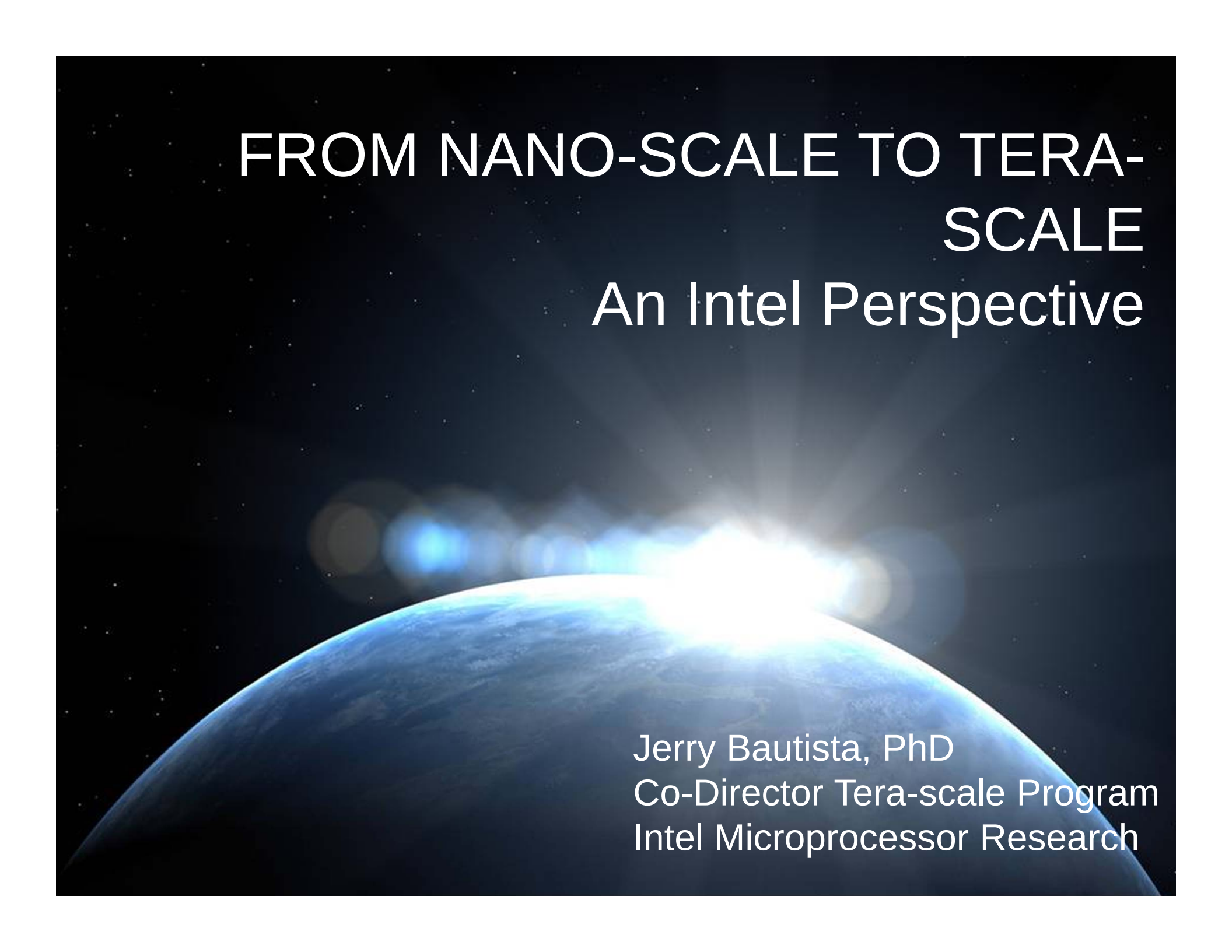




FROM NANO-SCALE TO TERA- SCALE

An Intel Perspective

Jerry Bautista, PhD
Co-Director Tera-scale Program
Intel Microprocessor Research

Corporate Technology Group (CTG) Strategic Objectives

Conduct world-class research

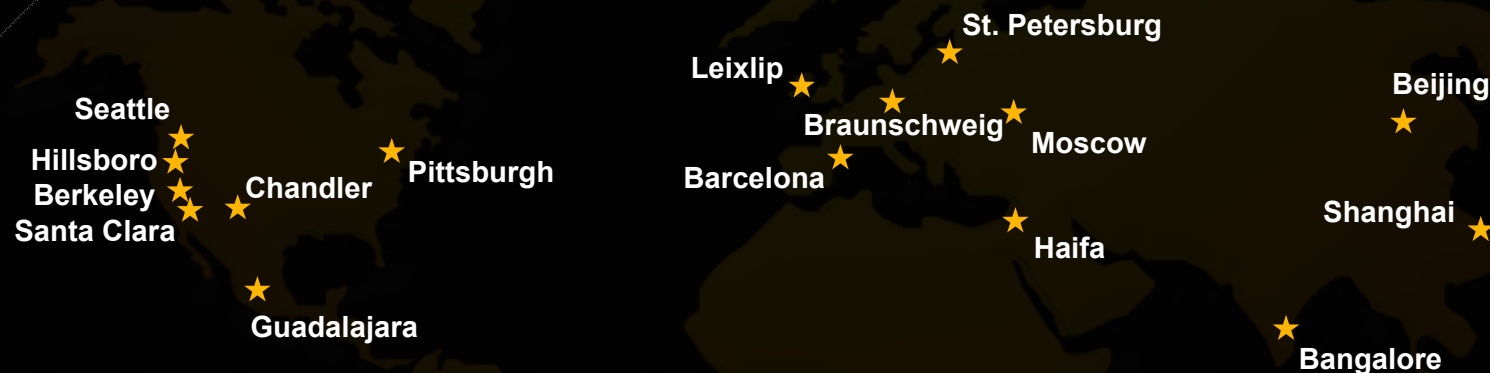
Deliver innovative technologies from
concept to product adoption

