



The Path To Exascale – Challenges and Opportunities

EMIT Workshop, Manchester, 30th July 2015

Gaurav Kaul – Solutions Architect

John Swinburne – HPC Architect



Legal Notices and Disclaimers

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. 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, life-sustaining, critical control or safety systems, or in nuclear facility applications.

Intel products 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.

Intel may make changes to dates, specifications, product descriptions, and plans referenced in this document at any time, without notice.

This document may contain information on products in the design phase of development. The information herein is subject to change without notice. Do not finalize a design with this information.

Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined." Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them.

Intel Corporation or its subsidiaries in the United States and other countries may have patents or pending patent applications, trademarks, copyrights, or other intellectual property rights that relate to the presented subject matter. The furnishing of documents and other materials and information does not provide any license, express or implied, by estoppel or otherwise, to any such patents, trademarks, copyrights, or other intellectual property rights.

Wireless connectivity and some features may require you to purchase additional software, services or external hardware.

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

Intel, the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

Copyright © 2014 Intel Corporation. All rights reserved.

Legal Disclaimer & Optimization Notice

INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS". NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO THIS INFORMATION INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

Copyright© 2014, Intel Corporation. All rights reserved. Intel, the Intel logo, Atom, Xeon, Xeon Phi, Core, VTune, and Cilk are trademarks of Intel Corporation in the U.S. and other countries.

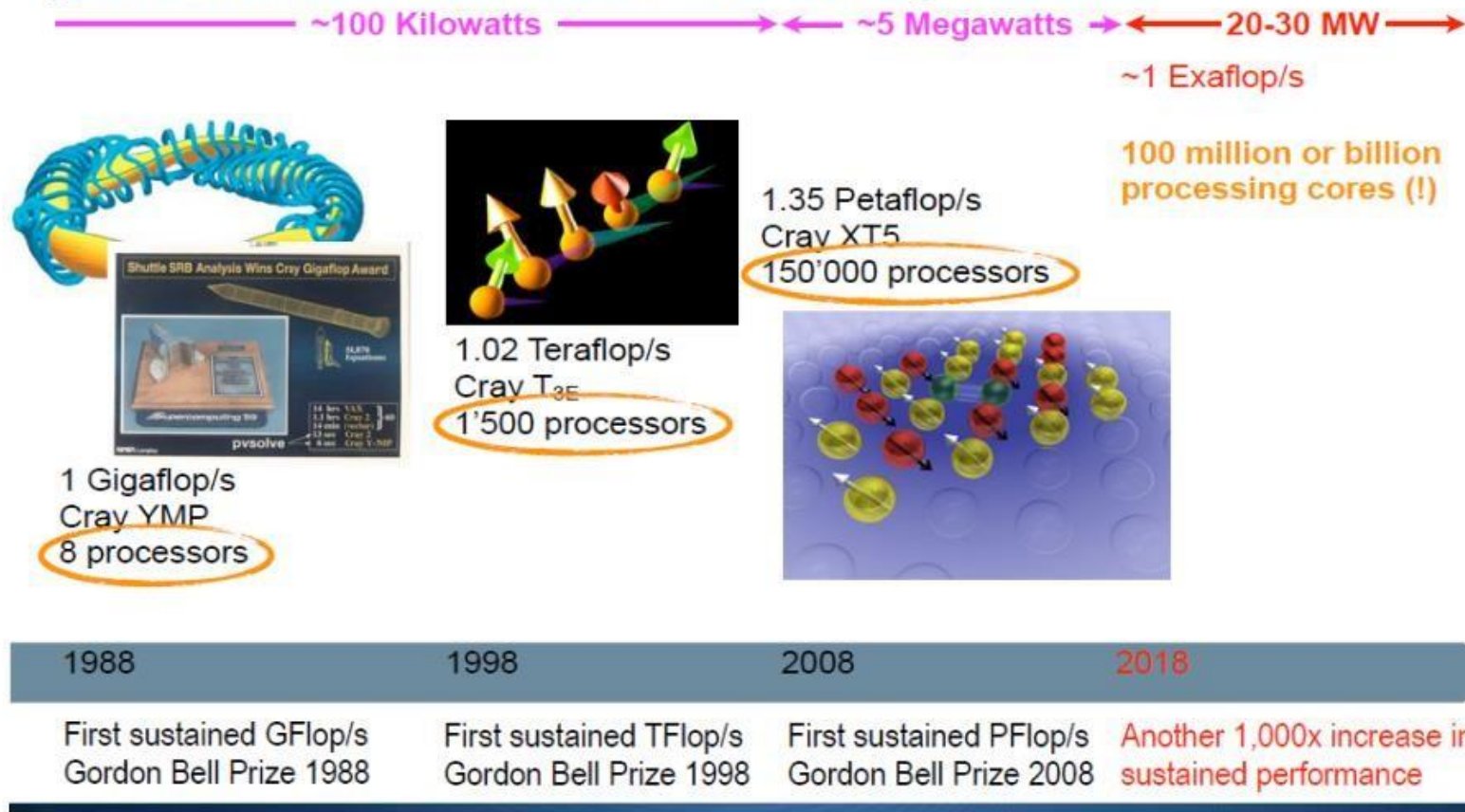
Optimization Notice

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

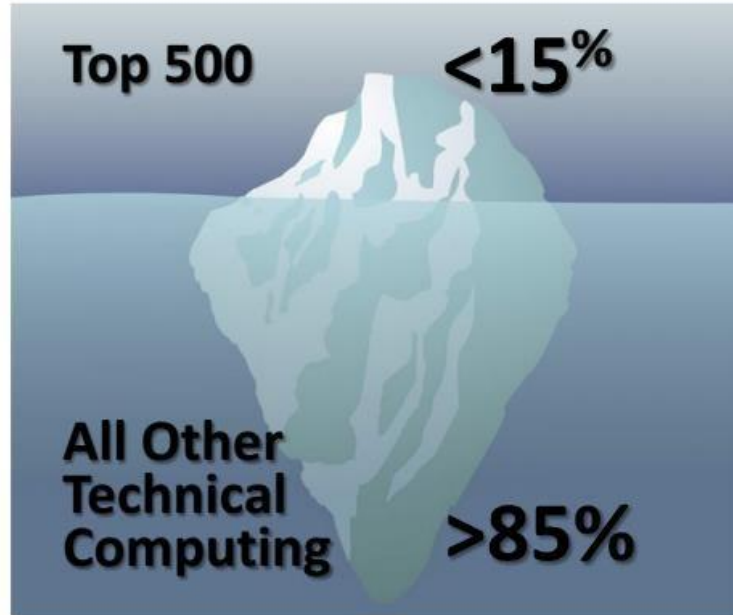
Notice revision #20110804

A bit of History

Computer performance and application performance increase $\sim 10^3$ every decade



The Top 500 Waterfall



% of sockets sold

Source: Top500.org and Intel Estimate of Top500 sockets as % of sum of analysts reports of HPC and branded Workstations sockets. Performance waterfall timelines based on TOP500.org statistics (#1-#500) and Intel estimate (#500 to projected Intel Knights Landing)
Other brands and names are the property of their respective owners.

