Leveraging heterogeneous systems and deep memory hierarchies for brain tissue modeling

Tier 1 ALCF Theta Early Science Program

Fabien Delalandre
Blue Brain Project, HPC Team Manager
• Blue Brain Project (BBP) & Human Brain Project (BBP) introduction

• On node portable performance optimization

• Future R&D Directions
• ~90 people & 25 nationalities, expecting 110-120 people in next 2-4 years (~20M budget/year)

• Operates heterogeneous infrastructure (4 rack Blue Gene/Q, clusters, storage, private cloud, volunteer computing, desktops, ...)

• More than 50 applications written in python, java, C/C++, ... organized in complex workflows

• Extensive software engineering practices & infrastructures (more than just continuous integration ...
Matterhorn

Help Yourself!
Blue Brain Project Development Strategy

• Pioneers strategy to integrate fragmented biological knowledge into **unifying models of brain tissue**

• Develops **unique (super)computer-based platform** for building, simulating & evaluating unifying brain models

• Use platform for simulation-based research: Tested on a *rat’s neocortical microcircuit* → **see e.g. Markram et al. 2015 (Cell)**

• Simulation core of the European Human Brain Project
• 111,700 neurons/mm³
• 31,000 neurons
• 55 morphological types
• 13 excitatory & 42 inhibitory m-types
• 31,000 neurons/mm³
• Excitatory to inhibitory neuron ratio of 86:14%
• 346 m of axon
• 211 m of dendrites
• Maximum branch order of m-types:

| Excitatory | 24 | 35 |
| Inhibitory | 50 | 17 |

• 0.63 synapses/mm³
• 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

| Excitatory | 4.3 | 8.5 |
| Inhibitory | 6.0 | 11.0 |

• 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:

| Uniform | 8 | 6 | 6 |

• 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 micrcircuit
• Mean conductance per synapse: 0.85 nS for excitatory & 0.66 nS for inhibitory synapses
• Total conductance in a single neuron is 971 nS

| 697 nS | 274 nS |

• Extrinsic to intrinsic synapse ratio of 75:25%
• 2258 possible synaptic pathways
• 3025 viable synaptic pathways
• 600 intra-laminar pathways
• 1658 inter-laminar pathways
• Mean synapses/connection

100 µm

ø 0.28 mm²
• 31'000 neurons
• 207 morpho-electrical types
• 31'628 types of connections
• 40 million intrinsic synapses
• 141 million extrinsic synapses

Markram et al., Cell 2015
https://bbp.epfl.ch/nmc-portal
Co-designed Heterogeneous ICT Facility

Relevant publications:
- Migliore et al, 2006
- Druckmann et al, 2007
- Druckmann et al, 2008
- Hines et al, 2008 a & b
- Kozloski et al, 2008
- Hay et al, 2011
- Hines et al, 2011
- Ranjan et al, 2011
- Lasserre et al, 2012
- Druckmann et al, 2012
- Hill et al, 2012
- Tauheed et al, 2012
- Hernando et al, 2012
- Ramaswamy et al, 2012
- Reimann et al, 2013
- Schürmann et al, 2014
- Delalondre et al, 2014
- Devresse et al, 2015
- Kumbhar et al, 2016
- …
Co-designed Heterogeneous ICT Facility

Relevant publications:
- Migliore et al., 2006
- Druckmann et al., 2007
- Druckmann et al., 2008
- Hines et al., 2008 a & b
- Kozloski et al., 2008
- Hay et al., 2011
- Hines et al., 2011
- Ranjan et al., 2011
- Lasserre et al., 2012
- Druckmann et al., 2012
- Hill et al., 2012
- Tauheed et al., 2012
- Hernando et al., 2012
- Ramaswamy et al., 2012
- Reimann et al., 2013
- Schürmann et al., 2014
- Delalondre et al., 2014
- Devresse et al., 2015
- Kumbhar et al., 2016
- ...
(European) Human Brain Project

- Building infrastructure dedicated to Neuroscience

- EU Flagship project initiated by EPFL/BBP gathering more than 100 universities/labs

- Started in September 2013 with budget/proposal submission every 2 years (Total budget expected over 10 years ~ 1 billion euros)
Building Infrastructure Dedicated to Neuroscience

PaaS/IaaS
- Neuroinformatics
- Brain Simulation
- High Performance Analytics and Computing
- Medical informatics
- Neuromorphic Computing
- Neurorobotics
(European) Human Brain Project

Building Infrastructure Dedicated to Neuroscience

PaaS/IaaS

- Neuroinformatics
- Brain Simulation
- High Performance Analytics and Computing
- Medical informatics
- Neuromorphic Computing
- Neurorobotics

Accessible through HBP Web-based Collaboratory (SaaS)
Building Infrastructure Dedicated to Neuroscience

PaaS/IaaS

- Neuroinformatics
- Brain Simulation
- High Performance Analytics and Computing
- Medical informatics
- Neuromorphic Computing
- Neurorobotics

Accessible through HBP Web-based Collaboratory (SaaS)

PLATFORM RELEASE
30 March 2016

https://www.youtube.com/watch?v=2XXz2quUWFQ
• Blue Brain Project (BBP) & Human Brain Project (BBP) introduction

• On node portable performance optimization

• Future R&D Directions
Portable On Node Optimization

• What scientific problem we are trying to solve?