Collaborate with industry via
standards, alliances and evangelism

Engage worldwide for the best
research and technology



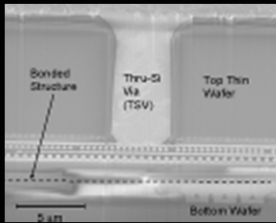
Research at Intel



- Nearly 1000 researchers
- 15 locations worldwide
- Innovative research models



Future Technology Challenges



**MEMORY
AND
STORAGE**

**WIRELESS
AND
COMMS**

**ENERGY-
EFFICIENT
PLATFORMS**

**NEW
PLATFORM
CLASSES**

**TERA-
SCALE
COMPUTING**

10X

LOWER
POWER

NEW DEMANDS

HIGHER
PERFORMANCE

10X



Future Technology Challenges



WIRELESS
AND
COMMS

ENERGY-
EFFICIENT
PLATFORMS

NEW
PLATFORM
CLASSES

TERA-
SCALE
COMPUTING

10X LOWER
POWER

NEW DEMANDS

HIGHER
PERFORMANCE 10X



Nanodevice R&D

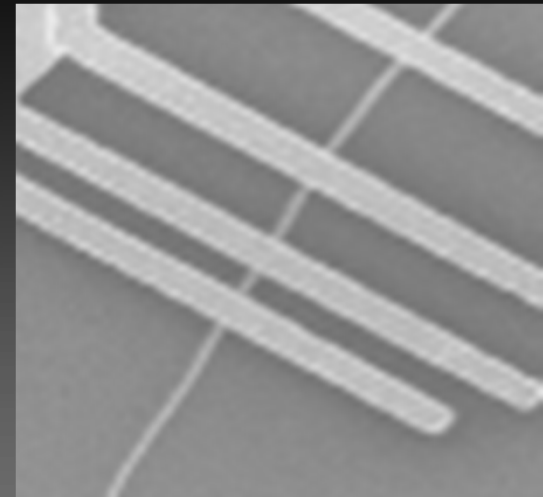
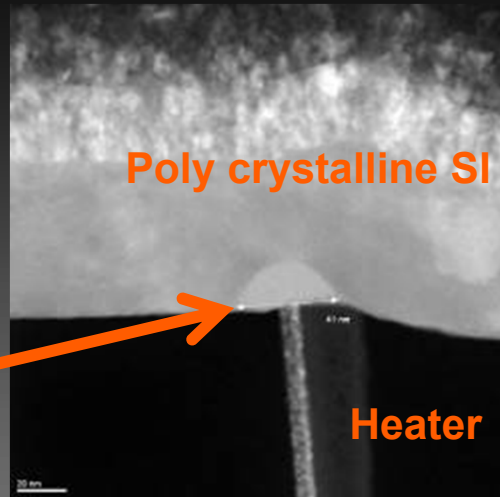
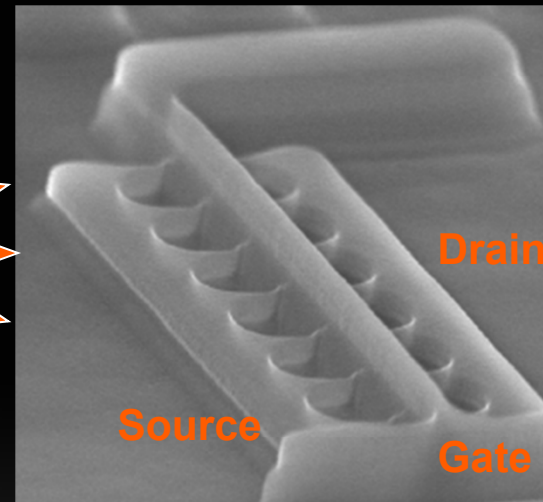
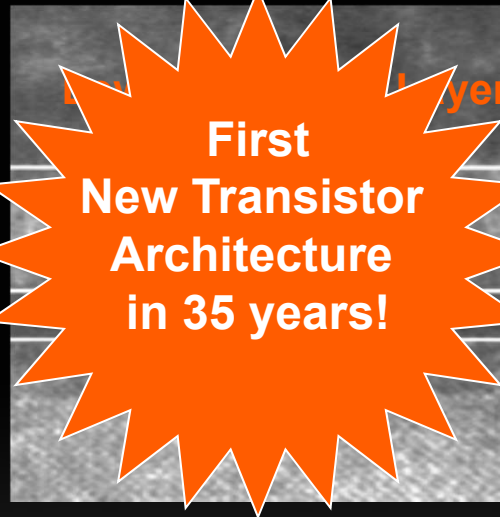
**High-K Metal
Gate
Transistor
at 45 nm**

**First
New Transistor
Architecture
in 35 years!**

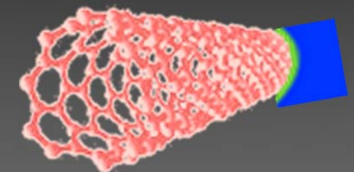
**Non-planar
Tri-Gate
Transistor
Beyond 32 nm**

**Phase Change
Memory (PCM)
Below 45 nm**

**Chalcogenide
Storage**



**Carbon
Nanotube
Transistor**

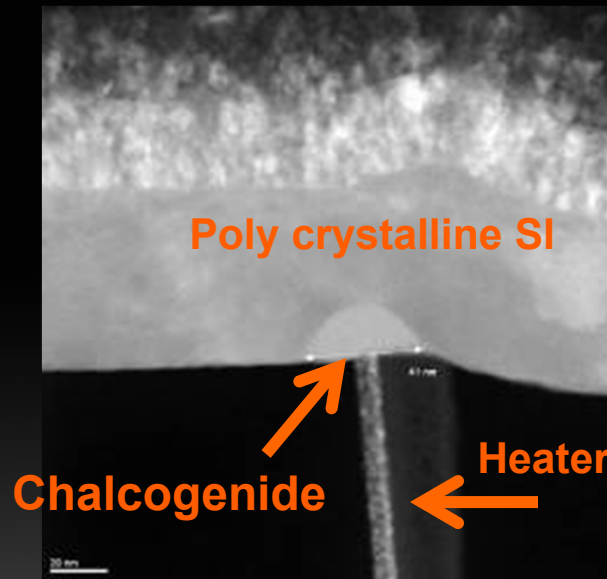


Source: Intel



Phase Change Memory

PCM Cell

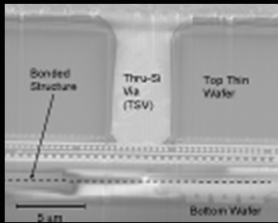


Source: Intel

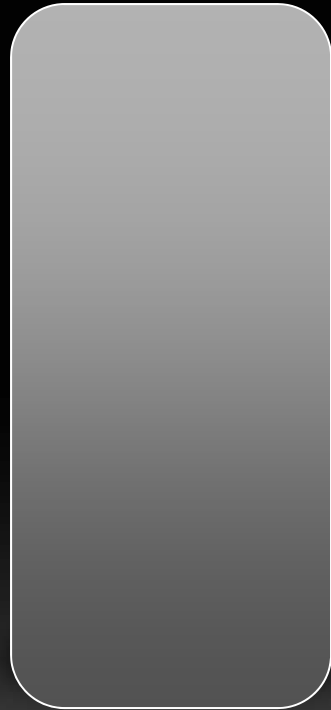
- Non volatile
- Bit alterable
- Very fast and low power
- Scalable well below 45 nm
- High endurance (>100M writes)
- Partners
 - Ovonyx Since 2000
 - R&D with ST since 2003



