# Aim High

## Intel Technical Update

## *SOS11*

June 12, 2007

David Barkai, Ph.D.
HPC Comp Scientist

# Risk Factors

Today's presentations contain forward-looking statements. All statements made that are not historical facts are subject to a number of risks and uncertainties, and actual results may differ materially. Please refer to our most recent Earnings Release and our most recent Form 10-Q or 10-K filing available on our website for more information on the risk factors that could cause actual results to differ.

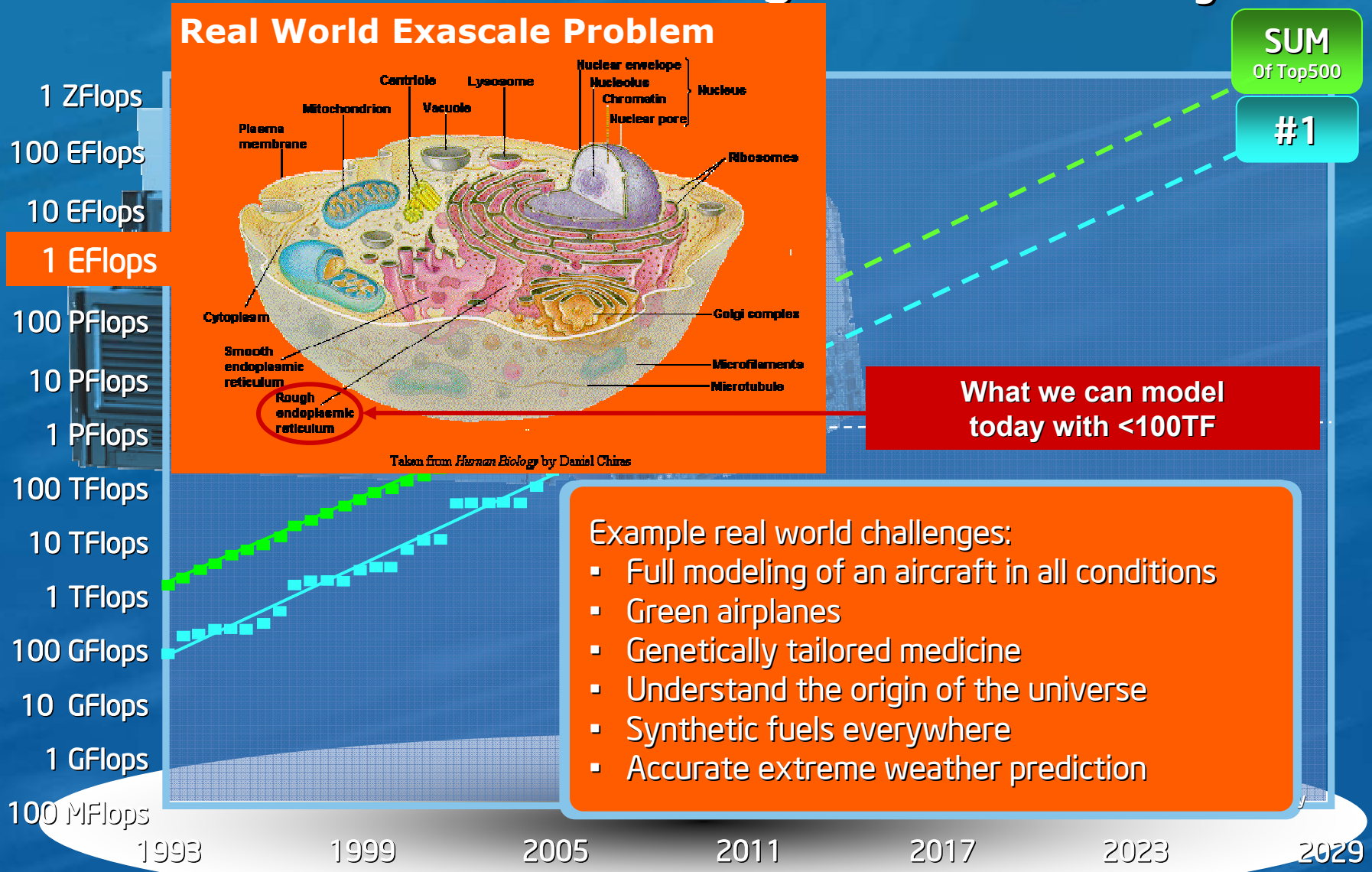
Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing. For more information on performance tests and on the performance of Intel products, visit Intel Performance Benchmark Limitations (http://www.intel.com/performance/resources/limits.htm).

# Real World Problems Driving Petascale & Beyond

## Real World Exascale Problem




Taken from *Human Biology* by Daniel Chiras




What we can model today with <100TF

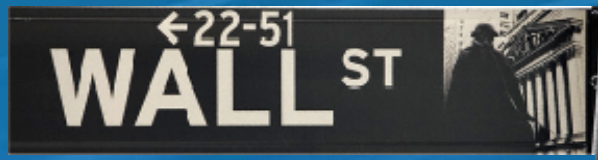
Example real world challenges:

- Full modeling of an aircraft in all conditions
- Green airplanes
- Genetically tailored medicine
- Understand the origin of the universe
- Synthetic fuels everywhere
- Accurate extreme weather prediction

# Supercomputing Datacenters Today

## *Supercomputers Solve Complex Problems*

- 60% of US profits are in financial services
- As many as ten top tier firms with IT Budgets > $2B/year
- Performance Needs:
  - 60,000 Transactions Per Second
  - 30ms average latency
  - 32GB real time database
- Good Scale out Application
  - Application is amenable to parallelization both data & thread
  - Requires Double Precision FP, high memory BW, …




**Sample Application: Collateralized Debt Obligation (CDO) Monte Carlo Simulation**

400,000 underlying assets (400 branches, ~1,000 loans each)

Monte Carlo Simulation considers external factors and their interrelationships

400B scenarios (simplified).

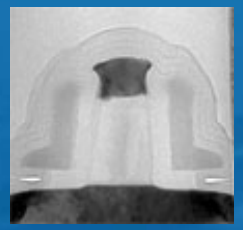
- For each asset there are 1M paths. 1,000 prepay * 1,000 loss
- Future goal to compare CDO to another CDO deal (CDO^2) 160T scenarios (400B^2). Considering CDO^3

**Wall Street:**
***Building 2-4 Data Centers Per Company by 2010***

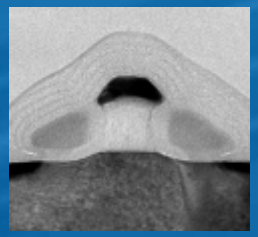
Source: Intel

# Silicon Future

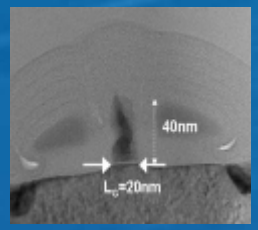
90nm
2003



65nm
2005

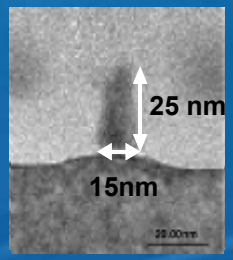


45nm
2007

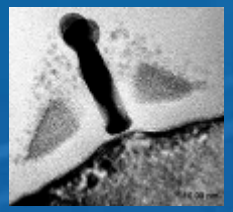



2005 – 2012

32nm
2009




22nm
2011



2013 – 2017

16nm
2013

11nm
2015

8nm
2017

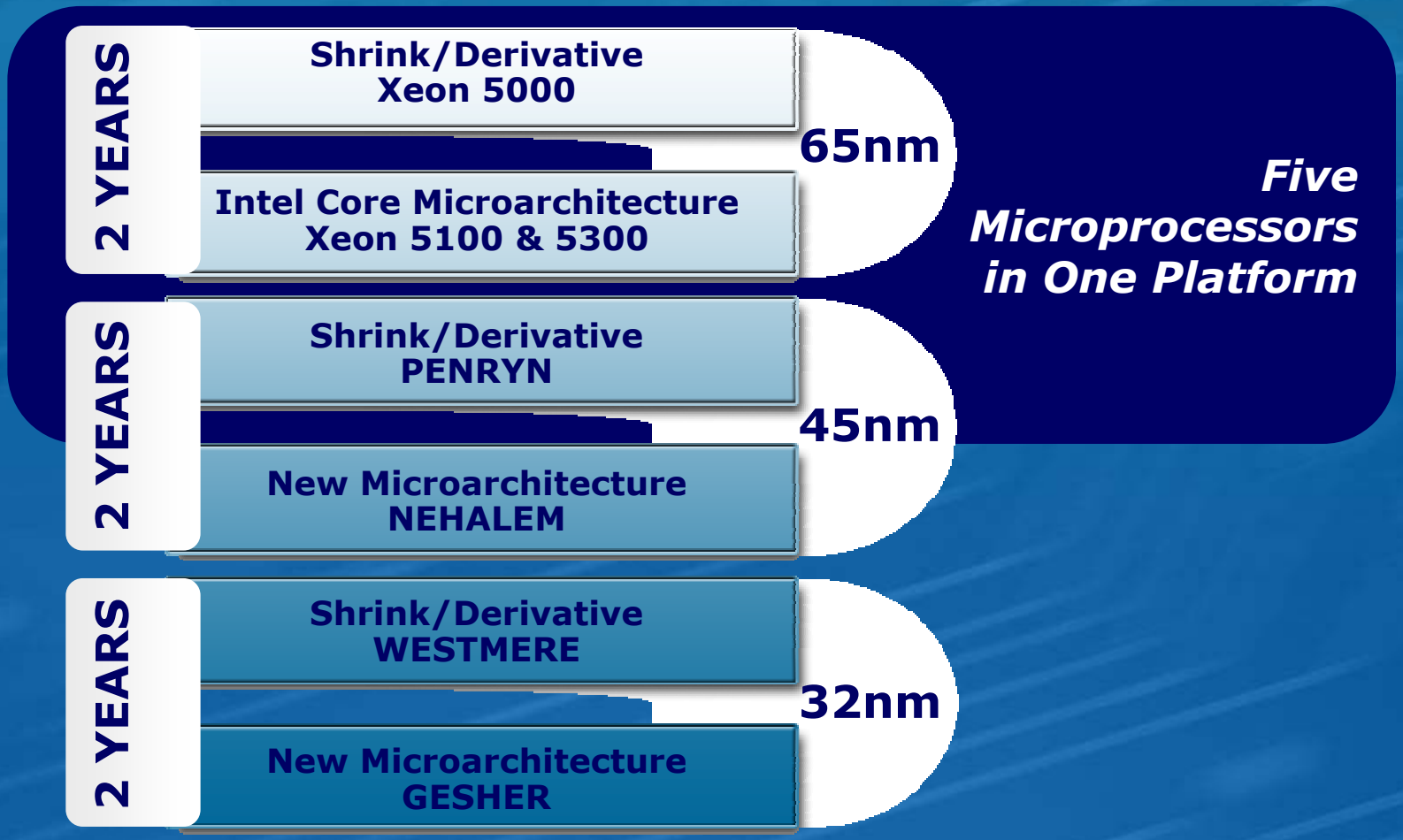
New Intel technology generation every 2 years
Intel R&D technologies drive this pace well into the next decade

Roadmap

Research

intel

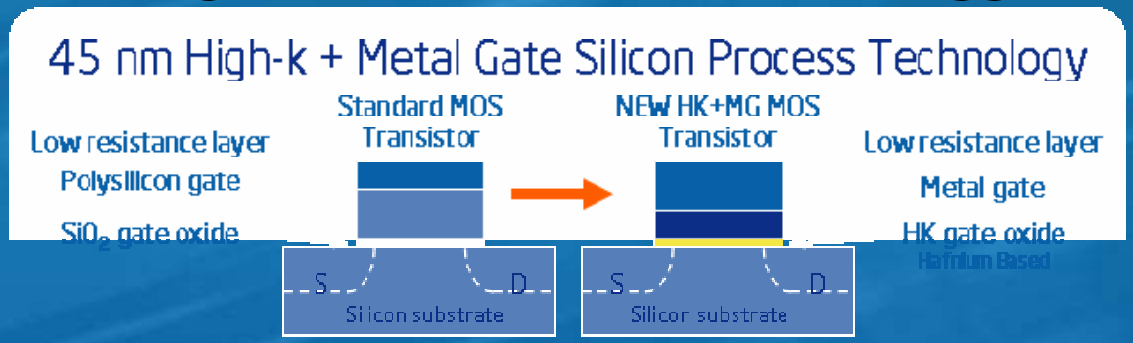
# Intel Design & Process Cadence






All dates, product descriptions, availability and plans are forecasts and subject to change without notice.

