BioDynaMo
Biology Dynamics Modeller

The Anticipated Challenges of Running Biomedical Simulations in the Cloud

Early-Career Researchers in Medical Applications
Short Talks on Computing and Simulation

Lukas Breitwieser

CERN openlab

CERN

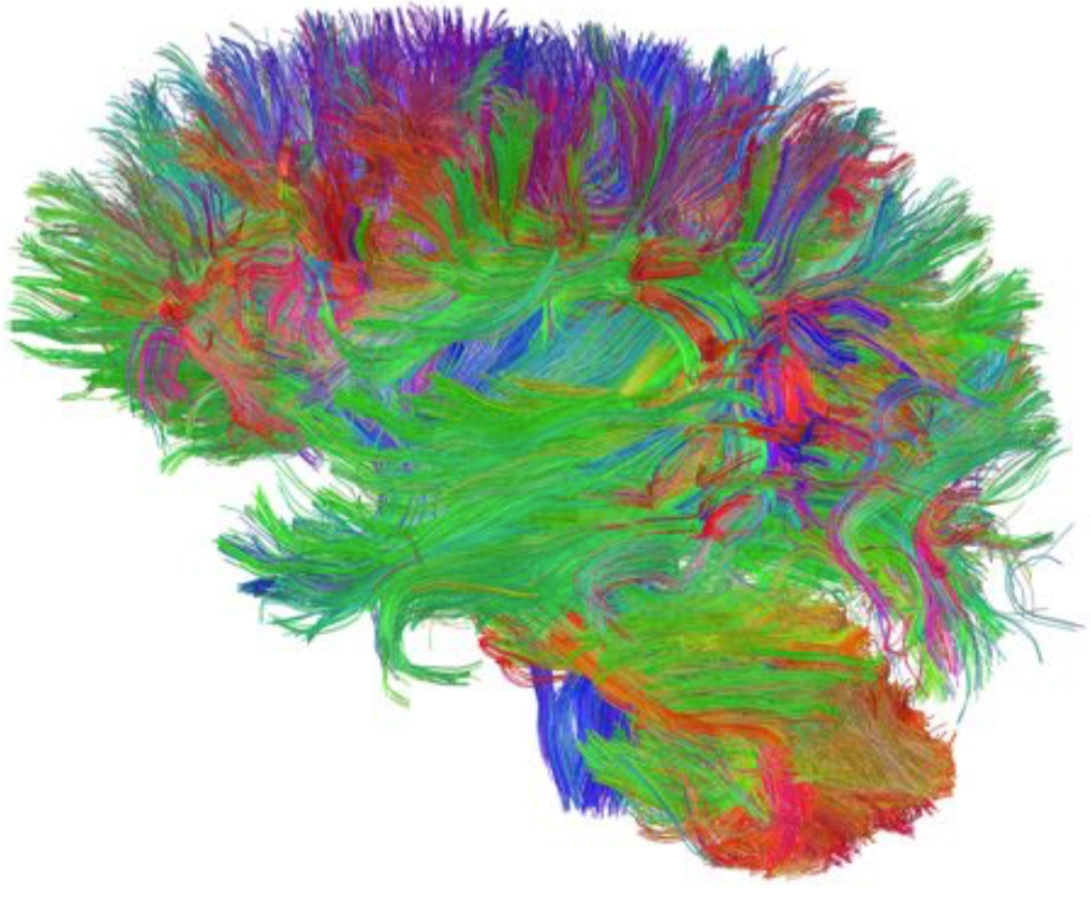
CERN openlab

Newcastle University

ETH zürich

intel

Understanding Developmental Diseases



Executive Summary

- **Motivation:**

  Run *tightly-coupled* HPC workloads in the cloud

  - Widely accessible, cost-effective
- **Problem:**

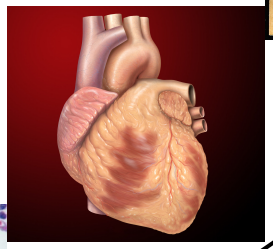
  Data exchange between servers will be a bottleneck
- **Key Idea:**

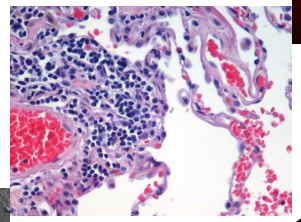
  Exploit inherent simulation characteristics to reduce data volume

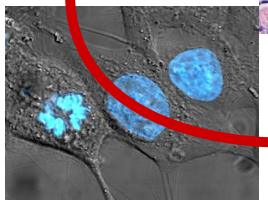
Outline

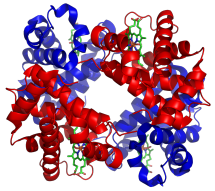
1. Biological Simulation Basics
2. Distributed Runtime
3. HPC on Cloud
4. Data Movement Minimization

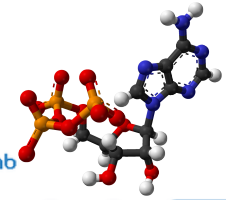
From Atoms to Organisms...

