CHEP 2012

Computing in High Energy and Nuclear Physics

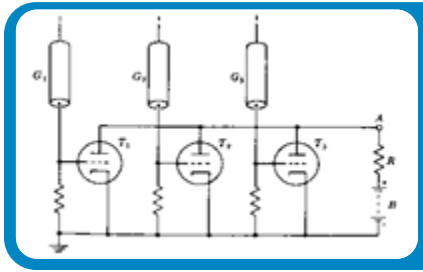
Forrest Norrod

Vice President and General Manager, Servers

Advancing human progress through basic knowledge
advances technological innovations

One might ask …

whether basic circuits in computers might have been found by people who wanted to build computers.

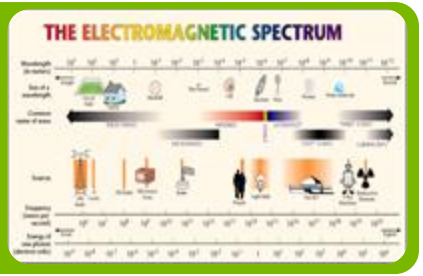


As it happens, they were discovered in the thirties by physicists dealing with the counting of nuclear particles because they were interested in nuclear physics. [1]

Or whether, in an urge to provide better communication, one might have found electromagnetic waves.

They weren't found that way.

They were found by Hertz who emphasized the beauty of physics and who based his work on the theoretical considerations of Maxwell. [1]