Performance Waterfall*

#1 Top500 System to Single Socket

6-8 years

#1 to #500

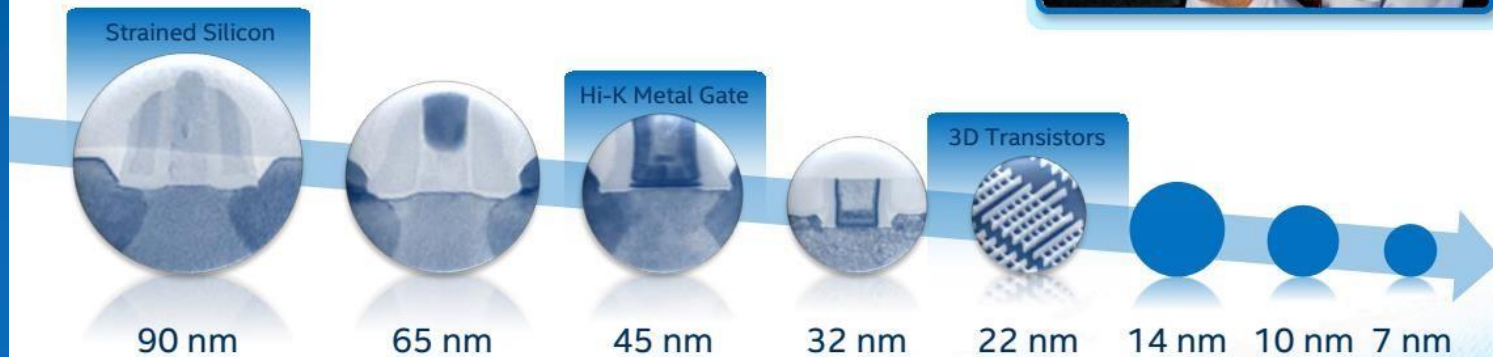
~9 years

#500 to Single Socket

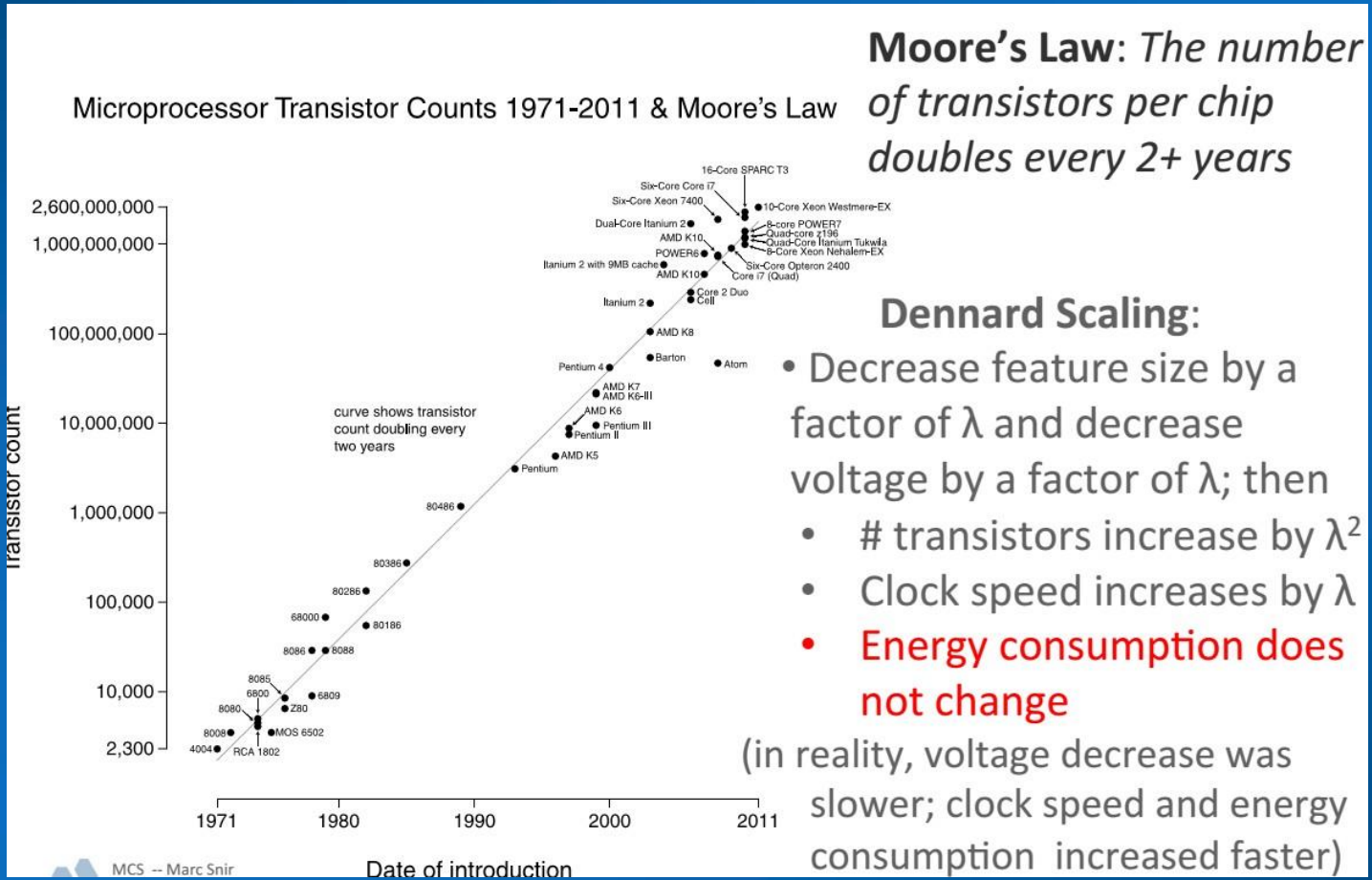
**plus.....similar waterfalls for other capabilities in areas like fabrics, storage, software, ...*

50 years of Moore's Law

Enabling new devices with higher functionality and complexity while controlling power, cost, and size



Moore and Dennard Scaling



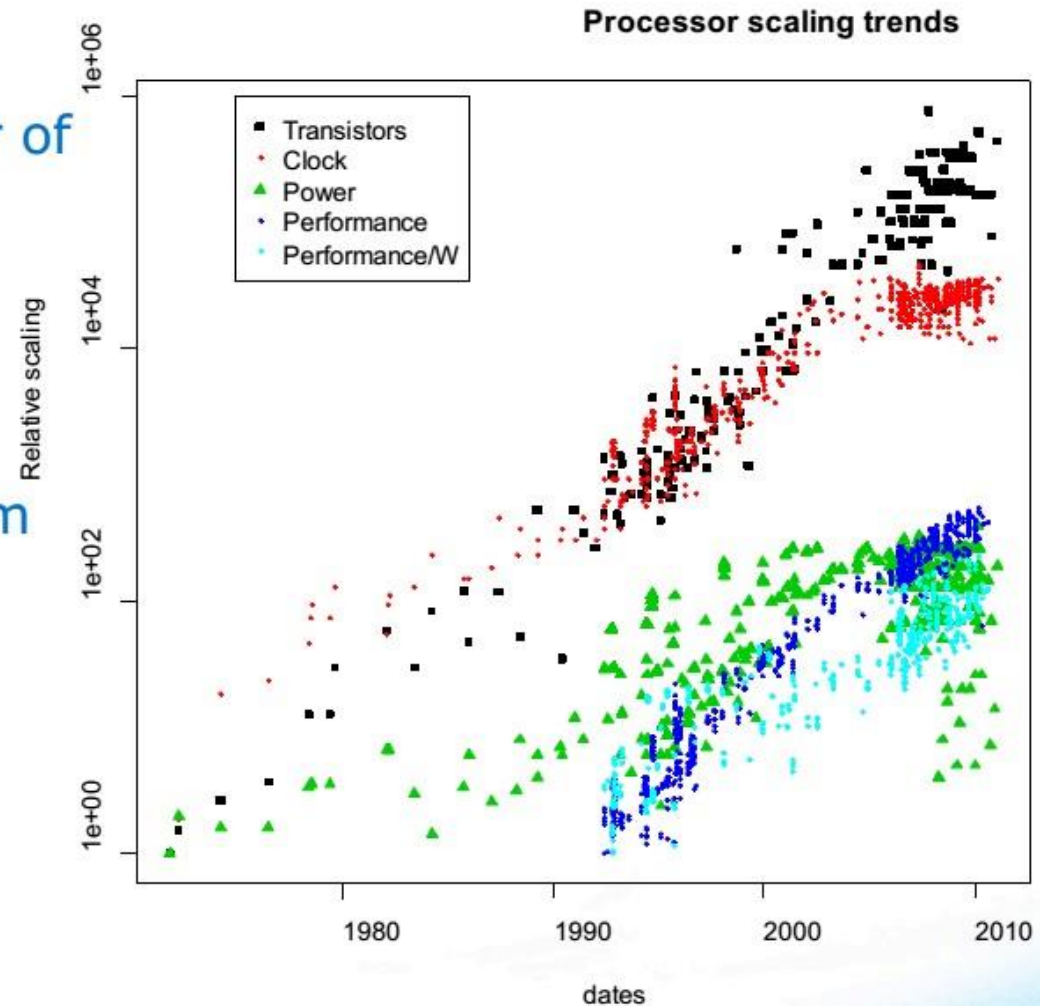
Current Processor Performance Trends

After ~2004 only the number of transistors continues to increase exponentially

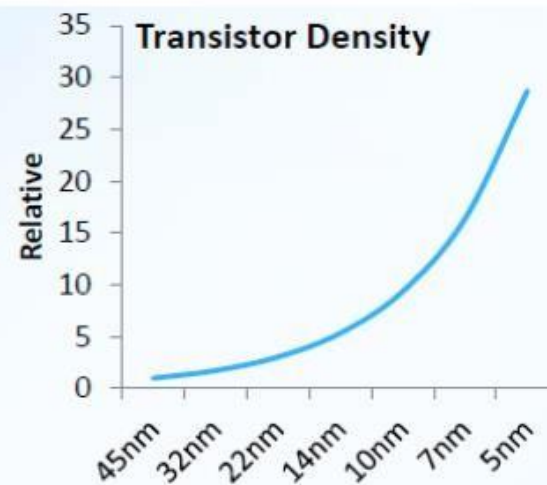
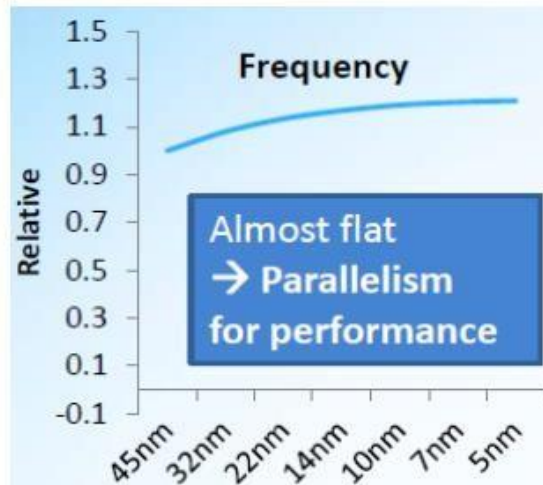
We have hit limits in

- Power
- Instruction level parallelism
- Clock speed

Single core scalar performance is now only growing slowly



Technology Scaling Outlook



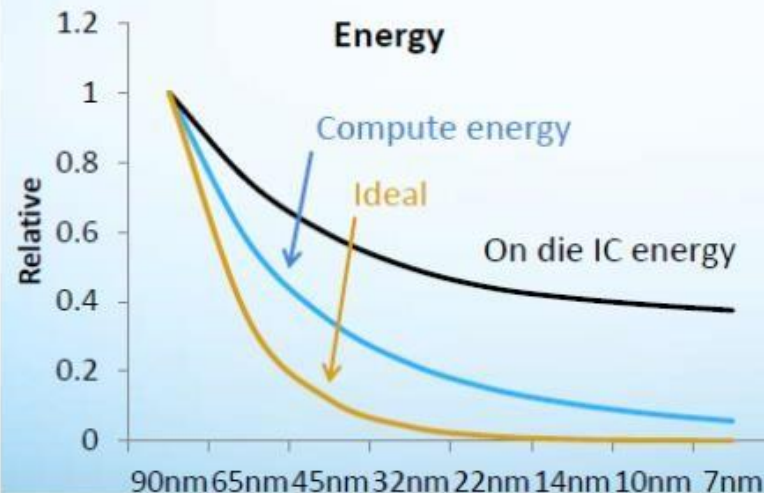
→ Lots of transistors!

Pollack's rule:

Single-thread perf.

