Power-Performance Uplift Using Hetero General compute & Domain Specific Areas

Feb 23, 2011

Gans Srinivasa
Intel Corp

Thanks to:
Ravi Iyer, Scott Hahn, Bhushan Chitlur
Doug Carmean, Pranav Mehta, Alon Naveh, Henry Gabb

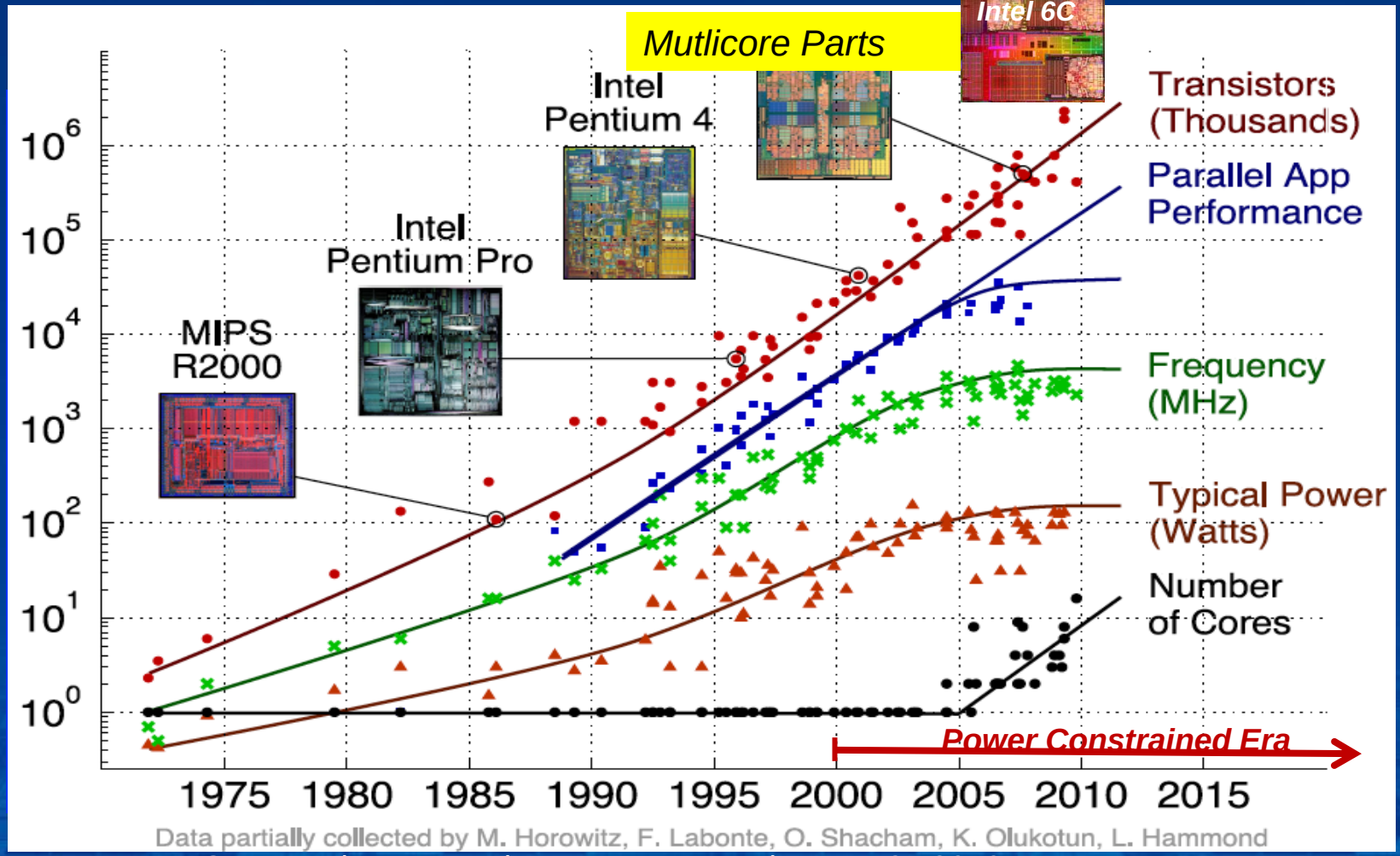
Legal Disclaimer

- *INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL® PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. INTEL PRODUCTS ARE NOT INTENDED FOR USE IN MEDICAL, LIFE SAVING, OR LIFE SUSTAINING APPLICATIONS.*
- *Intel may make changes to specifications and product descriptions at any time, without notice.*
- *All products, dates, and figures specified are preliminary based on current expectations, and are subject to change without notice.*
- *Intel, processors, chipsets, and desktop boards may contain design defects or errors known as errata, which may cause the product to deviate from published specifications. Current characterized errata are available on request.*
- *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.*
- *Intel, Intel Inside, Xeon, and the Intel logo are trademarks of Intel Corporation in the United States and other countries.*
- **Other names and brands may be claimed as the property of others.*
- *Copyright © 2011 Intel Corporation.*

Agenda

- Power-Performance
  - Servers to Clients
- Heterogeneous Architecture
- The Future Challenges

