

HPC & Big Data

THE TIME HAS COME FOR A SCALABLE FRAMEWORK

**Barry Davis, General Manager, High Performance Fabrics Operation
Data Center Group, Intel Corporation**

Legal Disclaimer

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.

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 does not control or audit the design or implementation of third party benchmarks or Web sites referenced in this document. Intel encourages all of its customers to visit the referenced Web sites or others where similar performance benchmarks are reported and confirm whether the referenced benchmarks are accurate and reflect performance of systems available for purchase.

Intel processor numbers are not a measure of performance. Processor numbers differentiate features within each processor family, not across different processor families. See www.intel.com/products/processor_number for details.

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.

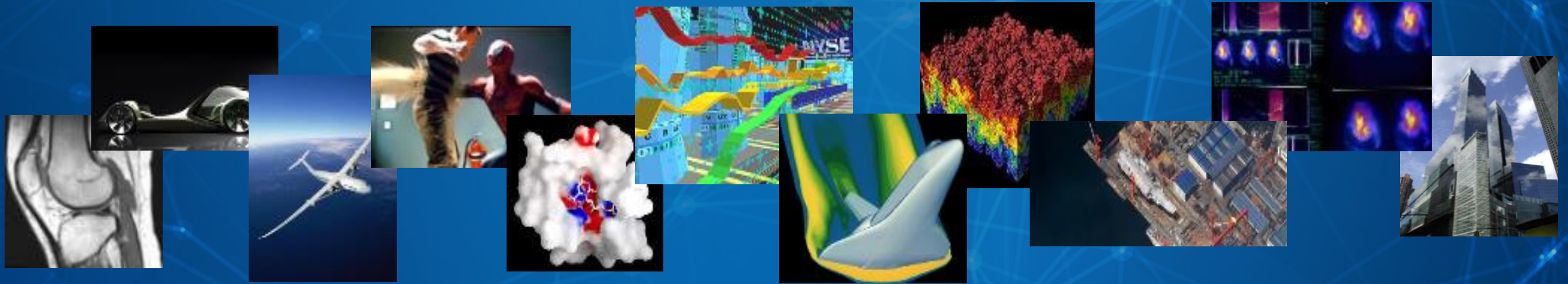
Intel, Intel Xeon, Intel Core microarchitecture, and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

Copyright © 2011, Intel Corporation. All rights reserved.

HPC As We Know It - Solves Problems of Every Scale

...Computing to generate insight through simulation, prediction, visualization or analysis of systems and data ...



HPC has become a critical global competitive differentiator

“to be a superpower you must lead in supercomputing...” CTO, Dawning (Sugon)

HPC is highly correlated to economic competitiveness, scient. advances, & national prestige

... “To out-compute is to out-compete,” Council on Competitiveness

In a worldwide IDC study, 97% of companies that adopted supercomputing said they could no longer compete or survive without it

Big Data – We All Know What This Is



Big Data – Volume/Variety/Velocity

3 V's of Big Data

Just having large amounts of data does not make it Big Data. Big Data is defined by the 3 V's.

● VOLUME ● VARIETY ● VELOCITY

3 V's

VOLUME

Petabytes, not gigabytes

“The sum of all the words you will hear in your lifetime”