$\sim \sqrt{\text{Transistor count}}$

→ Manycore for efficient
use of transistors

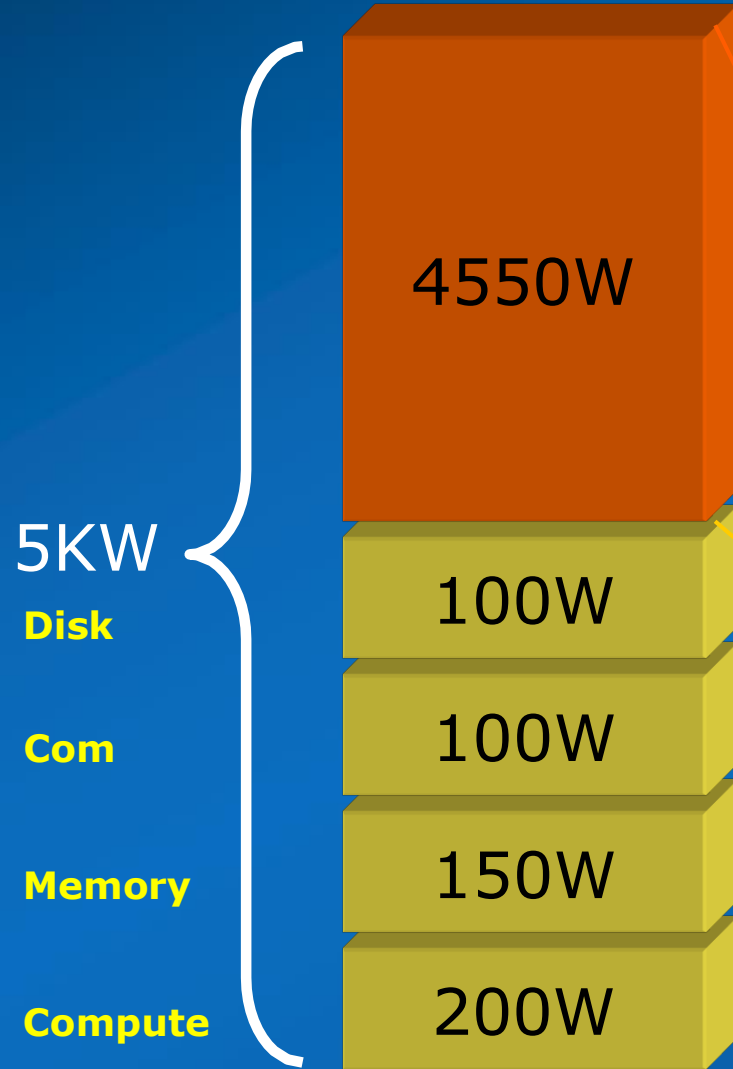


Transistor density will scale better than power
→ Overprovisioned, energy-limited hardware
→ Dynamic execution, fine-grained power
management

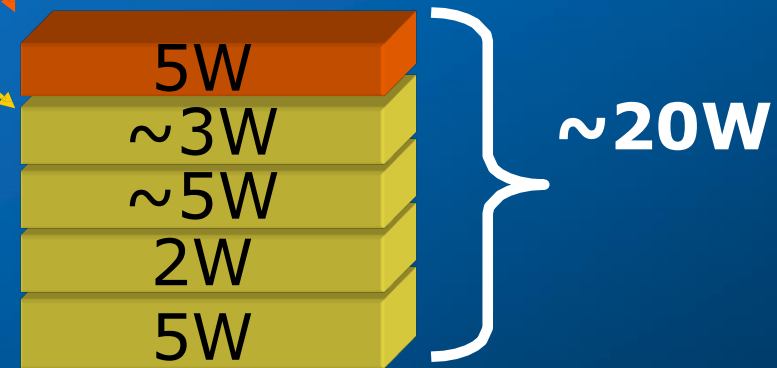
Signaling power will not scale as well as logic
→ Data movements will dominate energy

The Power & Energy Challenge

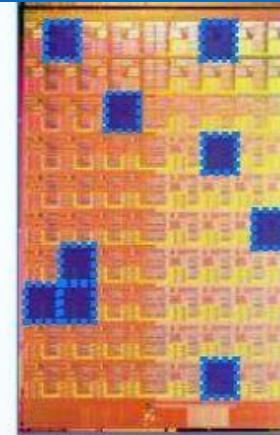
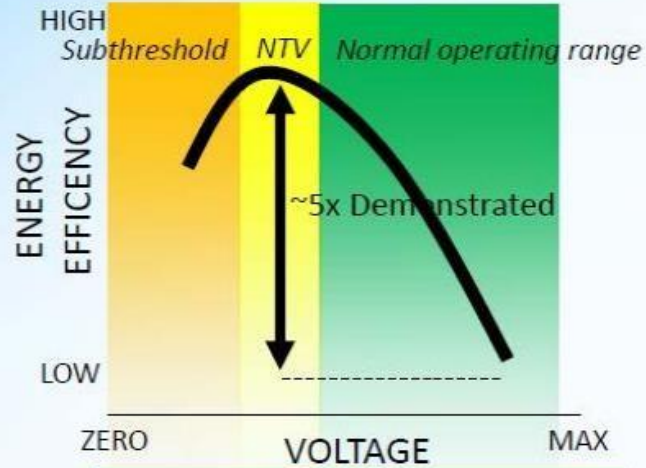
TFLOP Machine today



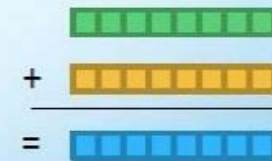
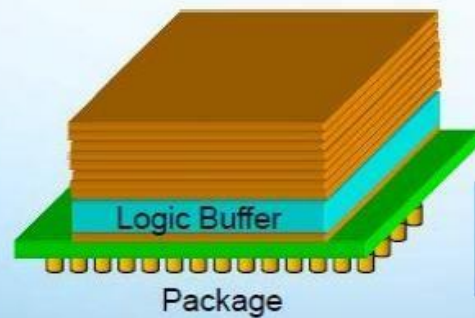
TFLOP Machine then
With Exa Technology



Promising Technologies



Fine-grain power management



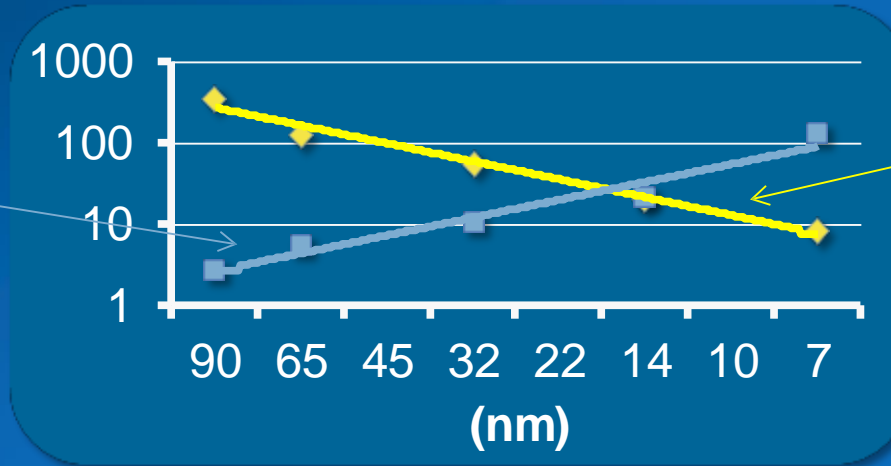
Specialized circuits (SIMD, encryption, ...)

Rethink System Level Architecture



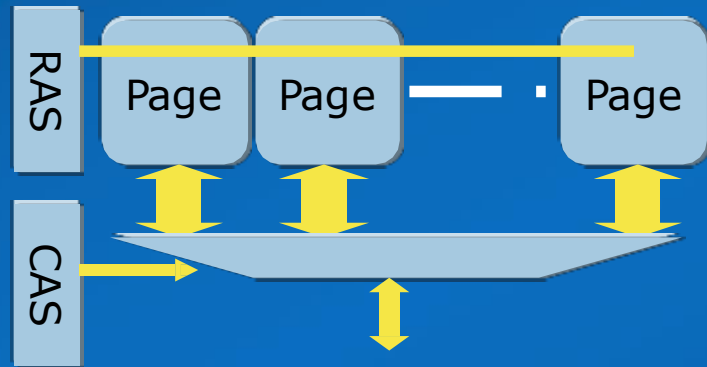
Revise DRAM Architecture

① Need exponentially increasing BW (GB/sec)



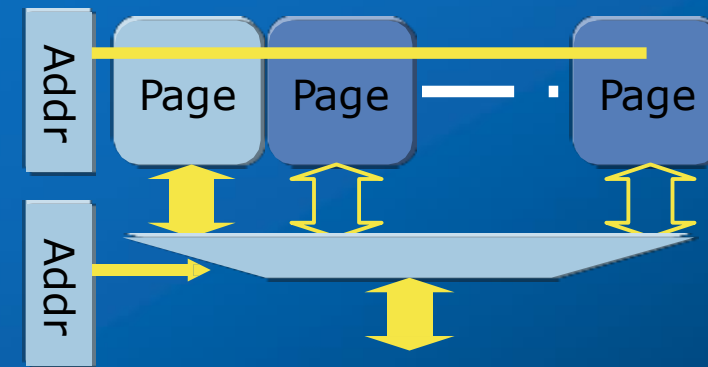
② Need exponentially decreasing energy (pJ/bit)

Traditional DRAM



Activates many pages
Lots of reads and writes (refresh)
Small amount of read data is used
Requires small number of pins

New DRAM architecture



Activates few pages
Read and write (refresh) what is needed
All read data is used
Requires large number of IO's (3D)

3D-Integration of DRAM and Logic

Logic Buffer Chip

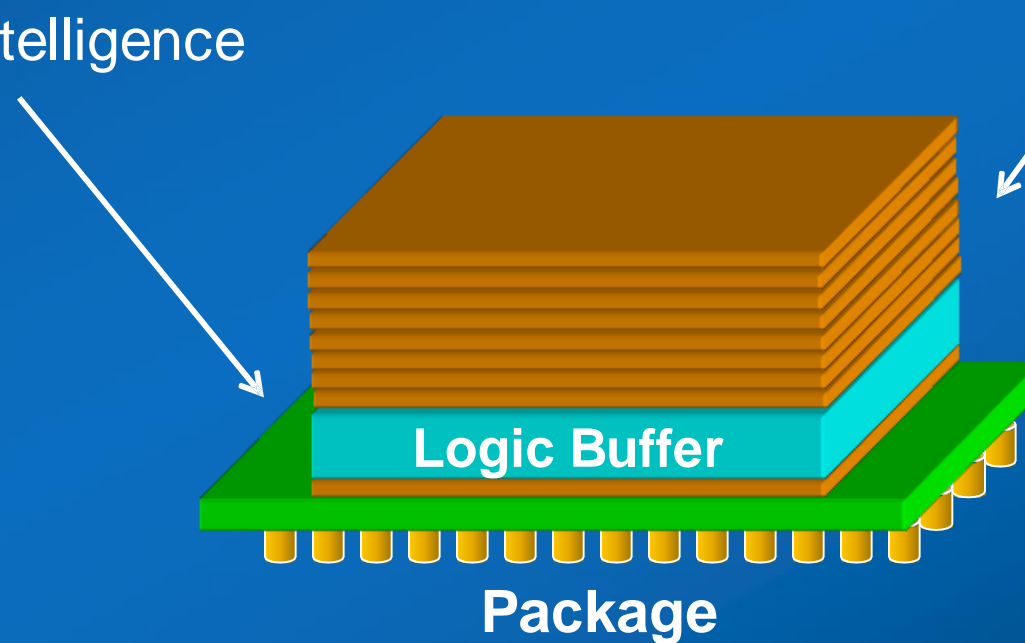
Technology optimized for:

- High speed signaling
- Energy efficient logic circuits
- Implement intelligence

DRAM Stack

Technology optimized for:

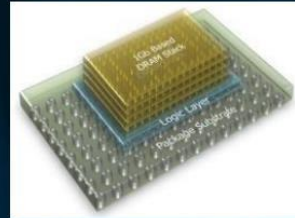
- Memory density
- Lower cost



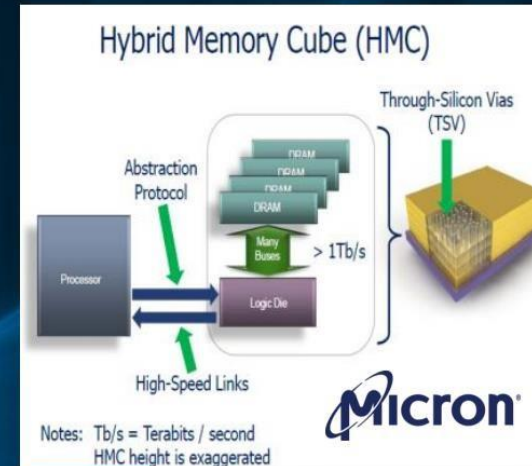
3D Integration provides best of both worlds

DRAM Scaling Using 3D Memory

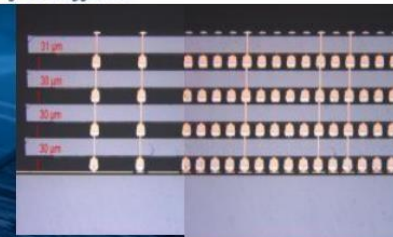
1Tb/s HMC DRAM Prototype



- 3D integration technology
- 1Gb DRAM Array
- 512 MB total DRAM/cube
- 128GB/s Bandwidth
- <10 pJ/bit energy



	Bandwidth	Energy Efficiency
DDR-3 (Today)	10.66 GB/Sec	50-75 pJ/bit
Hybrid Memory Cube	128 GB/Sec	8 pJ/bit



10X higher bandwidth, 10X lower energy



Source: Micron

Needs a Paradigm Shift

Past and present priorities—

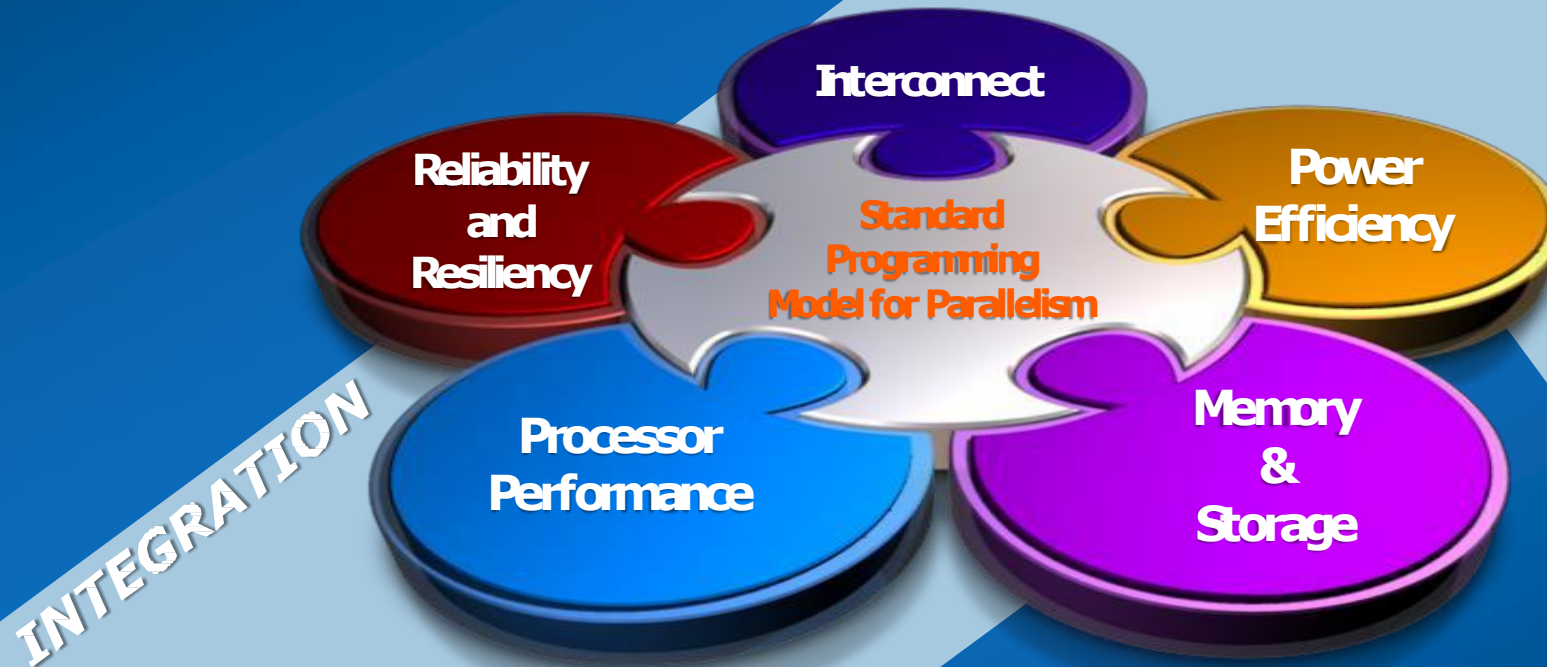
Single thread performance	Frequency
Programming productivity	Legacy, compatibility Architecture features for productivity
Constraints	(1) Cost (2) Reasonable Power/Energy

Future priorities—

Throughput performance	Parallelism
Power/Energy	Architecture features for energy Simplicity
Constraints	(1) Programming productivity (2) Cost

**Evaluate each (old) architecture feature with
new priorities**

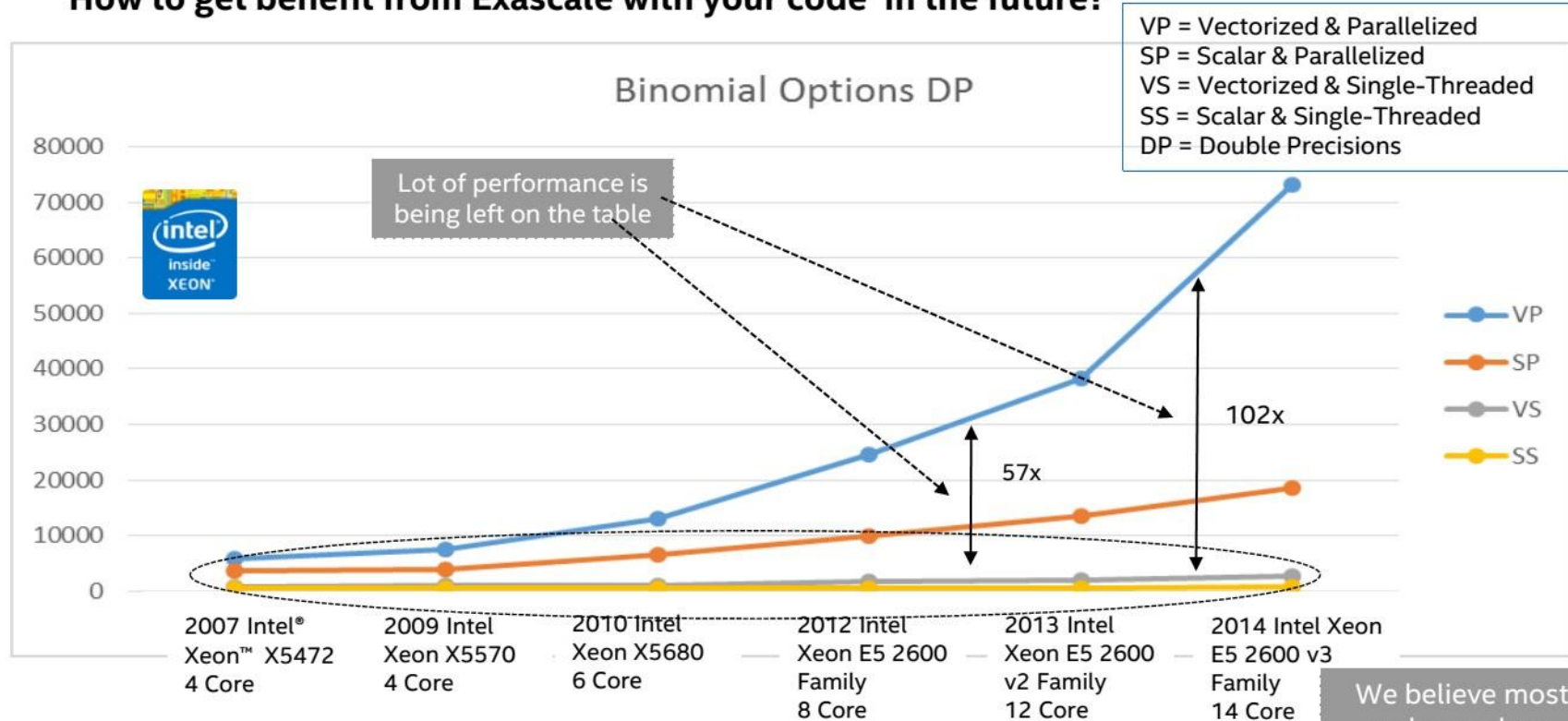
Intel: Investing to Remove 6 Bottlenecks