Energy Efficient Computing: Overall processor Power Transition to Multicore



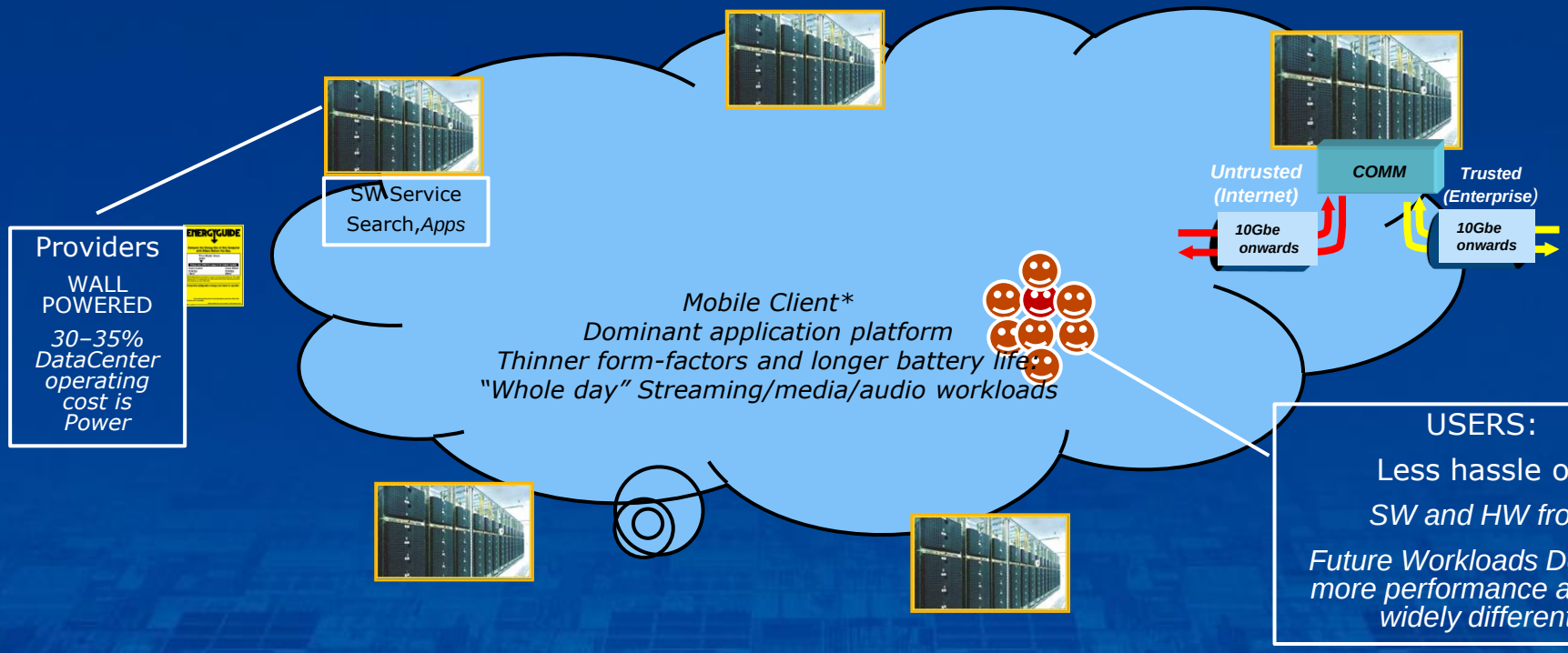
UC PAR Lab Presentation – Krste Asanovic – May24,2010

What is Next?

Trends in "The Cloud"

Rapid Cloud Growth: Reduces the hassle for users and providers.

- User needs only a browser: No worry about software installation, patches etc
- HW & Service Providers: Focus on Total Cost of Ownership
- SW Providers: Apps in controlled env and no shrink warp worries



Trend: Cloud needs more Energy efficiency, more communication bandwidth, lower TCO.

Today we have homogenous Multicore. Scaling will be limited by power and area.

Power efficient small cores have emerged and show potential for power efficient performance.

What could we do spanning Servers to Clients/Mobile Devices?

* http://parlab.eecs.berkeley.edu/publication/137; http://parlab.eecs.berkeley.edu/

