HPC on the Cloud – A Norwegian Perspective

Nico Reissmann: nico.reismann@ntnu.no
Jan Christian Meyer: jan.christian.meyer@ntnu.no
* Einar Næss Jensen: einar.nass.jensen@ntnu.no

http://www.hpc.ntnu.no/
Who we are

- NTNU IT Department
- Providing HPC services for
  - NTNU researchers and students on local clusters
  - Sintef
  - National infrastructure and users as part of Sigma2
What is HPC? (High Performance Computing)

- It depends……
  - Lots of Memory
  - Blazing fast CPU
  - Millions of iops
  - Extreme Network
- All of the above x1000
- Basically:
  - When your laptop is not enough. (my laptop have been running Fluent for 4 days and I cannot get any work done)
  - Or maybe you don’t even have laptop.
So you need a new Computer

- Blazing fast monster machine!!!
  - 2 x 28 cores CPU
  - 1.5TB memory
  - 2x nvme drives
  - 25Gbit ethernet
- .... basically a new laptop
- .... with same problems
We need HPC
Let's pool our resources together

- and build a really, really large computer (Abel, Vilje, Stallo, Hexagon on top100 2012)

<table>
<thead>
<tr>
<th></th>
<th>Year</th>
<th>Cores</th>
<th>GPU</th>
<th>Interconnect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idun (NTNU only)</td>
<td>2016</td>
<td>1700</td>
<td>62</td>
<td>FDR/EDR</td>
</tr>
<tr>
<td>Abel</td>
<td>2012</td>
<td>10000</td>
<td>X</td>
<td>FDR</td>
</tr>
<tr>
<td>Stallo</td>
<td>2012</td>
<td>10000</td>
<td>X</td>
<td>FDR</td>
</tr>
<tr>
<td>Hexagon</td>
<td>2014</td>
<td>22000</td>
<td>X</td>
<td>Gray-Gemini</td>
</tr>
<tr>
<td>Vilje</td>
<td>2012</td>
<td>22000</td>
<td>X</td>
<td>FDR</td>
</tr>
<tr>
<td>Fram</td>
<td>2017</td>
<td>30000</td>
<td>X</td>
<td>EDR</td>
</tr>
<tr>
<td>Saga</td>
<td>2019</td>
<td>9000</td>
<td>32</td>
<td>FDR</td>
</tr>
<tr>
<td>B1</td>
<td>2020</td>
<td>???</td>
<td>???</td>
<td>???</td>
</tr>
</tbody>
</table>
B1 (the next big thing)

- Procured by Sigma2
- Process started early 2018
- Set to replace Vilje and Hexagon
- Goal was to get at least 50k - 60k cores with high speed interconnect and at least 2.5 PB storage.
- Budget: ca 90-100million NKR
- Public announcement soon(ish)
Cloud evaluation

• All purchases related to IT infrastructure must do evaluation on whether a cloud-based solution is a viable alternative.
• Use Fram supercomputer as a reference and get public pricing from cloud vendors.
FRAM

- 1 Gbit ethernet for admin
- EDR Infiniband 100Gbps interconnect, RDMA capable
- 2.4 PB parallell filesystem storage
- 150TB transfer of data within cluster each month
- at least 100TB outbound each month

<table>
<thead>
<tr>
<th>Class</th>
<th>#nodes</th>
<th>Cores/nodes</th>
<th>Memory (GB)</th>
<th>Disk (GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute A</td>
<td>940</td>
<td>32</td>
<td>64</td>
<td>120</td>
</tr>
<tr>
<td>Compute B</td>
<td>8</td>
<td>32</td>
<td>512</td>
<td>2x960</td>
</tr>
<tr>
<td>Compute C</td>
<td>2</td>
<td>28</td>
<td>6000</td>
<td>14400</td>
</tr>
<tr>
<td>Frontend</td>
<td>10</td>
<td>32</td>
<td>128</td>
<td>2x800</td>
</tr>
</tbody>
</table>
Feature comparison

- Many cloud vendors. Investigated “the usual suspects”, Google (Cloud), Azure, Amazon, IBM (Softlayer). Only public available pricing considered.

<table>
<thead>
<tr>
<th></th>
<th>Compute A</th>
<th>Compute B</th>
<th>Compute C</th>
<th>Frontend</th>
<th>Storage</th>
<th>Network</th>
<th>RDMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fram</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Cloud</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>AWZ</td>
<td>yes</td>
<td>unclear</td>
<td>no</td>
<td>unclear</td>
<td>unclear</td>
<td>unclear</td>
<td>no</td>
</tr>
<tr>
<td>Azure</td>
<td>unclear</td>
<td>unclear</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>unclear</td>
<td>yes</td>
</tr>
<tr>
<td>Softlayer</td>
<td>yes</td>
<td>unclear</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>
Monthly cost comparison in Mnok

- The cost of Fram includes total HW cost, electricity, cooling, housing
- assuming four years life span
- When cloud alternative missing, assuming closes match
- The only alternative with RDMA alternative is 7 times more expensive

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Fram</th>
<th>Cloud</th>
<th>Aws</th>
<th>Azure</th>
<th>Softlayer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>1,687</td>
<td>6,8311</td>
<td>7,6003</td>
<td>12,3383</td>
<td>11,3959</td>
</tr>
</tbody>
</table>
Should we do HPC in the cloud?

- Probably. But only for specific workload
  - “HPC”
  - long running, single core (or single node)
  - small input/output datasets
  - Not requiring expensive hardware
  - where cloud competition is good.
  - remember that HPC is moving target.
Hidden costs

- Difficult to get exact pricing
- Asking for a deal will give huge discounts.
- Risk of “enslavement” to specific vendor
- GDPR and sensitive data?
and our hidden costs

- Takes more than two years for full process of procurement, installation and production
- Lots of resources required
- In April-May 2019, on average -250K cores
- Scientists have to wait ……..
HPC bonus (NTNU) Heat Capture
Campus for the future

- 5.0-5.2Gwh each year (2014)
- Estimated price per kWh: 0.68nkr/0.1$ (2014)
- We can do the same with Electricity/Power
- Heat is environmental waste, and needs to be recycled
- Heat, cooling, power, HPC consolidated in one building
More information

• Sigma2: http://www.sigma2.no/
• NTNU HPC: http://www.hpc.ntnu.no/

• einar.nass.jensen@ntnu.no