Impact on Applications

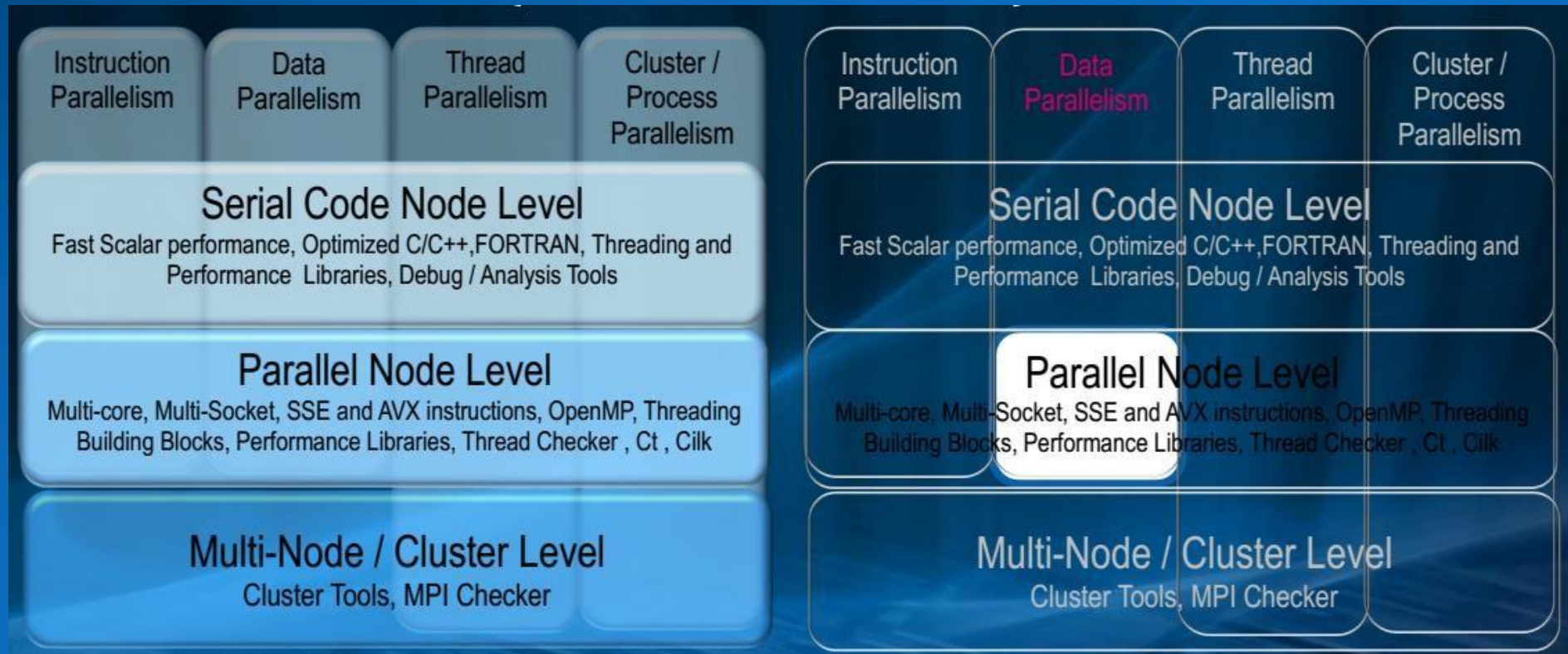
How Can I Achieve High Performance?

How to get benefit from Exascale with your code in the future?

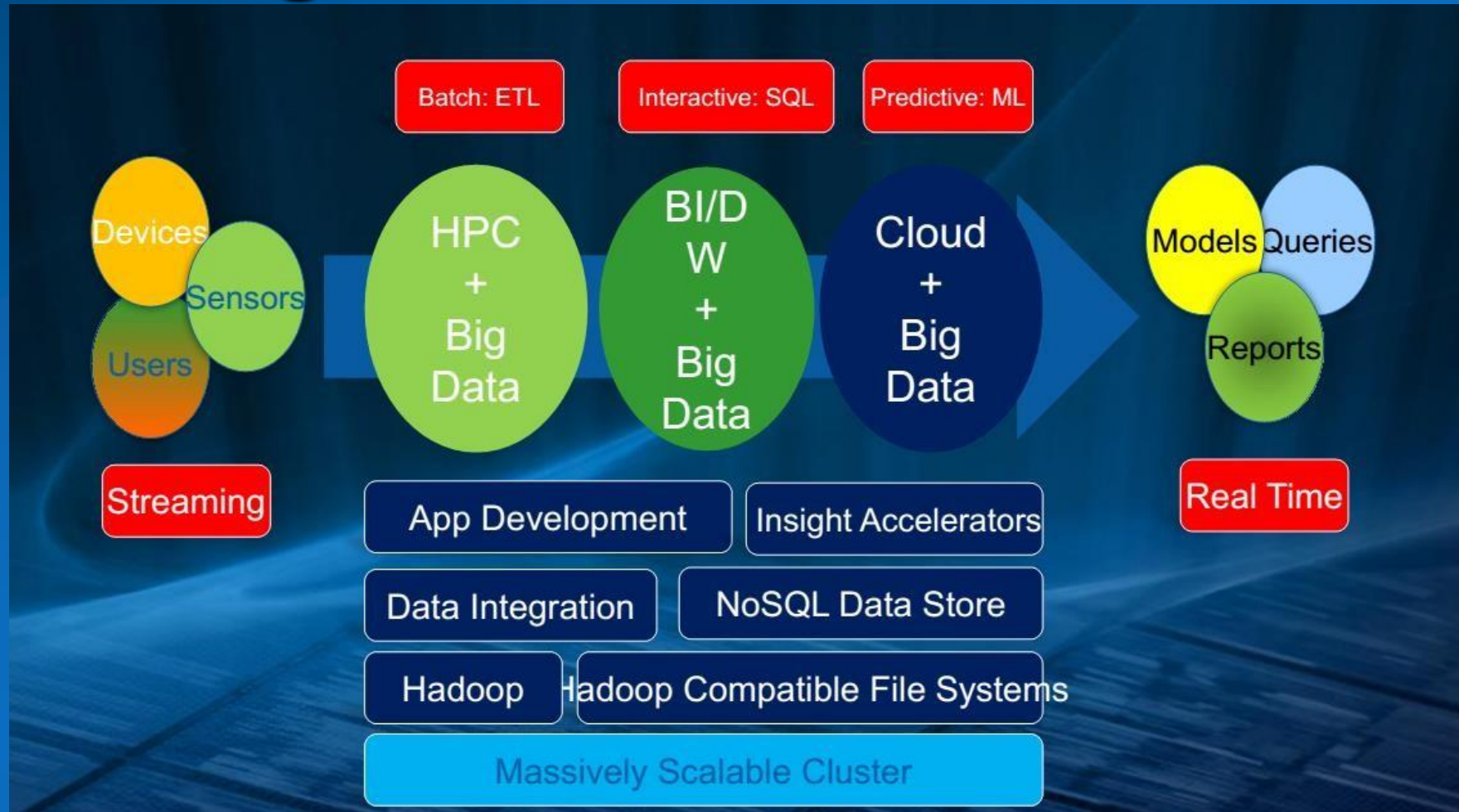


Modernization of your code is the solution

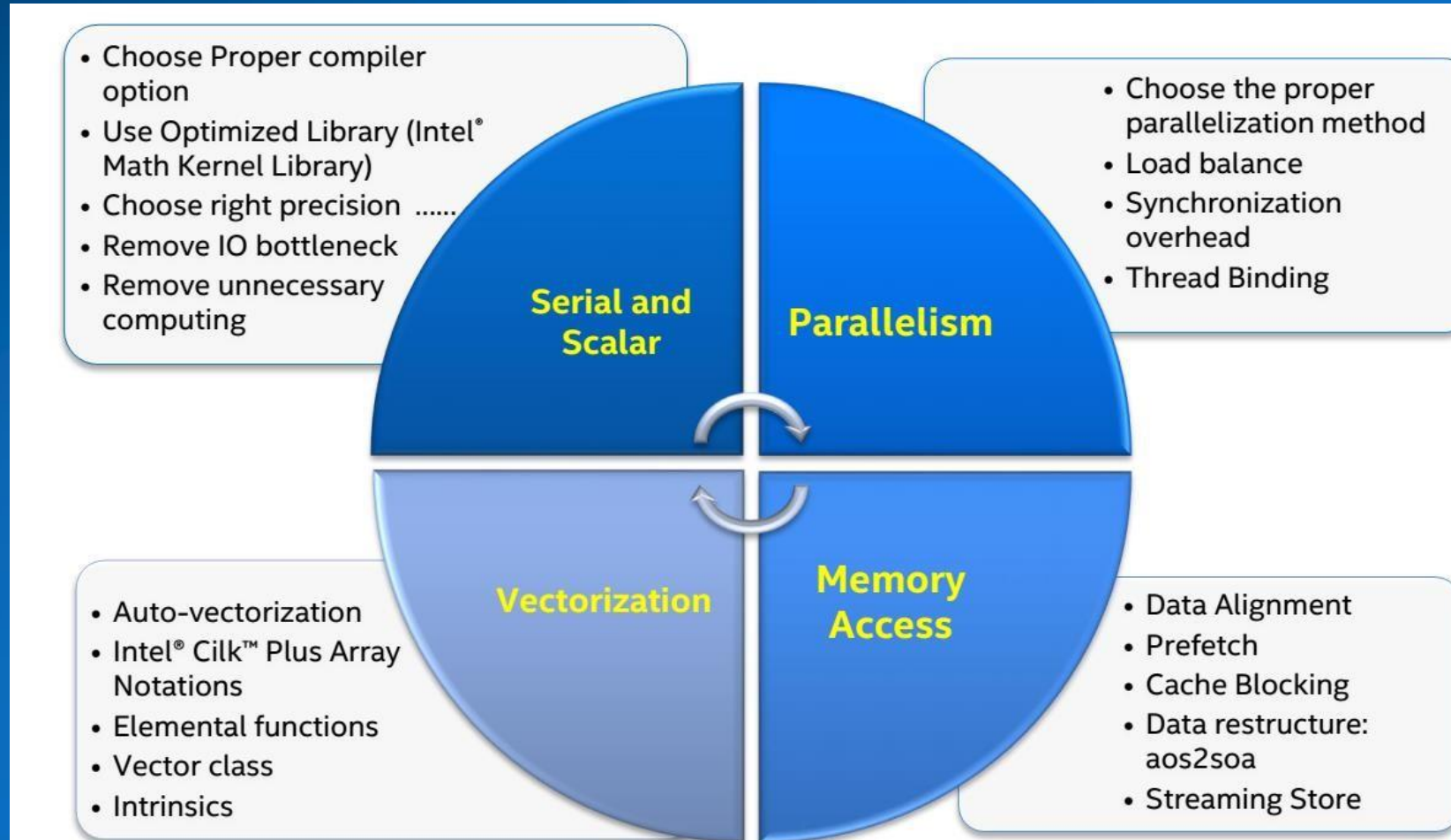
The Many Ways to Parallelism



And New Workloads will Emerge



Code Modernization – The 4D Approach



Intel® Xeon Phi™ Product Family

Based on Intel® Many Integrated Core (MIC) Architecture

“Meet **Knight's Landing: Intel's most powerful chip ever** is overflowing with cutting-edge technologies”

