



**MVA PICH**

MPI, PGAS and Hybrid MPI+PGAS Library

# **Bridging Neuroscience and HPC with MPI-LiFE**

**Shashank Gugnani**

The Ohio State University

E-mail: [gugnani.2@osu.edu](mailto:gugnani.2@osu.edu)

<http://web.cse.ohio-state.edu/~gugnani/>

# Neuroscience: Challenges and Opportunities

- Advancements in imaging technology
  - MRI, CAT Scan, etc.
- Higher resolution neuroimaging data
- How do we handle huge amounts of this data?
- Requirement for more efficient and faster algorithms
- *Can High Performance Computing help?*

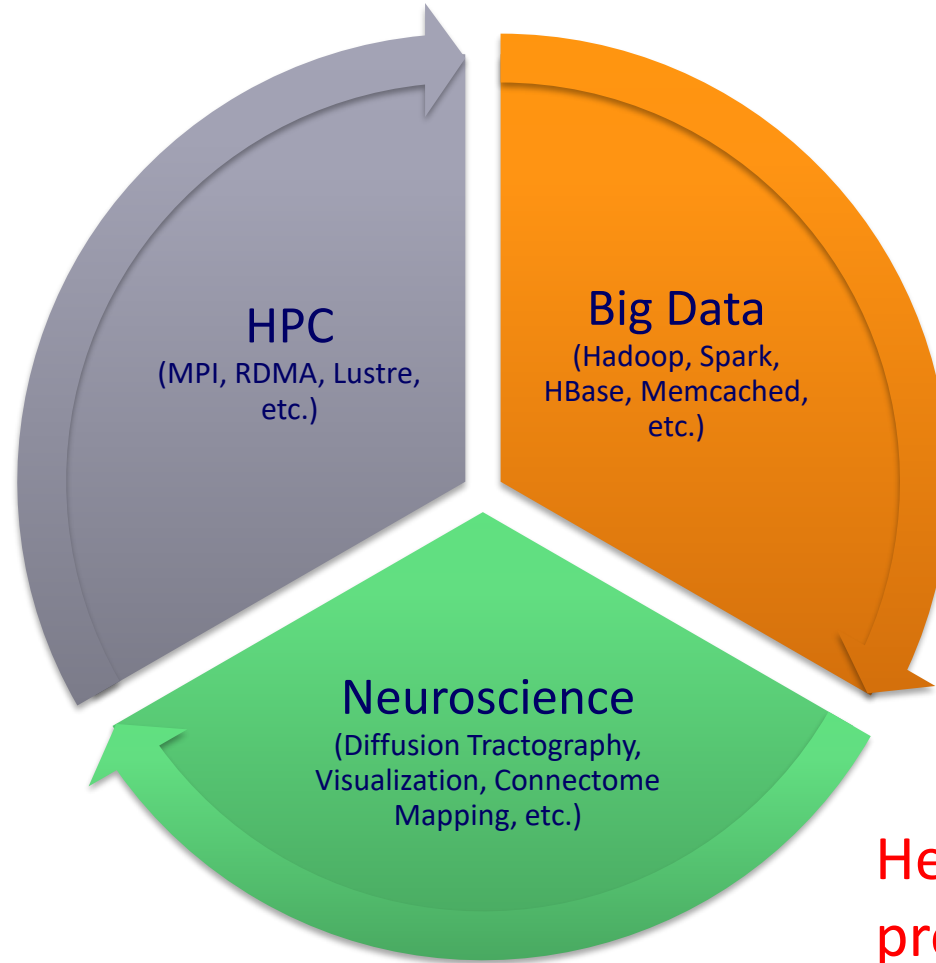
# Advanced Computational Neuroscience Network (ACNN)

- Develop a new generation of sustainable interdisciplinary **Neuroscience Big Data** research
- Collaboration between University of Michigan, Indiana University, Northwestern University, **Ohio State University**, Case Western Reserve University, and Washington University
- Funded by **NSF**
- <http://www.neurosciencenetwork.org/>

# Our Goal

- Accelerate Neuroscience Data
  - High-Performance Computing (HPC)
    - Different Programming Models (MPI, OpenMP and PGAS)
  - Big Data
    - Hadoop, Spark
  - Deep Learning
    - Caffe, Tensorflow

# Bridging HPC and Neuroscience



Convergence of HPC, Big Data, and NeuroScience!