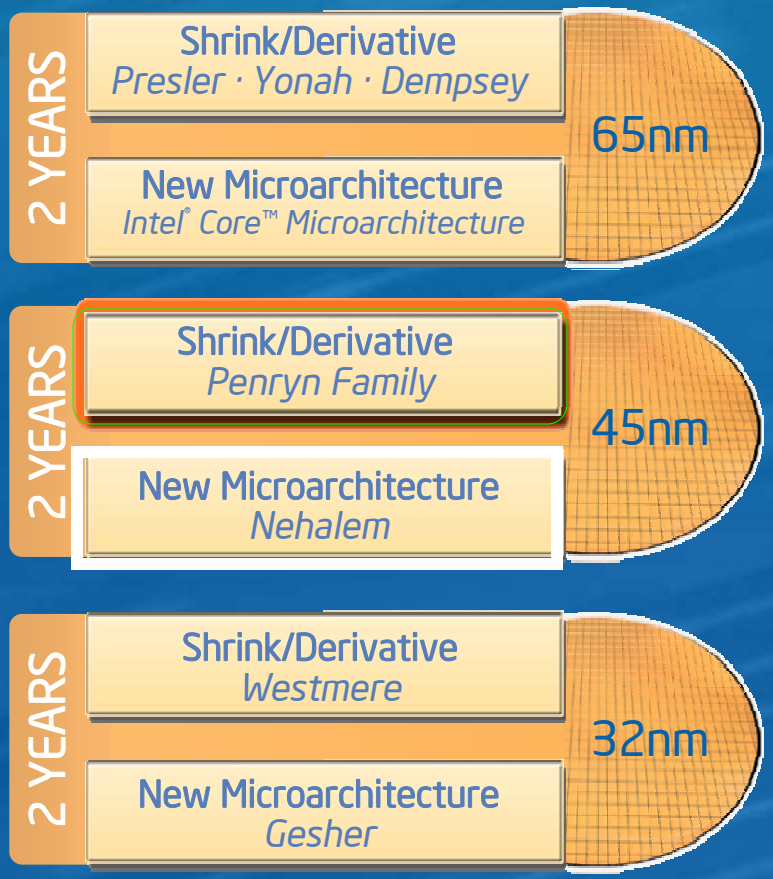
# Next Generation Intel 45nm High-k Process Technology




- ~2x larger transistor budget provides freedom to add new features and higher performance with cost effective die sizes
- >20% faster transistor switching speed delivers higher core speeds and increased instructions per clock
- Lower leakage current reduces power consumption and enables more capability and performance in a smaller power envelope than 65nm processors

**High-k + Metal Gate Transistors Provide Significant Performance Increase & Leakage Reduction, Ensuring Continuation of Moore's Law**

# Next, new 45nm microarchitecture




Increase performance per given clock cycle

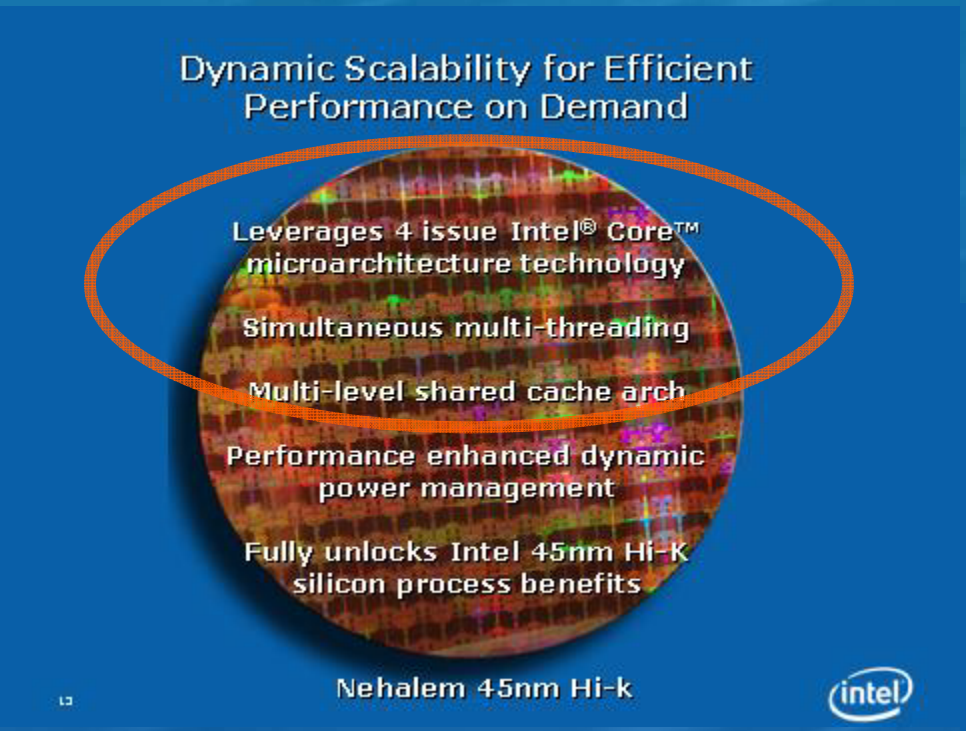
Increase processor frequencies

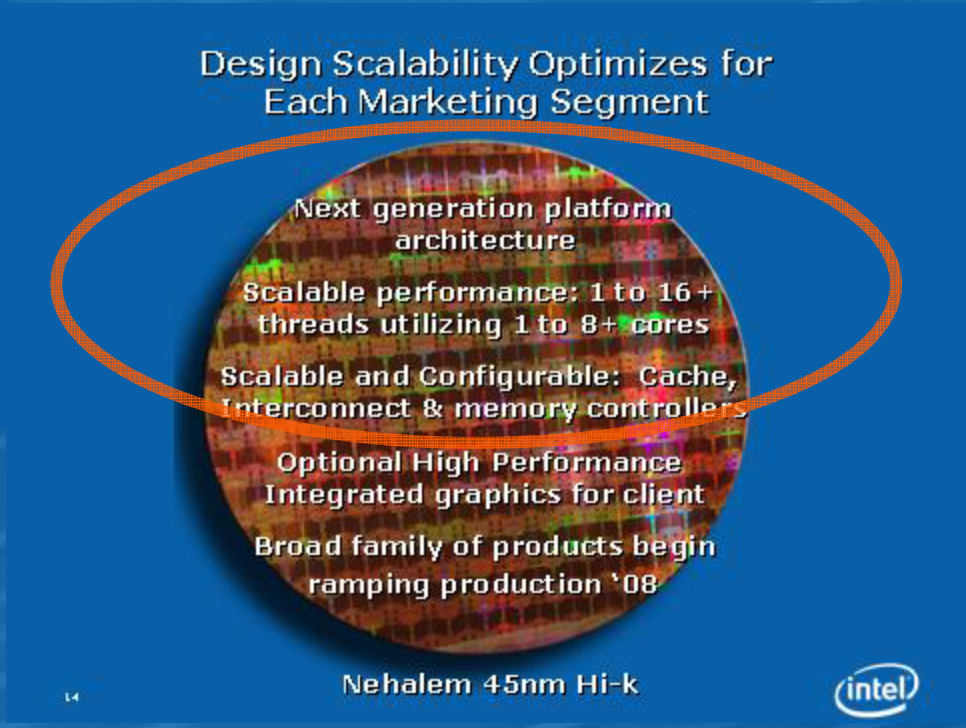
Extend energy efficiency

Deliver lead product for 45nm High k + metal gate process technology

Deliver optimized processors across each product segment and power envelope

# Nehalem: What We've Said Publicly



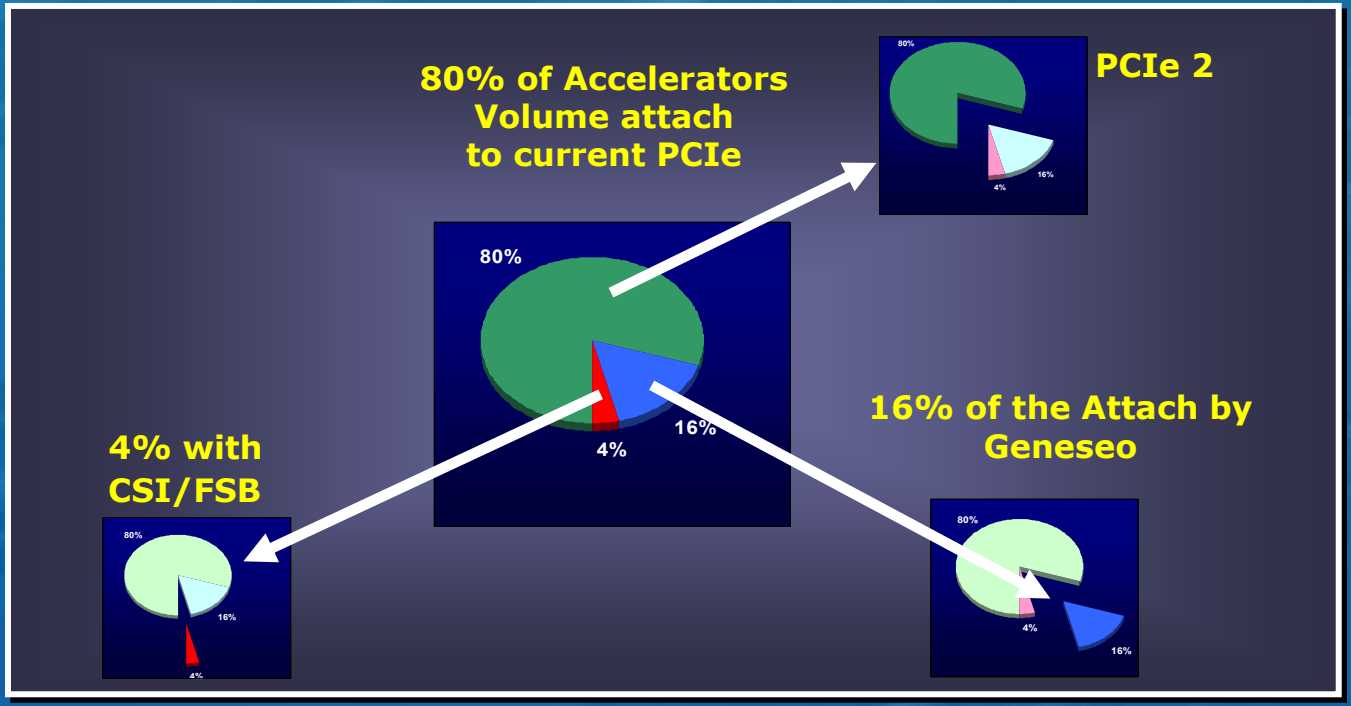



# High Density Intel Platforms

- Port Townsend – UP solution from EPSD
  - Best for bandwidth demanding applications
  - Bensley equivalent CPU density/rack (2 CPU/1U)
  - 1Tflop in 12U!
- Atoka – DP solution from SGI & Super Micro
  - Maximum compute density (4 CPU/1U)
  - >1TF in 15 nodes!