• What is our development workflow/Tools?
Biological Problem to Solve

Dendrites
Collect electrical signals

Cell body
Integrates incoming signals and generates outgoing signal to axon

Axon
Passes electrical signals to dendrites of another cell or to an effector cell

Figure 45-2b Biological Science, 2/e © 2001 Pearson Prentice Hall, Inc.
Biological Problem to Solve Today

Synapse
3.5k/Neuron

Ion Channel
3-5/Compartment

Action Potential
tiem.utk.edu

http://www.sailhome.org/Concerns/Excitotoxins.html

Illustration by J.P. Cartailler. Copyright 2007, Symmation LLC
Biological Problem to Solve Tomorrow

Synapse
10k/Neuron

Ion Channel
20/Compartment

http://www.sailhome.org/Concerns/Excitotoxins.html

Action Potential
tiem.utk.edu

Illustration by J.P. Cartailler.
Copyright 2007, Symmation LLC
Different Scales/Different Representations

Point Neuron (NEST)

Morphologically Detailed (NEURON/CoreNeuron)

3D Modeling (STEPS)
Different Scales/Different Representations

Point Neuron (NEST)

Morphologically Detailed (NEURON/CoreNeuron)

3D Modeling (STEPS)
Subset of Use Cases with Very Diverse Requirements

Hippocampus
Armando Romani [EPFL]

Structural Plasticity
Giuseppe Chindemi [EPFL]

Gap Junction
Oren Amsalem [HUJI]
Resolution Workflow

Ion Channel

Synapse

Mechanism Instance

Mechanism Type

Data Management Organization

EPFL/Blue Brain Project
Resolution Workflow

Ion Channel

Synapse

Per Neuron Algebraic System (O(350))
Kernels – 85% of Total Compute Time

Compute Bound – 20% of Kernels

```
/* loop2: this is second kernel with more compute due to exp and div */
#pragma for_vector_loop
for (i = 0; i < nodedcount; i++) {
    int idx = nt->node_index[i];
    double v = nt->vec_v[idx];
    p3[i] = data[ion1[i]];
    double qt = 2.9528826414121;
    double mAlpha = (0.002*(v+32.0)) / (1.0-{exp[-v-32.0]/6.0});
    double mBeta = (0.124*(v-32.0)) / (1.0-{exp[v+32.0]/6.0});
    double mInf = mAlpha/(mAlpha+mBeta);
    double mTau = (1.0/(mAlpha+mBeta))/qt;
    p1[i] = p1[i] + (1.0-exp(dt*(-1.0/mTau))) * (-mInf/mTau) / ((1.0/mTau)-p1[i]);
    double hAlpha = (-0.015*(v+50.0)) / (1.0-{exp[(v+50.0)/6.0]});
    double hBeta = (-0.015*(v-50.0)) / (1.0-{exp[-v-50.0]/6.0});
    double hInf = hAlpha / (hAlpha+hBeta);
    double hTau = (1.0/(hAlpha+hBeta)) / qt;
    p2[i] = p2[i] + (1.0-exp(dt*(-1.0/hTau))) * (-hInf/hTau) / ((1.0/hTau)-p2[i]);
    nt->vec_v[idx] = v;
    p5[i] += v;
}
```

Memory Bound – 80% of Kernels

```
/* loop1: this is one kernel with less compute / memory streaming */
#pragma for_vector_loop
for (i = 0; i < nodedcount; i++) {
    int idx = nt->node_index[i];
    double v = nt->vec_v[idx];
    p3[i] = data[ion1[i]];
    double gNaTs2 = p1[i] * p2[i] * p2[i] * p2[i] + p3[i];
    double ina = gNaTs2 * (v - p3[i]);
    /* no data hazard below */
    data[ion1[i]] += gNaTs2;
    data[ion2[i]] += ina;
    nt->vec_rhs[idx] -= ina;
    nt->vec_d[idx] += gNaTs2;
}
```
Compute/Bandwidth Analysis

Compute Bound
Compute/Bandwidth Analysis

Compute Bound

Memory Bound
Looking at ways to overlap kernels
Portable On Node Optimization

• What scientific problem we are trying to solve?

