What Supercomputer is this?

During this decade,

More money was spent on Disposal Diapers than on Supercomputer



**What Supercomputer  
is this?**



# Famous Jurassic Park Prop



# The best Co-Array machine



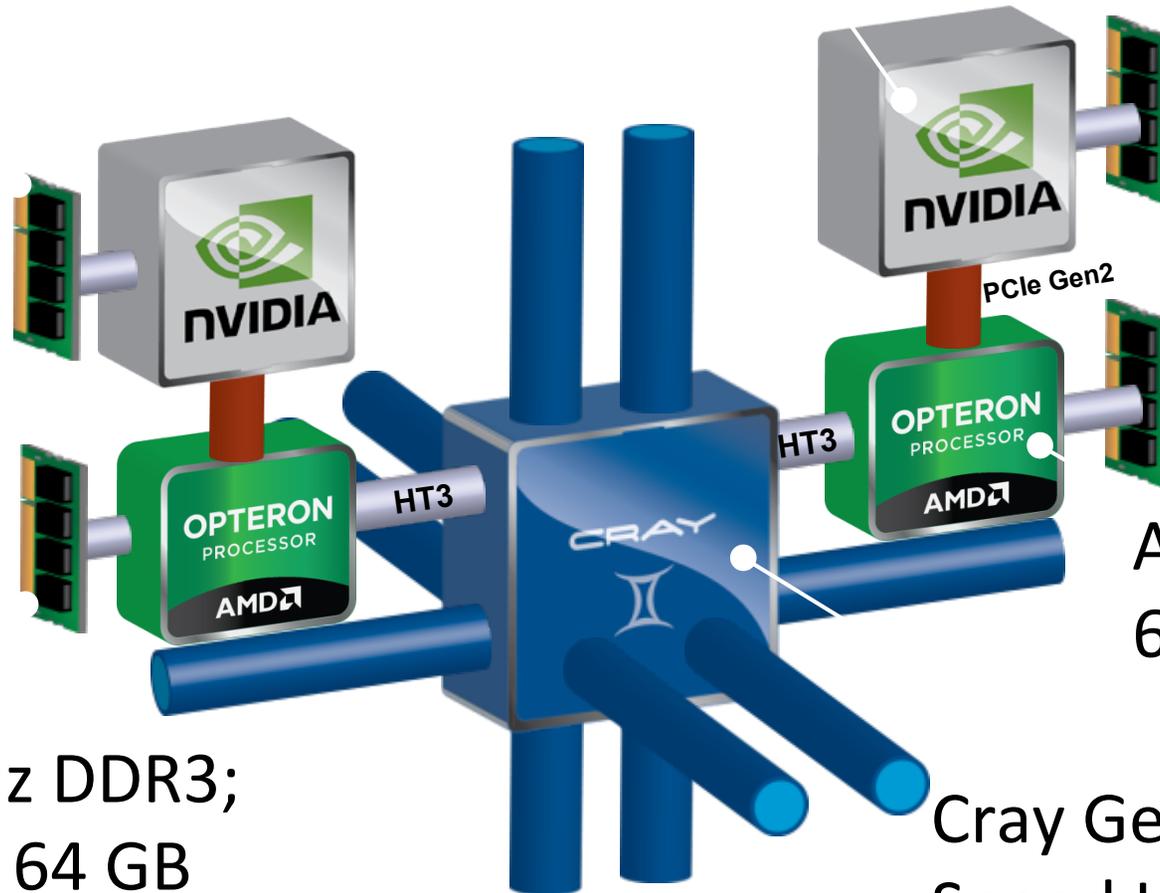
# The System that shot down a satellite



# Our Current Challenge

6GB GDDR5;  
138 GB/s

NVIDIA Tesla GPU  
with 665GF DFPF



AMD Series  
6200 CPU

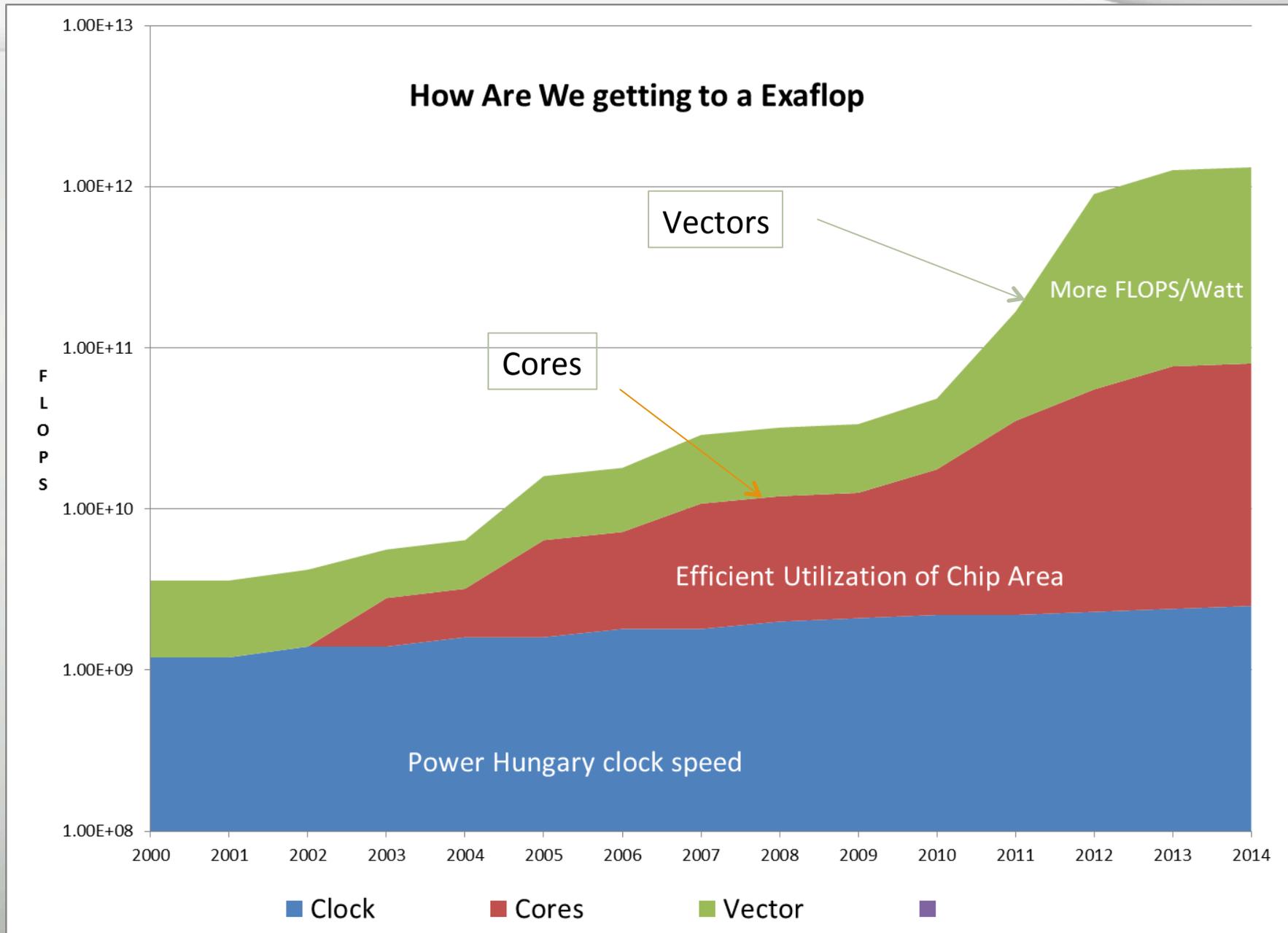
1600 MHz DDR3;  
16, 32 or 64 GB

Cray Gemini High  
Speed Interconnect

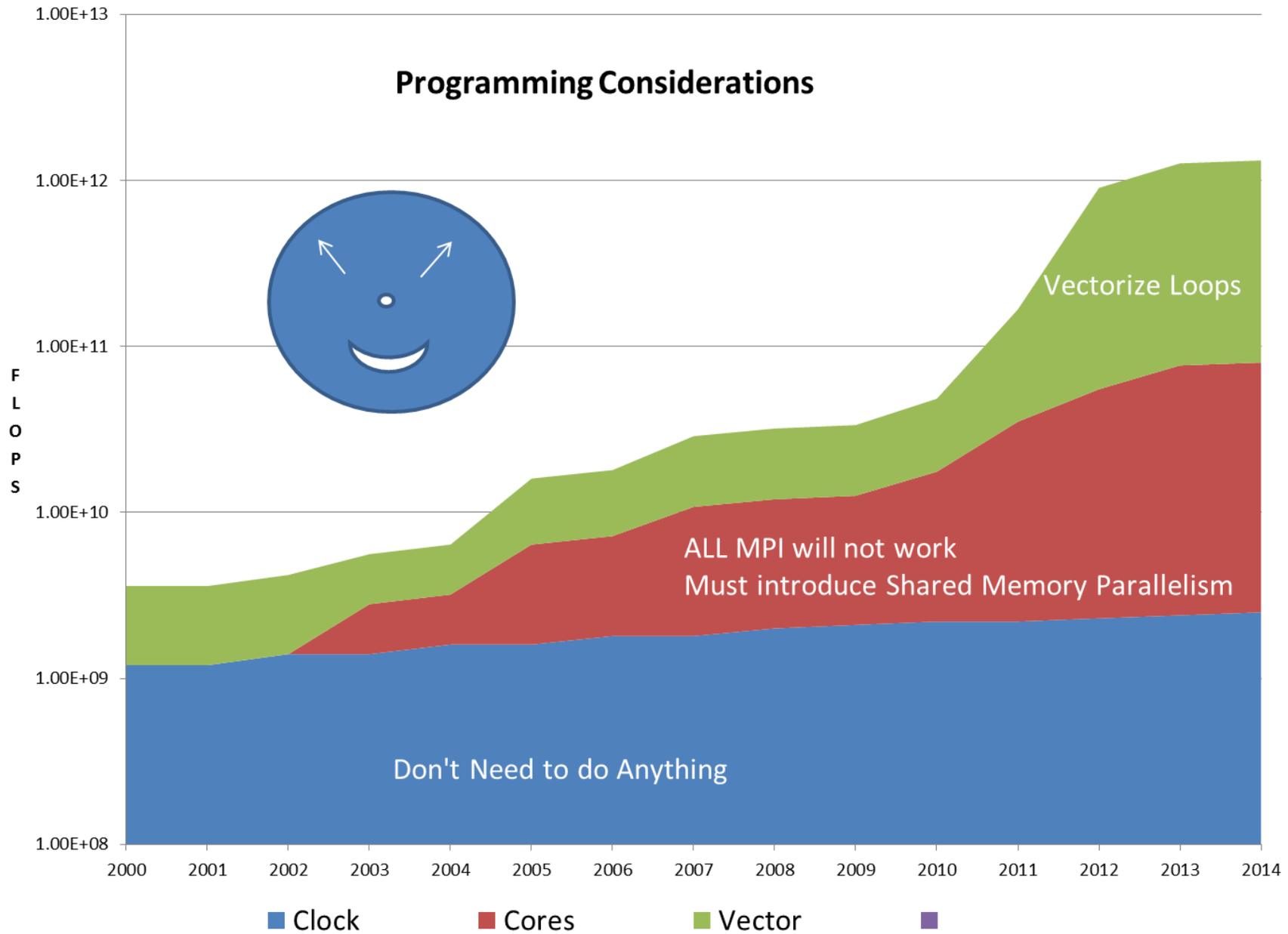