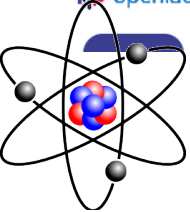












Organism

Organ

Tissue

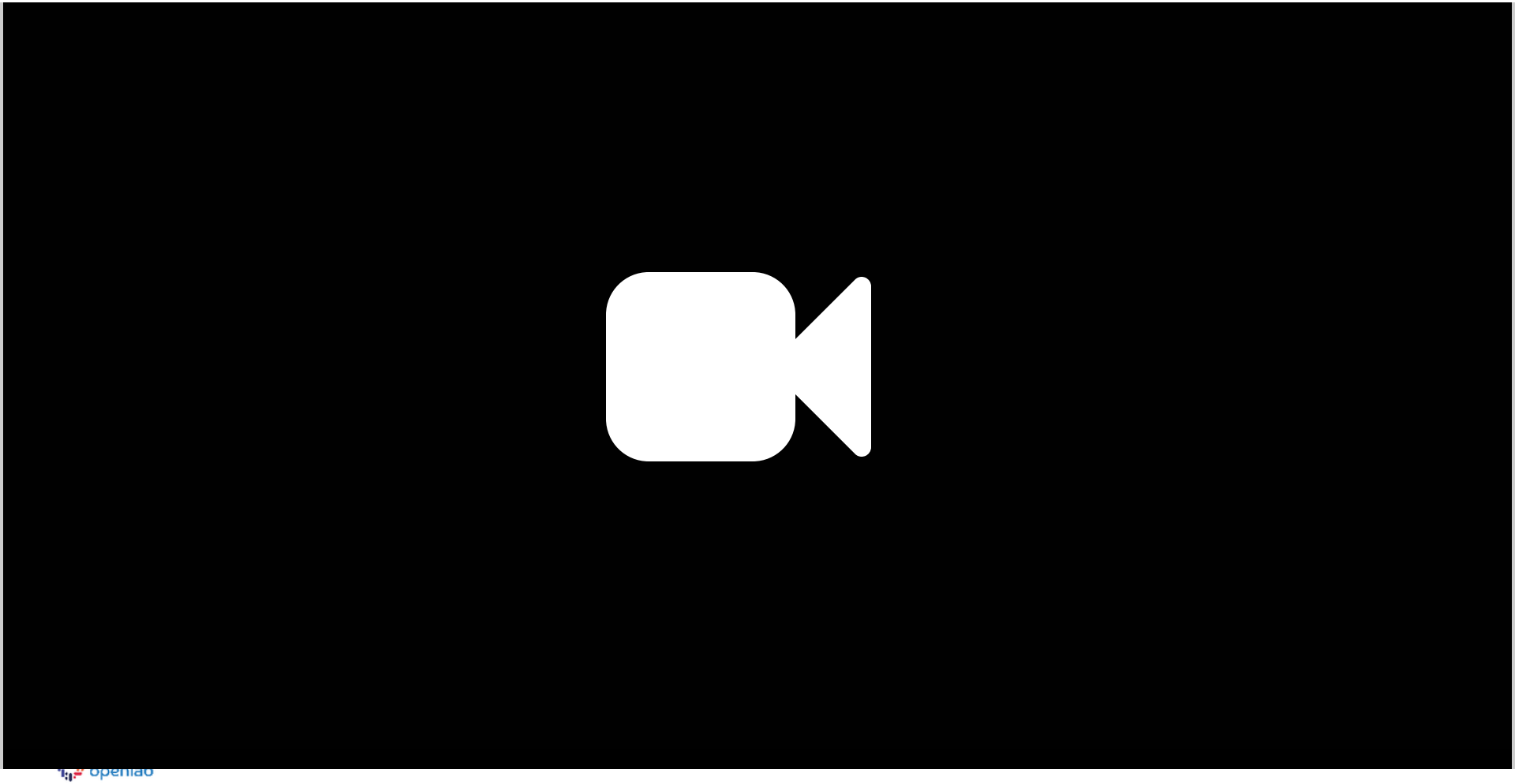
Cells

Macromolecule

Molecule

