HPC2N @ Umeå University

Introduction to HPC2N and Kebnekaise

Jerry Eriksson, Pedro Ojeda-May, and Birgitte Brydsö
Outline

• Short presentation of HPC2N
• HPC at a glance.
• HPC2N – Abisko, **Kebnekaise**
• HPC Programming models – how to develop your own code
  (separate slides packages)
  • Nvidia, GPU: OpenAcc, Cuda
  • Intel OpenMP
HPC2N - “HPC to North"

• A national center for Parallel and Scientific Computing

• Five partners:
  • Luleå University of Technology
  • Mid Sweden University
  • Swedish Institute of Space Physics
  • Swedish University of Agricultural Sciences - SLU
  • Umeå University

• Funded by the Swedish Research Council (VR) and its Meta-Center SNIC togheter with the partners.
From macro scale to micro scale

- Provides state-of-the-art resources and expertise for Swedish eScience
  - Scalable and parallel HPC
  - Large-scale storage facilities
  - Grid and cloud computing
  - Software and advanced support for eScience applications
  - International network for research and development

DFT computation, semi-stable, binding energy 15eV; Sven Öberg, LTU
Main areas of HPC2N users

- Biosciences and medicine
- Chemistry
- Computing science
- Engineering
- Materials science
- Mathematics and statistics
- Physics including space physics
Storage Levels @ HPC2N

Basically three types of storage are available at HPC2N:

- **Center Storage** - Parallel file system (fast discs)
  - Closely connected to our computing resources; Abisko and Kebnekaise

- **SweStore** - disk based (dCache)
  - part of SNIC Storage, responsible for national accessible storage

- **Tape Storage**
  - Backup
  - Long term storage
HPC2N Think Tank!

- User support (primary, advanced, tailored)
  - Research group meetings @ UmU
- User training and education program
- Workshops & Colloquia
- Research & Development - Technology transfer
- Provide various state-of-the-art HPC resources
HPC – Towards Exascale Computing

• **Moore's law**: the number of transistors in a chip doubles every second years.
• Parallel Computing:
  • Increase number of cores.
• Heterogenous clusters
  • Different processors and memories.
• Power efficiency!
HPC EcoSystems

- Application Level
  - Applications and Community Codes
  - FORTRAN, C, C++, and IDEs
  - Domain-specific Libraries
  - MPI/OpenMP + Accelerator Tools
  - Numerical Libraries
  - Performance and Debugging (such as PAPI)
  - Lustre (Parallel File System)
  - Batch Scheduler (such as SLURM)
  - System Monitoring Tools
- Middleware and Management
- System Software
- Cluster Hardware

Computational Science Ecosystem
- Linux OS variant
- Infiniband + Ethernet Switches
- SAN + Local Node Storage
- X86 Racks + GPUs or Accelerators
PRACE - Partnership for Advanced Computing in Europe

Tracing tools (GROMACS, 16 Cores)
Now to the clusters and programming models

A large amount of numbers and technical information will follow!!

Relax, you do not need to know everything in detail, and we offer training for those things you should know.
Abisko

• 332 nodes with a total of 15936 CPU cores.
• AMD Opteron 6238 (Interlagos)
• The 10 'fat' nodes have 512 GB RAM each, and the 322 'thin' nodes have 128 GB RAM each.
• (More details can be found on our web-pages)
Kebnekaise
Intels processors
Compute nodes

● 432 nodes
● Intel Broadwell (E5-2690v4)
● 2x14 cores/node
● 128GB memory
● Infiniband FDR
Large memory nodes

- 20 nodes
- Intel Broadwell (E7-8860v4)
- 4x18 cores/node
- 3TB memory
- Infiniband EDR
KNL - Intel Knights Landing

- 36 nodes
  - 68 cores
  - 1.4GHz (1.2GHz AVX)
- 192 GB memory - 16 GB MCDRAM
- Infiniband FDR
- *Installation in February*
Knights Landing Overview

Chip: 36 Tiles interconnected by 2D Mesh
Tile: 2 Cores + 2 VPU/core + 1 MB L2
Memory: MCDRAM: 16 GB on-package; High BW
         DDR4: 6 channels @ 2400 up to 384 GB
IO: 36 lanes PCIe Gen3, 4 lanes of DMI for chipset
Node: 1-Socket only
Fabric: Omni-Path on-package (not shown)

Vector Peak Perf: 3+TF DP and 6+TF SP Flops
Scalar Perf: ~3x over Knights Corner
Streams Triad (GB/s): MCDRAM: 400+; DDR: 90+

Source: Intel. All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice. All data are preliminary based on current expectations and are subject to change without notice. Binary compatible with Intel Xeon processors using Haswell, Intel processors, Intel Xeon processors, and Intel Xeon Phi. Performance numbers are based on STREAM4 key memory access patterns. Additional information and performance results have been obtained from Intel’s internal testing. Results may vary due to differences in systems configuration, software, operating system, and other factors.
## Intel Xeon Phi