Current Data Center

Source – NY Times, 06.14.2006

A single Layer 2 domain

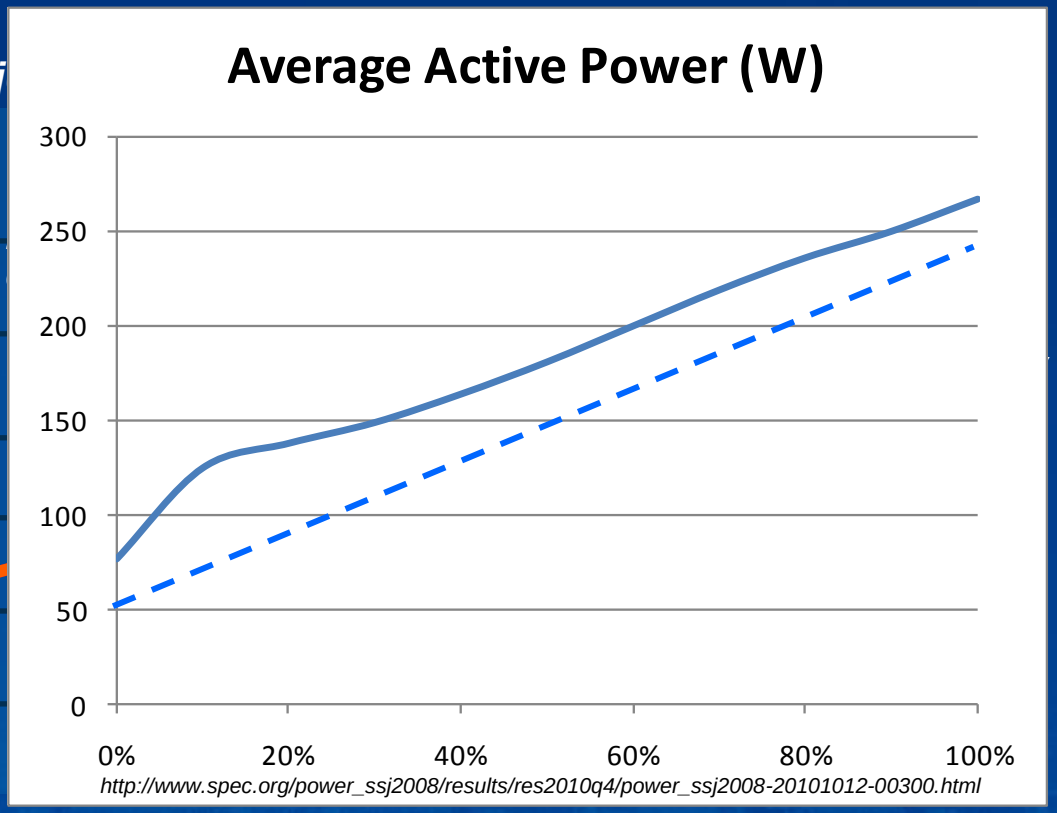
A Rack of 40 Servers

Proj

EP Xeon Server Node Today

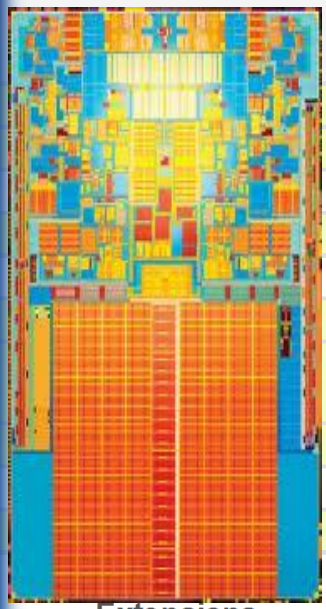
2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011

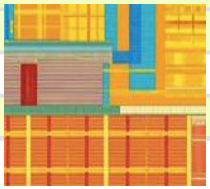
forecast



- Many modern day data centers operate around 20% load.
- Reducing idle to 50W @10c/kwh for 10000 nodes saves ¼ Millon $ /yr on energy cost. Potential is very high.
- How can we go beyond this and achieve higher power-performance efficiency?

Heterogeneous Space





	Intel® Xeon® Processor E5507 (4M Cache, 2.26 GHz, 4.80 GT/s Intel® QPI)	Intel® Atom™ processor D525 (1M Cache, 1.80 GHz)	Intel® Xeon® Processor X5690 (12M Cache, 3.46 GHz, 6.40 GT/s Intel® QPI)	Intel® Core™ Duo Processor T2300 (2M Cache, 1.66 GHz, 667 MHz FSB)
	Nehalem-EP	Pineview	Westmere-EP	Yonah
	Q1'10	Q2'10	Q1'11	
	E5507	D525	X5690	T2300
	4	2	6	2
	4	4	12	2
	2.26 GHz	1.8 GHz	3.46 GHz	1.66 GHz
	4 MB Intel® Smart Cache	1 MB L2 Cache	12 MB Intel® Smart Cache	2 MB L2 Cache
		DMI	QPI	FSB
			6.4 GT/s	667 MHz
		64-bit	64-bit	32-bit
Extensions		**SSE2, SSE3, SSSE3**	**SSE4.2**	
Lithography	**45 nm**	**45 nm**	**32 nm**	**65 nm**
		13 W	**130 W**	**31 W**
V ge		0.800V-1.175V	0.750V-1.350V	1.1625V - 1.30V
s		DDR3-800 (SODIMM only); DDR3-800/1066	DDR3-800/1066/1333	
ECC supported		No		No
Lithography	45 nm	45 nm	32 nm	65 nm
Processing Die Size		**87 mm2**		**90 mm2**
# of Processing Die Transistors		176 million		151 million

- Workloads are not homogeneous
- Energy efficiency increasingly important
- Frequency to Multicore play is running out of steam

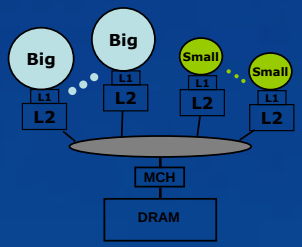
What is Next?