Future Technology Challenges



**MEMORY
AND
STORAGE**



**ENERGY-
EFFICIENT
PLATFORMS**



**NEW
PLATFORM
CLASSES**



**TERA-
SCALE
COMPUTING**

10X

LOWER
POWER

NEW DEMANDS

HIGHER
PERFORMANCE

10X

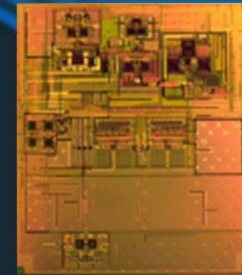


Digital Multi-Radio Technology

RF / Platform
Interference Mitigation

Digitally Enhanced
Radio RFIC

Today's
Notebook
Tomorrow's



7+ radios – 500cc
2 radios – 2500cc



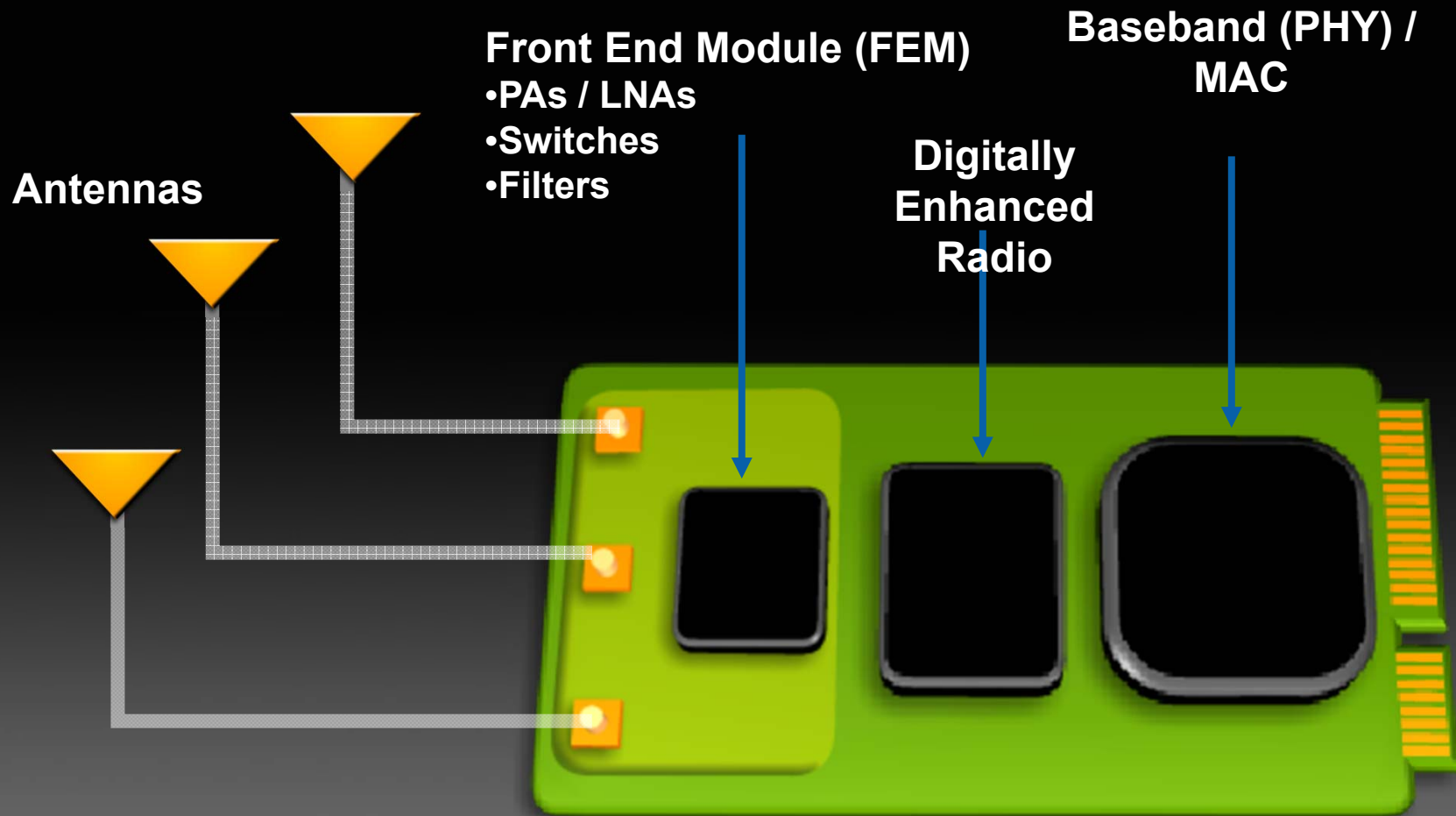
Scalable
Communication Core

Tunable Front-End
Modules & Antenna

Worldwide Regulatory

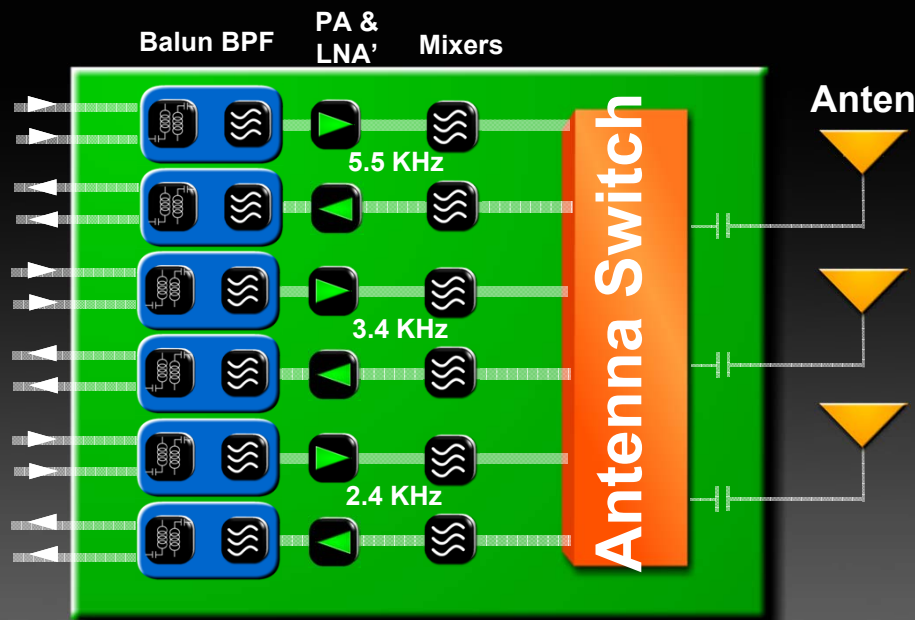


Digital Multi-Radio Components



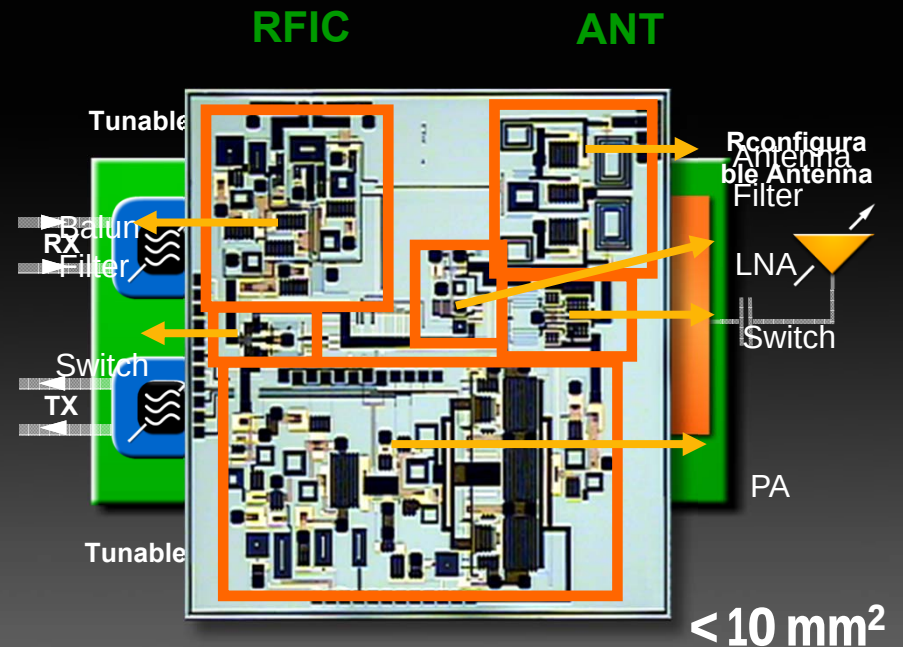
Conventional FEM v. Flexible FEM

Today's Conventional
Multi-Standard
(WiFi a/b/g) + WiMax FEM



\$5

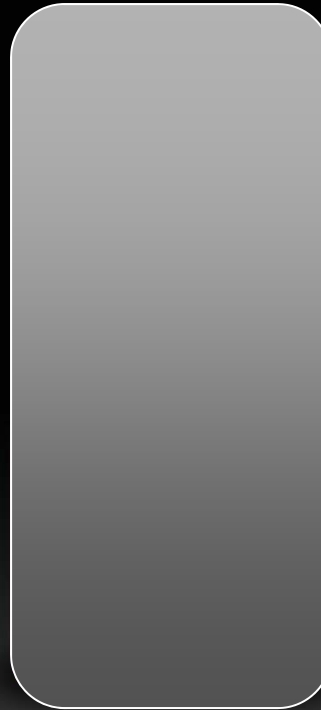
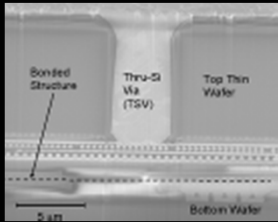
Next Generation
Conventional Multi
Standard
(WiFi a/b/g) + WiMax FEM