<table>
<thead>
<tr>
<th>Xeon Phi</th>
<th>Clock Speed</th>
<th>Cores/Threads</th>
<th>Peak DP TFLOPS</th>
<th>DDR4 Memory</th>
<th>MCDRAM Memory</th>
<th>TDP (Watts)</th>
<th>1K Tray Unit Price</th>
<th>$ / TFLOPS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Knights Landing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7290</td>
<td>1.5 GHz</td>
<td>72 / 288</td>
<td>3.46</td>
<td>384 GB</td>
<td>16 GB</td>
<td>7.2 GT/sec</td>
<td>245</td>
<td>$6,254</td>
</tr>
<tr>
<td>7250</td>
<td>1.4 GHz</td>
<td>68 / 272</td>
<td>3.05</td>
<td>384 GB</td>
<td>16 GB</td>
<td>7.2 GT/sec</td>
<td>215</td>
<td>$4,876</td>
</tr>
<tr>
<td>7230</td>
<td>1.3 GHz</td>
<td>64 / 256</td>
<td>2.86</td>
<td>384 GB</td>
<td>16 GB</td>
<td>7.2 GT/sec</td>
<td>215</td>
<td>$3,710</td>
</tr>
<tr>
<td>7210</td>
<td>1.3 GHz</td>
<td>64 / 256</td>
<td>2.66</td>
<td>384 GB</td>
<td>16 GB</td>
<td>6.4 GT/sec</td>
<td>215</td>
<td>$2,438</td>
</tr>
<tr>
<td><strong>Knights Corner</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7120P</td>
<td>1.24 GHz</td>
<td>61 / 61</td>
<td>1.21</td>
<td>30.5 MB</td>
<td>16 GB</td>
<td>5.5 GT/sec</td>
<td>300</td>
<td>$4,129</td>
</tr>
<tr>
<td>7120X</td>
<td>1.24 GHz</td>
<td>61 / 61</td>
<td>1.21</td>
<td>30.5 MB</td>
<td>16 GB</td>
<td>5.5 GT/sec</td>
<td>300</td>
<td>$4,129</td>
</tr>
<tr>
<td>5110P</td>
<td>1.05 GHz</td>
<td>60 / 60</td>
<td>1.01</td>
<td>30 MB</td>
<td>8 GB</td>
<td>5.0 GT/sec</td>
<td>225</td>
<td>$2,649</td>
</tr>
<tr>
<td>5120D</td>
<td>1.05 GHz</td>
<td>60 / 60</td>
<td>1.01</td>
<td>30 MB</td>
<td>8 GB</td>
<td>5.5 GT/sec</td>
<td>245</td>
<td>$2,759</td>
</tr>
<tr>
<td>3120A</td>
<td>1.10 GHz</td>
<td>57 / 57</td>
<td>1.0</td>
<td>28.5 MB</td>
<td>6 GB</td>
<td>5.0 GT/sec</td>
<td>300</td>
<td>$1,695</td>
</tr>
<tr>
<td>3120P</td>
<td>1.10 GHz</td>
<td>57 / 57</td>
<td>1.0</td>
<td>28.5 MB</td>
<td>6 GB</td>
<td>5.0 GT/sec</td>
<td>300</td>
<td>$1,695</td>
</tr>
</tbody>
</table>

General or special-purpose processor?
GPU nodes

- 32 nodes with 2x NVidia K80
- 4 nodes with 4x NVidia K80
- Intel Broadwell 2x14 cores (E5-2690v4)
- 128 GB memory
- Infiniband FDR
High Speed Interconnect

- Infiniband
- Three level fat tree structure
- FDR cards in nodes (leafs)
- EDR cards in large memory nodes
- EDR in switches
Kebnekaise in numbers

- 13 racks
- 544 nodes
- 17552 cores (of which 2448 cores are KNL-cores)
- 399360 CUDA cores (80 * 4992 cores/K80)
- More than 125TB memory (20*3TB + (432 + 36) * 128GB + 36*192GB)
- 66 switches (Infiniband, Access network, Management network)
Kebnekaise in numbers

- 83% of the system are standard and Large Memory nodes
- 7% GPU-nodes
- 7% KNL-nodes
- 4% Other nodes (login and management nodes, LNET-routers etc)
- 728 TFlops/s Peak performance
- 629 TFlops/s HPL (all parts)
- HPL: 86% of Peak performance

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard Nodes</strong></td>
<td></td>
<td>374 TFlops/s</td>
</tr>
<tr>
<td><strong>Large Memory Nodes</strong></td>
<td></td>
<td>34 TFlops/s</td>
</tr>
<tr>
<td><strong>2xGPU Nodes</strong></td>
<td></td>
<td>129 TFlops/s</td>
</tr>
<tr>
<td><strong>4xGPU Nodes</strong></td>
<td></td>
<td>30 TFlops/s</td>
</tr>
<tr>
<td><strong>KNL Nodes</strong></td>
<td></td>
<td>62 TFlops/s</td>
</tr>
<tr>
<td><strong>Total (All parts)</strong></td>
<td></td>
<td>629 Flops/s</td>
</tr>
</tbody>
</table>