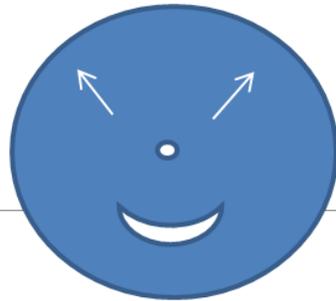
# Outline

- Future Architectural Directions
  - Chips are not being designed for HPC
  - Power consumption is a major concern
  - What is heterogeneous Computing?
- Programming implications
  - All MPI is not an option
  - OpenMP and OpenACC

## How Are We getting to a Exaflop



## Programming Considerations



# Potential System Architecture for Exaflop

Systems	2010	202?	Difference Today & 2018
System peak	2 Pflop/s	1 Eflop/s	O(1000)
Power	6 MW	~20 MW	
System memory	0.3 PB	32 - 64 PB [ .03 Bytes/Flop ]	O(100)
Node performance	125 GF	1,2 or 15TF	O(10) – O(100)
Node memory BW	25 GB/s [.20 Bytes/Flop]	2 - 4TB/s [ .002 Bytes/Flop ]	O(100)
Node concurrency	12	O(1k) or 10k	O(100) – O(1000)
Total Node Interconnect BW	3.5 GB/s	200-400GB/s (1:4 or 1:8 from memory BW)	O(100)
System size (nodes)	18,700	O(100,000) or O(1M)	O(10) – O(100)
Total concurrency	225,000	O(billion) [O(10) to O(100) for latency hiding]	O(10,000)
Storage	15 PB	500-1000 PB (>10x system memory is min)	O(10) – O(100)
IO	0.2 TB	60 TB/s (how long to drain the machine)	O(100)
MTTI	days	O(1 day)	- O(10)