\$1.50



Future Technology Challenges



**MEMORY
AND
STORAGE**

**WIRELESS
AND
COMMS**

**NEW
PLATFORM
CLASSES**

**TERA-
SCALE
COMPUTING**

10X

LOWER
POWER

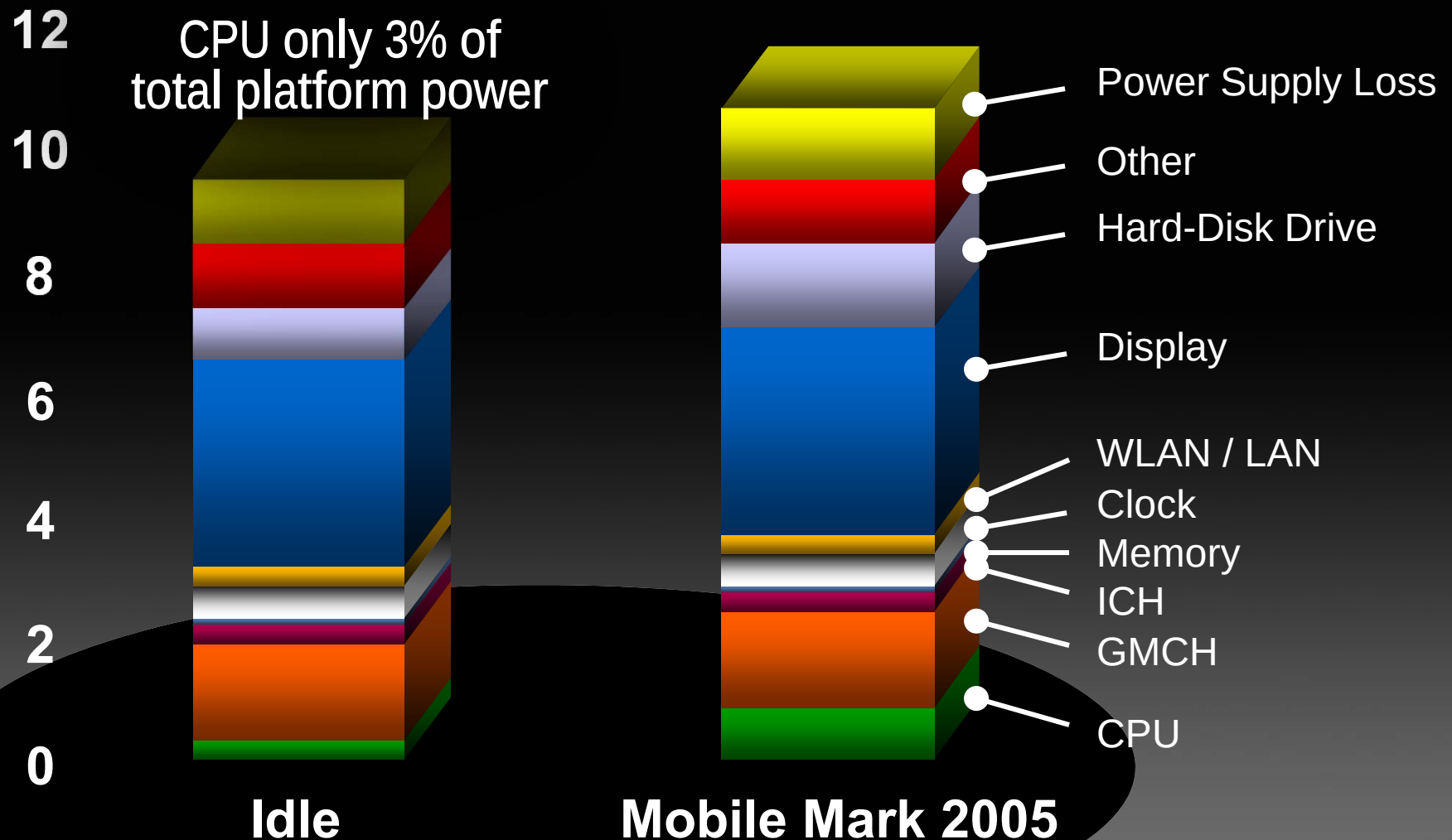
NEW DEMANDS

HIGHER
PERFORMANCE

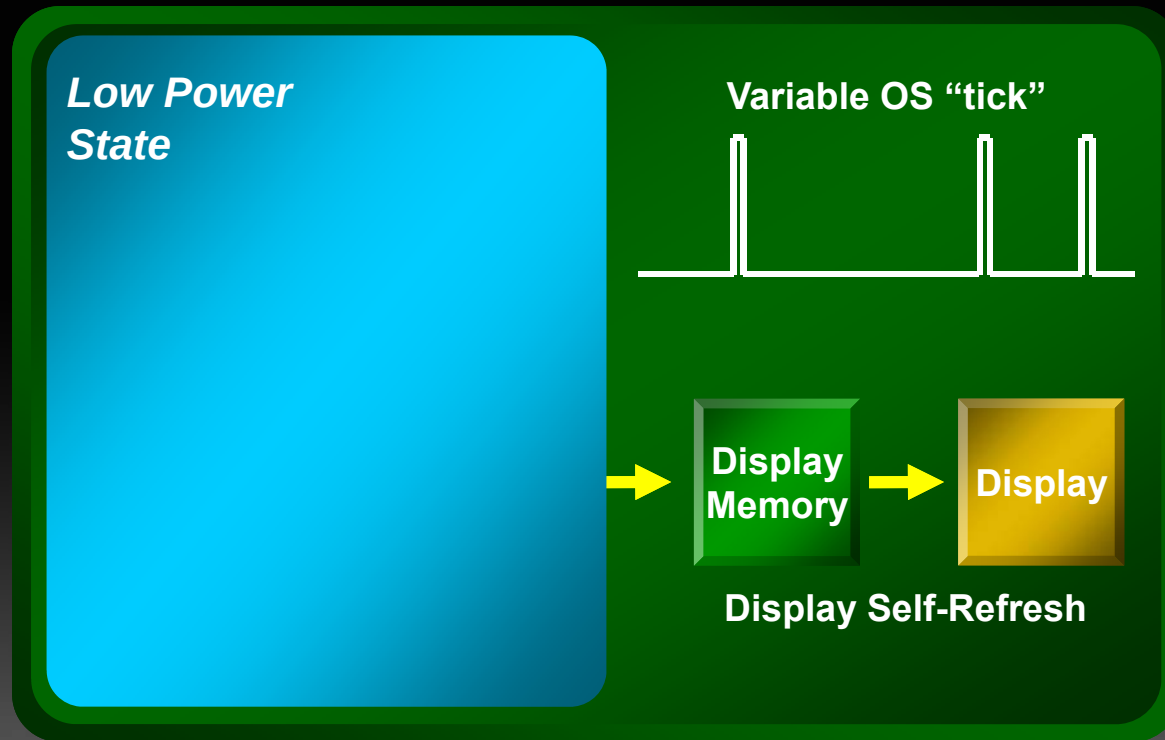
10X



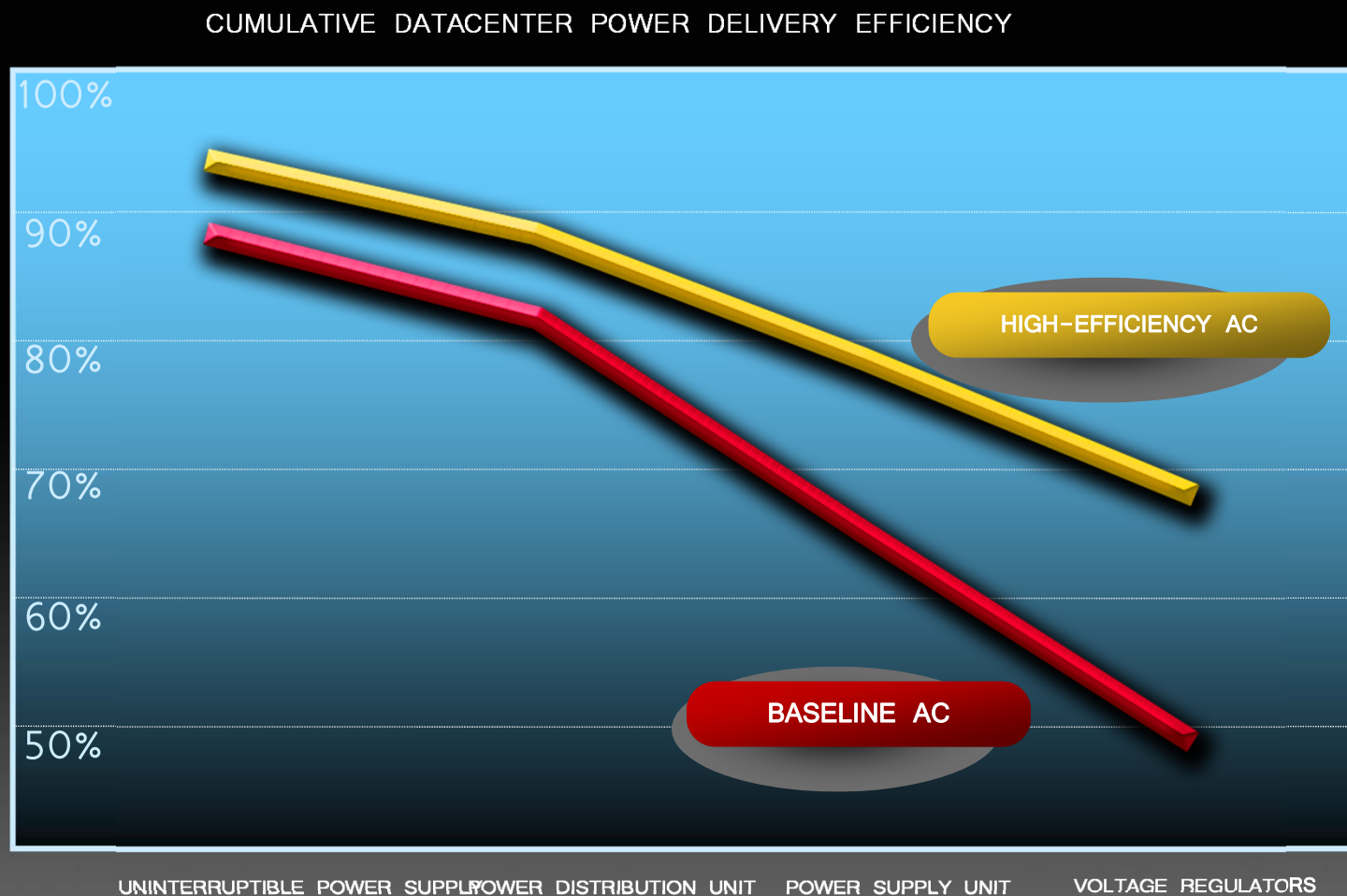
Energy Efficiency: Average vs. Idle