VARIETY

Typically unstructured

Tweets, chats, email, videos

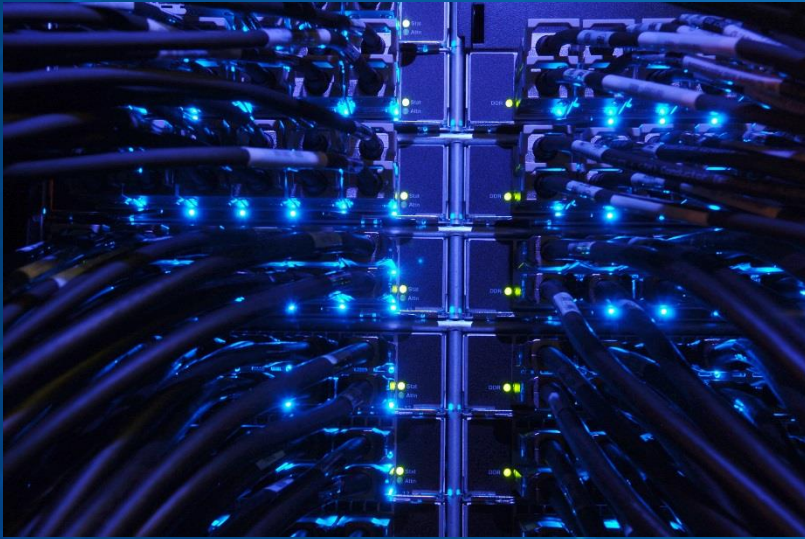
VELOCITY

Real-time or near-real-time

Continually loaded and/or analyzed

HPDA – High Performance Data Analysis

Where HPC Meets Big Data



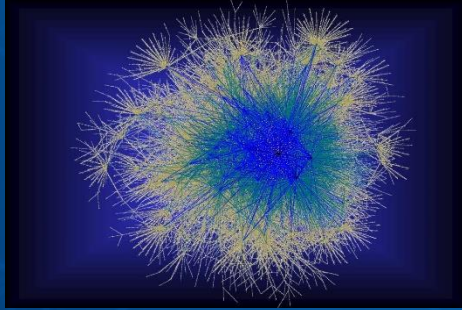
+



Driven By:

- Increasingly More Powerful HPC Systems Running Complex, Data-intensive Sims.
- Proliferation Of Data Gathering Sensors And Networks (More Data To Analyze)
- Growth & Improvement Of Modeling And Other Iterative Problem-solving Methods
- The Need For Real Time Solutions

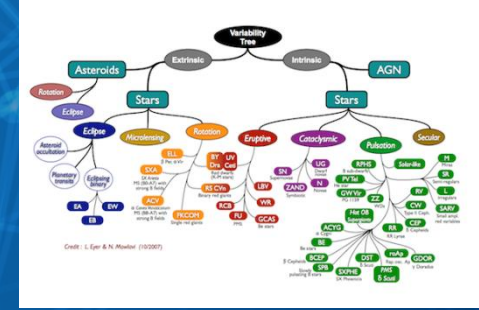
What Happens When HPC Collides With Big Data



Highly Complex



Highly Time Critical



Highly Variable

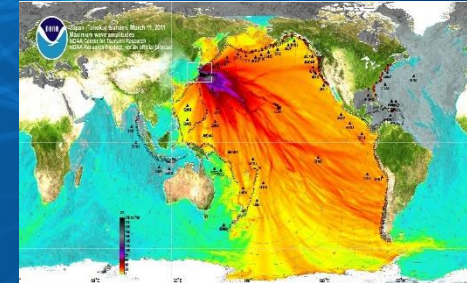
- **Graphing and mapping:** Greater accuracy in business forecasting
- **Pattern visualizations:** New views of complex data sets, enabling rapid identification of relationships for simple analysis
- **Scaling in-memory databases:** Complete data sets means benefit from real-time analytics while in motion
- **Meta-data:** The importance of metadata will jump dramatically resulting in enterprises finding new business insights and even revealing security threats

Some of Today's Use Cases

Weather

NOAA – HPC & big data analysis using HIRAM (Global Atmospheric Model) → greater accuracy in forecast/warning guidance to the public & weather forecasters.

in a run that took <4 hours.