**Two compelling alternatives for high density designs**

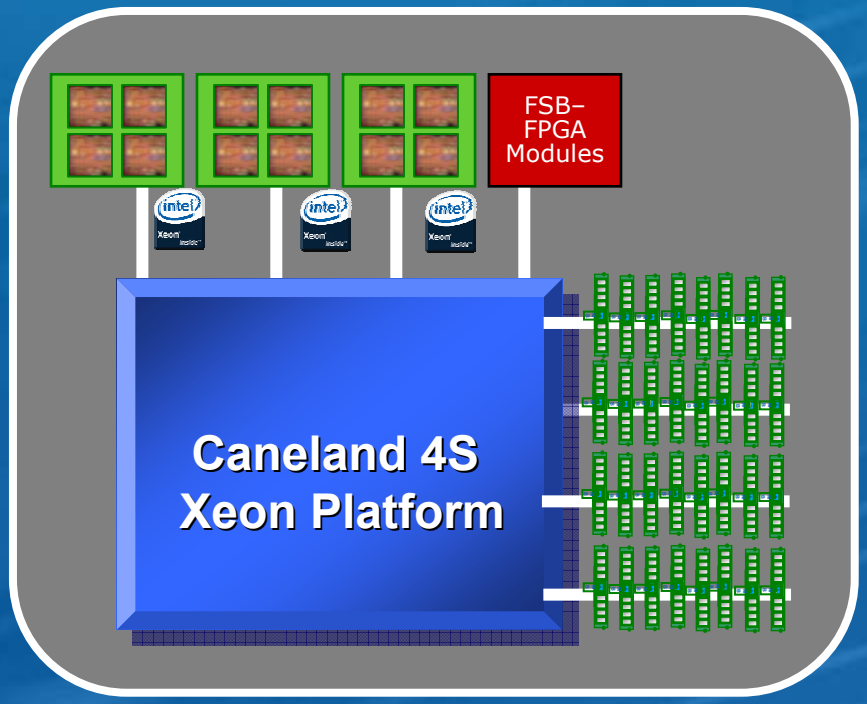
# Accelerator Attach Strategy



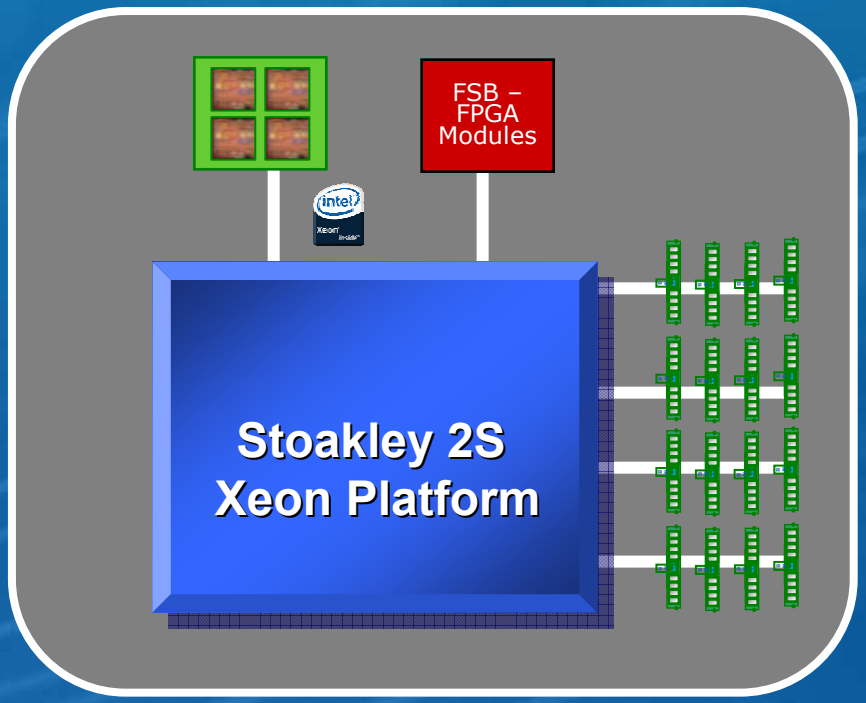

- Strategy
  - Ensure third-party innovation & differentiation on Intel platforms
  - Create common coprocessor API/Libraries framework strategy that encompasses all attach technologies
  - Work with the industry to innovate on ubiquitous PCI-E standard
  - Special licensing for FSB & CSI based accelerators, where appropriate

# What are FSB-FPGA Accelerators?

- FPGA Modules plugged into an Intel® Xeon® Processor socket & communicating using the Front Side Bus (FSB) Protocol




7000 Sequence




5000 Sequence

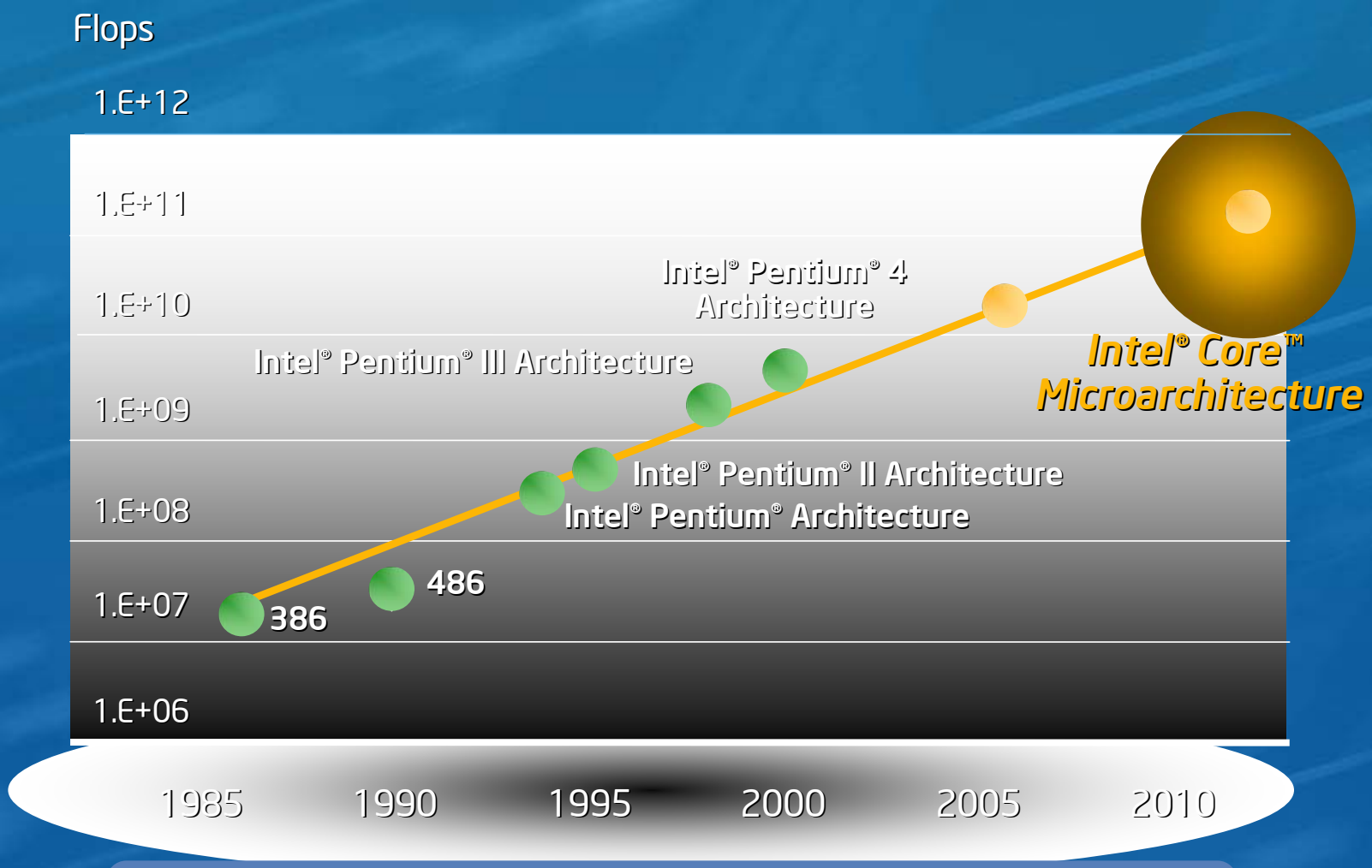
***Intel® Xeon® Server Platforms – DP & MP***

Let's Talk About ...