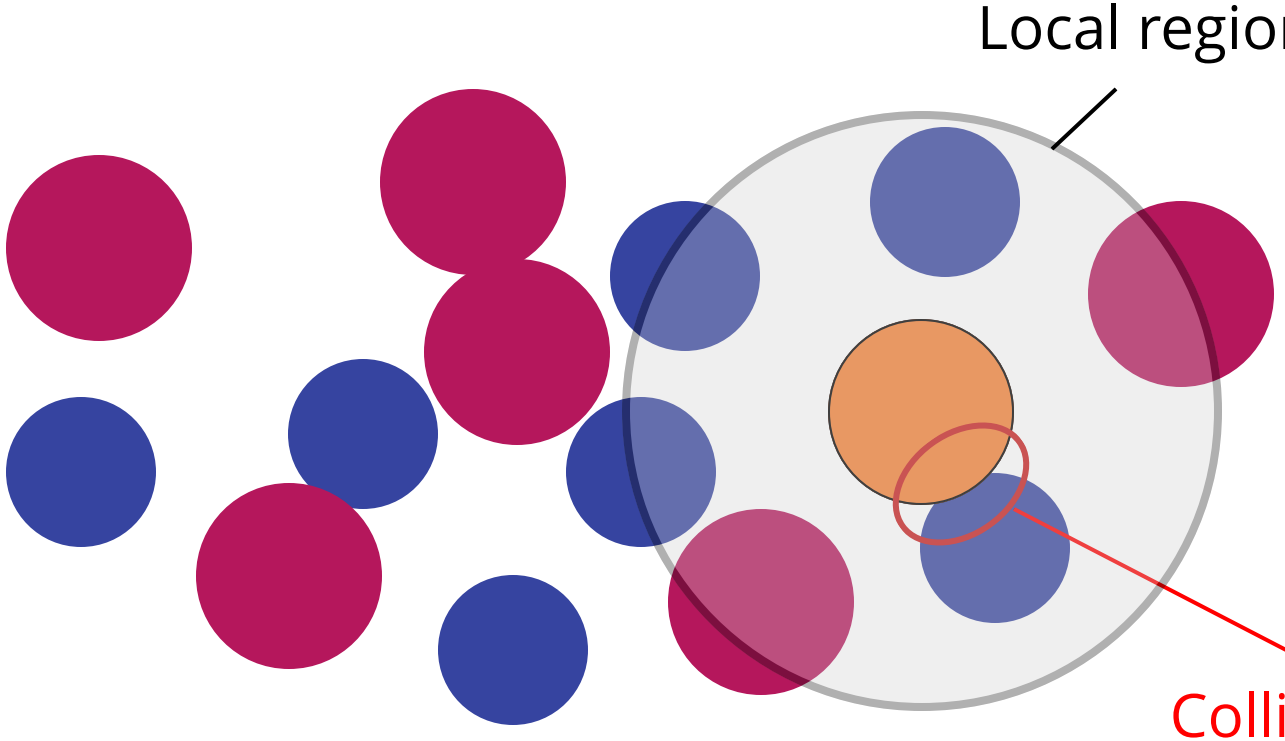
Atom

Demo: Tumor Growth



Agent-based simulations

Simulation object = *Agent*



Outline

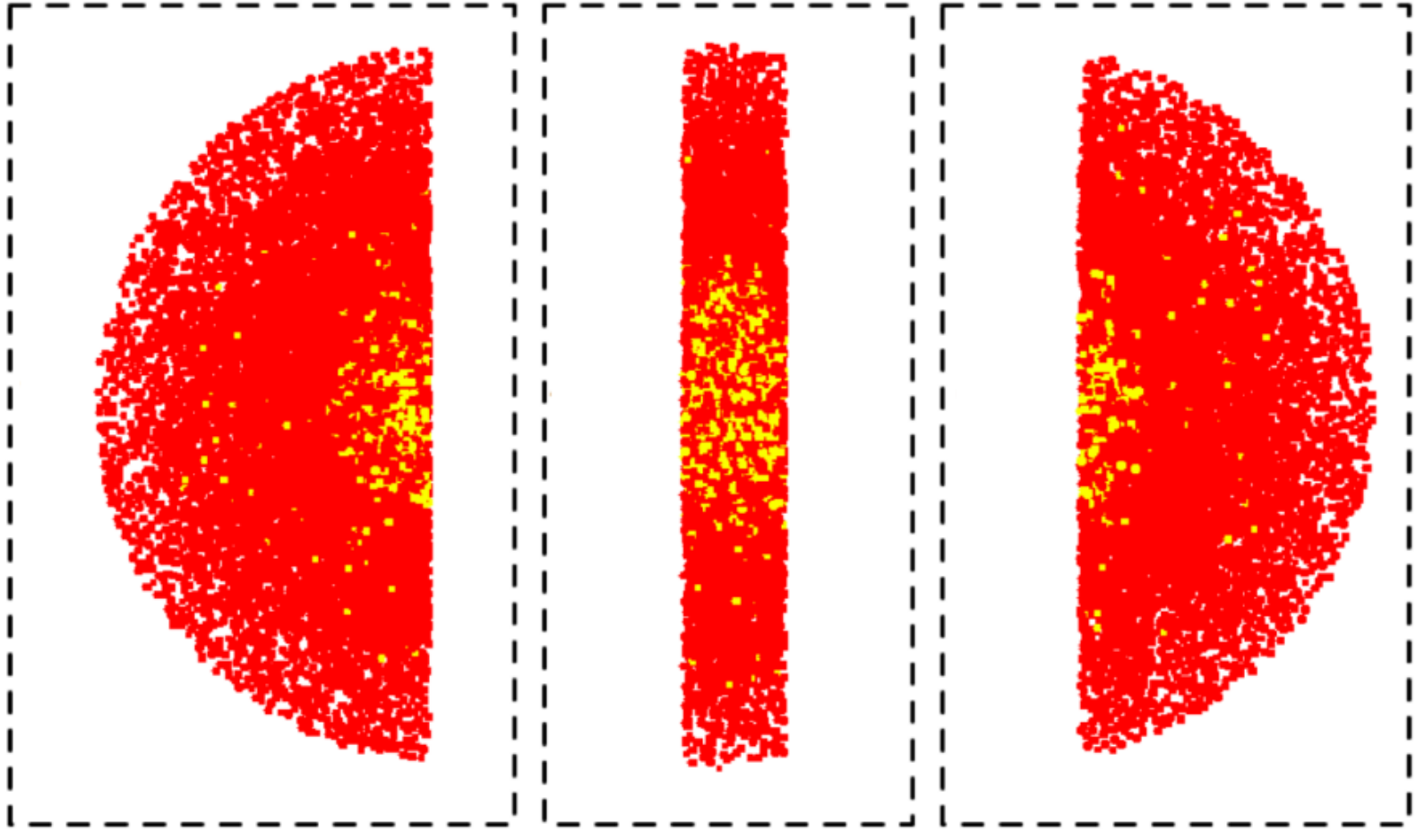