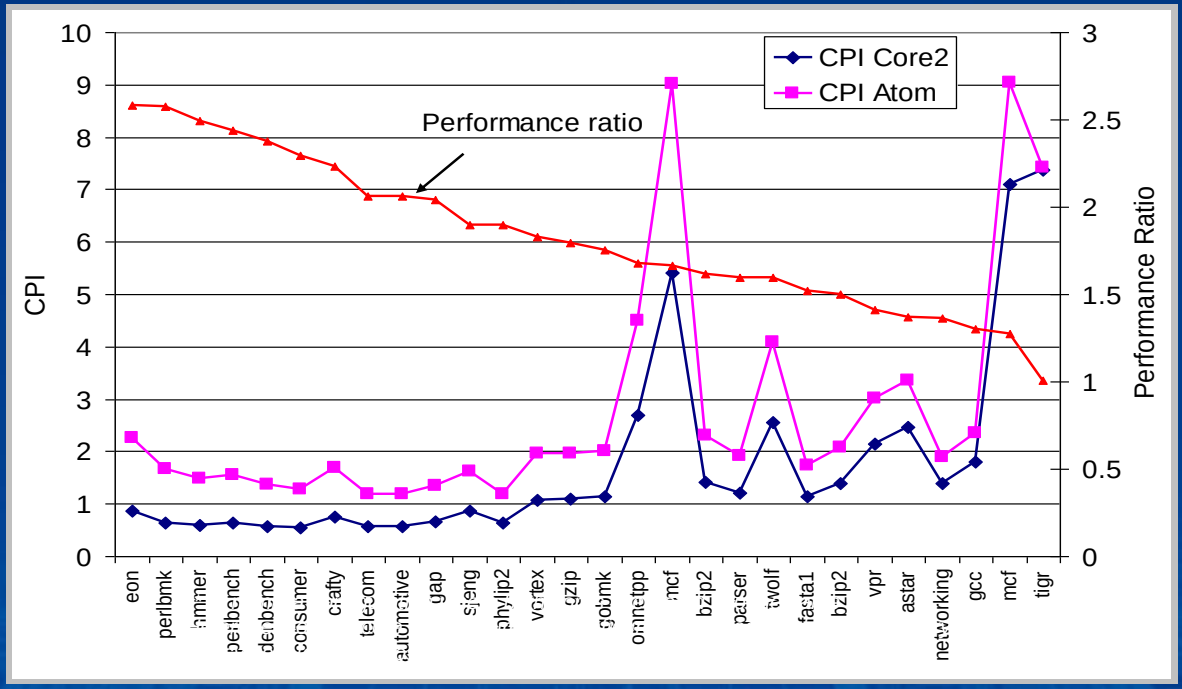
Heterogeneous architectures (mix of different cores and accelerators)

Big & Small cores Research

Recent measurements (SPECcpu and Bio workloads)

Comparison of Atom to Core 2 Duo



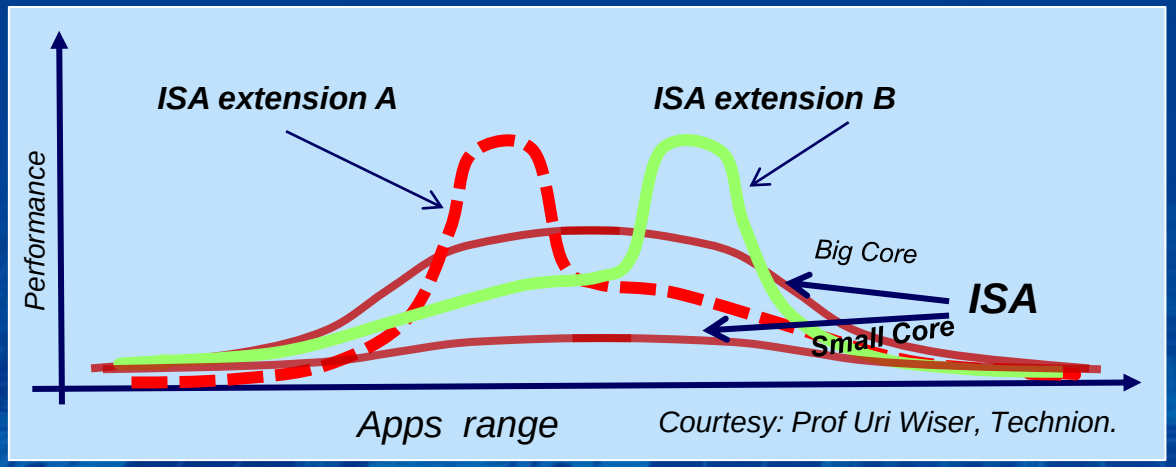


<u>Validation:</u>
Small Core Perf = 50% Big Core

Observations: Varying perf difference (1.03x – 2.63x) across applications
Shows the use case for heterogeneity
Knowledge of performance/power difference can help in scheduling

Hetero Core Solution Space

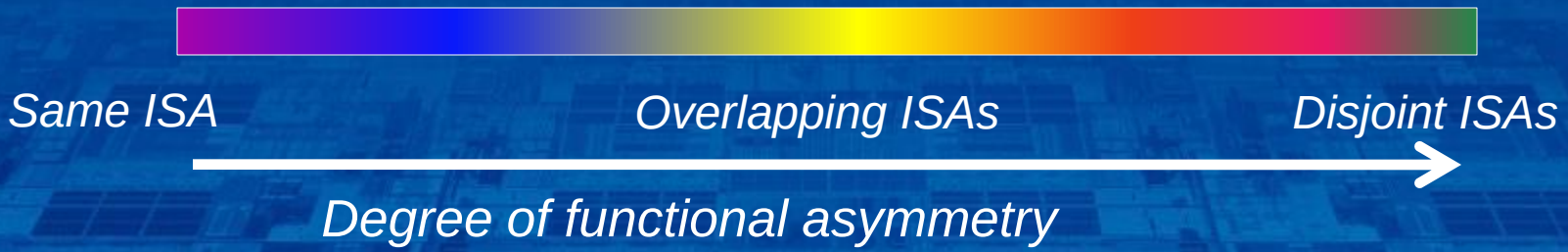
- Multiple compute elements that are tightly coupled
  - Big core, Atom, Smaller Cores and Hardware accelerators
  - Sharing cache coherency and common memory
  - Can be managed by a single OS or User level or Other means
- Power is key
  - Objective is to minimize energy per task
  - Ideally, each task is performed in most efficient compute element
  - Using a smaller core reduces power by an order of magnitude. Performance/Power improves.