# Challenges and Outlook in Building a Petascale Machine

intel

# Processor Performance

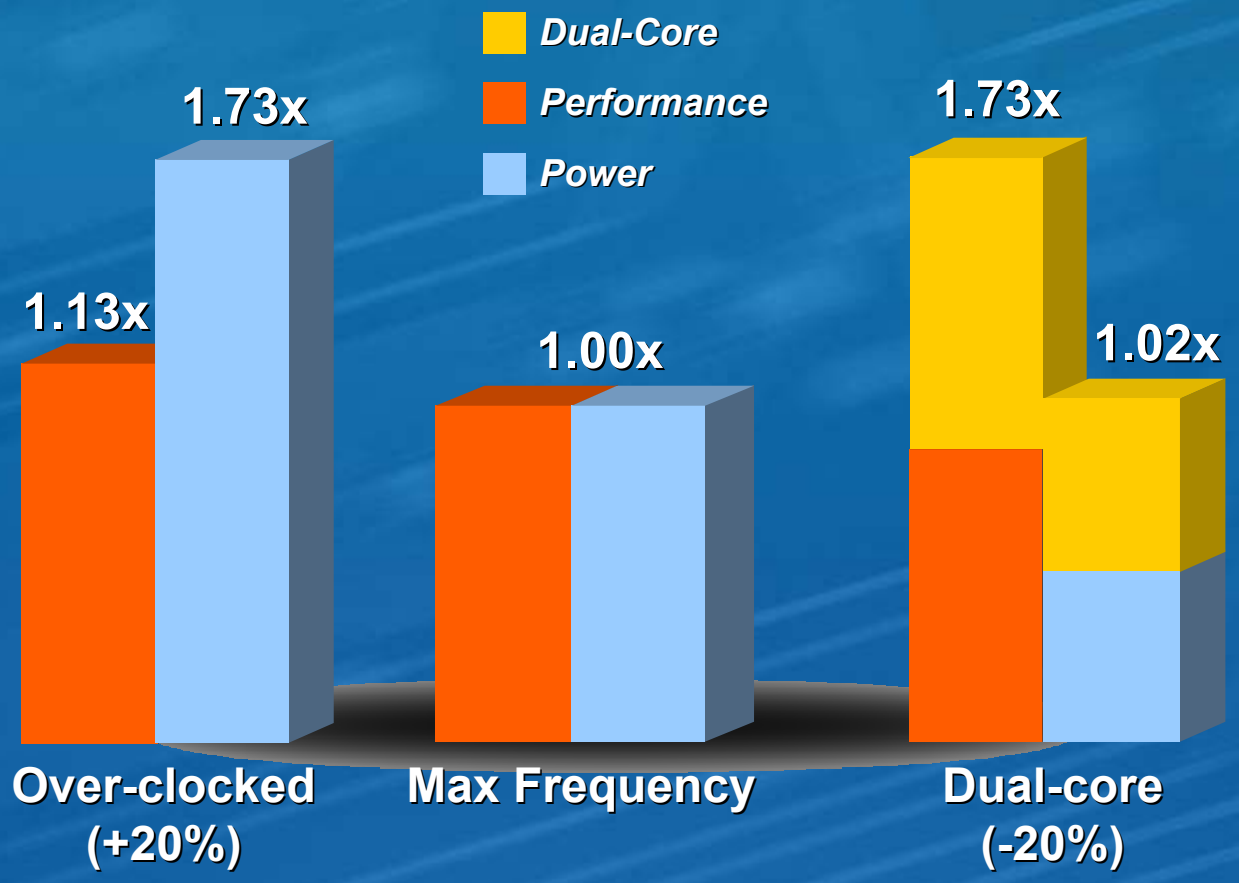



***Reaching Petascale with ~100,000 Processors in 2010****

Assuming approx. 100Glops processors

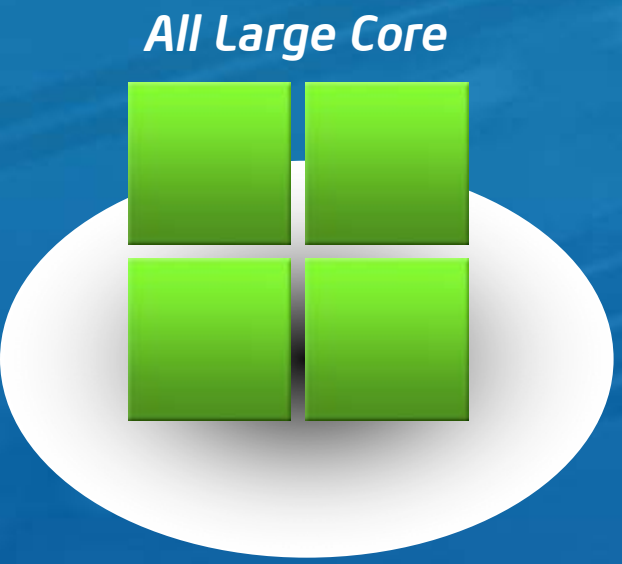
* Petascale assumes 10's of PF Peak Performance and 1PF Sustained Performance on HPC Applications.

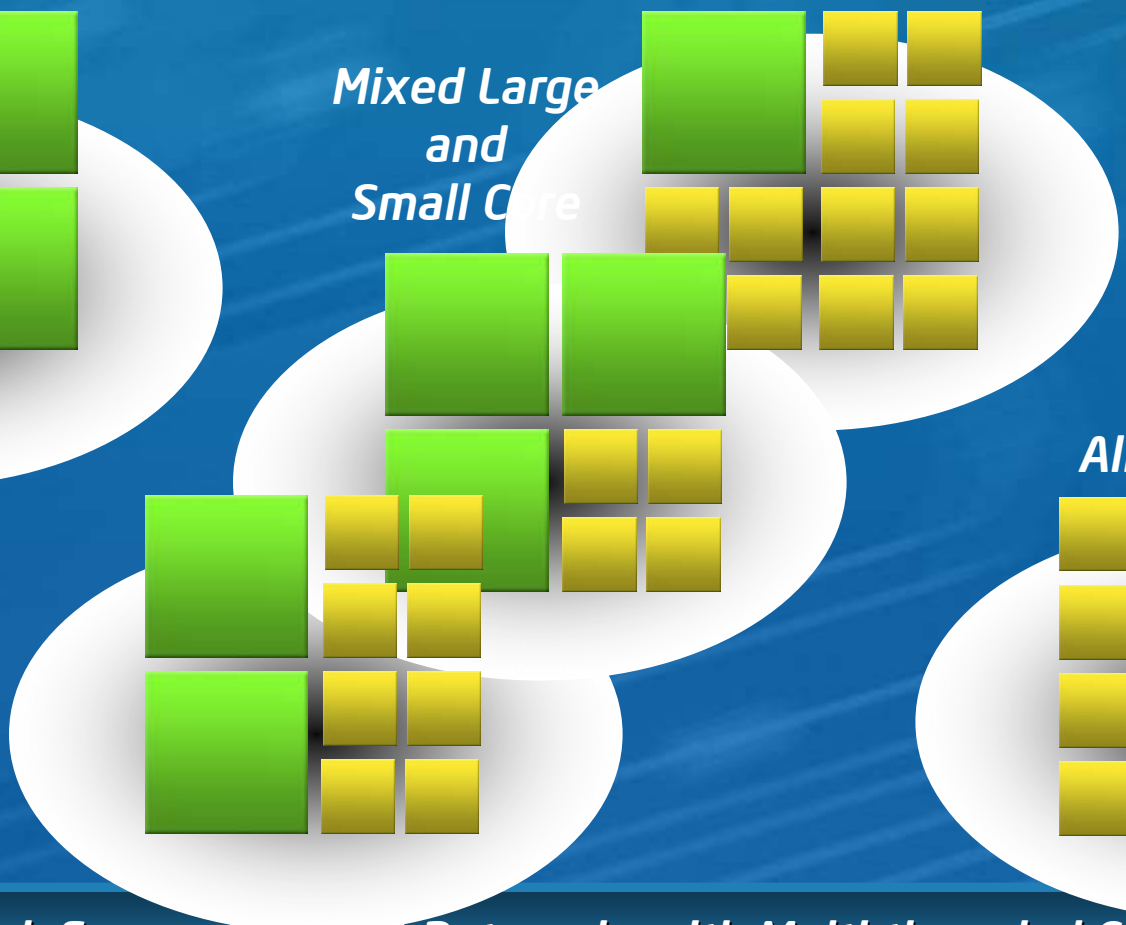
# Multi-Core
# Energy-Efficient Performance