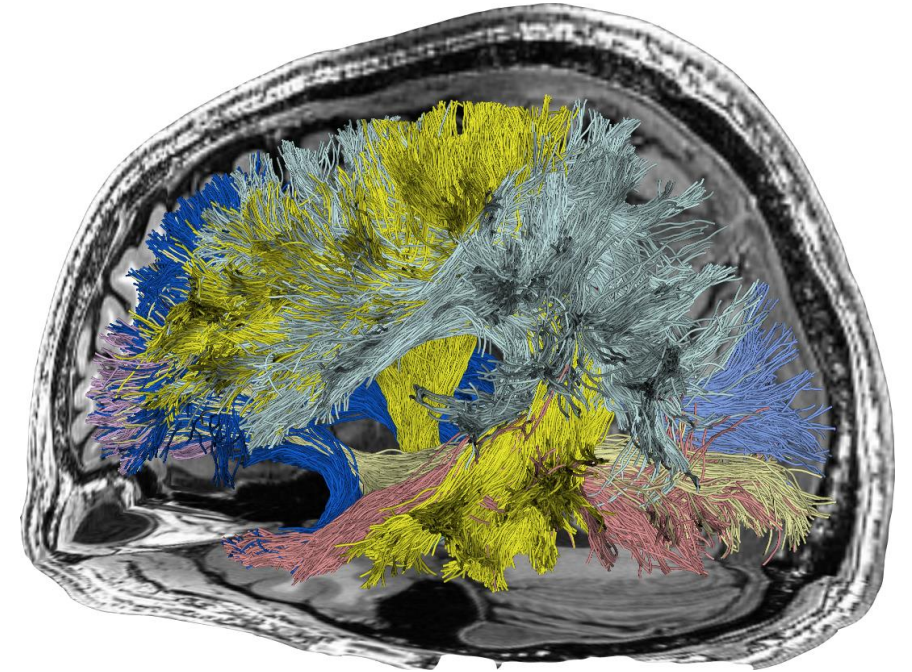
Heavy data analysis and processing requirements

# Overview of the NeuroHPC project

- Aimed at bringing HPC to the field of Neuroscience
- **MPI-LiFE**
  - Scalable and distributed tool for statistical evaluation of brain connectomes
  - Based on LiFE method in the Encode Toolbox
  - MPI-based distributed sparse multiway matrix multiplication
  - Exploits efficient communication in MVAPICH2
  - Highly optimized for virtualized environments using Docker
  - Flexibility to run on any multi-core laptop, desktop, server, and cluster
- **MPI-LiFE 0.9 is now available for download!** (11/10/2017)
- For more information, visit <http://neurohpc.cse.ohio-state.edu/>

# Brain-LiFE

- **Linear Fascicle Evaluation (LiFE<sup>1</sup>)**
  - an approach to predict diffusion measurements in brain connectomes
  - evaluate evidence supporting white-matter connectomes generated using MRI and computational tractography
  - show evidence for white-matter tracts and connections between brain areas
  - heavily used by computational neuroscientists



*Different colors show groups of white matter tracts connecting different portions of the brain*

<sup>1</sup><https://github.com/francopestilli/life>

# Brain-LiFE

- Originally written in MATLAB
  - Slow and inefficient
  - Convert to C to run outside MATLAB environment
  - Modularize and pipeline for distributed environment
- Based on Non-Negative Least Square (NNLS) constrained optimization
  - Uses large-scale sparse multiway matrix multiplication
    - Design, develop, and evaluate MPI-based algorithm
    - Dockerize solution



