High Performance Computing: Evolution or Re-Definition?

HPCS 2003
Sherbrooke, Quebec, Canada
May 11-14, 2003

David Barkai
HPC Computational Architect
Intel
What this talk is about

- Observations about the evolution of HPC
- Choices & Metrics
- Challenges
- From clusters to grids
- Convergence of technologies?
HPC Evolution

- Custom CPUs
- Custom memory
- Custom packaging
- Custom interconnects
- Custom OS

- COTS CPUs
- COTS memory
- Custom packaging
- Custom interconnects
- Custom OS

1980 ~$5 million/gigaflop

1990 ~$200K/gigaflop

2000 <$2K/gigaflop

From ~1GF → ~100’s GF → >10TF

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
The Enabling Hardware Technologies

**EUV Lithography**

**15nm TeraHertz Transistor**
(Performance 📈, Power 📉)

**300mm Wafers**

**Manufacturing (130nm)**

**52-megabit chip one square micron SRAM cell on 90nm**

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.*
Silicon by the end of the Decade

“30 gigahertz devices, 10 nanometer or less delivering a tera instruction of performance by 2010”\(^{(1)}\)

1) Pat Gelsinger, Intel CTO, Spring 2002 IDF

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
HPC Evolution

Vertical Model

- High Performance Computing
- Server Consolidation
- Workforce Productivity
- Technology Refresh

- Verticals:
  - Energy
  - Manufacturing
  - Communications
  - Digital Media
  - Financial Services
  - Retail & CPG
  - Government & Education
  - Healthcare
  - Life Sciences

- Additional Services:
  - CRM, SCM, EAI, ERP
  - System Management
  - Security, Mobility/Occasionally Connected

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
**Mutual Interests and Objectives**

- Technology Leadership
- Solution Enabler
- Market Maker
- Key Strategic Influencer
- Products

- Technology Adopters
- Grand Challenge Problems
- Market Mover & Influencer
- Early Innovators
- Computational Needs

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.*
iA Solution Deployments over the last 18mos

Industry
• Western GECO
• GCC
• Shell
• Saudi Aramco
• Tensor Geophysical
• WETA
• Pemex
• Google
• Inpharmatica
• Syrxx
• Immunex
• MDS Proteomics
• Pixar
• British Petroleum
• DaimlerChrysler
• SAS
• Ifineon
• Volvo

Academic
• SUNY Buffalo
• Denmark Scientific
• Mississippi State
• Louisiana State
• Brookhaven
• Clemson
• Utah
• Cornell
• Toronto
• Virginia Polytech
• Ohio State
• Tsinghua Uni.
• Imperial College
• National U of Singapore
• Swinburne
• Stanford Linear Accelerator
• South Hampton
• Valencia
• Oxford
• Johns Hopkins
• Cal Tech
• Belfast
• Zeijing
• Princeton

Government
• Tera-Grid
• Los Alamos National Lab
• Lawrence Livermore National Lab
• NOAA
• Pacific Northwest National Lab
• Sandia National Lab
• NCSA
• Classified Defense Sites
• China Atmospheric Center
• CERN
• China Atmosphere
• Seoul National University Grid
• Ohio Supercomputing Center
• Inria

…and many others

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
The Growing Popularity of Cluster Computing

The number of clusters in the TOP500 has grown to nearly 20 percent, with a total of 93 systems, 56 of them are Intel Architecture based.

Los Alamos builds largest Infiniband cluster

Science and Engineering News

11/28/02

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
The Community Triangle

HPC Community

HPC-Focused Computational Scientists

Fellow-travelers: OEM, IHV, ISV

Intel

World-wide HPC team chartered with being the conduit between Intel and the HPC user community

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
Intel® Architecture Server Processor Roadmap

**Performance, RAS, Scalability**

- Intel® Itanium® architecture
- 1 GHz
- 0.18µ
- 3MB IL3 cache

- Intel® Itanium® 2 processor platform compatible
- 0.13µ
- 6MB IL3 cache
- Intel® Itanium® 2 processor platform compatible

- Second half of decade could see:
  - <65nm process
  - ~1 billion transistors
  - Multi-threading on Itanium Processor Family

- Possible uses for 1Bn transistors:
  - Very large L3 caches
  - Multiple processor cores with shared cache

- Future Directions

- Deerfield
  - ~60w
  - ~1.5MB
  - 10Gb SerDes

- PCI-Express

- Future Directions

- Second half of decade could see:
  - <65nm process
  - ~1 billion transistors
  - Multi-threading on Itanium Processor Family

- Possible uses for 1Bn transistors:
  - Very large L3 caches
  - Multiple processor cores with shared cache

- All dates specified are target dates, are provided for planning purposes only and are subject to change.