PCWorld – Jun 23, 2014



2013

Knights Corner

Intel® Xeon Phi™
x100 Product Family

22 nm process
Coproprocessor
Over 1 TF DP Peak
Up to 61 Cores
Up to 16GB GDDR5



2016

Knights Landing

Intel® Xeon Phi™
x200 Product Family

14 nm process
Processor &
Coproprocessor
Over 3 TF DP Peak
Up to 72 Cores
On Package High-
Bandwidth Memory
3X Single-Thread
Out-of-order core

Future

Knights Hill

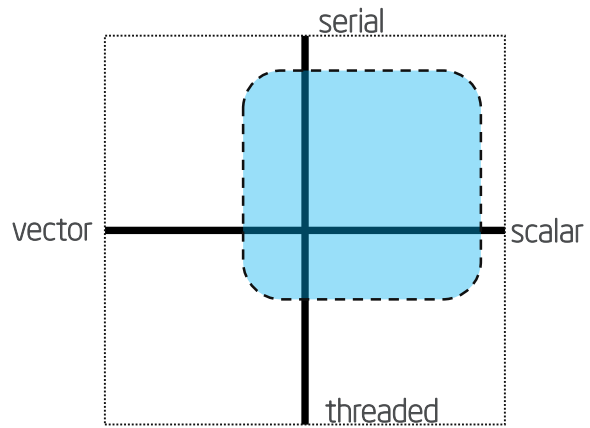
Next generation of
the Intel® MIC
Architecture Product
Line

In planning

*Per Intel's announced products or planning process for future products

No Positioning Change

Knights Landing Targeted for Highly-Vectorizable, Parallel Apps



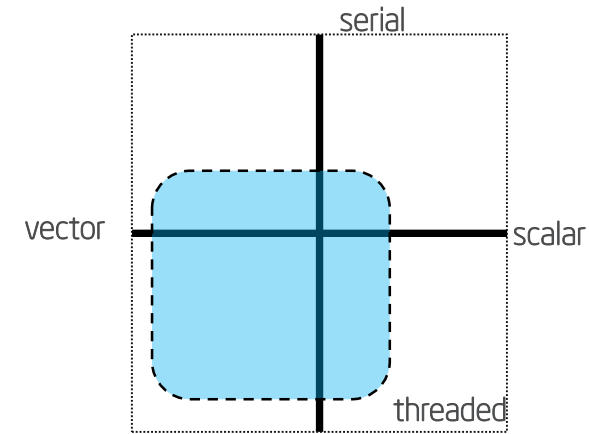
Most Commonly Used Parallel Processor*

Parallel, Fast Serial
Multicore + Vector
Leadership Today and Tomorrow

Single Source Code Optimization



+

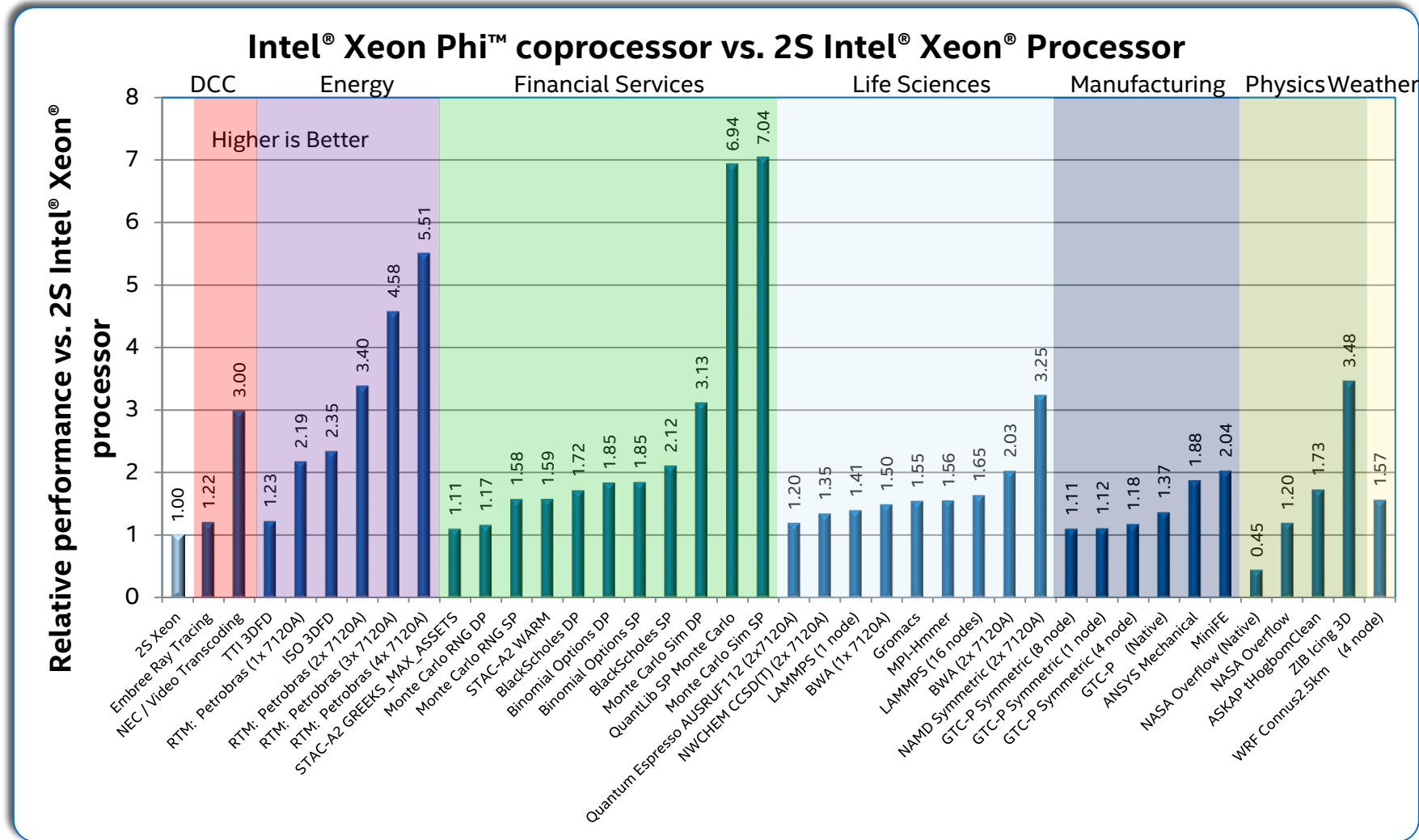


Optimized for Highly-Vectorizable Parallel Apps

Many Core
Support for 512 bit vectors
Higher memory bandwidth
Common SW programming

*Based on highest volume CPU in the IDC HPC Qview Q1'13

Is Xeon Phi compelling vs Xeon?



Intel Measured Results: Different hardware architectures may require different source code. Results are based on Intel's best efforts to use code optimized to run best on all architectures and perform the same work. Future code optimizations may result in different results.

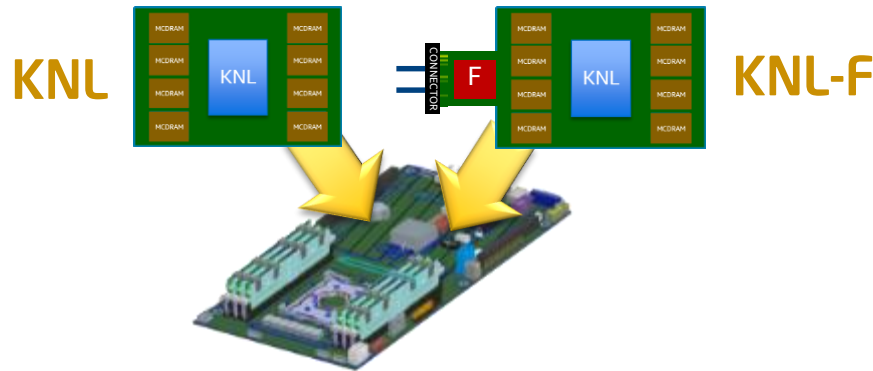
“Rifle shot” approach targeted with customers in FSI, Oil & Gas, and Life Sciences based on affinity/readiness

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. Source: Intel measured as of Q1 2014 Configuration Details: Please slide speaker notes. For more information go to <http://www.intel.com/performance>

❖ Xeon = Intel® Xeon® processor
❖ Xeon Phi = Intel® Xeon Phi™ coprocessor

Three Knights Landing Products

Knights Landing Processor *"Self-boot" Intel® Xeon Phi™ processor platform*



KNL and KNL-F Processors:

- Knights Landing IS the host processor
- Boots standard off-the-shelf OS's

Benefits:

- Higher performance density for highly parallel applications²
- Reduced system power consumption²
- Higher perf/Watt & perf/\$\$³

Knights Landing Coprocessor

Requires Intel® Xeon® processor host



Knights Landing Coprocessor:

- Solution for general purpose servers and workstations

Benefits:

- Targeted for applications with larger sections of serial work¹
- Upgrade path from Knights Corner as PCIe card

¹ Projections based on early product definition and as compared to prior generation Intel® Xeon Phi™ Coprocessors

² Based on Intel internal analysis. Lower power based on power consumption estimates between (2) HCAs compared to 15W additional power for KNL-F. Higher density based on removal of PCIe slots and associated HCAs populated in those slots.

³ Results have been estimated based on internal Intel analysis and are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance. ² Results based on internal Intel analysis using estimated theoretical Flops/s for KNL processors, along with estimated system power consumption and component pricing in the 2015 timeframe. See backup for complete system configurations.

