Optimizing HPC resource allocation through monitoring

Alexandre Beche <alexandre.beche@epfl.ch>
1. Context
2. Job monitoring
3. Beyond Slurm monitoring
Context
Blue Brain Project

Host Institution
Ecole Polytechnique Fédérale de Lausanne (EPFL)

Director
Henry Markram

Co-Directors
Sean Hill, Felix Schürmann

Team today
~85 scientists, engineers & staff

Timeline
2005 founded at EPFL
2011/2012 ETH Board funding
2013-2016 Swiss National Research Infrastructure

Main International Collaborations
Switzerland (CSCS, CERN)
Israel (HUJI)
USA (Yale, ANL, Allen Brain)
Spain (UPM)
Saudi Arabia (KAUST)
Europe (HBP)
**Neuronal anatomy**

- $\sim$ 2 mm thick
- 55 morphological types
  - 13 excitatory & 42 inhibitory m-types
  - 31,000 neurons
  - 111,700 neurons/mm$^3$
  - Excitatory to inhibitory neuron ratio of 86:14%
- 346 m of axon
- 211 m of dendrites
- Maximum branch order of m-types:

<table>
<thead>
<tr>
<th>Type</th>
<th>Axon</th>
<th>Dendrites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excitatory</td>
<td>24</td>
<td>35</td>
</tr>
<tr>
<td>Inhibitory</td>
<td>50</td>
<td>17</td>
</tr>
</tbody>
</table>

**Synaptic anatomy**

- 0.63 synapses/mm$^3$
- Extrinsic to intrinsic synapse ratio of 75:25%
- 3025 possible synaptic pathways
- 2258 viable synaptic pathways
- 664 excitatory pathways
- 1594 inhibitory pathways
- 600 intra-laminar pathways
- 1658 inter-laminar pathways
- Mean synapses/connection: 4.3 Exc, 8.5 Inh

**Neuronal physiology**

- 11 electrical types
- 207 morpho-electrical types
- 13 HH type ion channel models
- bAP & EPSP attenuation for 207 morpho-electrical types
- Ion channel density distribution profiles:

<table>
<thead>
<tr>
<th>Type</th>
<th>Axon</th>
<th>Soma</th>
<th>Dendrites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>8</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

**Synaptic physiology**

- 6 synapse types
- 207 synaptomes
- Space clamp corrected synaptic conductances for 607 pathways
  - The per synapse conductance of 1.5 nS for connections between L5TTPCs is the highest in the microcircuit
  - Mean conductance per synapse: 0.85 nS for excitatory & 0.66 nS for inhibitory synapses
  - Total conductance in a single neuron is 971 nS

- 80 authors
- Joint effort between computer and neuro-scientists
- Reproducible work
- Extensible

Markram et al, Cell 2015

https://bbp.epfl.ch/nmc-portal
HPC Today’s Infrastructure

Elastic Compute
- Visualization and analysis
- Web services
- SW development
- Continuous Integration
- Continuous Deployment

Production HPC
- Model development
- Reconstruction
- Simulation
- SW development

Systems Research
- Memory extension
- Application coupling
- Interactive supercomputing
- Reproducibility

IBM BlueGene/Q
- 4 Compute Racks
- 4096 Compute Nodes
- 5D CN torus
- 0.8 PF/s peak
- 64TB Flash
- 4.2 PB GPFS storage

IBM BlueGene Active Storage
- 64 IONodes
- 3D torus
- 128TB Flash
- Linux

GVA

IaaS
- 4 nodes
- Viz & Analytics
- 36 nodes

CEPH

NetApp

CEPH

GPFS

Viz & Analytics
- 12 nodes
- 14 nodes

Viz & Analytics

IBM BlueGene Active Storage
HPC Resources usage

**Facts:**

- 70 Users
- 35 Projects
- 3 Clusters

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Daily core hours available</th>
<th>Daily Job submitted</th>
<th>Active user*</th>
</tr>
</thead>
<tbody>
<tr>
<td>BlueGene</td>
<td>1572864</td>
<td>103</td>
<td>20</td>
</tr>
<tr>
<td>Lugano cluster</td>
<td>13824</td>
<td>632</td>
<td>53</td>
</tr>
<tr>
<td>Geneva cluster</td>
<td>4512</td>
<td>259</td>
<td>20</td>
</tr>
</tbody>
</table>