*Other brands and names are the property of their respective owners

© Copyright 2002 Intel Corporation. All Rights Reserved.
More than Silicon

Innovating at all levels of the platform

*Other names and brands are property of their respective owners.

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
Blade Clusters - The Next Wave?

- 88 RackBlades, 88 Nodes
- Up to 176 Intel® Xeon™ Processors
- Up to 704GB of DDR Memory
- GigE & Myrinet™ options

500 GFLOPS

Low Voltage Intel® Xeon™ processor

<table>
<thead>
<tr>
<th>Product Number</th>
<th>Core Speed (GHz)</th>
<th>L2 Cache</th>
<th>External Bus Speed (MHz)</th>
<th>Thermal Design Power (Max)</th>
<th>Voltage</th>
<th>T_case</th>
<th>Package</th>
</tr>
</thead>
<tbody>
<tr>
<td>RK80532EC025512</td>
<td>1.6</td>
<td>512K</td>
<td>400</td>
<td>30.0W</td>
<td>1.3V</td>
<td>81C</td>
<td>604 pin FC-mPGA-2P</td>
</tr>
</tbody>
</table>

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
Some choices of cluster interconnect

<table>
<thead>
<tr>
<th>Interconnect</th>
<th>Topology</th>
<th>MPI Latency</th>
<th>Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myrinet*</td>
<td>Fat Tree</td>
<td>7 us</td>
<td>243 MB/s</td>
</tr>
<tr>
<td>Dolphin* SCI</td>
<td>2D/3D Torus</td>
<td>&lt; 4 us</td>
<td>320 MB/s</td>
</tr>
<tr>
<td>Quadrics*</td>
<td>Fat Tree</td>
<td>&lt; 5 us</td>
<td>340 MB/s</td>
</tr>
<tr>
<td>InfiniBand* 4x</td>
<td>Various</td>
<td>7.5 us</td>
<td>805 MB/s</td>
</tr>
</tbody>
</table>

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
Software Development Tools

**SW Products**

- Compilers
- Performance Libraries
- VTune™ Performance Analyzer
- Intel® Threading Tools

**Developer Services**

[Intel's Developer Site](www.intel.com/ids)

*Other brands and names are the property of their respective owners*

© Copyright 2002 Intel Corporation. All Rights Reserved.
Choices, Metrics

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
More than Silicon – Broad Choice of Solutions and Services

Choices, Metrics

*Other brands and names are the property of their respective owners

© Copyright 2002 Intel Corporation. All Rights Reserved.
Solutions Market Development

- Solutions Channels
- Solutions Enabling
- Products
- Architecture / Standards

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
Differentiation of HPC

- Computationally intensive
- Large scale apps
- Random, dynamic data access patterns
- Pre-early adopters
- Expert users
Architectural implications

- Fast computational units
  - Not just floating point
- Large memory with high bandwidth & low latency
  - Larger caches don’t always help
  - Maintain bandwidth/latency within n-way node
- Wide, low-latency interconnect for distributed applications

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
We’d like to believe that...

- The academic and public sector HPC serves as proving grounds for Industry / private sector
- High-end technologies experimented by HPC community will go main-stream in 1-2 ‘technology generations’

The Industry needs to invest in HPC because its community is technology trend setter and innovator-adopter

*Other brands and names are the property of their respective owners © Copyright 2002 Intel Corporation. All Rights Reserved.
Needs of apps drive selection metrics

- Cannot satisfy every job - Step 1 = Determine the apps/datasets that matter!
  - Not necessarily the most demanding..

- App “environment” factors; such as --
  - Distribution of workload by number and size of jobs
  - Distribution by required time to solution and size
  - Frequency of compilations, restarts, ..

- App characteristics; such as --
  - Data access patterns and size
  - Granularity of parallelism
  - Computations - per data accessed; float/integer mix
The Metrics form a multidimensional Space

Interconnect
- Bandwidth from node
- Latency
- Bisection bandwidth

Memory
- Bandwidth
- Latency
- Size

Operational Env
- Size
- Power / rack, total
- File System*

Computations
- Flop rate
- Integer rate
- Float/Integer mix

* File system choice can be thought as derived from app needs, or as an arch choice that drives others.

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
Deriving a Metric

- Find your part of a surface or a bounded subspace in the universe of the metrics
- Determine the relationships among parameters in your metrics space
- Examples:
  - Max B:F in node, but <10
  - B:F:I = 4:1:0.5 or 1:1:0.1
  - Min Watt/Flop
  - Etc., etc.
A Paradox?