1. Biological Simulation Basics

2. Distributed Runtime

3. HPC on Cloud

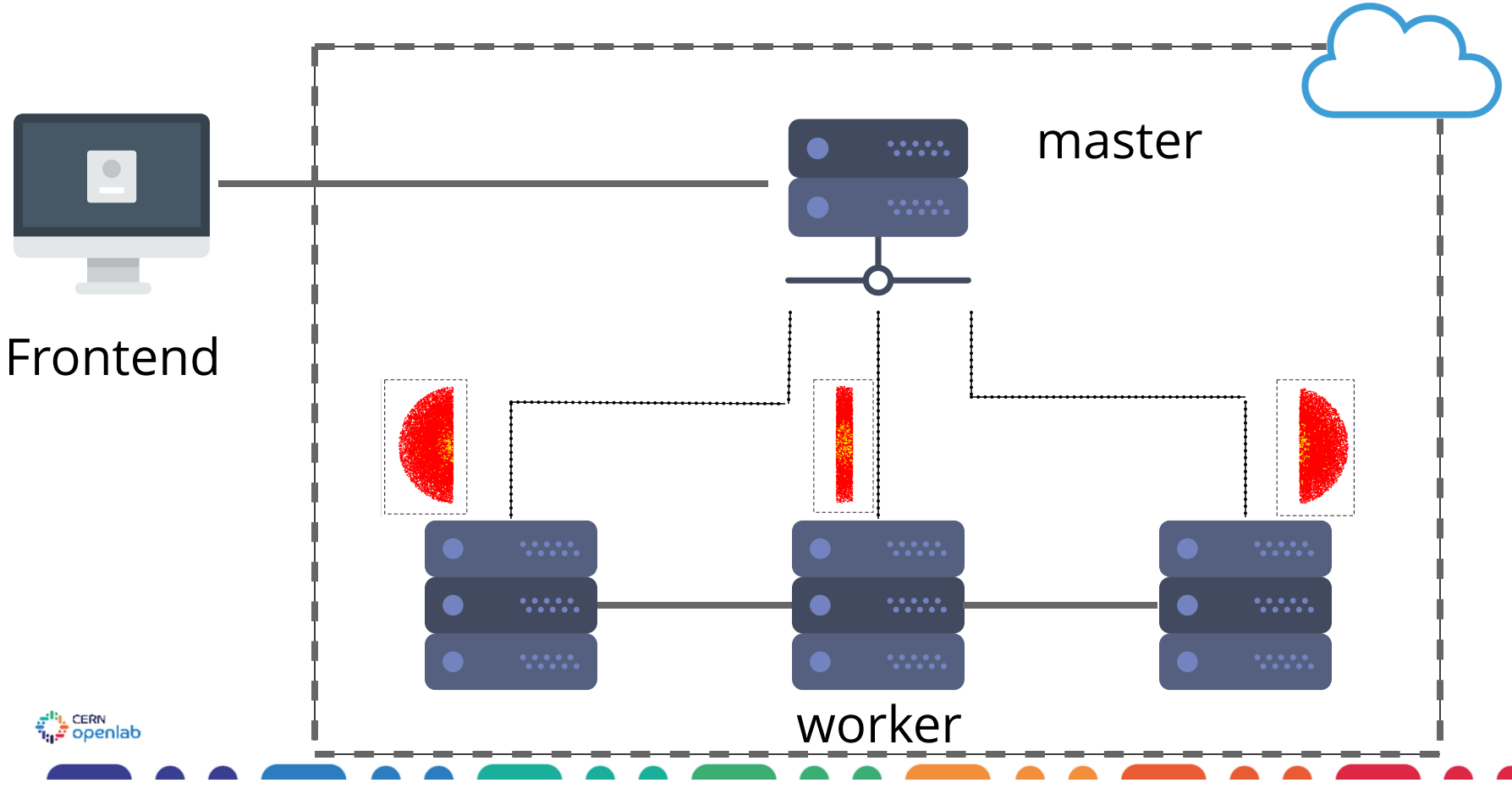
4. Data Movement Minimization