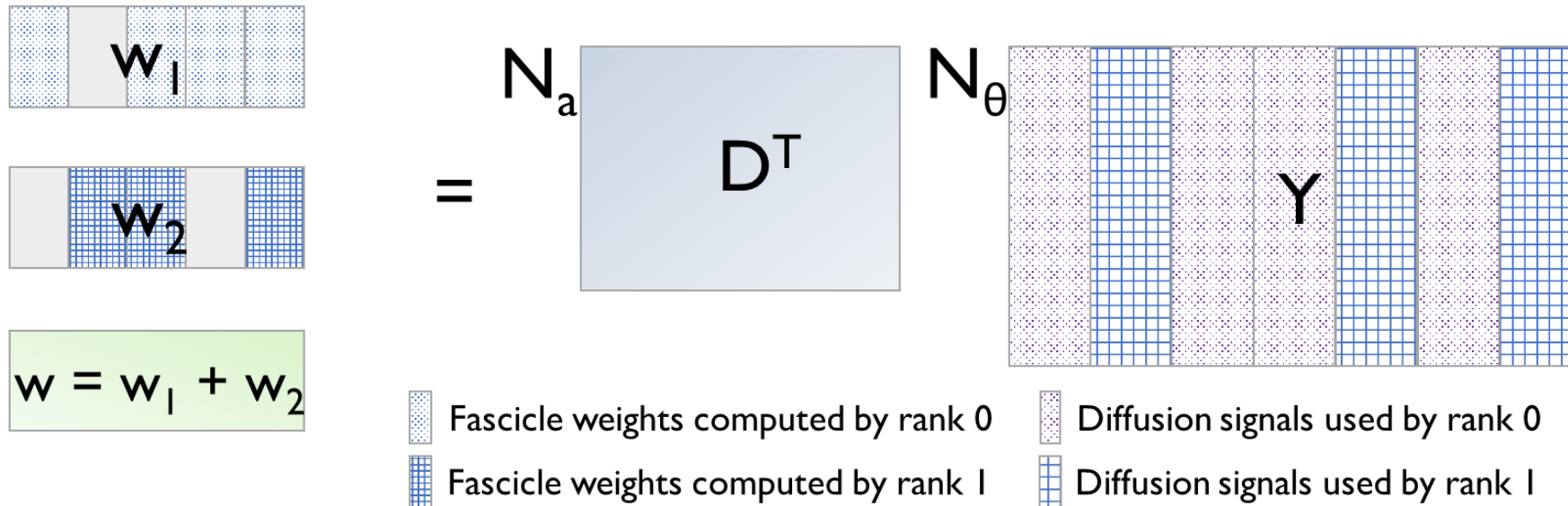
# MPI-LiFE: Initial Design using MVAPICH2

- Computationally intensive tasks are the computations of sparse multiway matrix by vector products
  - $w = M^T y$  and  $y = Mw$
  - Have been **parallelized** using **MPI** and **OpenMP** by dividing the task among multiple MPI processes
- Implementation uses **MVAPICH2**, from OSU team
  - Exploits efficient communication primitives to drastically improve performance
  - Uses high-performance networks to reduce communication latency

*S. Gugnani, X. Lu, F. Pestilli, C.F. Caiafa, and D. K. Panda, MPI-LiFE: Designing High-Performance Linear Fascicle Evaluation of Brain Connectome with MPI, HiPC'17*