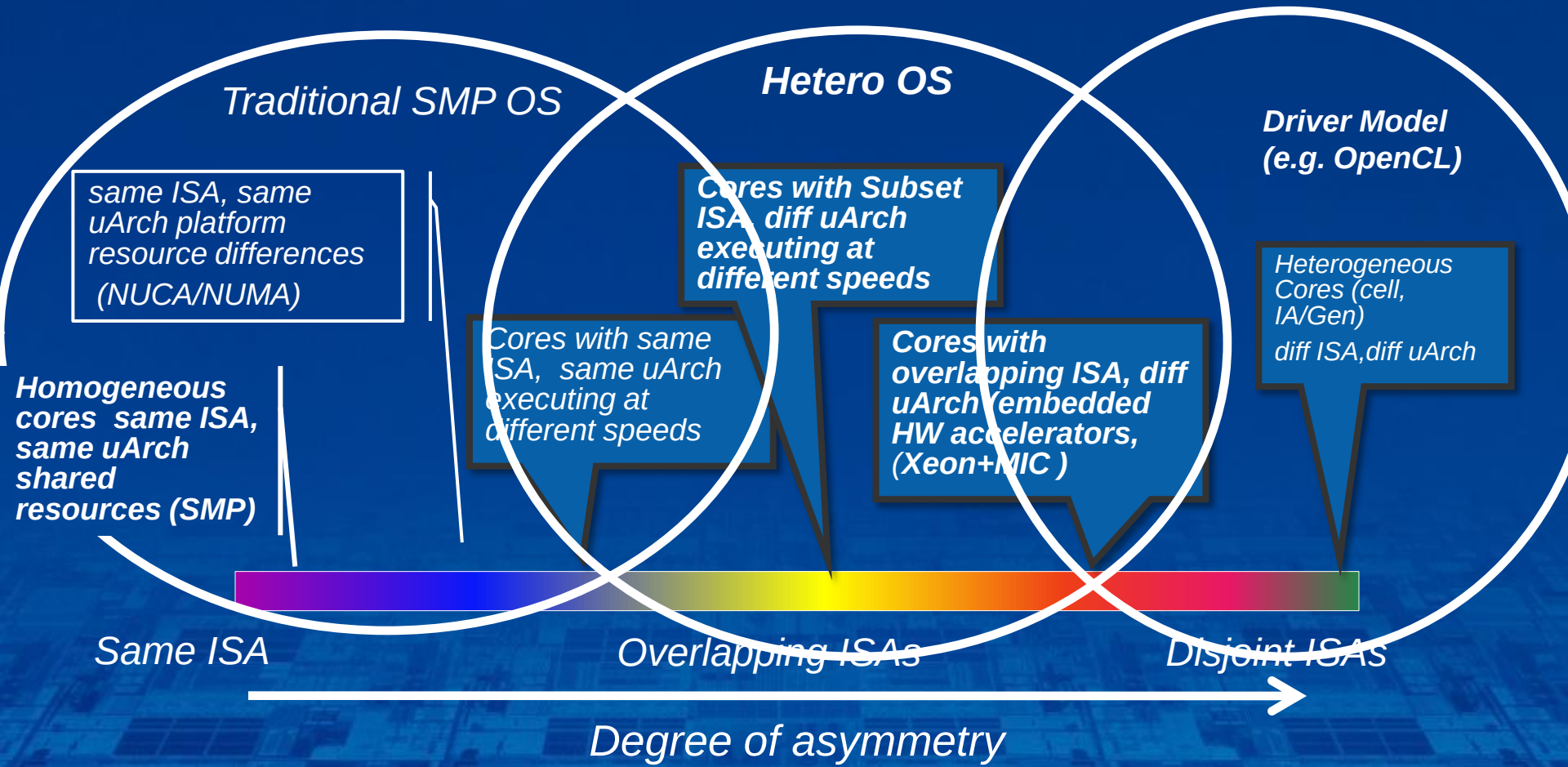
With all big cores, scalability is a problem because of power wall
With all small cores, it single thread performance that suffers
With heterogeneous (big + small), we can strike a power/performance balance and continue to scale

Heterogeneous Architecture Design Space

- Performance asymmetry
  - Cores have different performance and power
    - E.g., asymmetric cache sizes, clock speeds, uarch
  - Apps can run anywhere, but get different performance
- Functional asymmetry
  - Cores have different ISAs
  - Difference can be in many dimensions
    - Instructions, registers, data types, addressing modes, memory architecture, exception handling, I/O
  - Can have various degrees of difference