• What is our development workflow/Tools?
**Development Workflow & Tools**

**NeuroM(ini)app**
- 150-200 lines of code
- Used for prototyping

**CoreNeuron**
- 15-20k lines of code
- Optimized kernels

**NEURON**
- 200k lines of code
- All functionalities
Performance Portability Through DSL

Write Kernel using high level DSL
NMODL

Code generation

Modeling Kernels
(C/C++/...) + OpenMP/OpenACC

Low level Framework
(C/C++)

EPFL/Blue Brain Project
Performance Portability Through DSL

Write Kernel using high level DSL NMODL

Modeling Kernels (C/C++/…) + OpenMP/OpenACC

A = B x exp(C) + D
Cyme (intrinsics on CPU)

Low level Framework (C/C++)

Code generation
Software Portability & Reproducibility

• Scientific reproducibility **is a major issue**!

• Modules or equivalent too fragile/unpredictable

• RPM/Debian packages update break development environment

• Docker/Containers come with problems today...
Application Status

- Intel x86, KNC, Nvidia GPU, IBM Blue Gene/Q & ARM supported
- Weak scaling to full MIRA with less than 20% loss of efficiency
- All kernel performance profiled/classified
- Nix & Continuous integration/deployment moving to production (installation at BBP & JSC)
- New challenging use cases ... (Real time, ...)

EPFL/Blue Brain Project
Outline

• Blue Brain Project (BBP) & Human Brain Project (BBP) introduction

• On node portable performance optimization

• Future R&D Directions
R&D - Memory Intensive Computing

- Co-design Heterogeneous Infrastructure
- Development Workflow
- Leveraging Deep Memory Hierarchy
- Scientific Workflows
Open Questions

- Is this a system node or an eco-system?
- How can I best map kernels to hardware component?
- How can I explore workflow opportunities?
- How can I build a system which fits my requirements?
- How can I continuously collect application requirements?
NeuroM(ini)app
150-200 lines of code
Used for prototyping

Optimized App.
15-20k lines of code
Optimized kernels

200k lines of code
All functionalities
R&D - Co-Designed Heterogeneous Infrastructure

Dynamic Exascale Entry Platform (DEEP)

HBP Precommerical Procurement (PCP)
R&D - Co-Designed Heterogeneous Infrastructure

**NeuroM(ini)app**
- 150-200 lines of code
- Used for prototyping

**Optimized App.**
- 15-20k lines of code
- Optimized kernels

**Sc. App.**
- 200k lines of code
- All functionalities
NeuroM(init)App
150-200 lines of code
Used for prototyping

Optimized App.
15-20k lines of code
Optimized kernels

200k lines of code
All functionalities

1. Explore workflow overlapping opportunities
2. Explore new hardware (FPGA) for parts of workflow (decoupling)
3. Explore run times (HPX)
NeuroM(initial) app


150-200 lines of code
Used for prototyping

Optimized kernels
200k lines of code

All functionalities

Good ... but very time consuming/Error-prone !!
NeuroM(init)app

150-200 lines of code
Used for prototyping

Optimizations

200k lines of code
Optimized kernels


200k lines of code
All functionalities

SaaS for Automated Performance Analysis/Modeling of Annotated Kernels along with Performance History
Write Kernel using high level DSL NMODL

Low level Framework (C/C++)

Code generation

Modeling Kernels (C/C++/...) + OpenMP/OpenACC

R&D - Development Workflow & Tools

EPFL/Blue Brain Project
Write Kernel using high level DSL NMODL

Low level Framework (C/C++)

Code generation

Modeling Kernels (C/C++/...) + OpenMP/OpenACC/OpenCL

EPFL/Blue Brain Project
Write Kernel using high level DSL NMODL

Low level Framework (C/C++)

Modeling Kernels (C/C++/...) + OpenMP/OpenACC/OpenCL

R&D - Development Workflow & Tools

Code generation

LLVM
Deep Memory Hierarchy/Eco-System View

Blue Gene Active Storage (BGAS) – Active Result Buffer

EPFL/Blue Brain Project
Deep Memory Hierarchy/Eco-System View

Blue Gene Active Storage (BGAS) – Model Container

Simulator

Analysis & Viz

EPFL/Blue Brain Project
Deep Memory Hierarchy/System View

Support Complex Workflows Minimizing Data Movement

Static Workflow
MPI/OpenMP for Building/Simulation
Spark for Analysis
All data stays on node
Deep Memory Hierarchy/System View

Support Complex Workflows Minimizing Data Movement

Static Workflow
- MPI/OpenMP for Building/Simulation
- Spark for Analysis
- All data stays on node

Dynamic Workflow
- Simulation/Analysis/Visualization competing for resources
  - On node run time / ZeroMq
Distributed Key Value Store

• Extend application memory by offering STL Map-like interface to additional stores
• Implementation should hide data coalescing, movement & prefetching to application
• Possible HPX run time integration to move data between various stores
• Wish to provide hints to compilers/run time

Increase Usability of Systems

Single sign-on

Jupyter Notebook

Requires “Meta”-scheduler & Data Container+Nix

EPFL/Blue Brain Project