Display Self-Refresh Extended Idle Mode



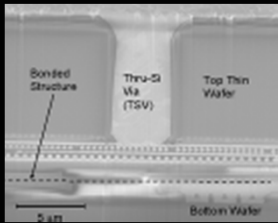
Power Delivery Inefficiency



Source: Intel



Future Technology Challenges



MEMORY
AND
STORAGE

WIRELESS
AND
COMMS

ENERGY-
EFFICIENT
PLATFORMS

TERA-SCALE
COMPUTING

10X

LOWER
POWER

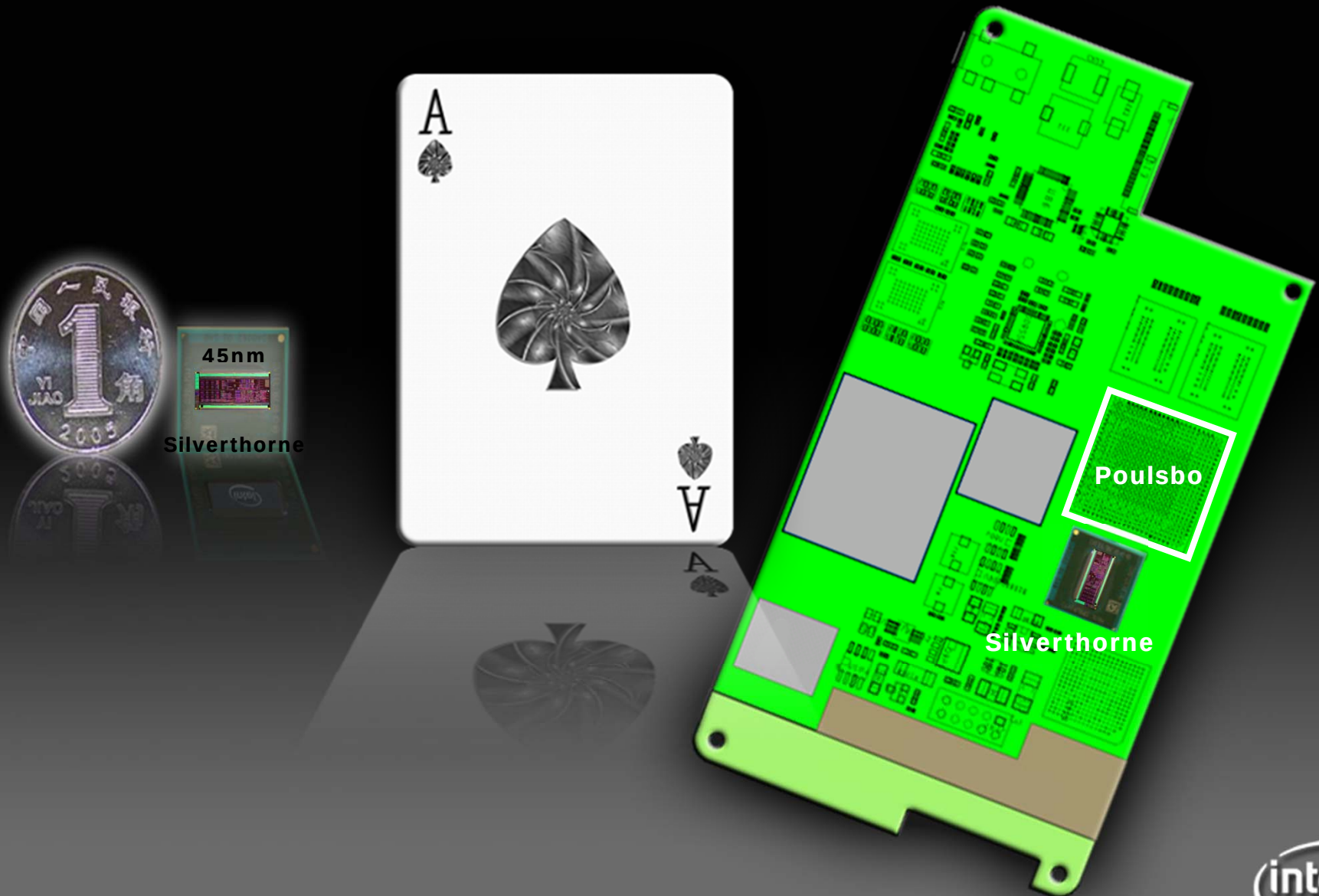
NEW DEMANDS

HIGHER
PERFORMANCE

10X



Enabling Full Internet in Your Pocket