A Paradigm Shift for Highly-Parallel

Server Processor and Integration are Keys to Future



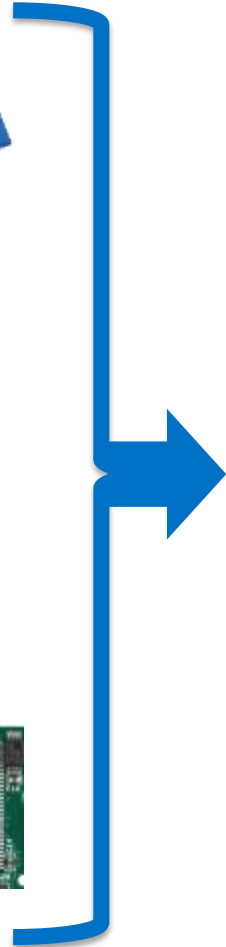
Coprocessor



Fabric



Memory



Server Processor

Memory Bandwidth

~500 GB/s STREAM

Memory Capacity

Over 25x* KNC

Resiliency

Systems scalable to >100 PF

Power Efficiency

Over 25% better than card¹

I/O

Up to 50 GB/s BW with int fabric

Cost

Less costly than discrete parts¹

Flexibility

Limitless configurations

Density

3+ KNL with fabric in 1U³

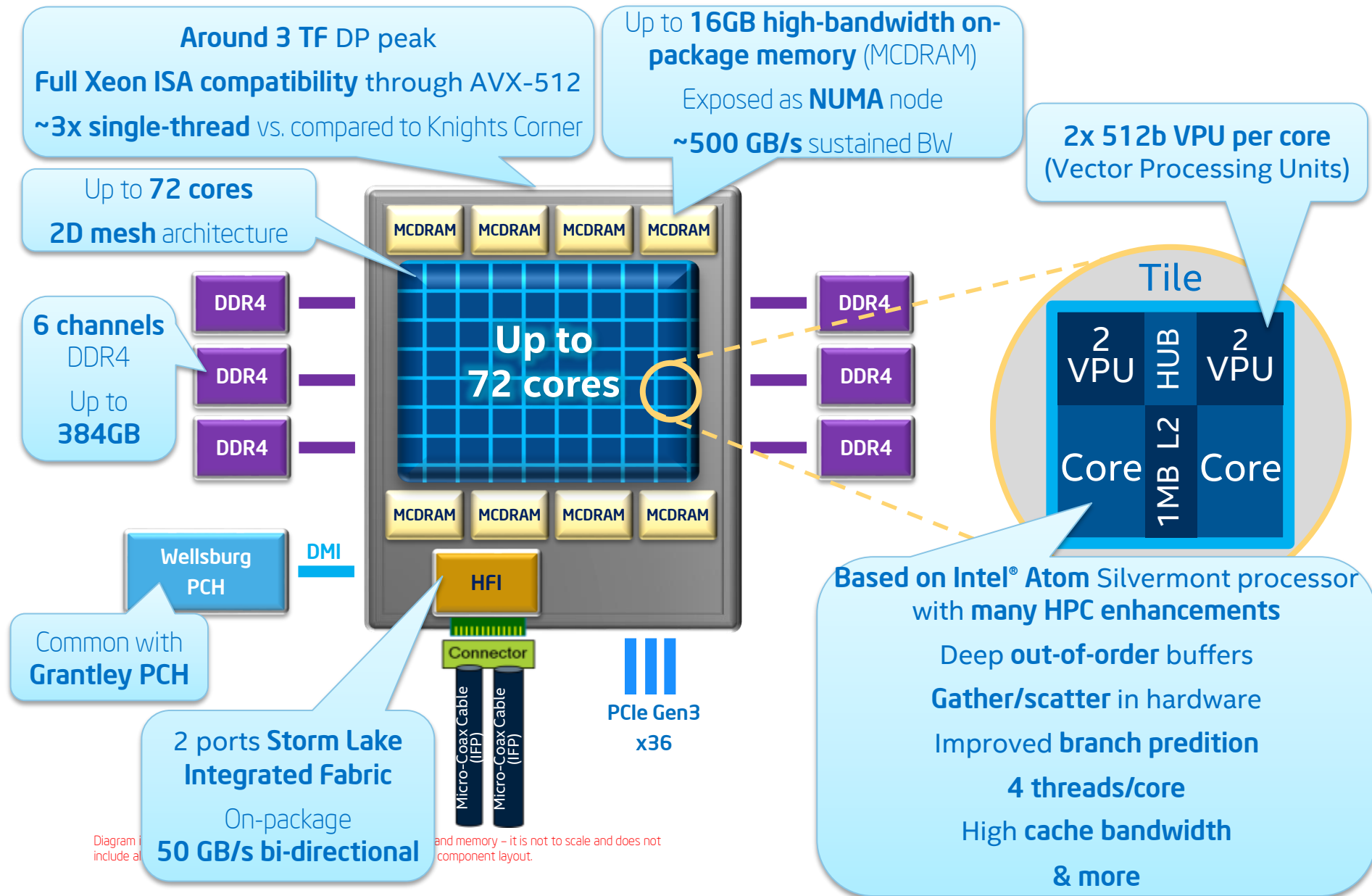
^{*}Comparison to 1st Generation Intel® Xeon Phi™ 7120P Coprocessor (formerly codenamed Knights Corner)

¹Results based on internal Intel analysis using estimated power consumption and projected component pricing in the 2015 timeframe. This analysis is provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance.

²Comparison to a discrete Knights Landing processor and discrete fabric component.

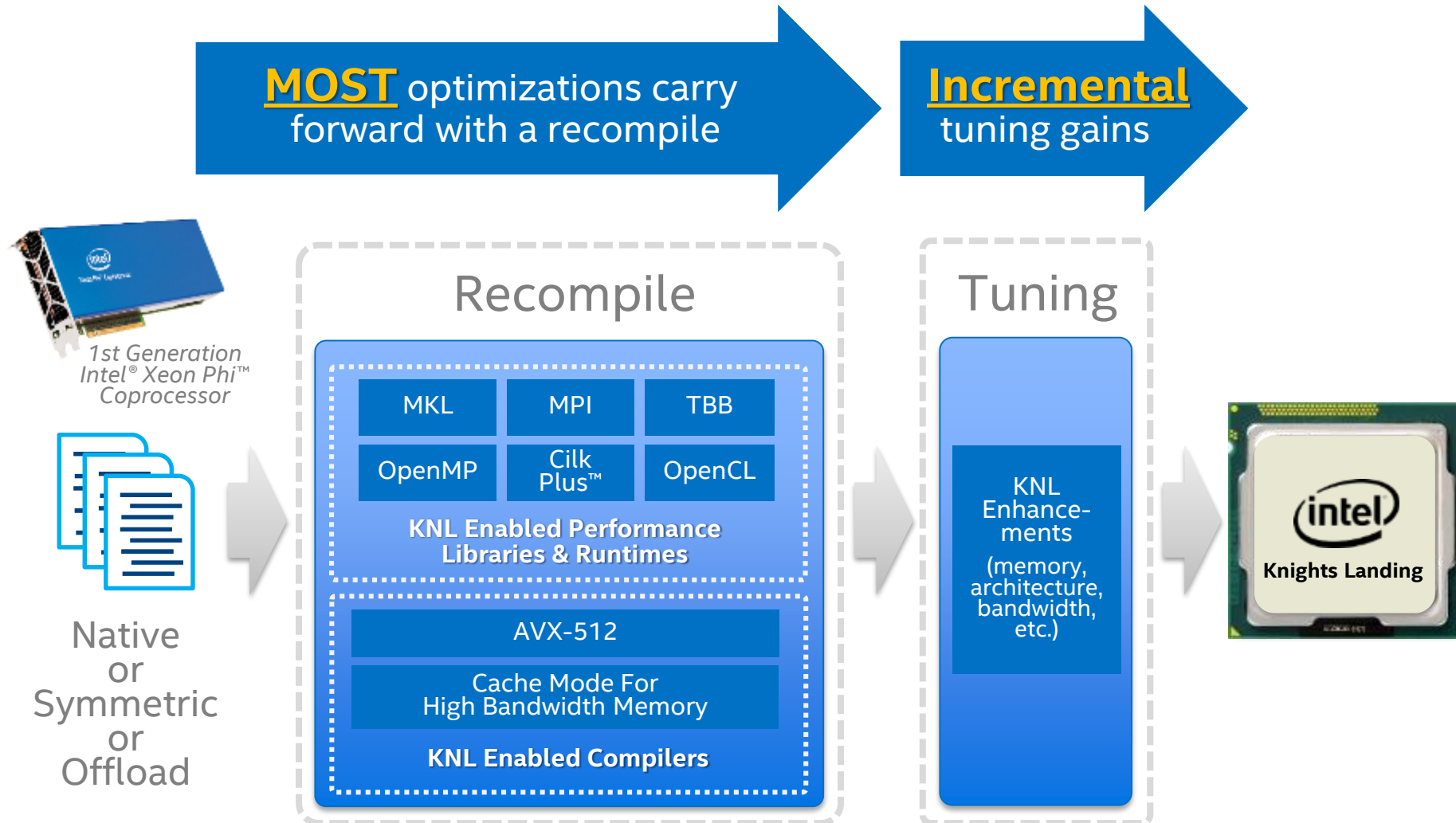
³Theoretical density for air-cooled system; other cooling solutions and configurations will enable lower or higher density.

Knights Landing Architectural Diagram



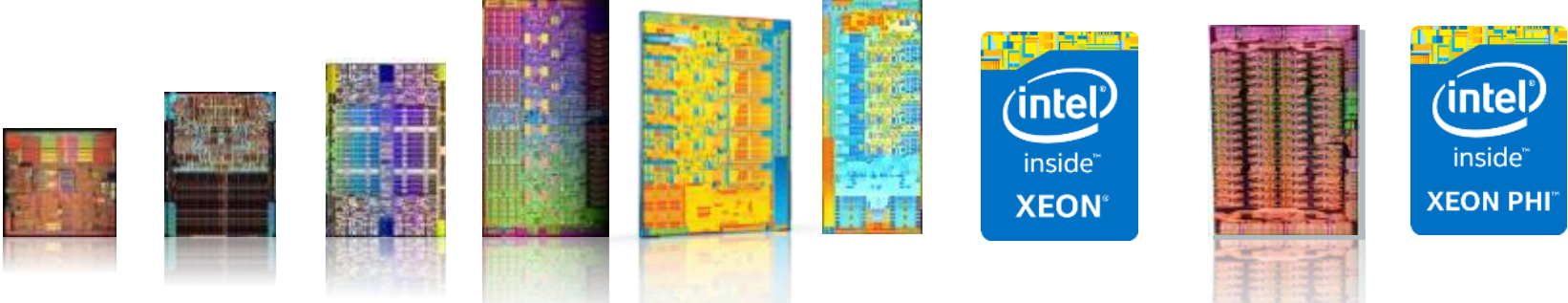
Today's Parallel Investment Carries Forward