Domain-Decomposition

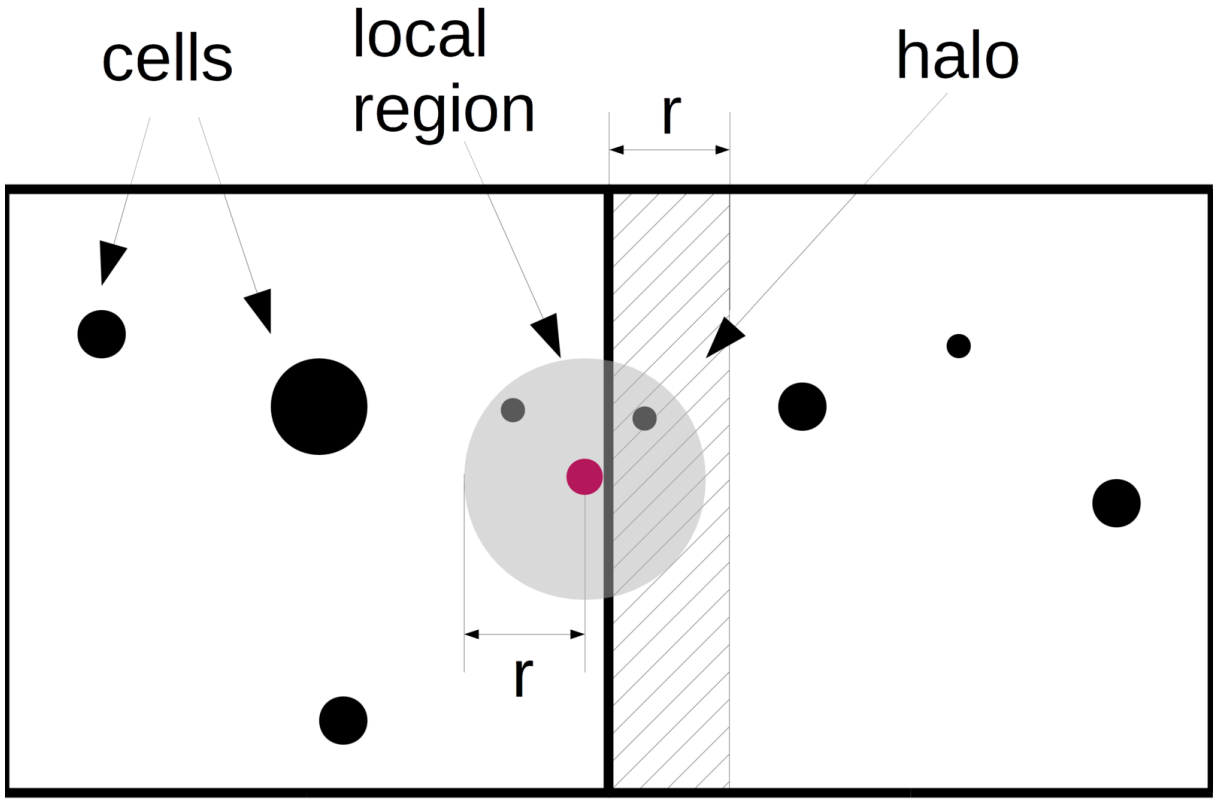


Hauri, Andreas. *Self-construction in the context of cortical growth*. Diss. 2013.