Fraud detection

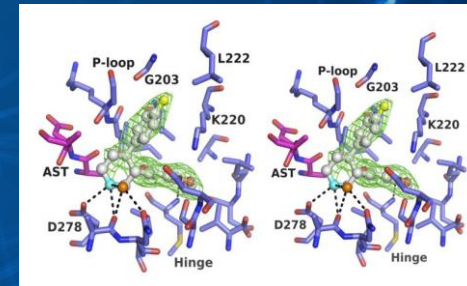
PayPal – Using HPC & Big Data analysis, PayPal analyzed suspicious patterns in over 13M transactions/day, saving \$710M in fraud during the first year

US Postal Service – Using MCDB & TimesTen app on an HPC system, the USPS processed 4B transactions in <6 hours resulting in real time fraud detection & rev. protection



Drug Design

Novartis – Using the Schrodinger drug discovery app in conjunction with HPC, Novartis tested 21 million small molecule drug candidates in a run that took <4 hours.



Financial Services

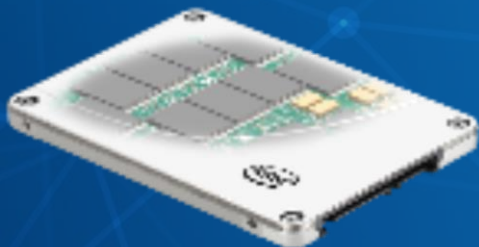
Company X – Using HPC and analytics development, Company X increased their efficiency in updating credit ratings by 30%.



What's Needed For Great HPDA?

(Hint...Same as What's Needed for Great HPC)

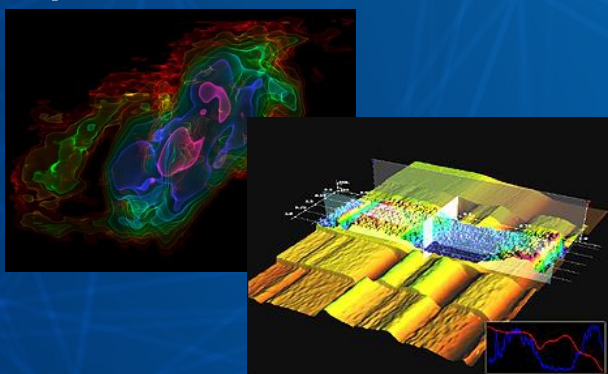
Low Latency, Fast Access Storage



High Throughput, Low Latency Interconnect



Complex Data Visualization



Multicore Processing



High-performance, Large Scale Data Storage

HPC is Entering New Era – An Era of Change

Intel Innovations Overcoming Architectural Challenges



Breaking Down “The Walls”

Memory | I/O | Storage
Energy Efficient Performance
Space | Resiliency |
Unoptimized Software