Sustained threading, vectorization, cache-blocking and more



Parallel is the Path Forward

Intel® Xeon® and Intel® Xeon Phi™ Product Families are both going parallel



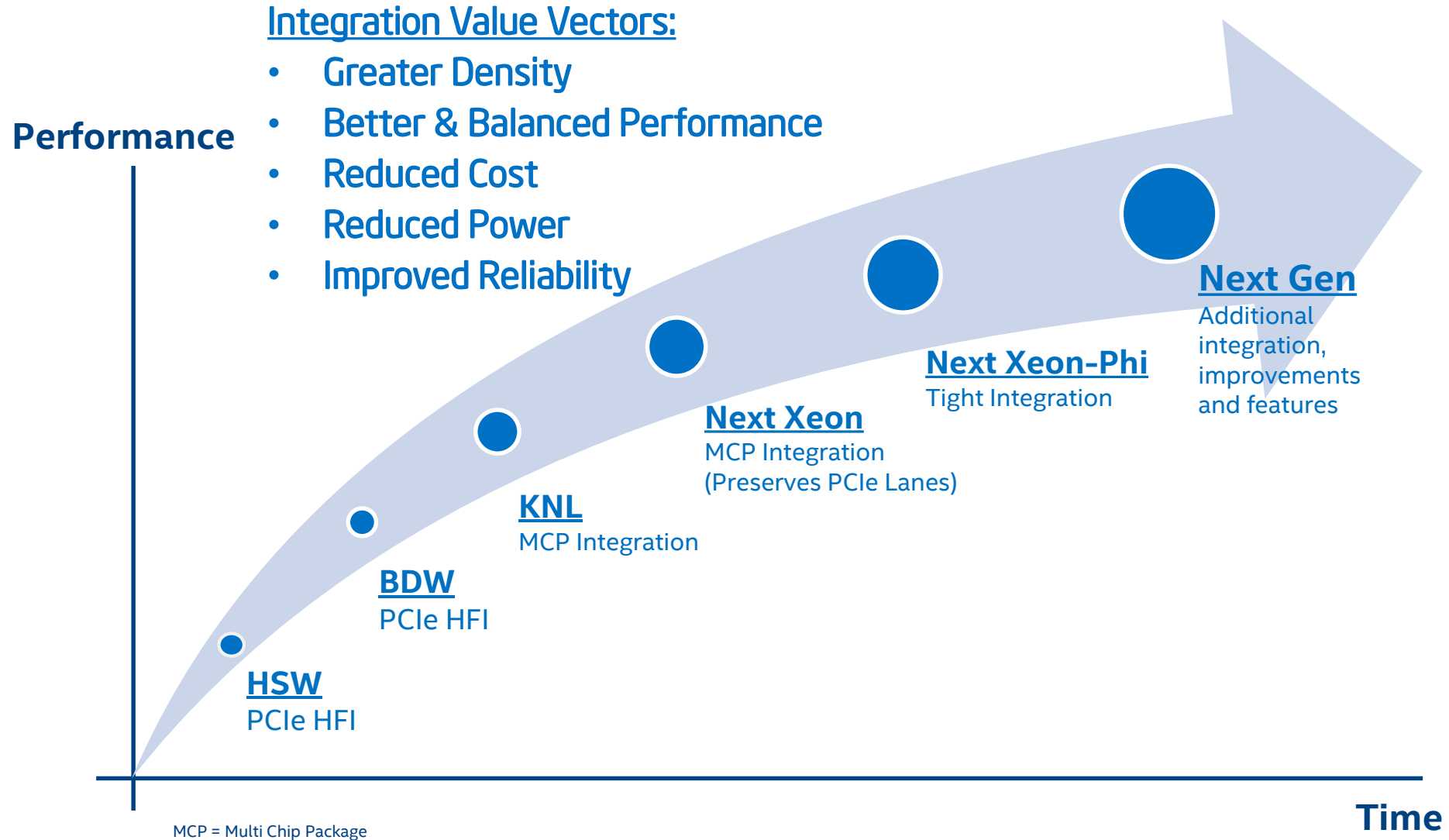
	Intel® Xeon® processor 64-bit	Intel® Xeon® processor 5100 series	Intel® Xeon® processor 5500 series	Intel® Xeon® processor 5600 series	Intel® Xeon® processor code-named Sandy Bridge EP	Intel® Xeon® processor code-named Ivy Bridge EP 4S ¹	~	Future Intel® Xeon® processor	~	Intel® Xeon Phi™ coprocessor Knights Corner	Intel® Xeon Phi™ processor & coprocessor Knights Landing ¹
Core(s)	1	2	4	6	8	12	~	<i>tbd</i>	~	57-61	Up to 72
Threads	2	2	8	12	16	24	~	<i>tbd</i>	~	228-244	Up to 288
SIMD Width	128	128	128	128	256	256	~	512	~	512	512

More cores → More Threads → Wider vectors

*Product specification for launched and shipped products available on ark.intel.com.

1. Not launched or in planning.

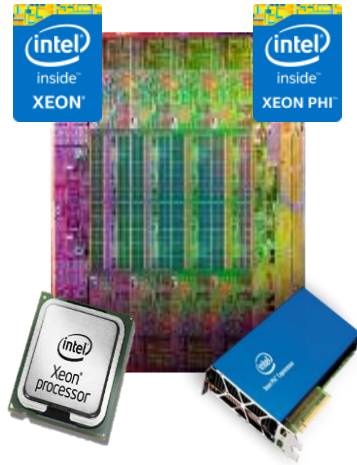
Intel® Omni-Path Fabric - CPU Integration



Intel's Technical Computing Portfolio

Technologies & Products

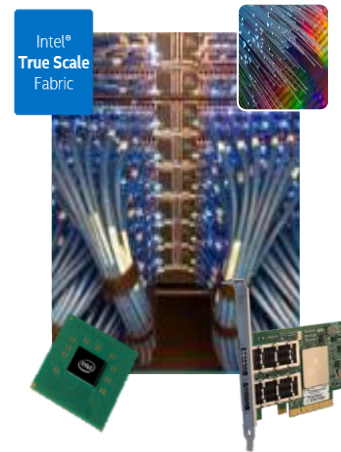
Compute Processing



Systems & Boards



Network & Fabric



I/O & Storage

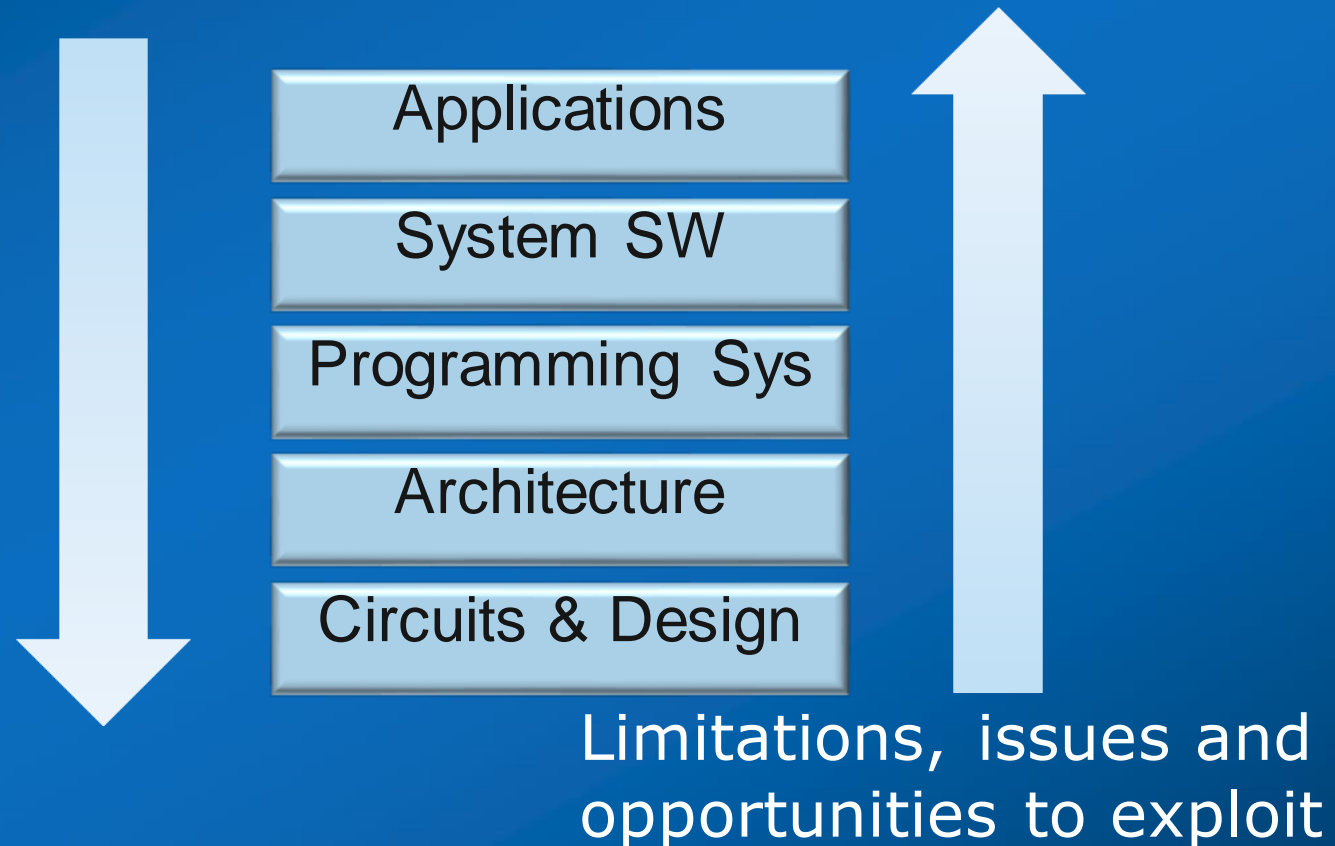


Software & Services



HW-SW Co-design

Applications and SW stack
provide guidance for efficient
system design



What will matter in 10 years

	Now	2025
Perf/\$	Linpack, Real Applications	Real Applications
Perf/Watt	Limited by worst case application	All applications will be able to run at chosen power level. Dynamic, optimal energy management.
Reliability	Use of file system checkpoint restart (spinning disks)	Transparent hardware and system software recovery. Checkpoints in non-mechanical media.
Big Data	Parallel IO	New storage paradigm

SW Challenges

Execution model

Programming model

1. Extreme parallelism (1000X due to Exa, additional 4X due to NTV)
2. Data locality—reduce data movement
3. Intelligent scheduling—move thread to data if necessary
4. Fine grain resource management (introspective)
5. Applications and algorithms incorporate paradigm change

Summary

- Exascale will be there by 2022 or so
- “Business as usual” (riding on Moore’s Law and commodity technology) is becoming increasingly harder
- Supercomputers are becoming more “special purpose”
 - Expect most/all supercomputers to use floating point accelerators in a few years; more specialized accelerators to follow
- Can continue to push performance to zetascale
 - Will need to think of supercomputers as unique facilities, such as particle accelerators – not clusters of PCs
- *Supercomputing will become much more interesting*