[1] C.H. Llewellyn Smith, former Director-General of CERN

Physicists' needs are outpacing Engineers delivery

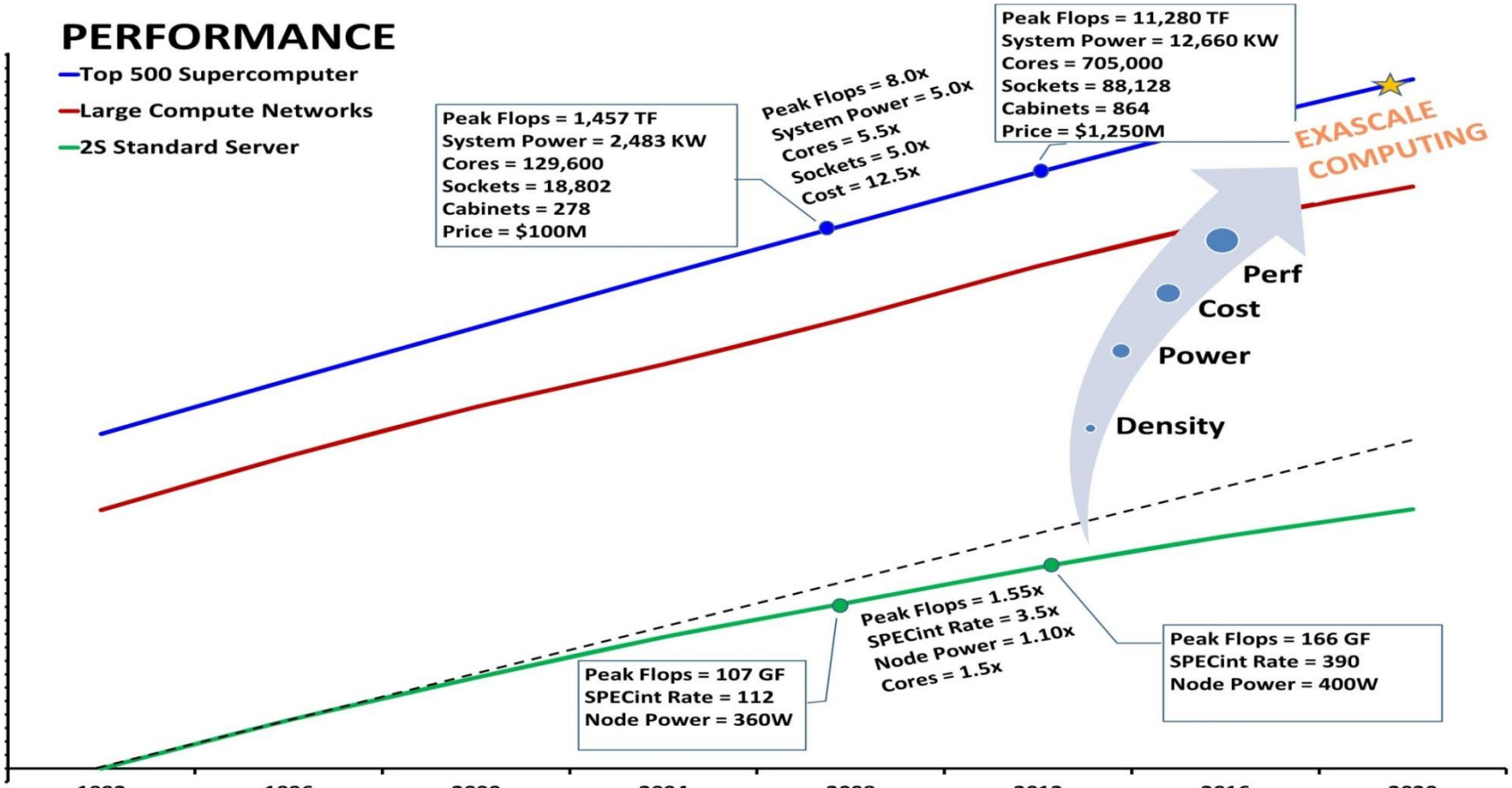
Source: www.top500.org

2007 goal was to achieve
Exascale by 2015
So far … off by 100x

Major challenges

- Power
- Reliability/Resiliency
- Concurrency & locality (drives efficiency)
- COST

Performance Trajectories



Source: ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems

What Levers do We Have?

- Challenge: Sustain performance trajectory without massive increases in cost, power, real estate, and unreliability
- Solutions: No single answer, must intelligently turn "Architectural Knobs"

$$\text{Performance} = (\text{Freq}) \times \left(\frac{\text{Cores}}{\text{Socket}}\right) \times (\#\ \text{of Sockets}) \times \left(\frac{\text{Inst or Ops}}{\text{Core} \times \text{Clock}}\right) \times (\text{Efficiency})$$