# MPI-based multiway matrix multiplication

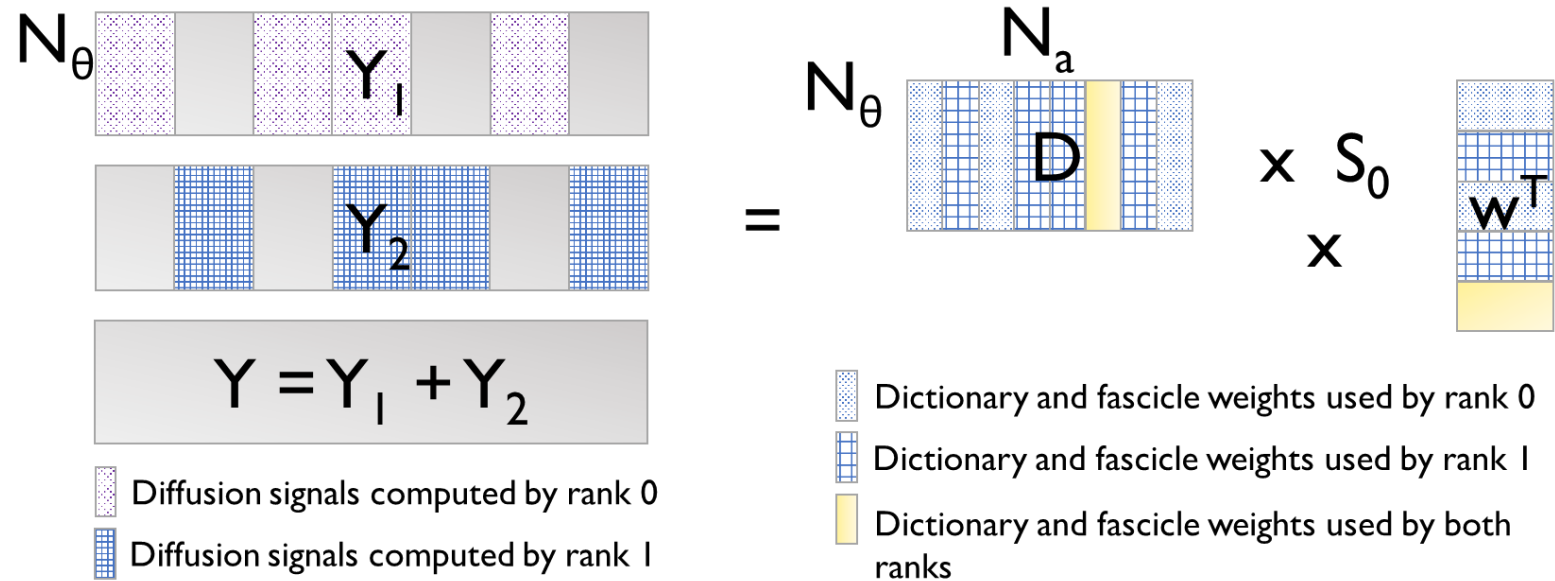
- $w = M^T y$ 
  - Distribution of data using MPI\_Bcast, MPI\_Scatter
  - Gathering of results using MPI\_Gather
  - Use of OpenMP to parallelize computation within an MPI process