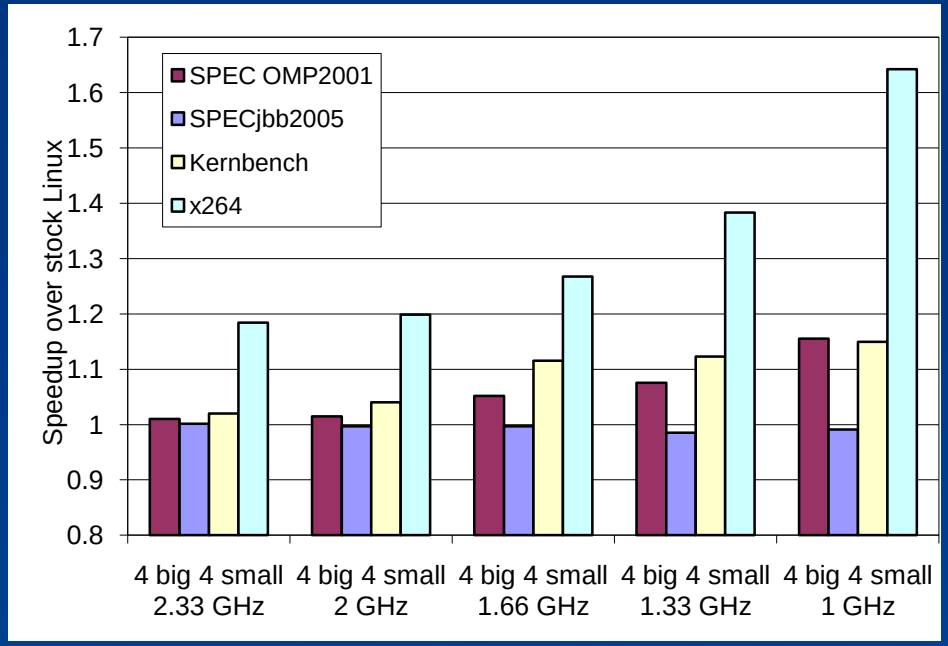


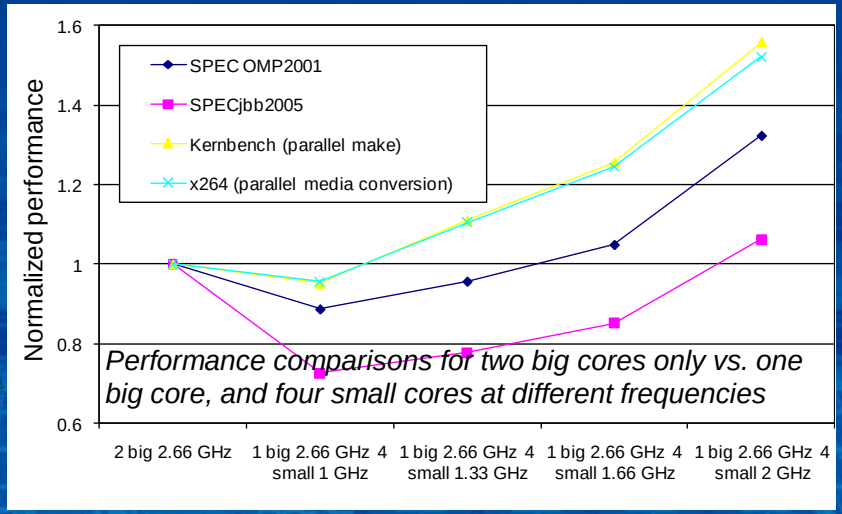
Heterogeneous Architecture Space



HeteroOS Research work @ Intel

- Cost effective solution for ST and MT
- Various scheduling options examined and continue the work
- Fault and migrate to support Instruction based asymmetry



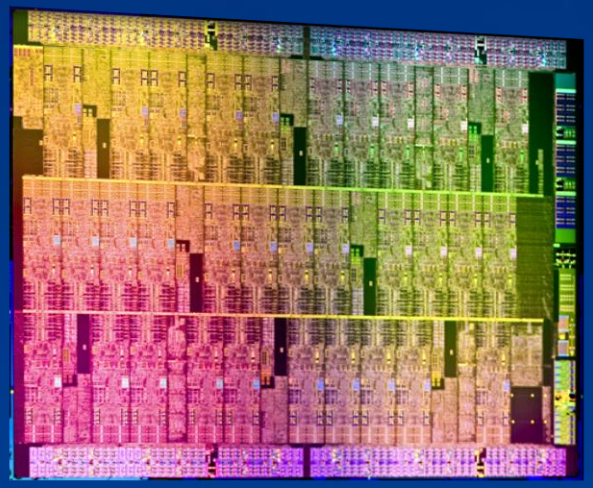


Reference: Operating System Support fro Overlapping ISA Heterogeneous Multi-core Architectures – HPCA 2010

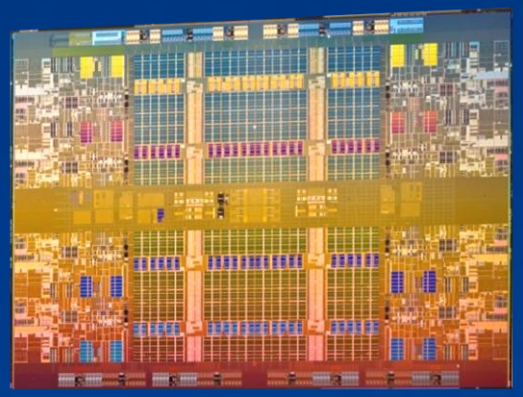
- Scott Hahn etal Intel corp.