Hardware Performance

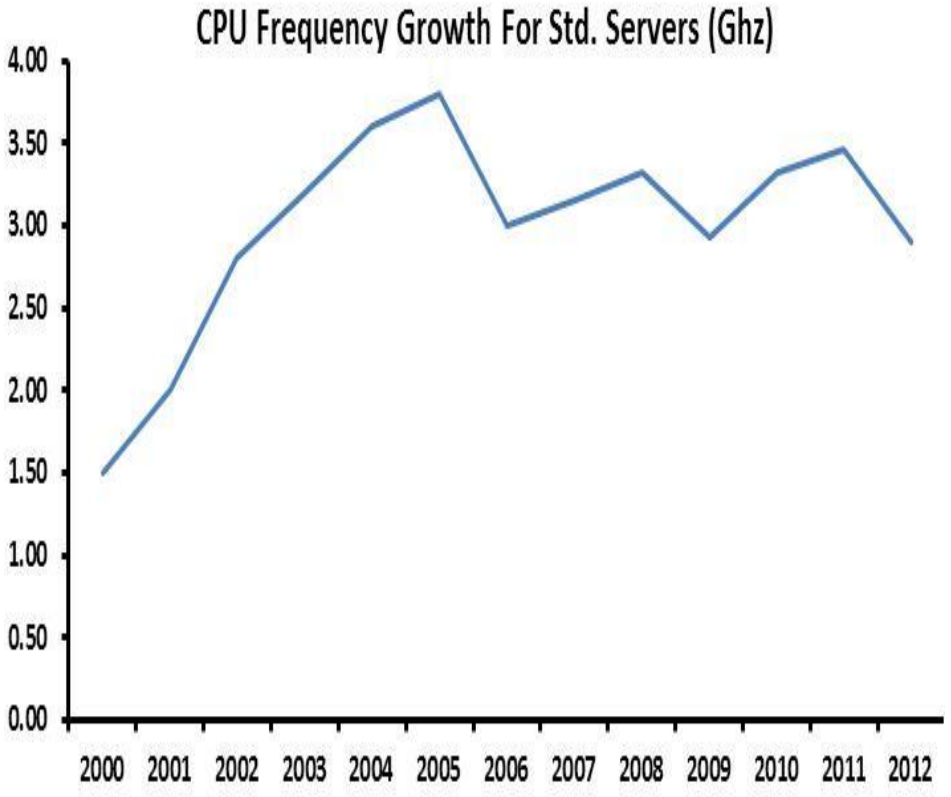
What you can really get

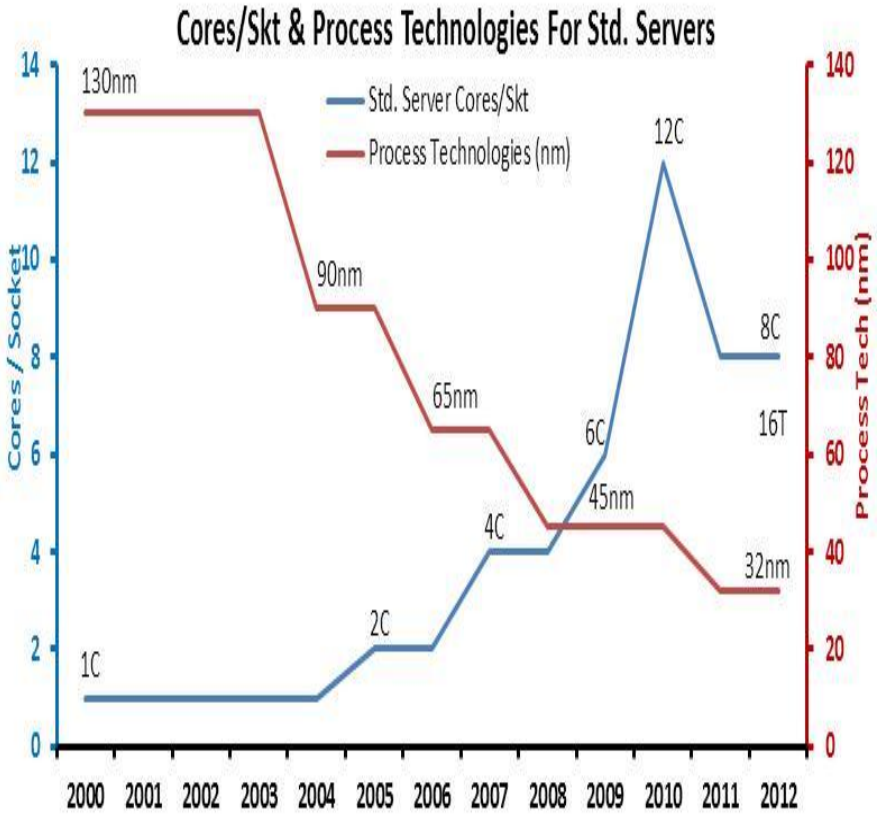
Architecture Knob #1: frequency

Frequency stagnant for a decade

Unlikely to change anytime soon

- Thermal density challenges
- Macro power & cooling concerns
- No longer a main focus of IC industry





Architecture Knob #2:
cores / socket

Moore's law is fine ...

130nm -> 22nm LOTS of Transistors

Cores/socket growth in standard CPUs tapering off ...

Physical "uncore" problem: very costly to feed the powerful beast

Mismatch with DRAM speeds

Huge growth in CPU skt size, pins, memory (500 to 2000 pins in last 10 years, 3500+ pin beasts on horizon)