Distributed Runtime



Border Region



Outline

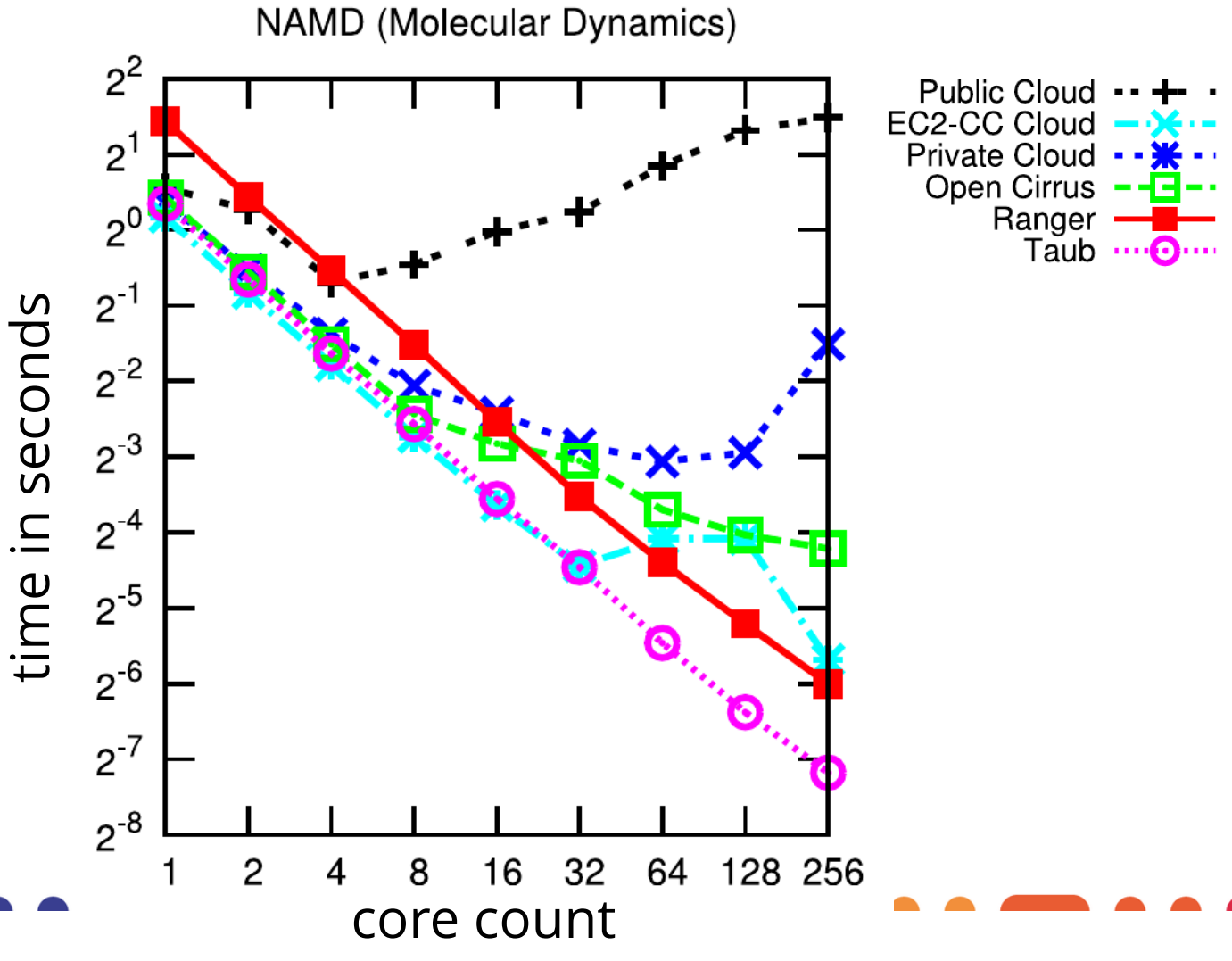
1. Biological Simulation Basics

2. Distributed Runtime

3. HPC on Cloud

4. Data Movement Minimization

NAMD Scaling



A. Gupta *et al.*, "Evaluating and Improving the Performance and Scheduling of HPC Applications in Cloud,"

Performance Issues in the Cloud

- Poor network performance compared to supercomputers
- Virtualization
- Resource contention with other tenants

Outline

1. Biological Simulation Basics
2. Distributed Runtime
3. HPC on Cloud
4. Data Movement Minimization

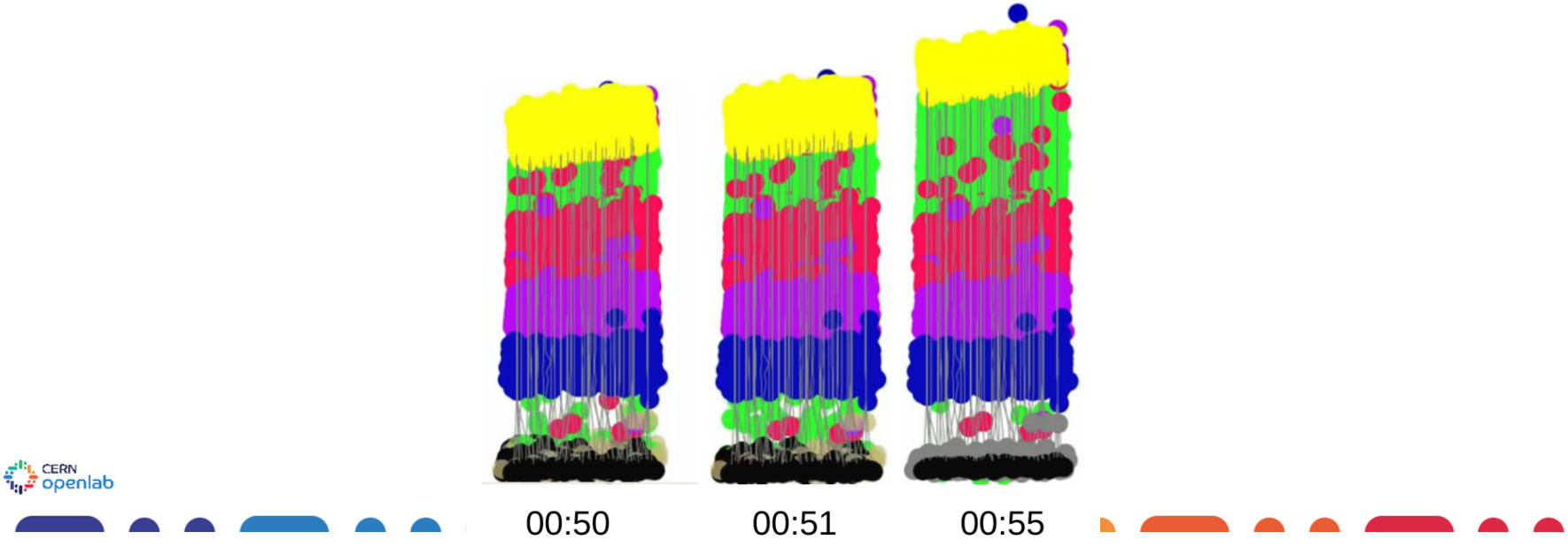
Key Observations

- Some regions and data members are static
- Changes are incremental
- Values might be predicted
- Communication can be replaced with additional computation

Static Regions

- **Definition:** simulation objects whose values do not change along the time dimension