- The choices in cluster components allow us to tailor a cluster to very specific needs
  - Examples: Vis, Storage clusters
- COTS make HPC cluster affordable at the department or project level
- As a result, the cluster can be as specific as we wish..

Commodity components enable (drive?) custom solutions

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
Prediction

- Special Purpose Clusters will be common
  - More correctly, “tailored” or “tunable” clusters may become the norm
  - Is “SPC” or “TC” in our future?
Not all is done..

- A reminder that existing challenges only get worse as we march on to the implications of Moore’s Law
- The next frontier(s)
  - I/O and interconnect keeping pace with processor speed and memory size
  - Managing clusters - quality and performance of monitoring, maintenance, effective use

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
Finding Solutions to the Challenges requires an Ecosystem

HPC Program Office
More than Silicon

Intel Building Blocks
- Microprocessors
- Chipsets
- Interconnects
- Platform Enabling
- Software Suite

Intel Programs
- 13 major hardware vendors
- Regional Leaders
- SI Programs
- iSV enabling

Intel WW Services
- Dedicated HPC Personnel
- End User Developer Enabling
- Infrastructure (WA, NM, VG)
- Premier Technical Support
- HW Design Support

Joint Solutions Centers
- Channel
- Oil
- Finance
- Life Sciences

Intel Portfolio Investments
- SCALI*, United Devices*
- Portfolio Matchmaking

Intel Training
- Worldwide Software College
- Worldwide Training
- Cluster Recipes
- Case Studies/Blue Prints

Strategic Collaborations
- Cornell Theory Center* (CTC)
- CERN
- NCSA
- Western GECO*
- European Virtual Grid Center*
- Daresbury Benchmark Center*
- Singapore BioIT Grid*

* Other names and brands may be claimed as the property of others.
Beyond Clusters - Grids

The Changing Models of High Performance Computing

Traditional HPC Architecture

Proprietary
RISC
Vector
Custom

Current HPC Architecture

Applications
OS
Middleware
Hardware

Standard Based Clusters SMP

Future HPC Architecture

Grid

Cluster, SMP, Blades

Rich Client

Shared Resources

Distributed Applications

From clusters to Grids

*All names and brands are property of their respective owners.

© Copyright 2002, Intel Corporation. All rights reserved.
Solutions Adoption Curve

From clusters to Grids

**Grid**
- GOAL = Select the Right Opportunities
- GOAL = Select the Right Partners
- GOAL = Extend Ecosystem Participation
- GOAL = Solid EU Proof Points

**Clusters**
- GOAL = Extend Ecosystem Participation
- GOAL = Solid EU Proof Points

**SMP**
- GOAL = Ecosystem scale

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
The duality of Grid Computing

- **Grid Computing as an networked ensemble of data centers**
  - As originally understood by the people who coined the term (and formed the Grid Forum)

- **Grid computing as harnessing resources on desktops**
  - Activity that was originally known as “distributed computing” (.. On the Internet)
  - Proprietary software is being replaced by Globus
Intersection of Technologies

- Is desktop grid computing (DGC) HPC? Is it cluster computing?
  - DGC can, and is, employed to solve problems of the size we call “HPC problems”
  - DGC employs, more and more, the same resource sharing tools used for cluster-based grids

- DGC is interesting - it builds on both P2P and clusters technologies
  - P2P: identification/authentication, intermittent connectivity
  - Cluster: scheduling, monitoring, ..
Fodder for Thought

- DGC, with P2P foundation and cluster-based traditional grid tools, can define a special class of clusters
- P2P technologies enable types of apps other than computations and data handling
  - Content distribution, team collaboration
- Should the concept of cluster be extended to include these other P2P app types?
  - After all, they involved coordinated use of multiple and networked collection of systems
- Should “HPC” be extended to these app types
  - And then really be a superset of HPTC
Not mere Semantics

- Defining the scope of “clusters” and “HPC” matters to their development
  - If technologies are shared it makes sense for all relevant practitioners to collaborate in their development
  - For that to happen we’d need some sort of a conceptual umbrella ..

- Clusters and HPC can be thought as occupying a spectrum - from dedicated high-end clustered SMPs to loosely coupled Internet-connected desktops
Concluding Remarks

- The economics of HPC have changed, and it drives exciting opportunities
- In an interesting twist commoditization opens the door to specialization
- The next big challenges are Interconnect (hardware) and Cluster Mgmt (mostly software)
- Intel works with the Industry and with the user community to enable HPC for discovery and science
- The “beyond clusters” examination can get the computer industry to pay closer attention to HPC

*Other brands and names are the property of their respective owners
© Copyright 2002 Intel Corporation. All Rights Reserved.
Thank you

Questions?