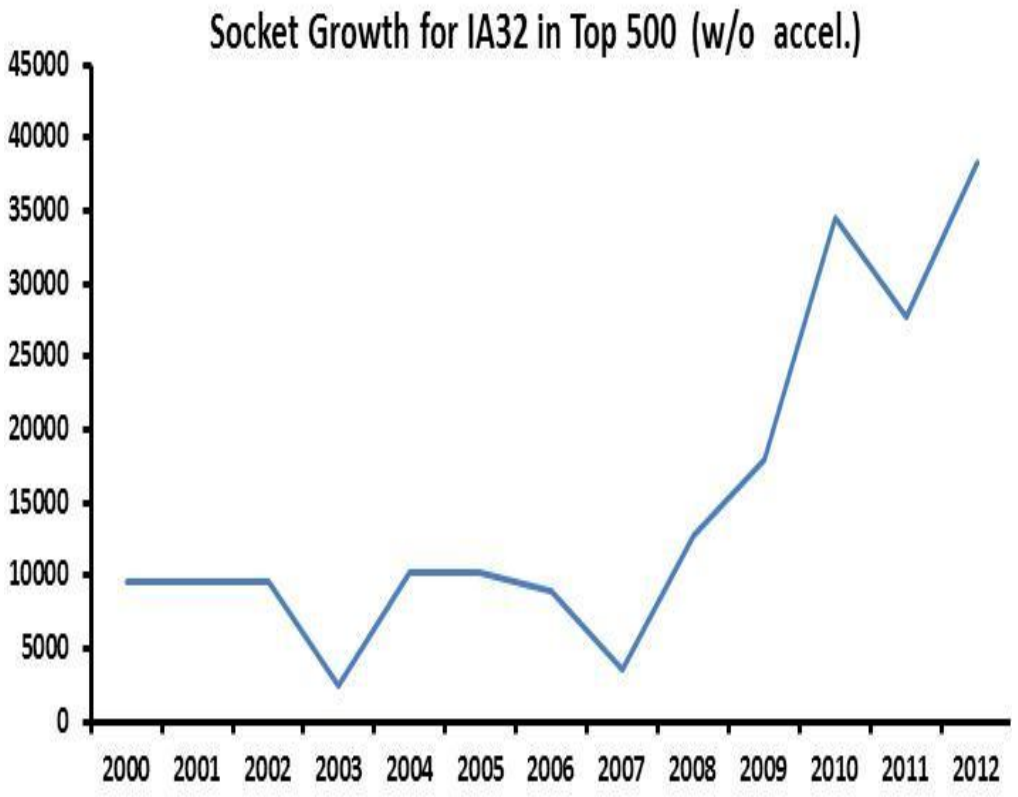
Architecture knob #3:
of sockets

Heaviest recent growth

- Despite accelerators
- "Easiest" knob to turn

But, Massive implications

- Network performance
- Resiliency & Stability
- Density
- Power/cooling costs



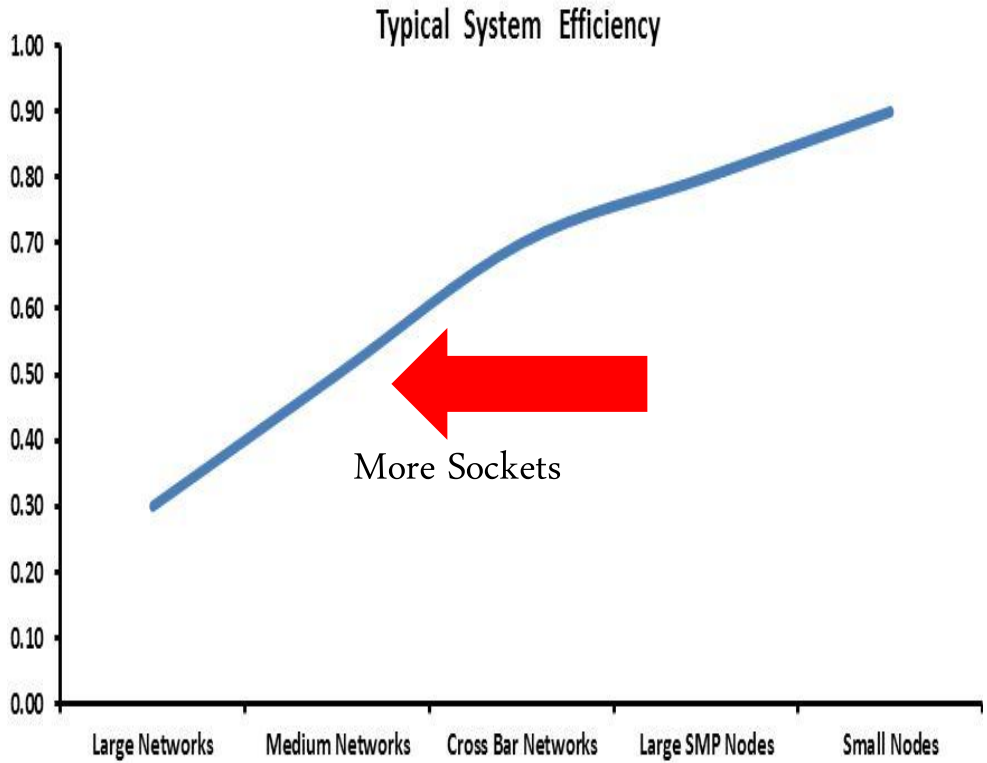
Architecture knob #5: efficiency

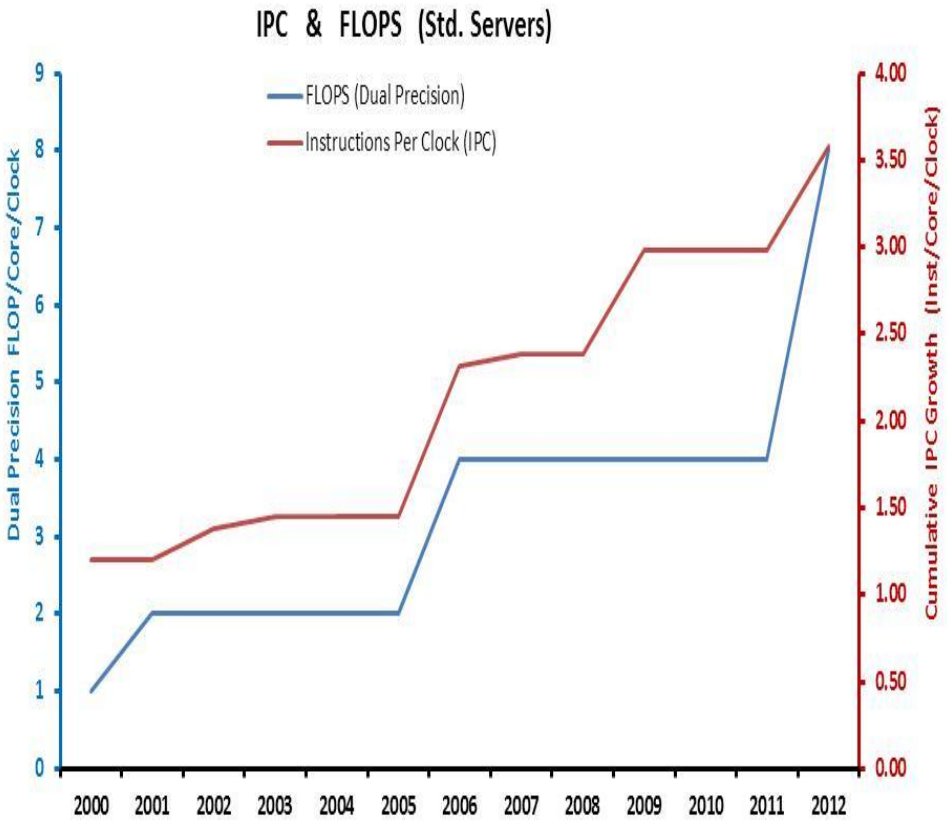
Large compute networks generally = lower system efficiency

Particularly as more CPUs & infrastructure are added

Efficiency driven by Locality, storage, network, CPU,.....and code quality

Typical values range: 0.3 – 0.9





Architecture knob #4:

Instructions per core/clock

IPC & FLOPS growth continues

uArch , replication of pipeline subunits, # of FPUs, FMA , BLAS libraries, others

ISA, AVX, SSEx, FPU Vector Width

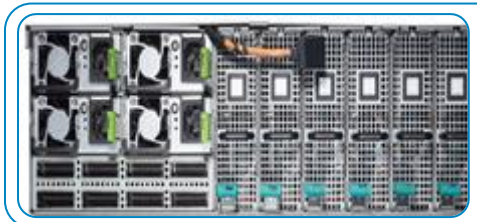
Direct FPGA-implementation of algorithms

Challenges: continual modification of software in order to take advantage of new instructions, new features, etc.