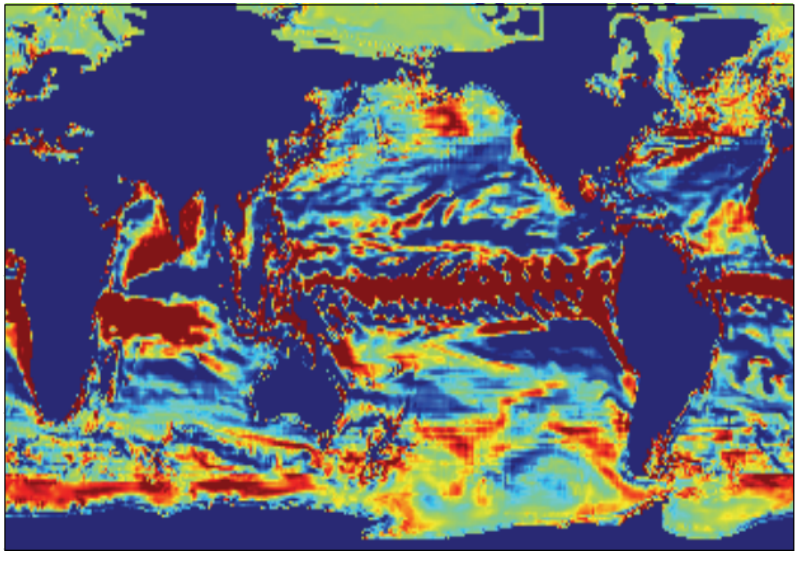
Example: growth of the cerebral cortex

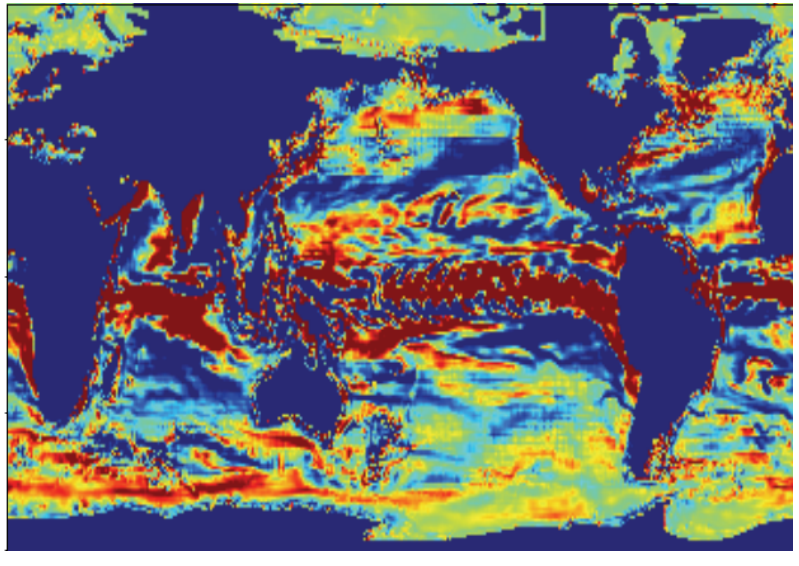


https://www.youtube.com/watch?v=9InvFfnAkus

Static Regions

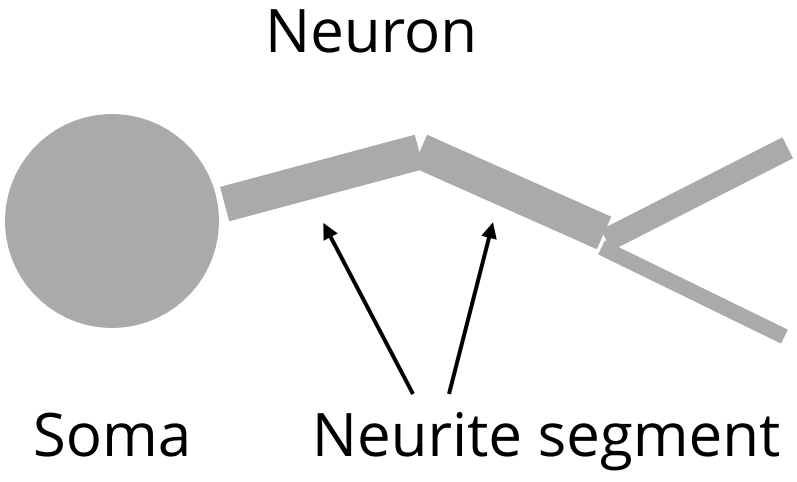
- Climate Simulation: continental data does not change during simulation of the ocean surface





S. Liu, X. Huang, Y. Ni, H. Fu, and G. Yang, “A High Performance Compression Method for Climate Data,” in 2014 IEEE International