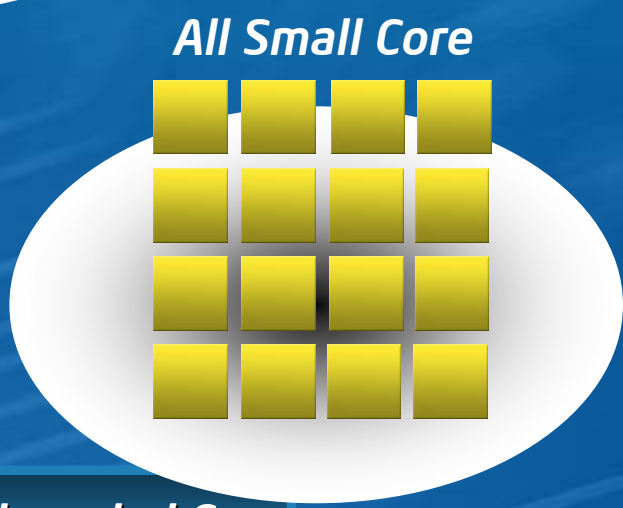



*Relative single-core frequency and Vcc*

# Multi-threaded Cores



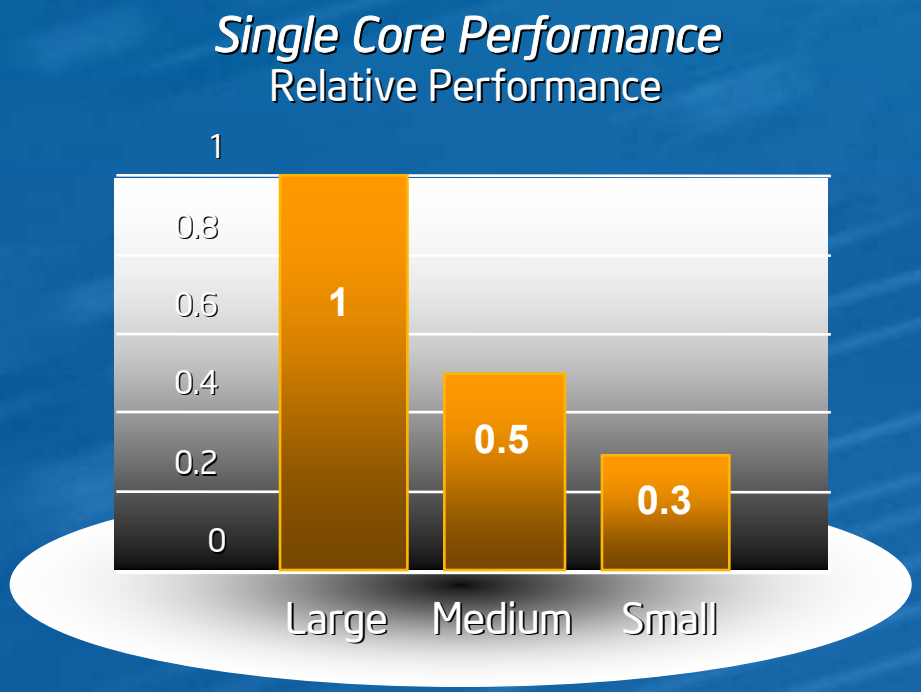



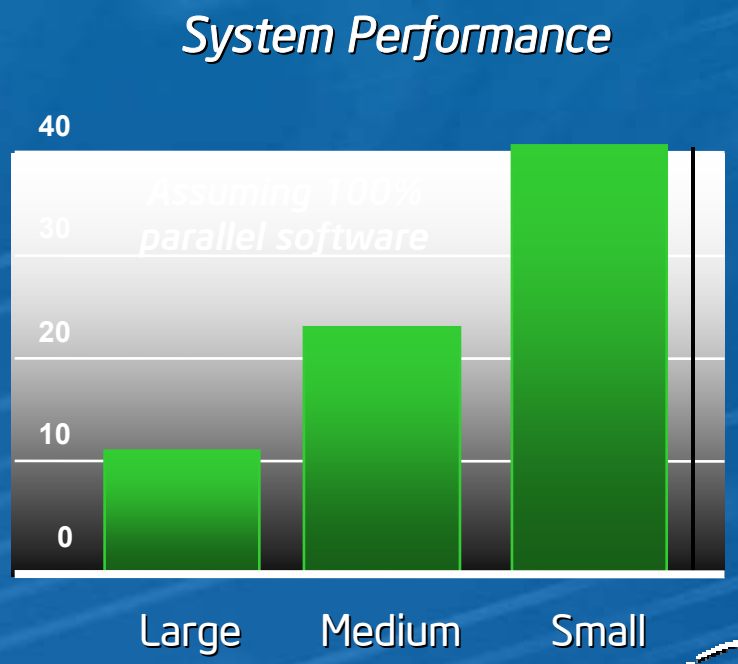

*Goal: Energy Efficient Petascale with Multi-threaded Cores*

Note: the above pictures don't represent any current or future Intel products

# Increasing Throughput through Parallelism

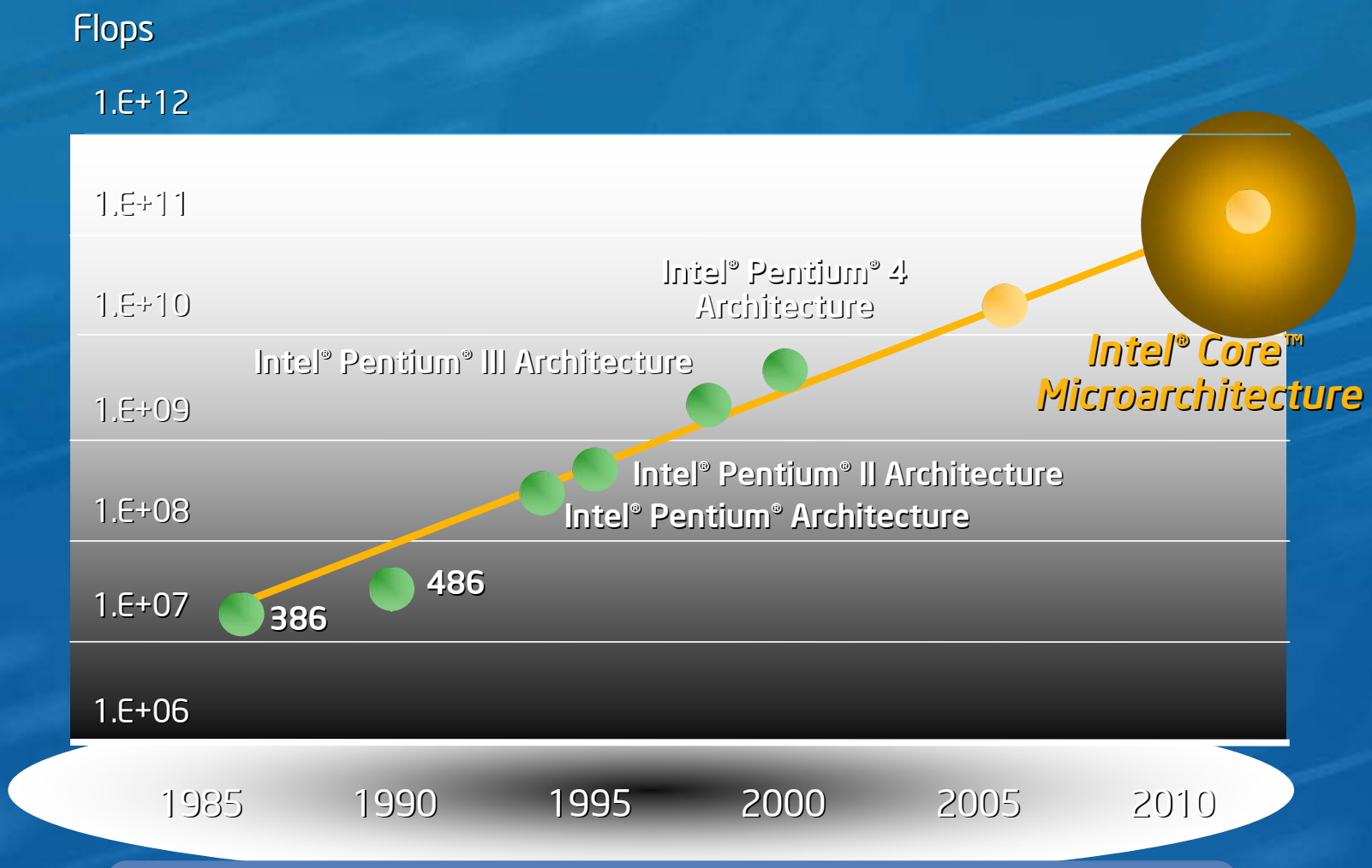
***Amdahl's Law: Parallel Speedup = 1/(Serial% + (1-Serial%)/N*)***




* N = number of cores

# Processor Performance



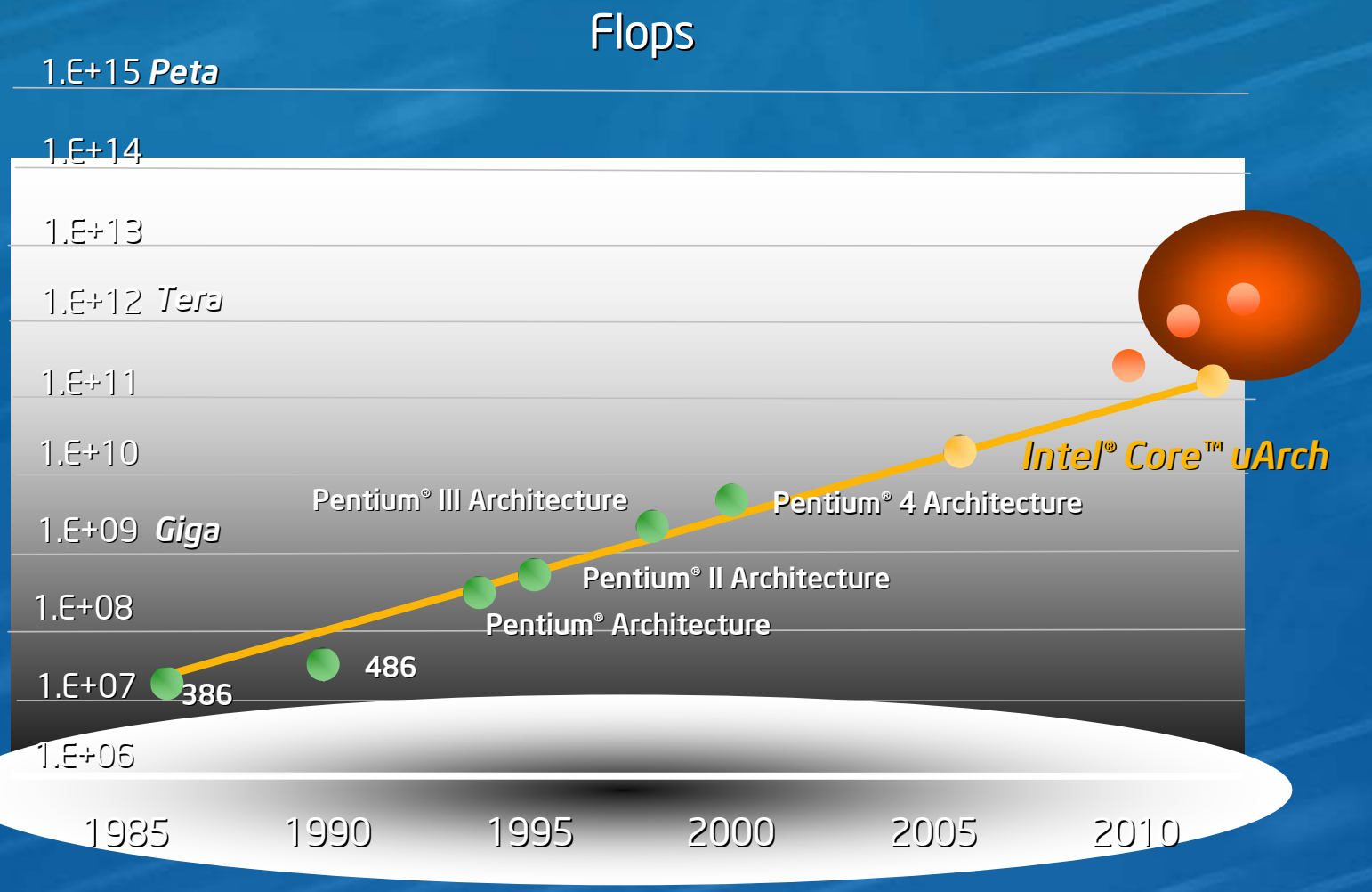

***Reaching Petascale with ~100,000 Processors in 2010****

Assuming approx. 100Glops processors

* Petascale assumes 10's of PF Peak Performance and 1PF Sustained Performance on HPC Applications.