So what can we do to make this possible and affordable?

Driving IPC

Today ***Hybrid of CPU:GPU***



2008: NVIDIA GPU

2012 / 13: AMD APU & INTEL MIC

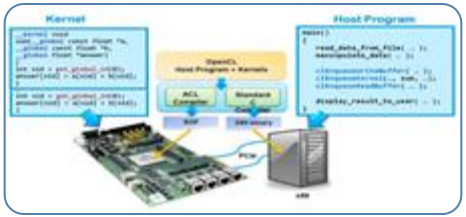
Tomorrow ***Accelerator hybrids FPGA, SOC ...***

FPGAs: performance and perf/watt advantages

- > 4k DSP cores in 28nm

System-on-Chip implementations

And ***Open programming***



OpenCL: ease programming model

Enablement on top of CPUs, GPUS, FPGAs - hiding the complexity of silicon design

Coupled with FPGAs - could change computational landscape

Scaling Sockets

Tomorrow ***New CPU Core Contender ARM***

ARM: potential to disrupt perf/$$, perf/watt model

Standard ISA and Open ecosystem

ARM+Native acceleration SOCs possible

Innovative ***packaging***

Shared Infrastructure evolving

Hybrid nodes

Concurrent serviceability

Highest efficiency for power and cooling

Disruptive ***integrated designs***

Extending design to facility

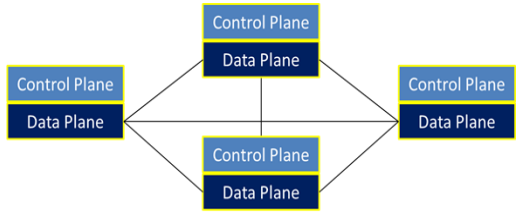
Modularized compute/storage/facility optimization