Static Data Members

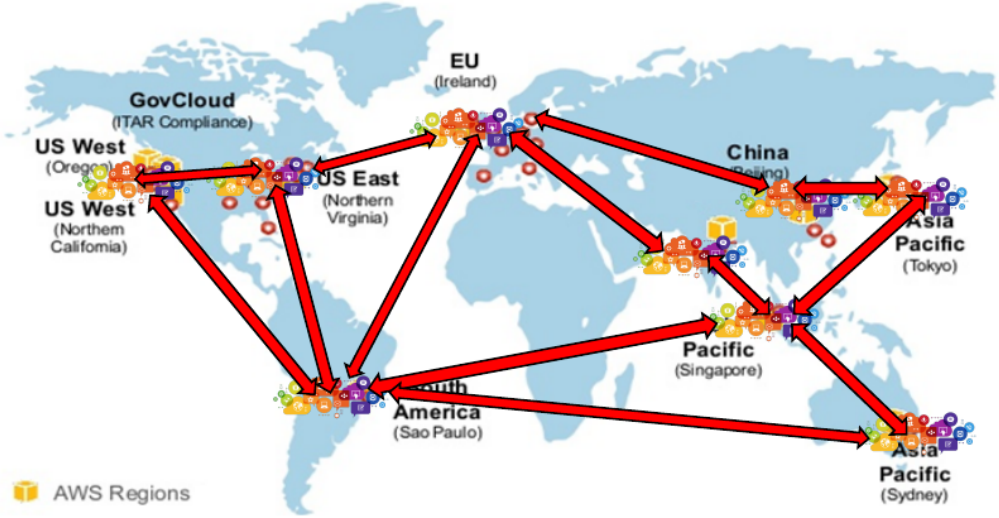



```
class NeuriteSegment {
  std::array<double, 3> position_;
  NeuriteSegment* parent_;
  ...
};
```

Incremental Changes

- **Observation**: Change between two time steps might be insignificant
- **Idea:** Communicate only significant updates

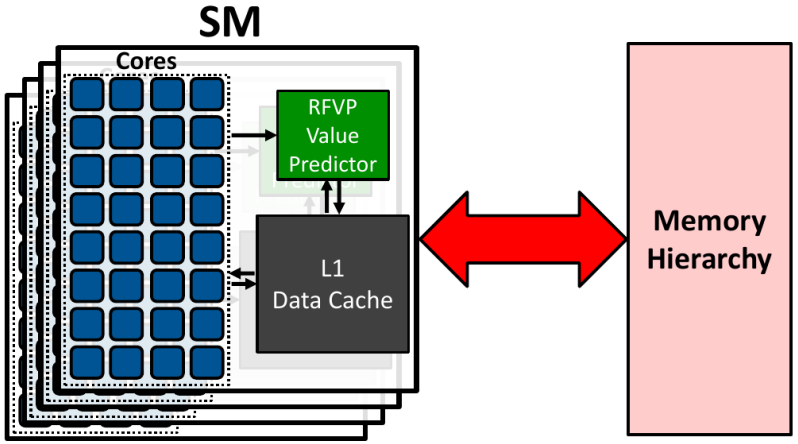
- **Example**: GAIA: Geo-distributed machine learning approachir



K. Hsieh et al., “Gaia: Geo-Distributed Machine Learning Approaching LAN Speeds,” in 14th USENIX Symposium on Networked Systems Design and Implementation (NSDI 17).

Value Prediction

- **Idea:** Predict values for safe-to-approximate variables
- Inspiration from computer architecture:
  "Rollback free value prediction"

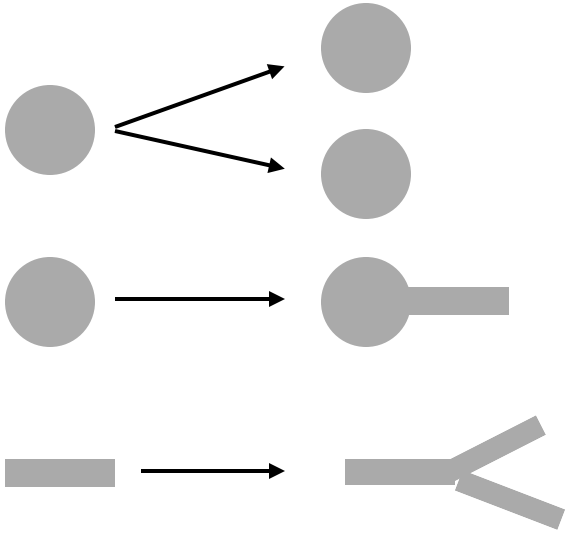



```
class NeuriteSegment {
  std::array<double, 3> position_;
  NeuriteSegment* parent_;
  ...
};
```


A. Yazdanbakhsh, G. Pekhimenko, B. Thwaites, H. Esmaeilzadeh, O. Mutlu, and T. C. Mowry, "RFVP: Rollback-free value prediction with safe-to-approximate loads," *ACM Transactions on Architecture and Code Optimization (TACO)*, vol. 12, no. 4, p. 62, 2016.

Computation vs Communication

- **Idea:** Recompute certain events on the destination server instead of transferring the results
- **Examples:**
  - Cell devision
  - Neurite
    - extension
    - bifurcation
    - ...



- Send event descriptor instead of whole new simulation object

Next Steps

- Develop the distributed runtime
- Verify that network properties are a bottleneck in the cloud
- Provide detailed analysis for different approaches
- Develop new ideas

QUESTIONS?

lukas.breitwieser@cern.ch

CERN openlab