* User who submitted at least one job over the last month
Facts:

<table>
<thead>
<tr>
<th></th>
<th>Daily core hours available</th>
<th>Daily Job submitted</th>
<th>Active user*</th>
</tr>
</thead>
<tbody>
<tr>
<td>BlueGene</td>
<td>1572864</td>
<td>103</td>
<td>20</td>
</tr>
<tr>
<td>Lugano cluster</td>
<td>13824</td>
<td>632</td>
<td>53</td>
</tr>
<tr>
<td>Geneva cluster</td>
<td>4512</td>
<td>259</td>
<td>20</td>
</tr>
</tbody>
</table>

* User who submitted at least one job over the last month

Observations:

- Clusters are scheduling 10x more jobs than bluegene
- Cluster nodes are shared (--exclusive is limited), not bluegene

Emphasis of the presentation will be put on HPC clusters
Job monitoring
Problem description

**Interactive HPC Clusters usage over time**

*Extracted from sreport*
Problem description

Interactive HPC Clusters usage over time

Symptoms:
User can’t get an allocation

Cause:
Cluster is fully allocated

Solution:
1) **Buy a bigger one 😊**
2) See if resources are optimally used
Slurm accounting DB

Knowledge of all jobs / step executed
• Average waiting time in the queue
• Submission rate by user / project

“sacct” data are indexed into ElasticSearch
• Near real-time (every 10 minutes)
• Analytics, web report generation

Limitations:
• Not natively aware of resources used by the job
BBP Monitoring infrastructure

Scalable and extensible framework
- Based on open source technologies
- Enable data collection and analysis
BBP Monitoring infrastructure

Scalable and extensible framework
- Based on open source technologies
- Enable data collection and analysis

Native host instrumentation
- 10 seconds resolution
- 250 metrics per node

Limitation: System metrics does not have any knowledge about workload
Non-optimal use case

Allocation details:
- Single nodes, all cores, batch partition

Job details:
- CPU bound

Job started
Job finished

Single-thread used
Un-used allocation
Benefits

Developers

• Analyze code from a system perspective
  – Non-intrusive monitoring / negligible (perf) overhead
• Detect code inefficiency / limiting resource
  – Non optimal parallelization
Benefits

Developers

• Analyze code from a system perspective
  – Non-intrusive monitoring / negligible (perf) overhead
• Detect code inefficiency / limiting resource
  – Non optimal parallelization

System admins

• Analyze system metrics with workload context
• Detect non-optimal allocation
  – Allocation bigger than execution time

Holistic view enabling cross team (competencies) debugging
Ongoing work

• Creating KPI out of the available metrics
  – Efficiency of a job (cpu seconds used /reserved)
Ongoing work / limitation

Ongoing work

• Creating KPI out of the available metrics
  – Efficiency of a job (cpu seconds used /reserved)

Limitations

• Mainly system metrics so far
  – Only memory are collected at cgroup level
• Missing infiniband metrics
• Job internals are hidden
Beyond Slurm monitoring
Focus has been put on monitoring the job in the infrastructure

No hint is given to the job internals

• Job entering in a given phase
Job workload context

Focus has been put on monitoring the job in the infrastructure

No hint is given to the job internals

- Job entering in a given phase

Providing a library to enhance job context

- Workload manager agnostic
- Not a profiling tool
- Lightweight way of sharing information from the job
- Allows to ship user-defined metadata
def count_element(self):
    lines = self._data.count()
@log_function_events
def count_element(self):
    lines = self.data.count()
Summary

Detection of non-optimal usage
• Un-used allocation
• Developers now have tools to understand job behaviors

Internal job monitoring
• Allows understanding which resources are consumed by section of job through user-defined metadata

Correlation of scattered information enable powerful analysis
BBP core services & HPC teams