Computation of  $w = M^T y$  using 2 MPI processes

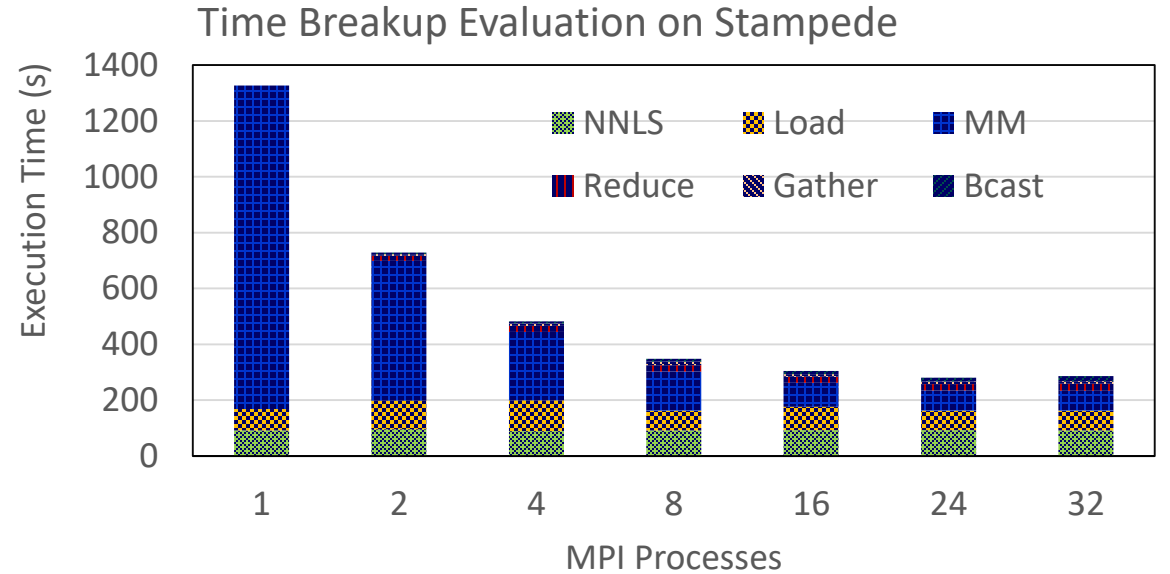
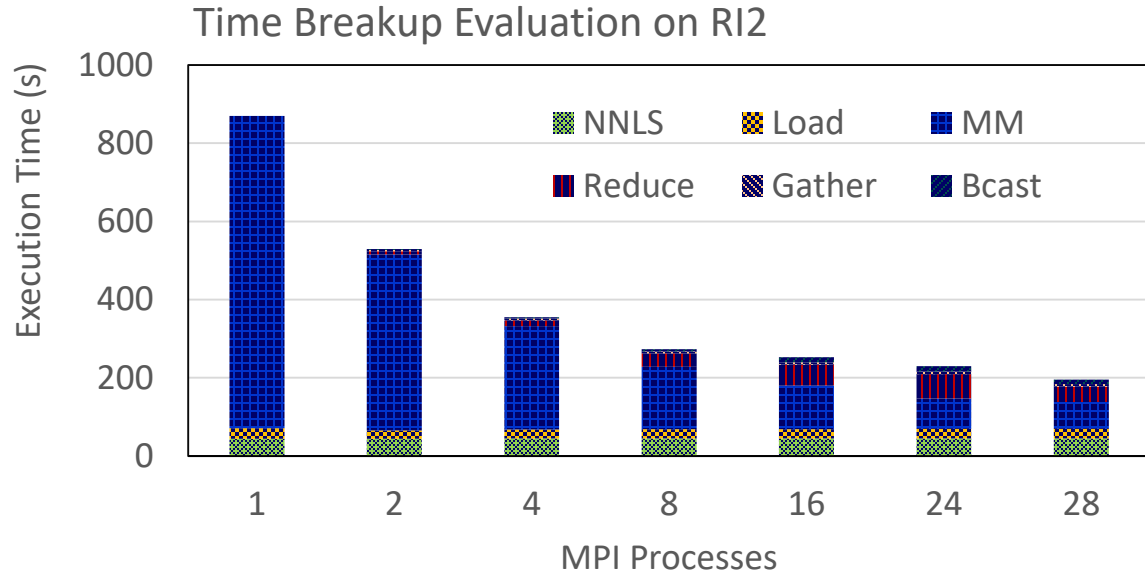
# MPI-based multiway matrix multiplication

- $y = Mw$ 
  - Distribution of data using MPI\_Bcast, MPI\_Scatter
  - Gathering of results using MPI\_Reduce
  - Use of OpenMP to parallelize computation within an MPI process