Fast and Efficient Data Mobility

Rapidly Growing
Big Data Analytics

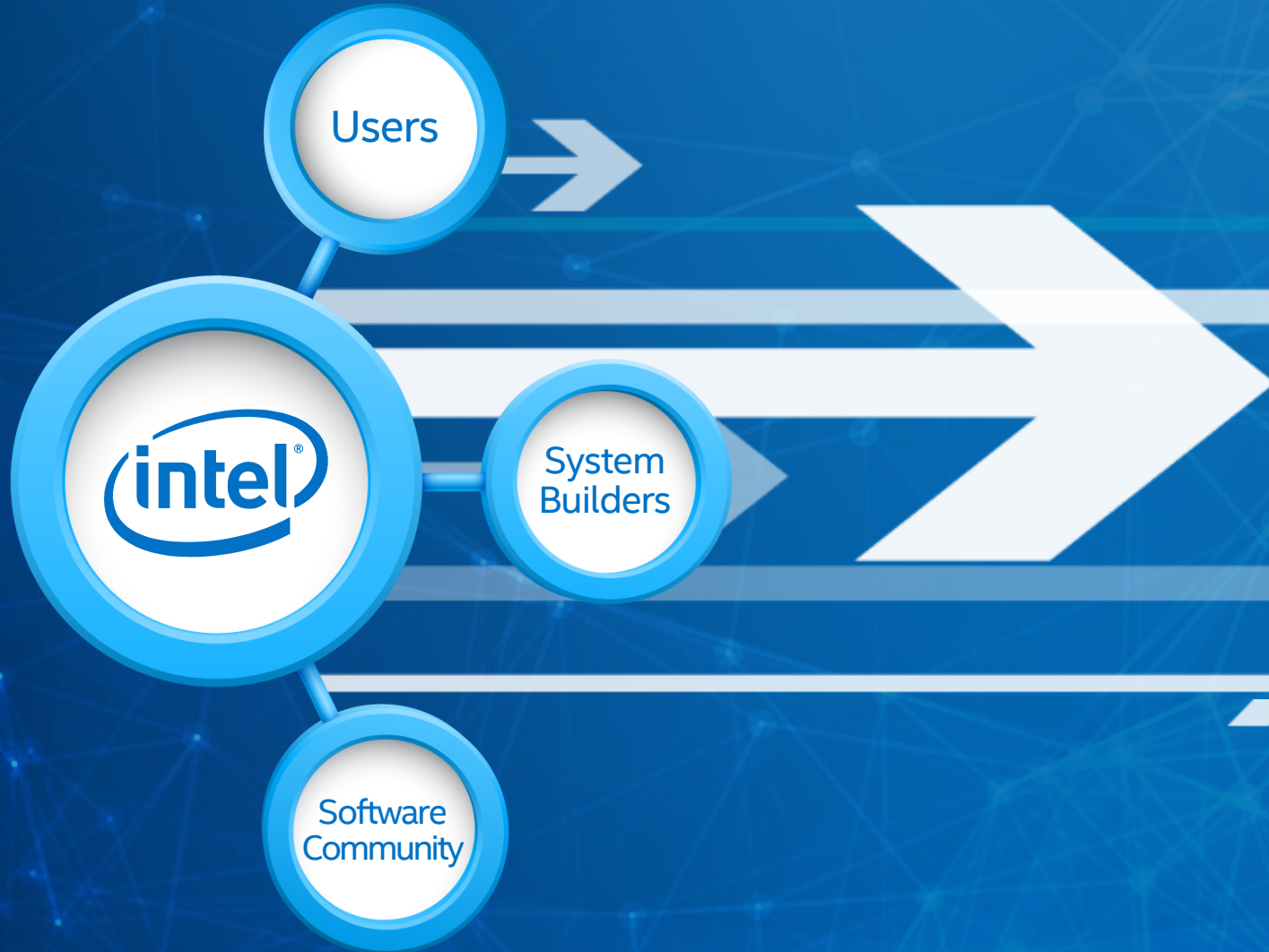


Extending HPC's Reach

Democratization at Every
Scale Cloud Access |
Exploration of New Parallel
Programming Models

Intel-led Collaboration: Unprecedented Breakthroughs

Brings innovations, holistic designs, and the means to deliver the full benefits to users