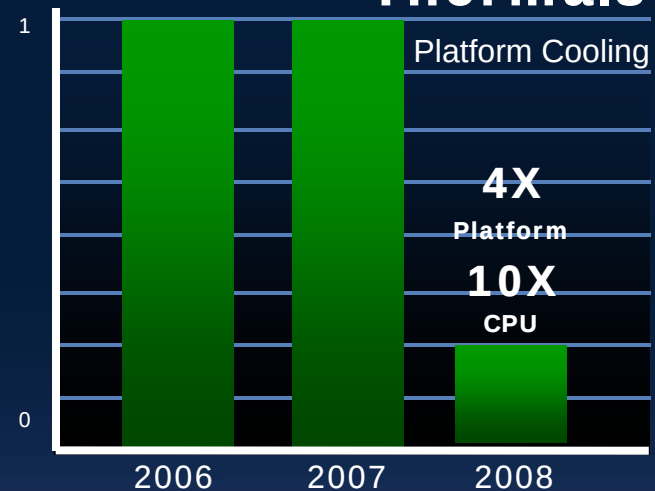
Mobil...ice



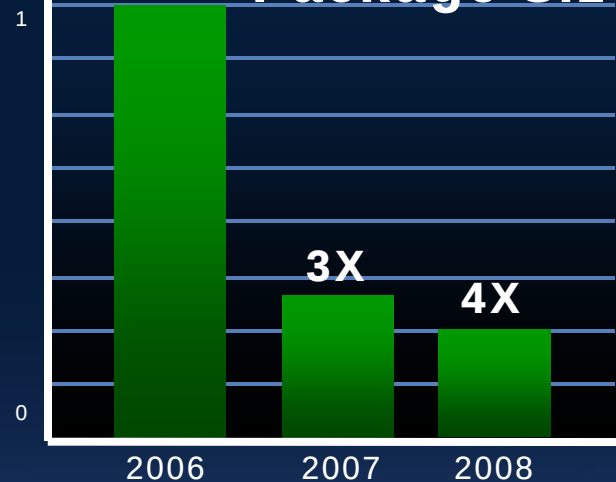
Enabling The Full Internet in Your Pocket

(Charts Shows CPU + Chipset*)

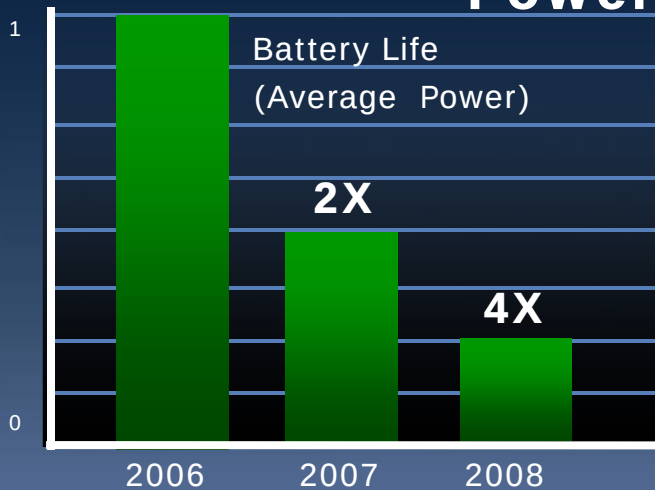
Thermals



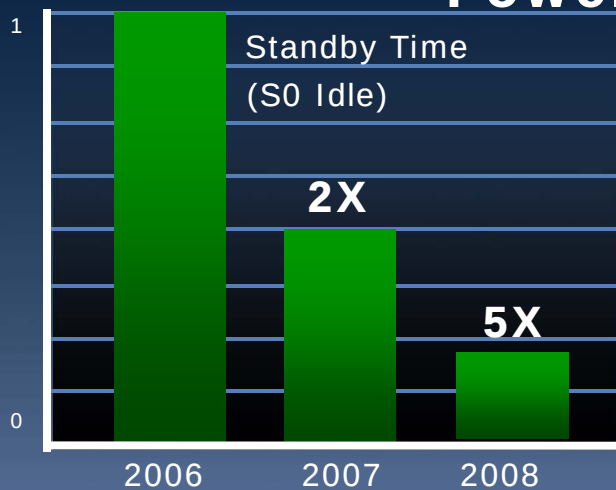
Package Size



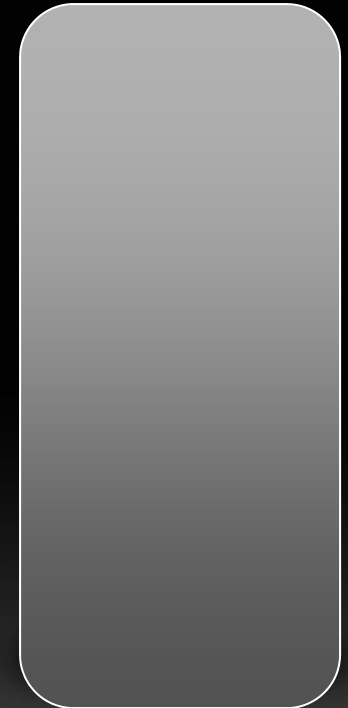
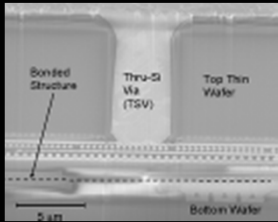
Power