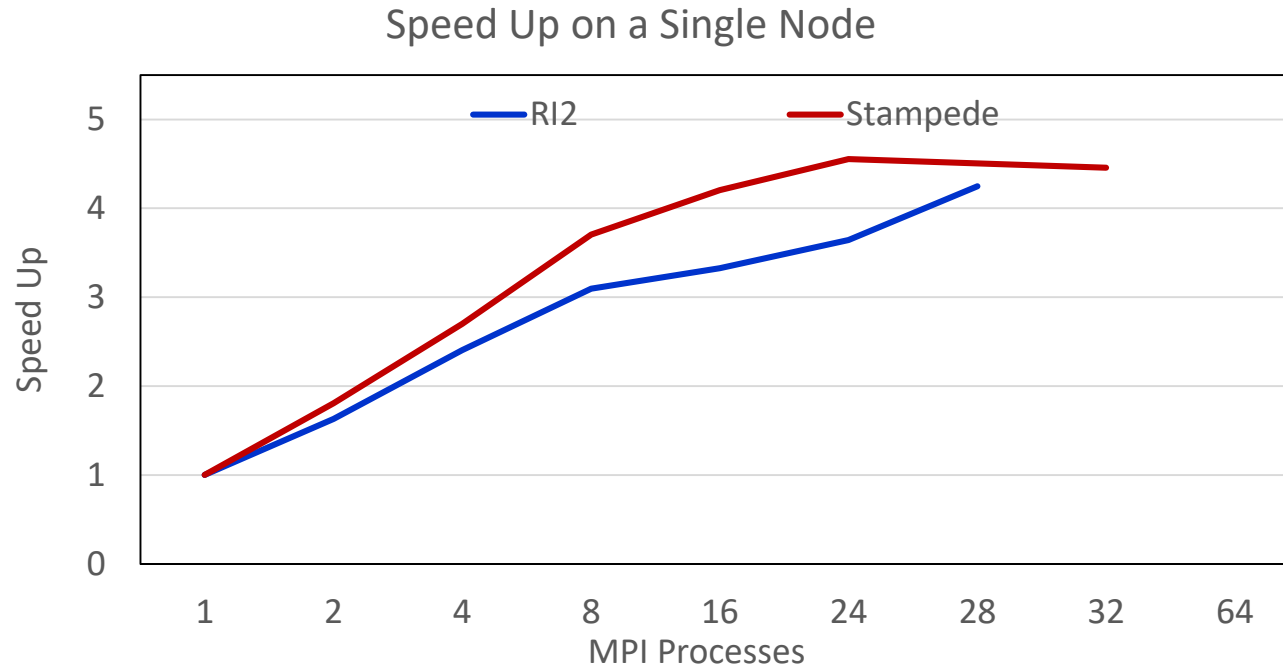
*Computation of  $y = Mw$  using 2 MPI processes*