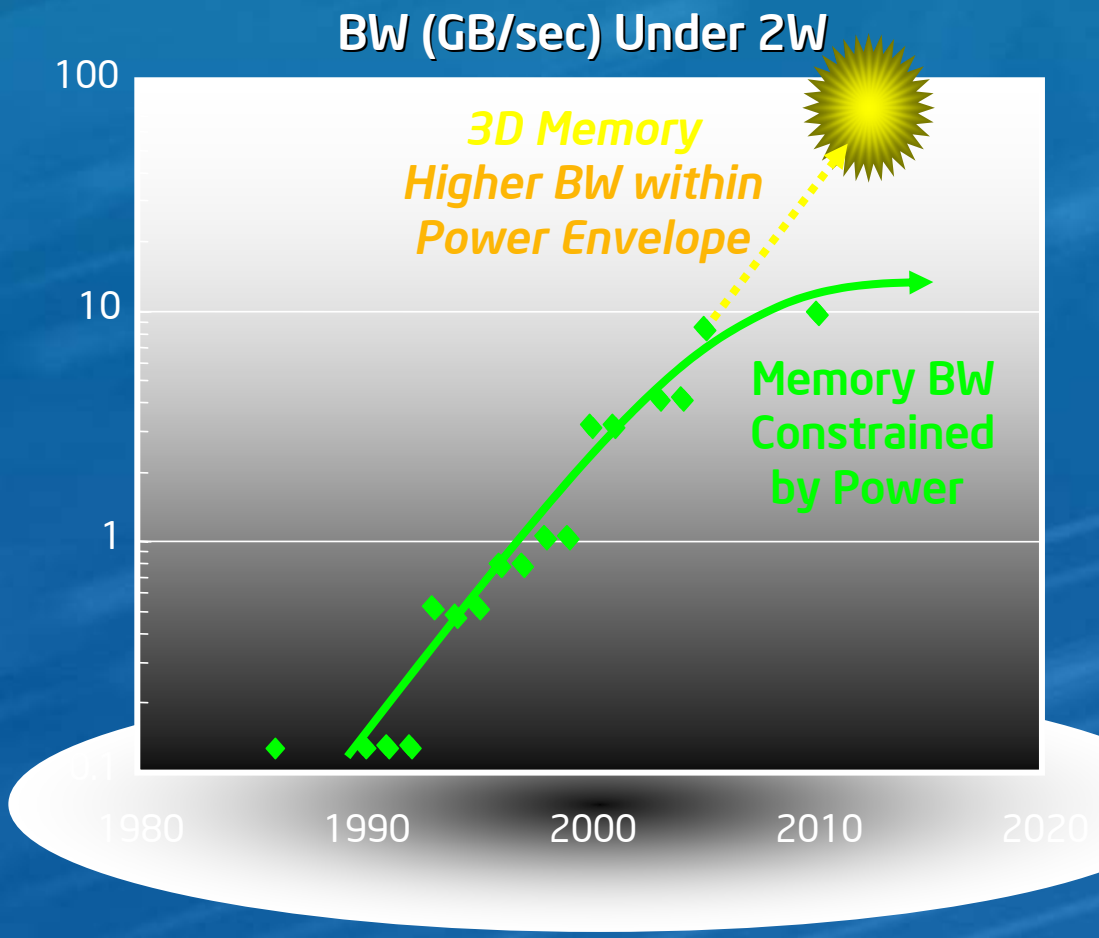
# Increasing Processor Performance

## *Through Multi-threaded Cores*




***Reaching Petascale with ~5,000 Processors***

# Increasing Memory Bandwidth
## *to Keep Pace*

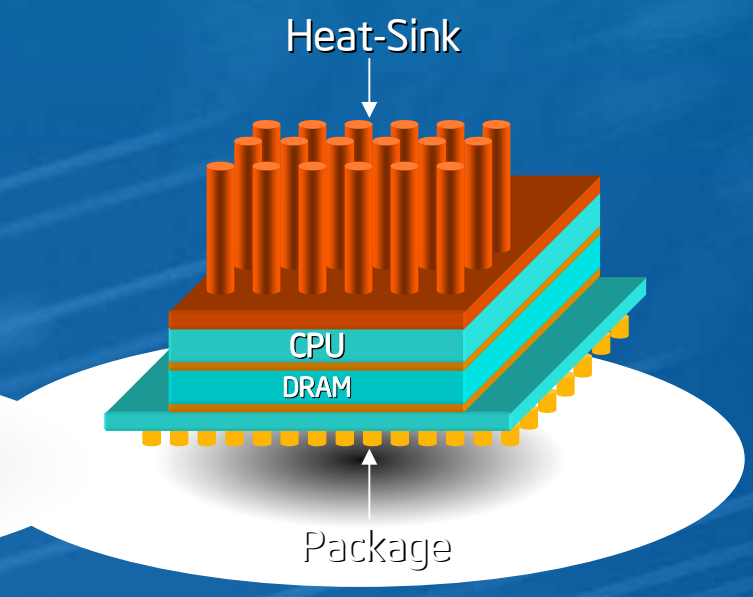



## *3D Memory Stacking*

*Power and IO Signals Go Through DRAM to CPU*

*Thin DRAM Die*

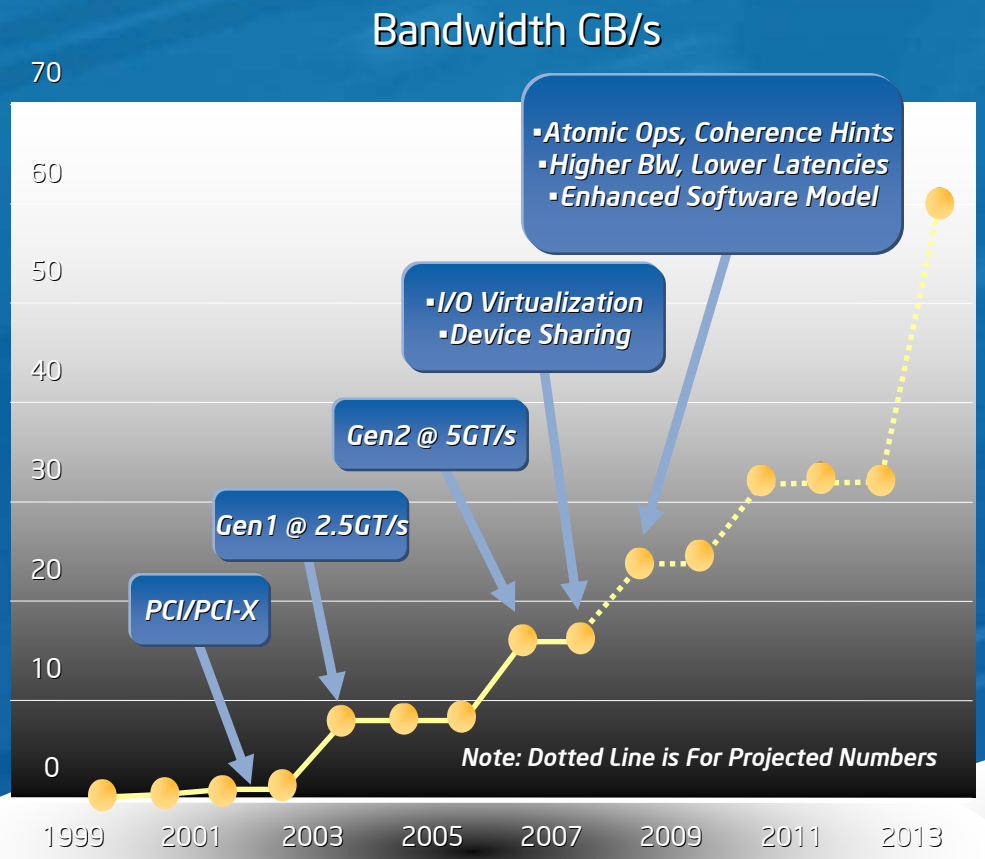
*Through DRAM Vias*



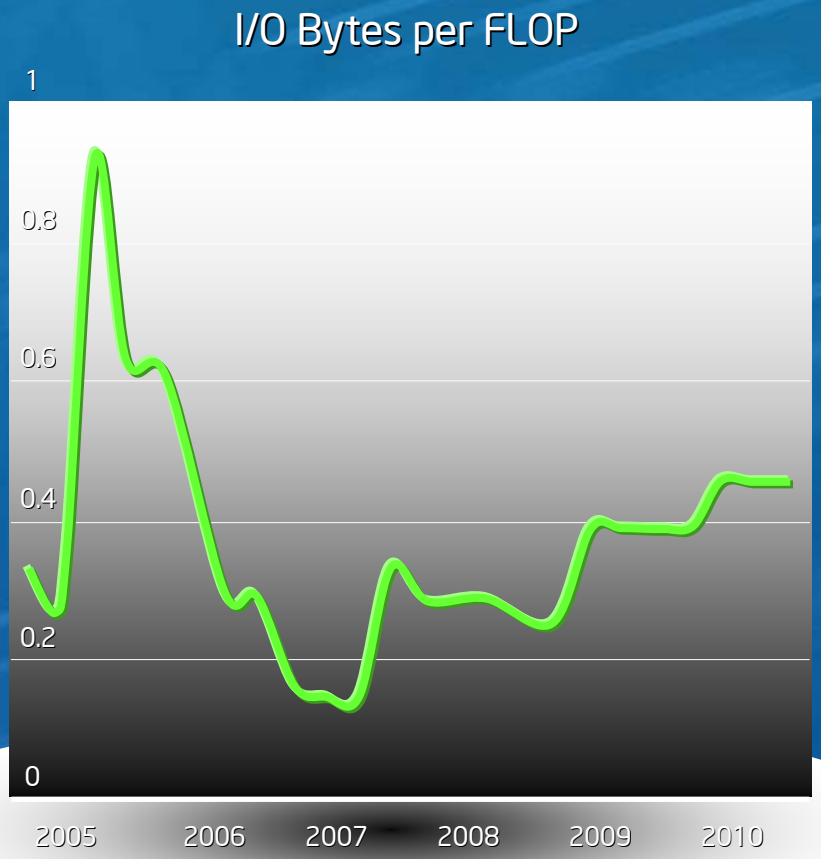

Source: Intel

# PCI Express to Meet I/O Demand

## *Performance, Bandwidth and Functionality*




*Based on x16 PCIe channel*




***Tracking Moore's Law***

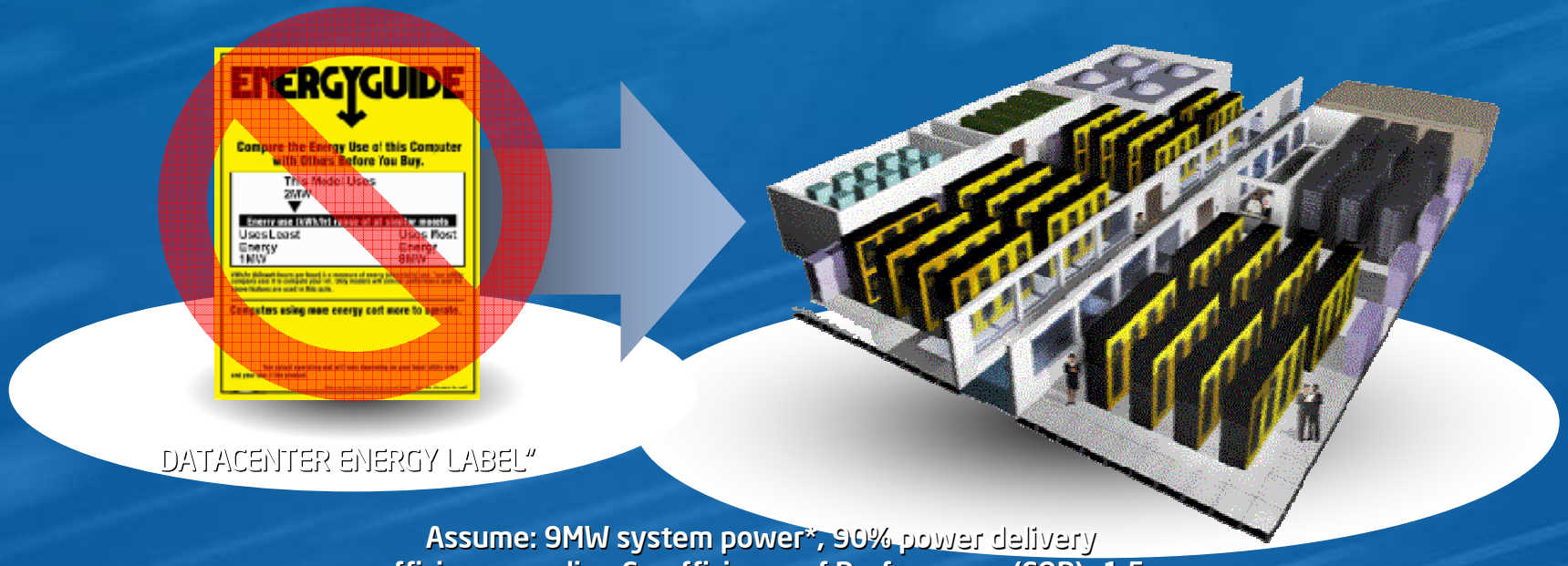
Source: Intel

*But* ... How about

# Power and Reliability

# Power and Cooling Cost Today

**10¢** *Kilowatt Hour* + **9** *Megawatt Datacenter* + *Power Delivery + Cooling* = **$14.6M** *Electricity Costs/Year*




Assume: 9MW system power*, 90% power delivery efficiency, cooling Co-efficiency of Performance (COP)=1.5