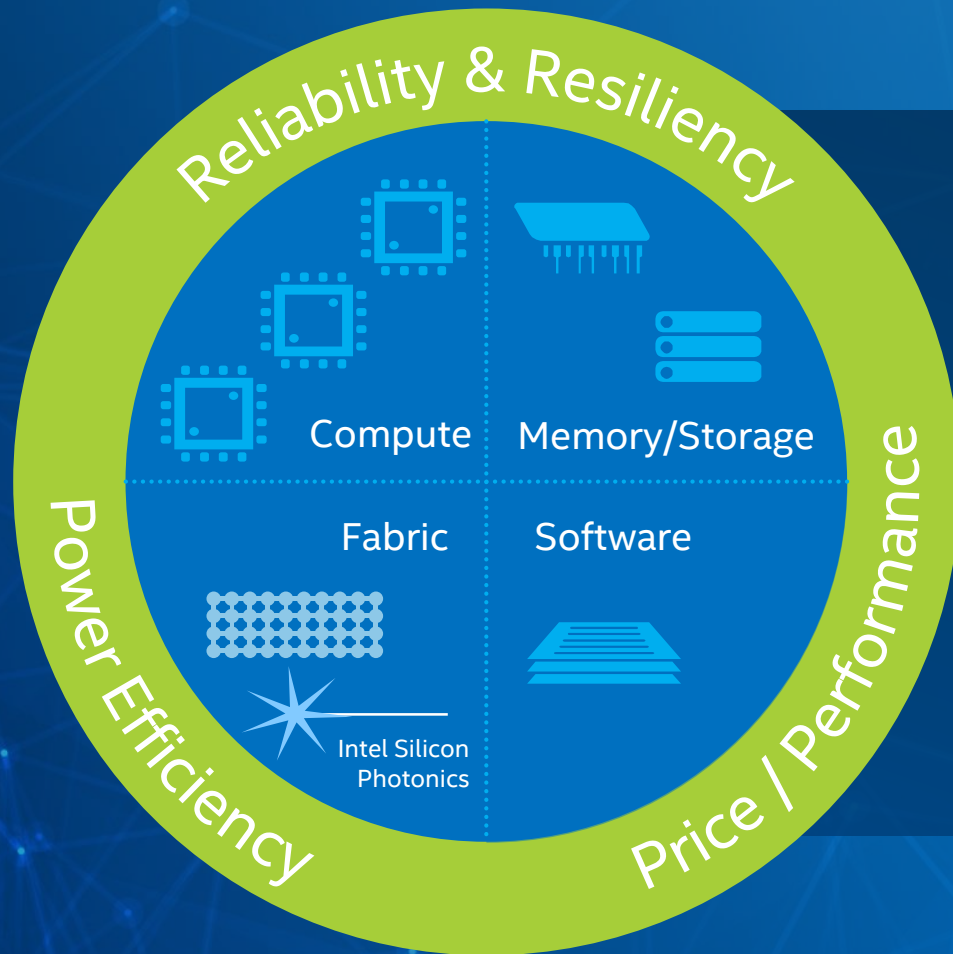
Expanding portfolio of game changing **technologies** in a **scalable system design framework**

Co-design approach that optimizes for overall workload performance, efficiency and reliability

Thriving, open, enabled, and innovating **ecosystem**

Intel's HPC Scalable System Framework

A Design Foundation Enabling Wide Range Of Highly Workload-optimized Solutions



Small clusters to Supercomputers

Compute and Data-Centric Computing

Standards-Based Programmability

Intel® Xeon® Processors
Intel® Xeon Phi™ Coprocessors
Intel® Xeon Phi™ Processors

Intel® True Scale Fabric
Intel® Omni-Path Fabric
Intel® Ethernet

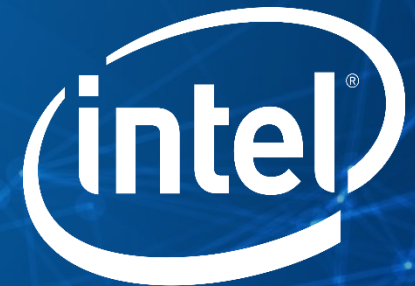
Intel® SSDs
Intel® Lustre*-based Solutions
Intel® Silicon Photonics Technology

Intel® Software Tools
Intel® Cluster Ready Program

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

The Most Advanced Supercomputer Ever Built

An Intel-led collaboration with ANL and Cray to accelerate discovery & innovation



Prime Contractor



Subcontractor

>180 PFLOPS

(option to increase up to 450 PF)

>50,000 nodes

13MW

2018 *delivery*

18X higher
performance*

>6X more energy
efficient*

Source: Argonne National Laboratory and Intel. *Versus ANL's current biggest system named MIRA (10PFs and 4.8MW)
Other names and brands may be claimed as the property of others.

Thank You