# Evaluation with MVAPICH2: Single Xeon Node



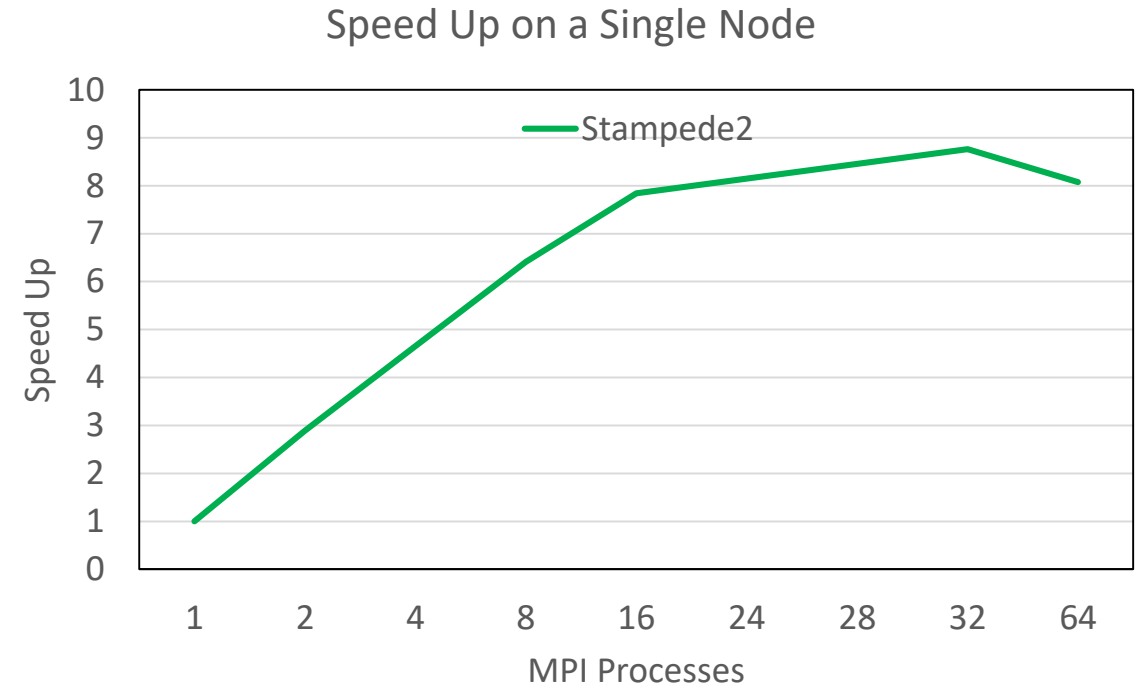
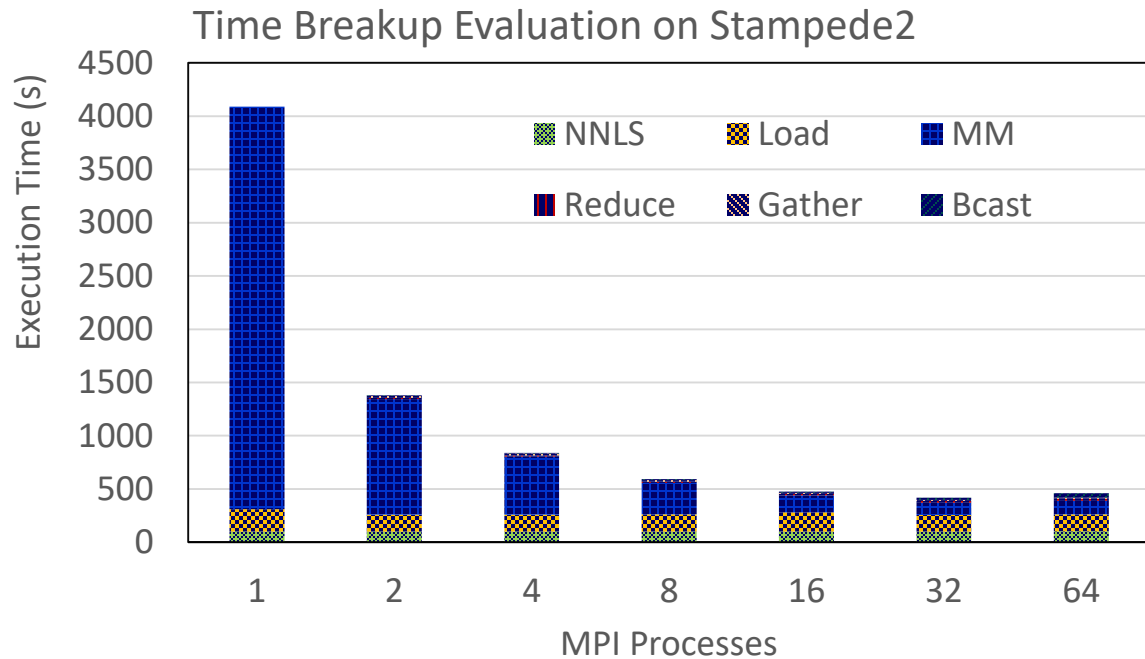
- Evaluation on
  - **OSU RI2** (Intel Broadwell CPUs, 28 cores, 128 GB memory per node)
  - **TACC Stampede** (Intel Sandy Bridge CPUs, 32 cores, 1 TB memory per node)
- Speed up of up to **4.2x** on RI2 and up to **4.5x** on Stampede