Power



Future Technology Challenges



**MEMORY
AND
STORAGE**

**WIRELESS
AND
COMMS**

**ENERGY-
EFFICIENT
PLATFORMS**

**NEW
PLATFORM
CLASSES**

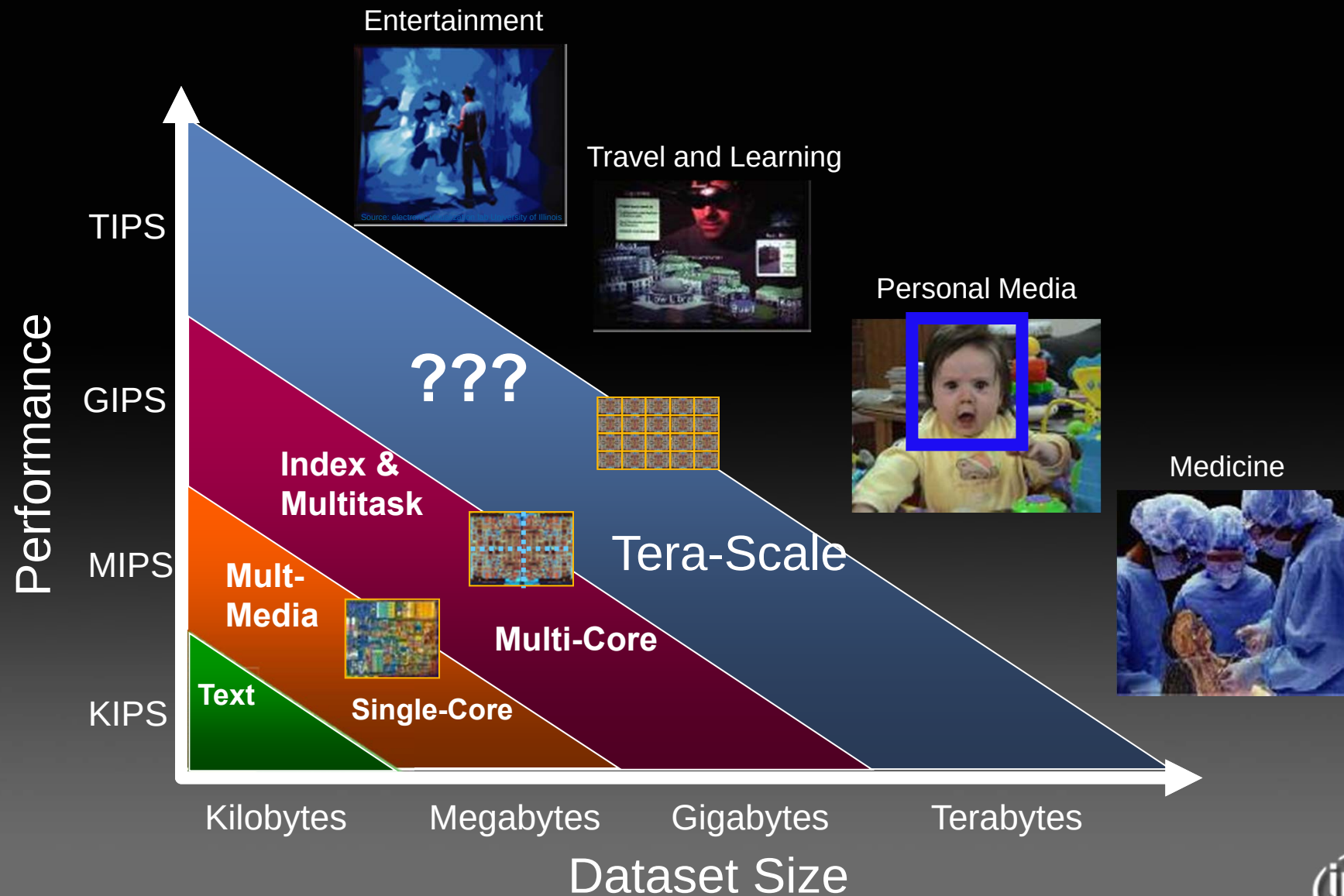
10X LOWER
POWER

NEW DEMANDS

HIGHER
PERFORMANCE **10X**



Future of Multi-Cores



Intel Tera-scale Computing Research

More than 100 projects worldwide

Microprocessor

Examples:

- Scalable memory
- Multi-core architectures
- Specialized cores
- Scalable fabrics
- Energy efficient circuits

Platform

Examples:

- Cache & Memory
- Virtualization
- Scaleable OS's
- I/O & Networking
- Silicon Photonics

Programming

Examples:

- New applications
- Workload analysis
- Transactional memory
- Languages & Compilers
- Libraries & Tools



www.intel.com/software/products

External collaborations
University Programs
Intel ® Press
Intel ® Software College



Intel Higher Education on track to impact 400 university curricula by end of 2007

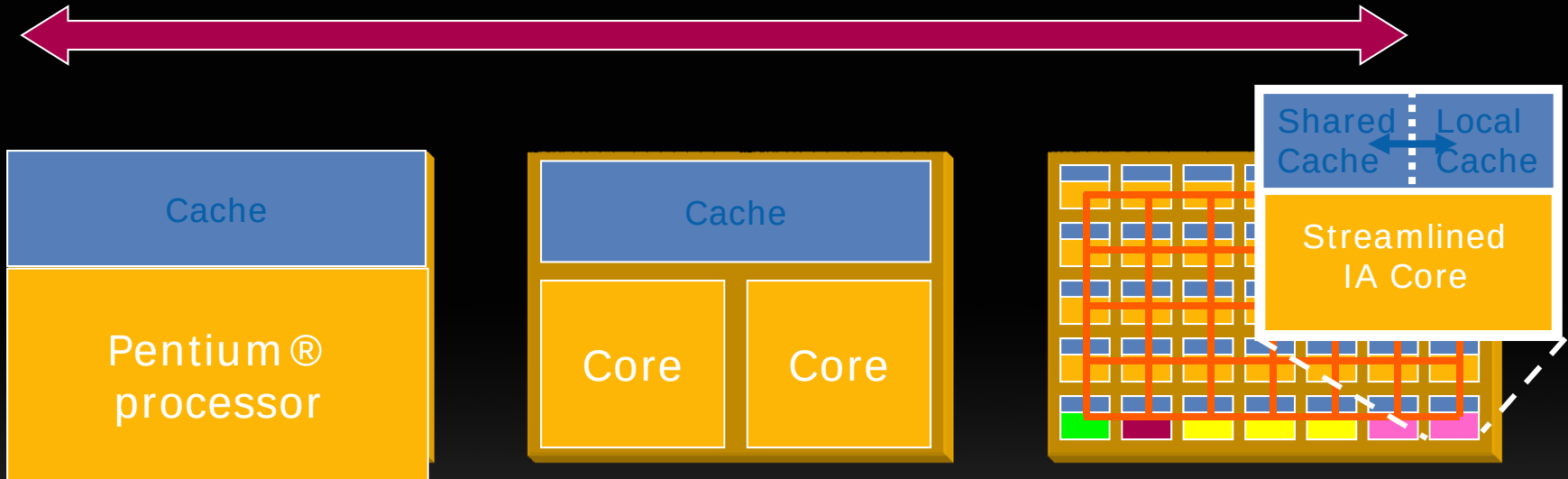
→ Focus on Parallel Programming



What Future Chips Might Look Like: From a few large cores to many simpler cores

Optimized for speed

Optimized for performance/watt



Pentium® processor era chips optimized for raw speed on single threads. Pipelined, out of order execution.

Today's chips use cores which balance single threaded and multi-threaded performance

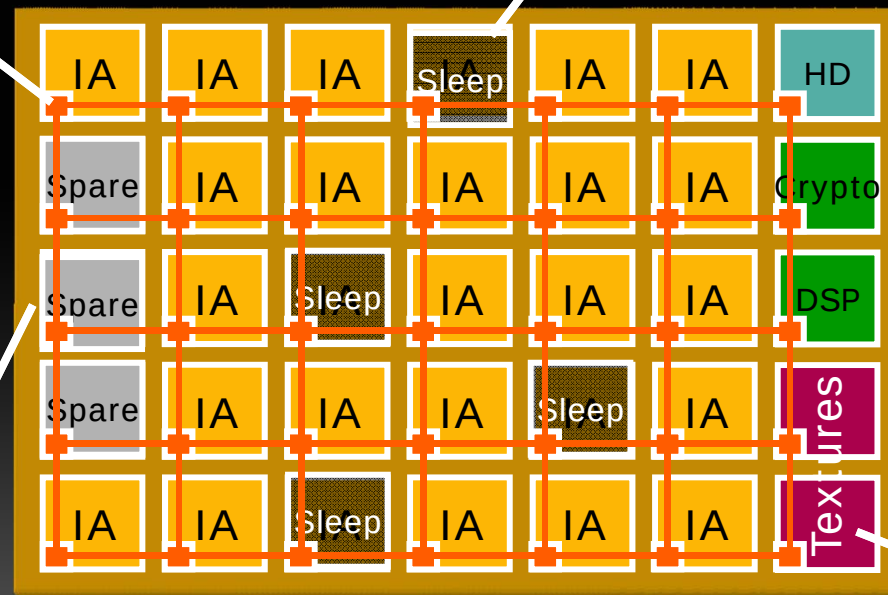
5-10 years: 10s-100s of IA cores, interconnect network, Some non-IA accelerators