*Source: HPC Wire "A Petaflop Before its Time," June 28, 2006

# Managing Power and Cooling Efficiency

**Silicon:**
Moore's law, Strained silicon, Transistor leakage control techniques, Clock gating

**Processor:**
Policy-based power allocation
Multi-threaded cores

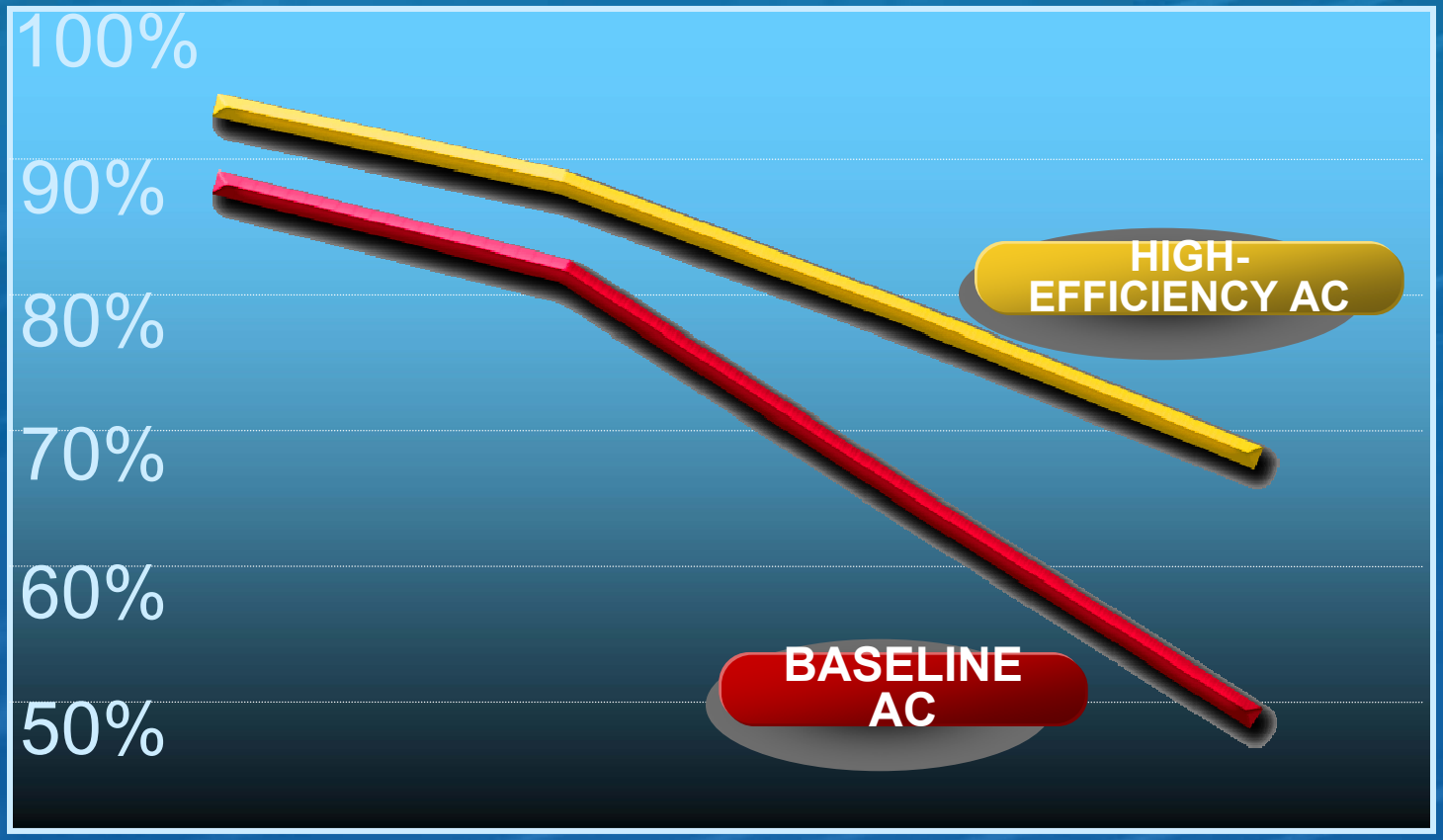
**System Power Delivery:**
Fine grain power management,
Ultra fine grain power management