# Evaluation with MVAPICH2: Single Xeon Node



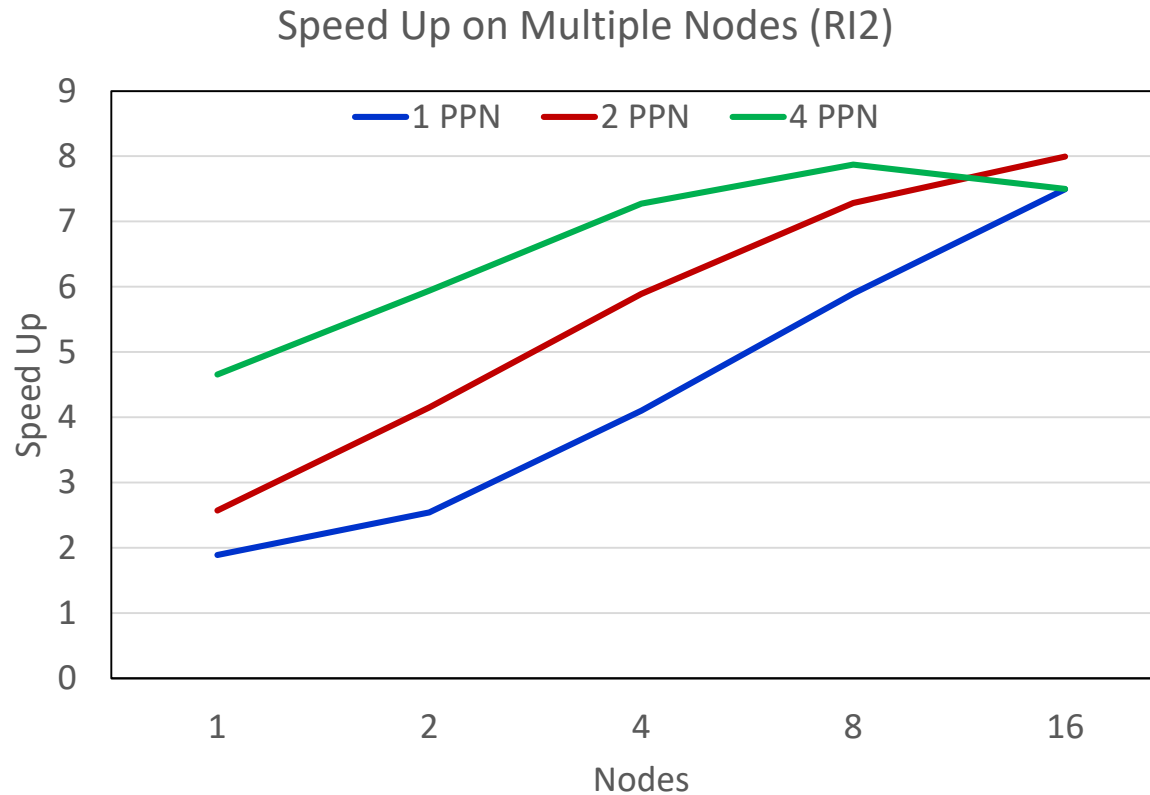
- Evaluation on
  - **OSU RI2** (Intel Broadwell CPUs, 28 cores, 128 GB memory per node)
  - **TACC Stampede** (Intel Sandy Bridge CPUs, 32 cores, 1 TB memory per node)
- Speed up of up to **4.2x** on RI2 and up to **4.5x** on Stampede

# Evaluation with MVAPICH2: Single KNL Node



- Evaluation on **TACC Stampede KNL** (Intel Xeon Phi KNL CPUs, 68 cores, 96 GB memory per node)
- Up to **8.7x** speed up

# Evaluation with MVAPICH2: Multi-Node



- Evaluation on **OSU RI2** (Intel Broadwell CPUs, 28 cores, 128 GB memory per node)
- Up to **8.1x** speed up on RI2

# MPI-LiFE: Continuing Work

- Maximum parallelization is currently being achieved with MPI for the current LiFE algorithm
- Exploiting enhanced and optimized designs for scale-up and scale-out
  - MPI+OpenMP
  - PGAS Models (OpenSHMEM)
  - GPGPUs
- The sequential algorithm may need to be re-designed
  - May lead to better parallelization

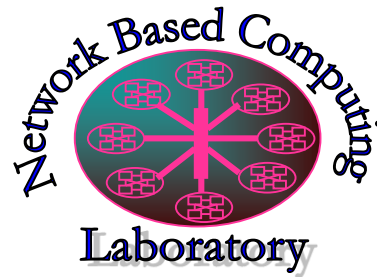


# Demo

- Demo on NSF Chameleon Cloud
  - 24 cores per node, Intel Haswell Processors
- Execution using dockerized version of MPI-LiFE
  - Available on docker hub: <https://hub.docker.com/r/neurohpc/neurohpc/>
- Analysis of tractography algorithm on dMRI scan
  - Demo run with 50 iterations

# Thanks!

[gugnani.2@osu.edu](mailto:gugnani.2@osu.edu)



Network-Based Computing Laboratory

<http://nowlab.cse.ohio-state.edu/>

The High Performance Neuroscience Project (NeuroHPC)

<http://neurohpc.cse.ohio-state.edu/>