Many Core and Multi-Core

Many Integrated Cores at 1-1.2 GHz



Multi-core Intel Xeon at 2.26-3.5 GHz



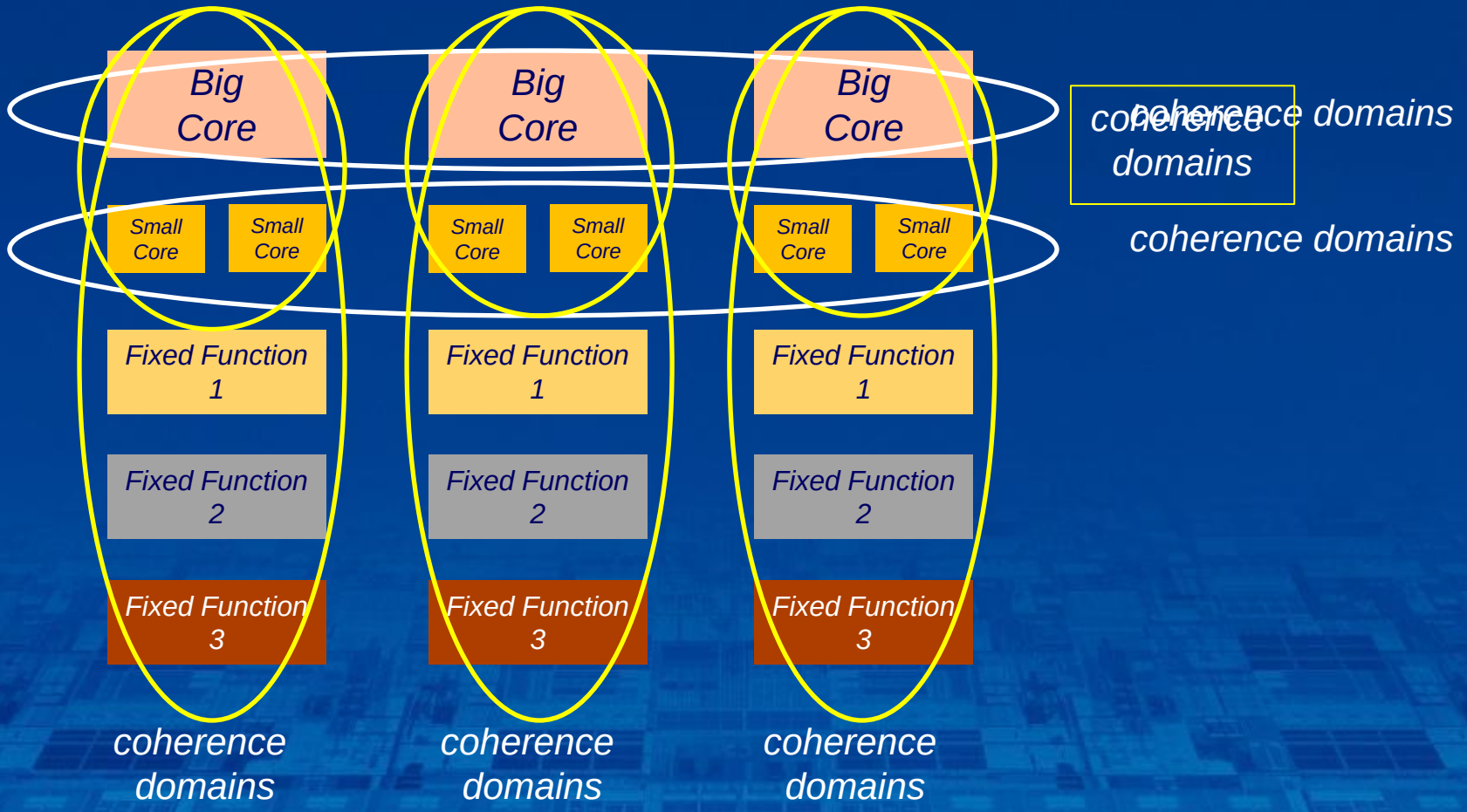
Die Size not to scale

In Intel® MIC, each core is smaller and lower power, has lower single thread performance, but higher aggregate performance

Many core relies on a high degree of parallelism to compensate for the lower speed of each individual core

Relatively few specialized applications today are highly parallel, but those applications will benefit from Intel® MIC