**Facilities:**
Air cooling and liquid cooling options
Vertical integration of cooling solutions

***Power Management: From Transistors to Facilities***

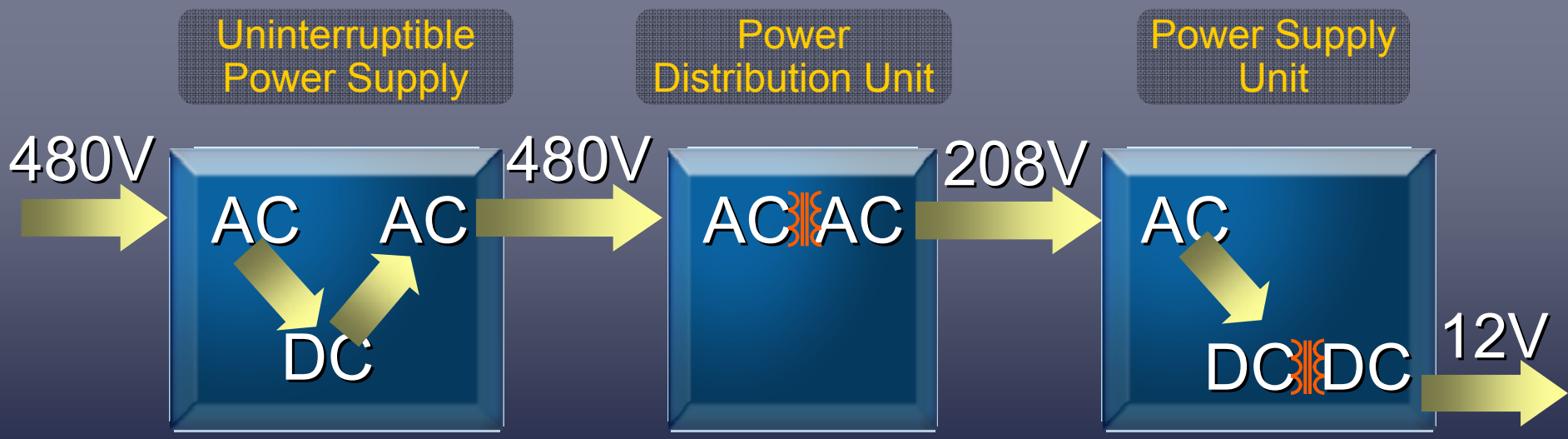
# DELIVERY INEFFICIENCY

CUMULATIVE DATACENTER POWER DELIVERY EFFICIENCY



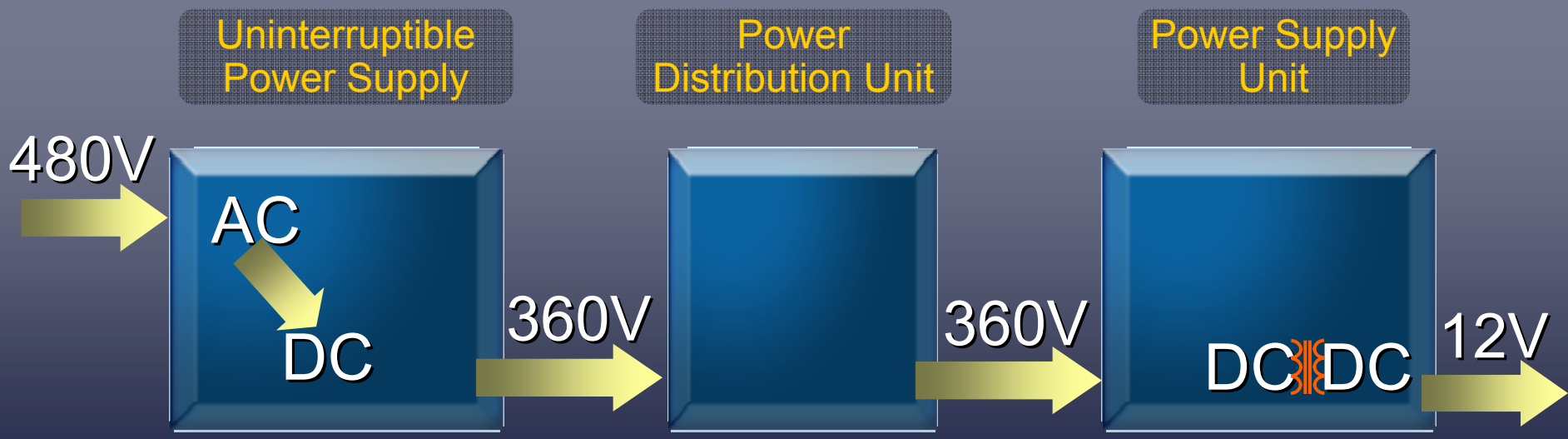

Source: Intel

# CONVERSION OVERKILL




CONVENTIONAL DATACENTER
AC POWER DISTRIBUTION

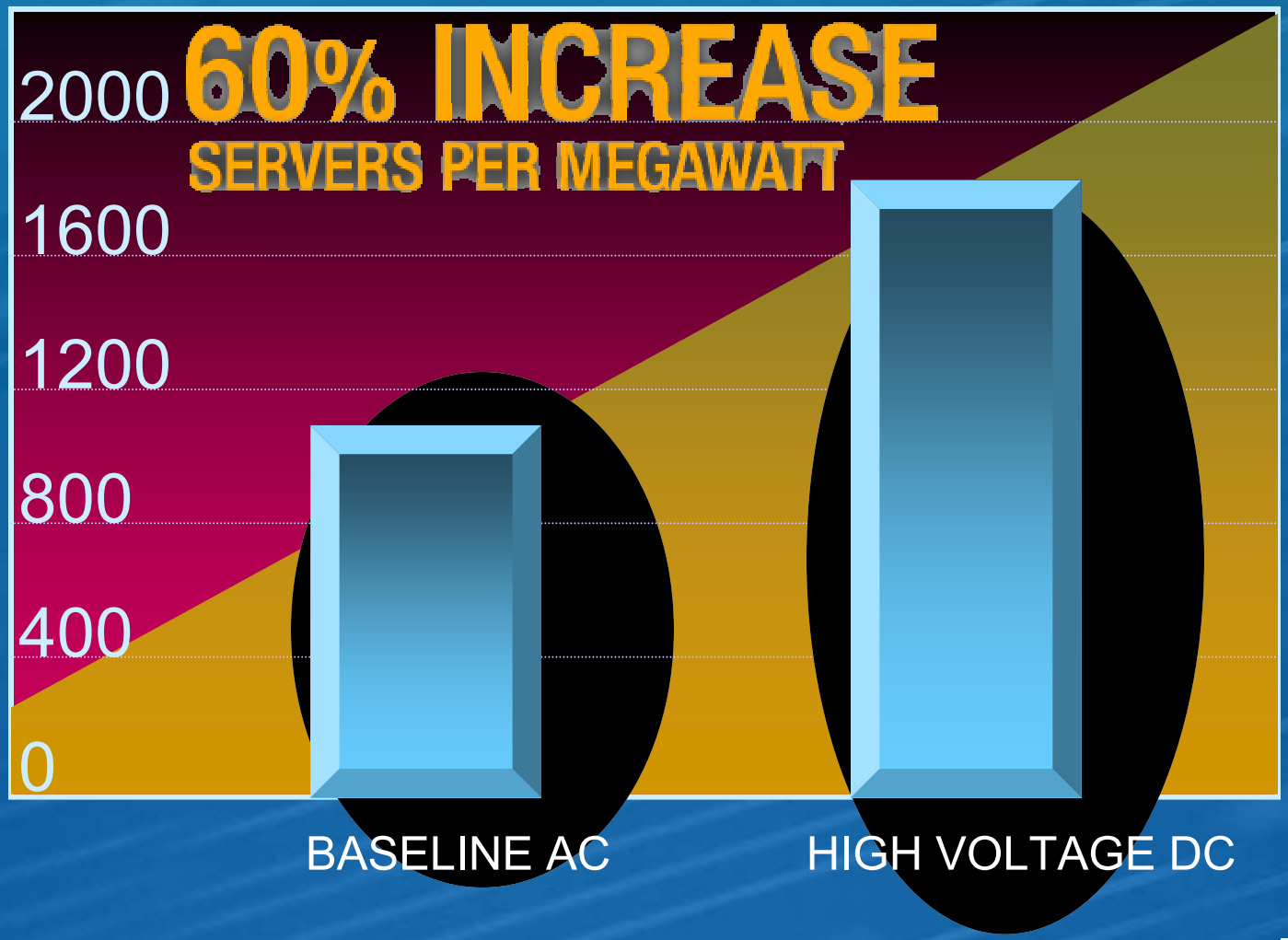
# SIMPLIFIED DISTRIBUTION




HIGH VOLTAGE DC POWER DISTRIBUTION

# HIGH VOLTAGE DC
# DATACENTER PROTOTYPE

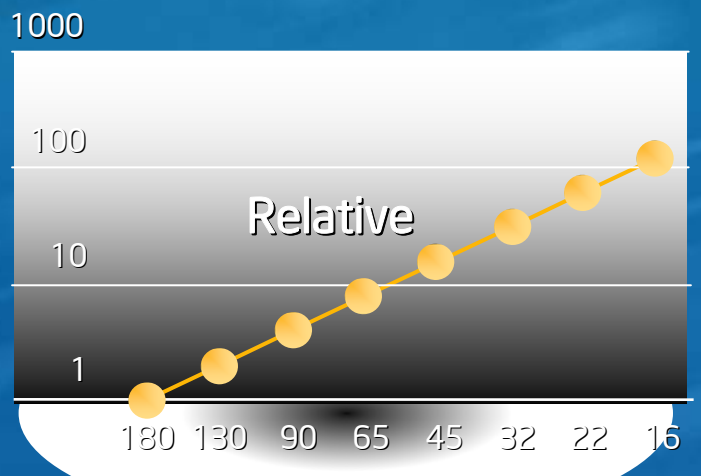
# EFFICIENCY REALIZED




Source: Intel

# Reliability Challenge
## *Billions of Transistors*

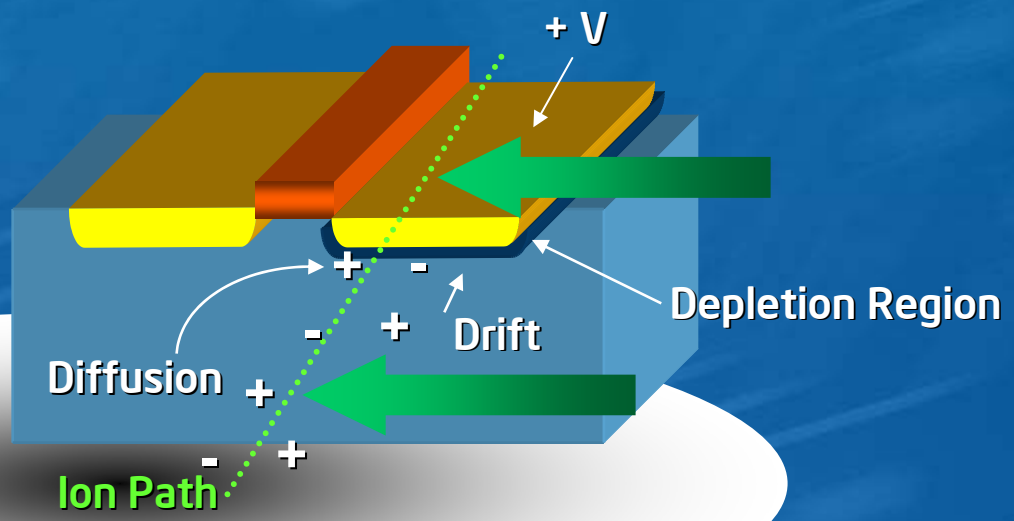
Soft Error FIT/Chip (Logic & Mem)




Assume: Chip size stays roughly the same with each generation.

- ***FIT/bit (mem cell)***: expected to be roughly constant
- ***Moore's law***: increasing the bit count exponentially: 2x every 2 years

***An exponential growth in FIT/chip***



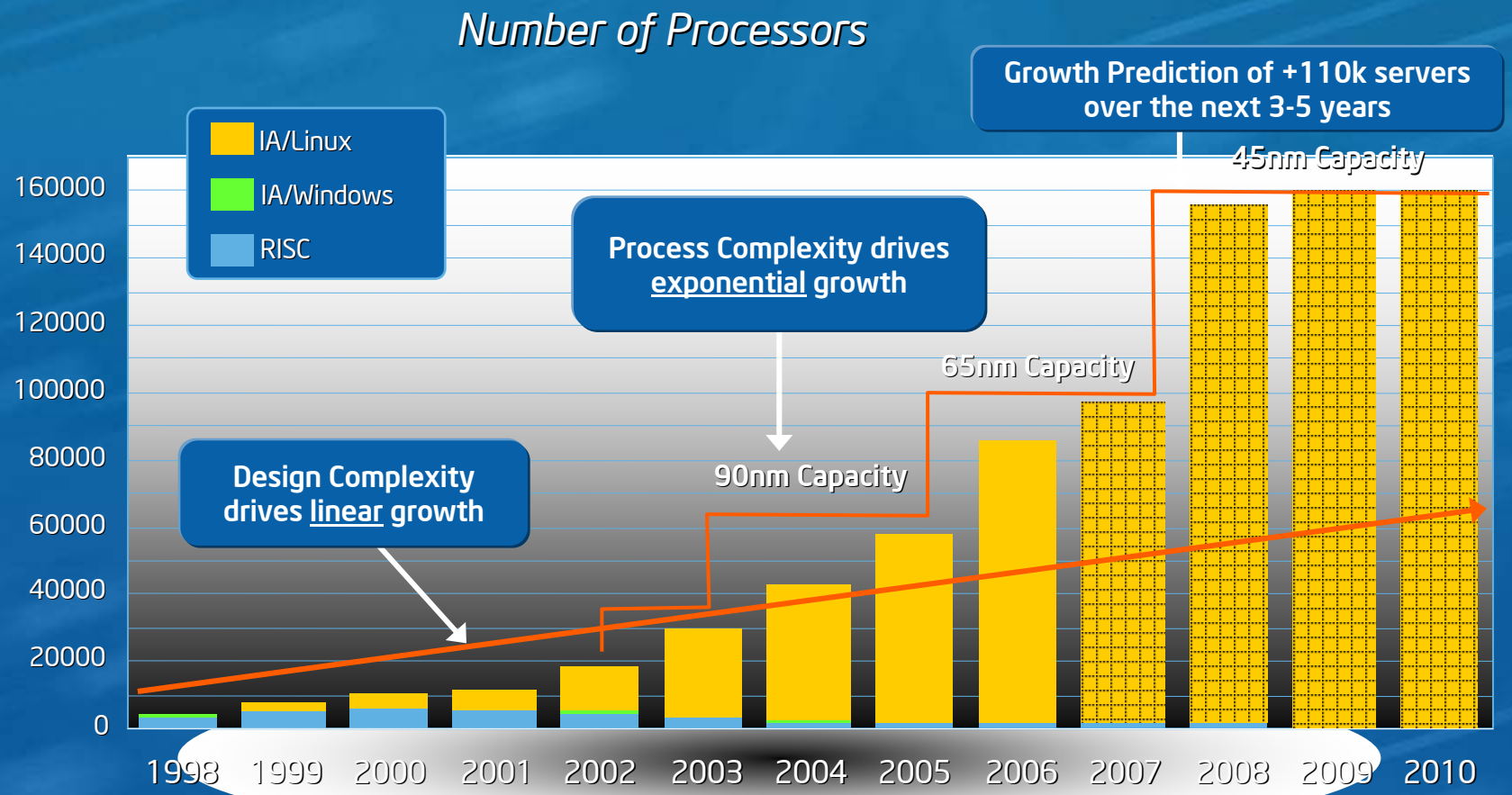

***Soft Errors or Single Event Upsets (SEU) are caused by charge collection following an energetic particle strike.***

***Soft Error: One of Many Challenges***

Source: Intel

# 100s & 1000s of Processors

## *An Example of Datacenter Growth …*




***Compute Capacity Growth is Accelerating***

Source: Intel

*Other names and brands may be claimed as the property of others

# Reliable Systems With Unreliable Components

*Architectural Techniques*

**Micro Solutions**

Parity
SECDED ECC
π bit

**Macro Solutions**

Lockstepping
Redundant multithreading (RMT)
Redundant multi-core CPU

*Circuit Techniques*

**Device Param Tuning**

**Rad-hard Cell Creation**

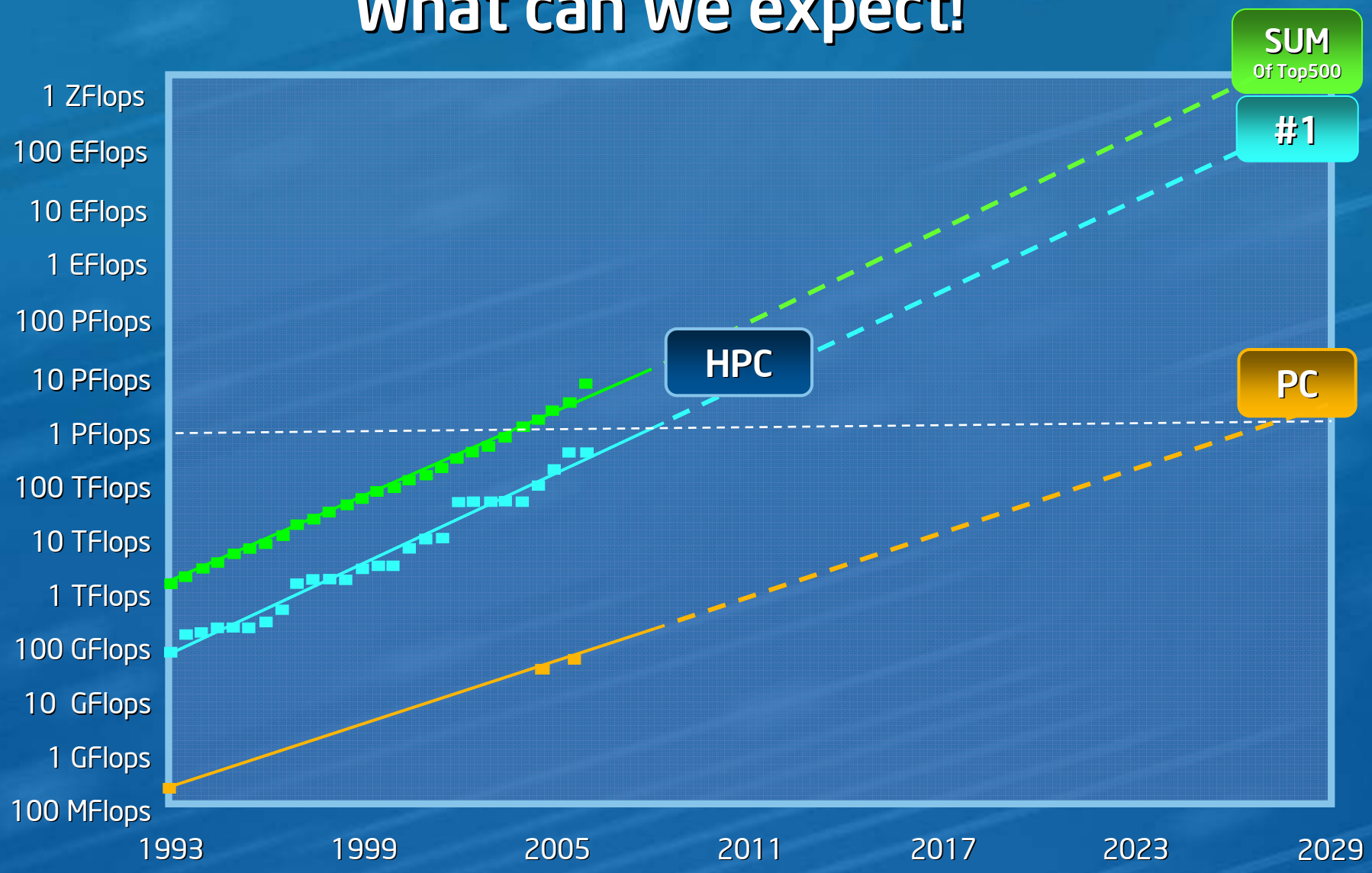
*Process Techniques*

**State-of-Art Processes**

**Reducing Single-Bit Soft Errors**

intel

# What can we expect!




Source: HPC - www.top500.org, June 2006, Intel

# HPC @ Intel