Other reasons for many cores

- **Integrated Network**

- Higher b/w & lower latency compared to SMP

- **Fine grained power management**
Voltage scaling at core level



- **Yield & Resiliency**

Spare cores & binning

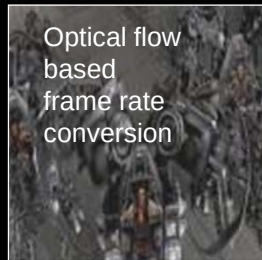
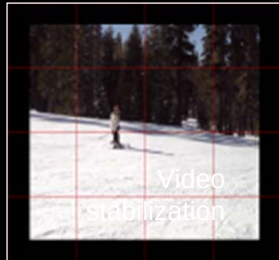
Hybridization

Integration of fixed function accelerators

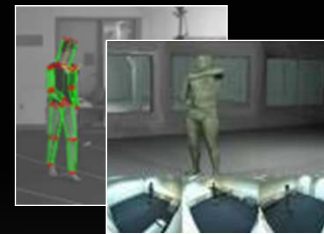
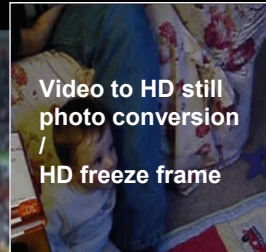
Improved virtualization another likely benefit of many core



What to do with the performance?



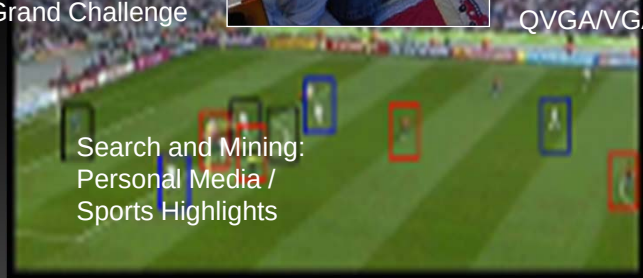
- Better Statistical Priors
- Combination of techniques
- Automated parameter selection
- Better optical flow
- ...



- Clothes
- Faces
- Multiple people
- Less Cameras
- Lighting



- Commercial content
- Personal content



10 Gigafllops

100 Gigafllops

Terafllops



Compute Requirements (for real-time performance @ 30 f/s)
Correlates with degree of difficulty (with quantum steps in between)

**These apps can use every
available FLOP!**



Ray Tracing – photo realism



Stats:

Render Time: 30 minutes

Processor: P4 3Ghz

RAM: 512MB

Renderer: Mental Ray

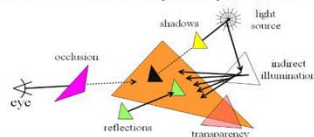
DOF: Simulated in Photoshop

~3000 FLOPs/Ray Segment

~1500 Bytes/Ray Segment

3D Graphics Rendering

- Today's Game Graphics Processing : Each triangle is processed independently in layers, foreground, background, shadows, etc.
- Ray Tracing – physically simulating light, compute intensive, collision of light rays with objects, reflections, refractions.
 - Can be used to detect body-to-body collisions as well

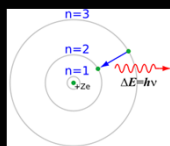


**Rendering consumes
~90 % of system
resources.**



Model-Based Computing

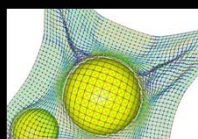
- Representing and processing data using digital “concepts”
 - Composed of mathematical rules and variables that approximate reality
 - Allow computers to recognize, manipulate, and represent things and ideas



Atomic Model

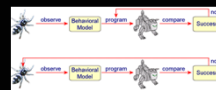


Body Model

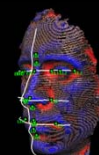


Cloth Model

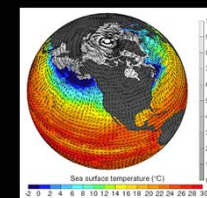
Financial Model



Behavioral Model



Facial Model

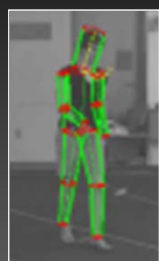


Climate Model

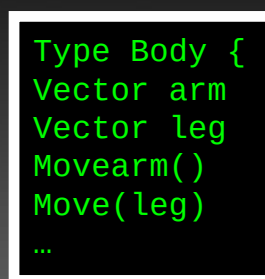
Example: Modeling body motion



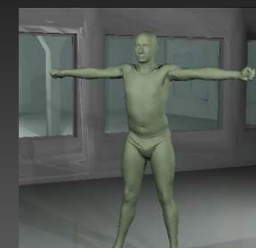
Raw Image



Analyze Image



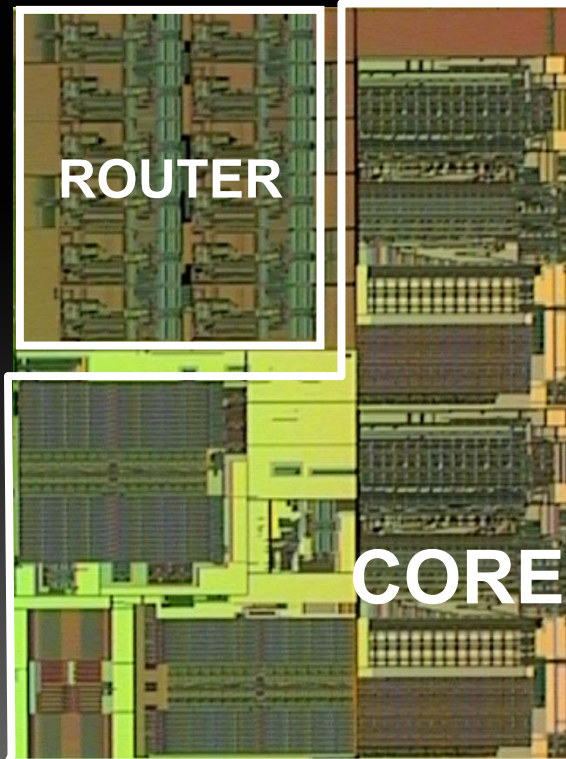
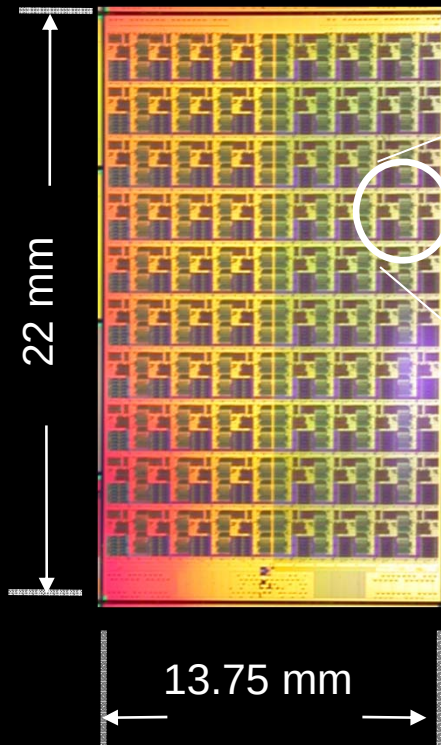
Express model



Display representation

Polaris: Tera-Scale Prototype

80 Cores
1 TFLOP at 62 Watts
256 GB/s bisection



Exa-scale Software Opportunities and Risks

Reduced Productivity in Parallel Programming

- Decomposition - Difficult to compose parallel programs
- Data races - Bugs increase exponentially with degree of parallelism
- Load Balance – Trouble keeping all processors busy

*Major challenges in enabling “easy” parallel programming
- a **major** focus area for Research*

EVERYTHING MATTERS

Energy. Performance. Storage. I/O. Mobility.
Security. Reliability. Applications.
Provisioning. Cost.