2000 nodes, 30 PB storage, 600 kW in 22 m^2

Ultimate in power efficiency, PUE as low as 1.02

Efficient Coordination

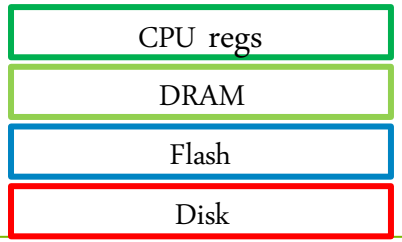
Redefining ***Networking***



Flatten & UnLock Network Topologies

- Software Defined (OpenFlow)
- Flat L2 networks
- East-West optimization
- 10G -> 40G -> 100G

Reinventing ***Memory Hierarchy***



Memory Access a principal driver of "stalls"

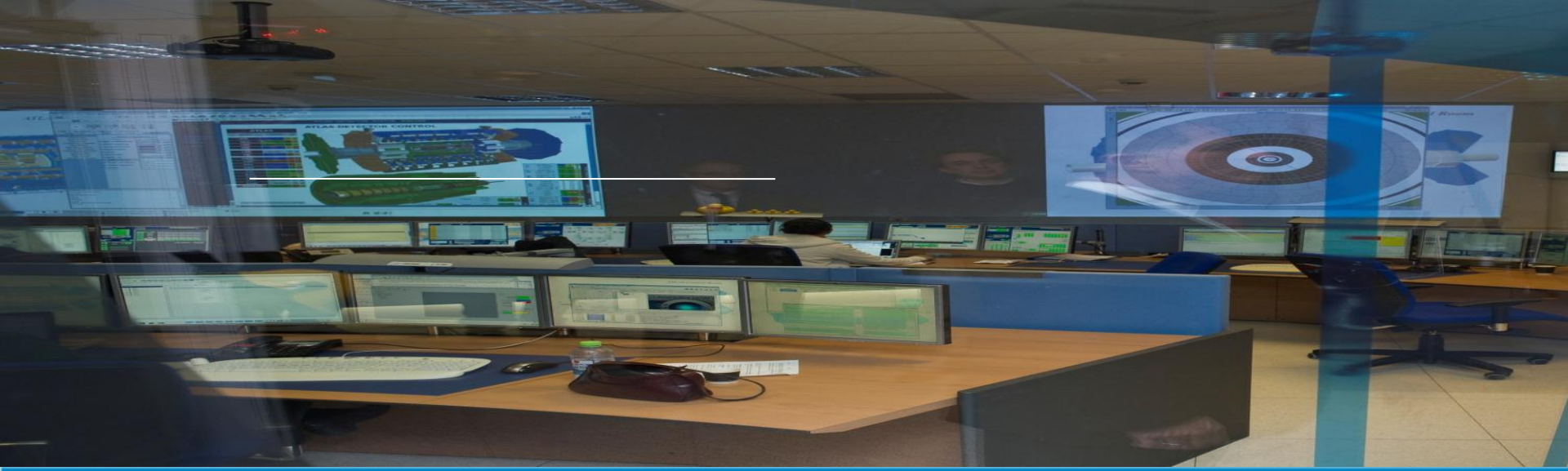
- CPU/DRAM stacking
- Flash as a new tier (1/2 way in between DRAM and Disk)
- Memory virtualization coming soon…

open ***collaboration***

Accelerating collaborative usage and models

- Transparent Infrastructure stacks
- Leverage "Cloud computing" investments
- Cloud usage for low-duty-cycle usage
- Openflow, OpenStack, Hadoop,…

We'll get there....but not in 2015



Gratia
Tak
Toda
Dankon
Tänan
Grazie
Dua-Netjer-en-ek
Thank-You
Bedankt
Obrigado
Merci-beaucoup
Sas-efharisto
Gracias
Spasibo
Danke-schön
Arigato
Dziekuje
Tack