Generalized Forms



SW challenges

Should be tolerant of coherence domains

Scheduling policies

- *Hierarchical?*
- *OS centered?*
- *User level orchestrated?*

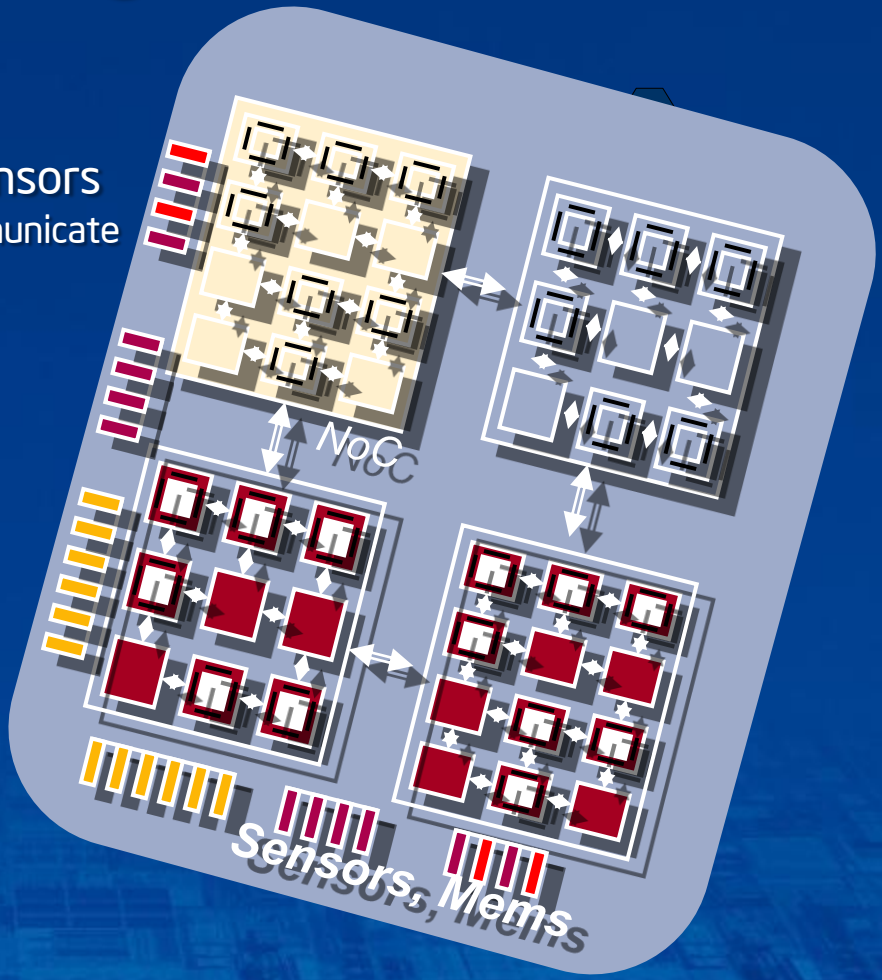
Heterogeneous friendly

- *Big core / Little core*
- *Fixed function*
- *Migratory (similar)*
- *Absence / Presence (dissimilar)*

Let us turn this into our opportunity & Innovate

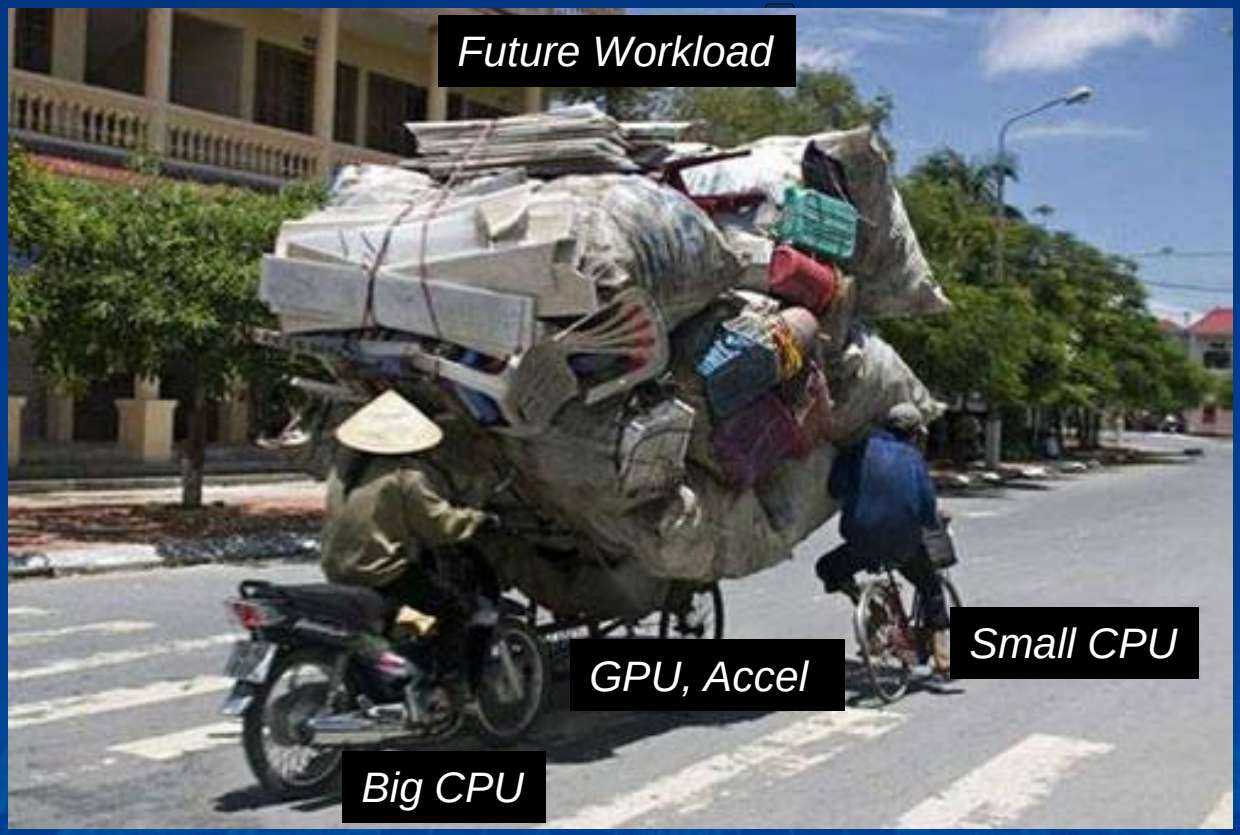
HW Challenges

- "On Chip Diversity": Various IP cores, MEMs, Sensors
  - Interconnect fabrics for heterogeneous region to communicate efficiently
  - Along with existing cores – need continuity
  - Cores communicating in a fault tolerant fashion
- Network On Chip communication
- Multiple Clock domains & Voltage islands
- SOC and 3D chip/platform will be the norm



High Level of integration

How Not To Do Hetero CMP



Let us get this mix right

More Discussion and state of current work on Friday