

# Phylanx Year 2 Kickoff Meeting

April 19, 2018

- Welcome
- Attendees: Hartmut Kaiser, Steven Brandt, Adrian Serio, Rod Tohid, Alireza Kheirkhahan, Shahrzad Shirzad, Parsa Amini, Bibek Wagle, Tianyi Zhang, Cheick Bamba, Thomas Heller, Kevin Huck, Muhammad Monil, Kate Isaacs, Kathryn Williams, Christopher Taylor, John Leidel, Bryant Lam, Chris Mineo, Indranil Banerjee, Stefan van Zwam
- Chris
  - What we have accomplished
  - What we want to accomplish
    - Algorithms
    - Distributed
    - Accelerators
    - APEX
    - Visualization
  - Future
    - Distributed Visualization?
    - Phylanx-Spark interface?
    - Phylanx-DataFrames Interface?
    - Phylanx-Streaming Support?
- Hartmut
  - Overview of Phylanx
    - The challenge- large datasets on distributed heterogeneous hardware
    - Goal- abstract away these obstacles for domain scientists
  - NumPy + HPC
  - Parts of Phylanx
  - John- What do you need in hardware
    - One cycle context switches
    - Global Address Resolution
    - Hardware Queues
    - Hardware atomics
- Monil
  - Performance comparison
  - Looking at NumPy, Dask, SciKit-Learn, TensorFlow, Phylanx, Spartan
  - TensorFlow took the longest, SciKit was the fastest
  - We need installing to be easy
    - SPAC and Containers
- Katy
  - Data Visualization
  - Current visualizations
  - I want to know what you want!
  - Discussion
    - Bryant- Can we see Hardware information?
- Stefan
  - Tiling

- Combinatorial Optimization (combo of Operations research, Mathematics, Computer science)
- Tiling problem posed by Phylanx
- Model
  - List tiling choices
  - Costs on edges
  - Hartmut- Do you take into consideration re-tiling?
    - John- BFS algorithm is a good one to look at for re-tiling
  - Thomas- sounds good for performance guided optimization
- Meta-Questions
- John- Are you considering pruning?
  - David Calahan
- Rod
  - AST Transformations
  - Parts of Phylanx
  - Interfaces to Phylanx
    - Python and PhySL front-ends
  - Thomas- Will you analyze implicit loops
    - Rod- We can look into this
  - Multi-precision optimizations
- Kevin
  - APEX
  - What is APEX
    - Parts of APEX
  - Views offered
    - OTF2
      - Vampir
    - Taskgraph View
  - Next Steps
    - Phylanx Policies
    - HPX Policies
    - Phylanx Counter Integration
    - Task Dependency Analysis
    - Updated Documentation
  - Example of automatic task analysis
    - Finds tasks “to blame” from idle regions
- Kate
  - Visualization in Support of Phylanx Development
  - Gantt Charts
  - Views of applications
    - Logical view
    - Real time view
  - Phylanx
    - Be able to view multiple levels
  - Integration into the workflow
    - Jupyter Notebook
  - Questions:
    - How do we organize data?

- How do we manage large domains?
    - How to we combine vis.?
    - What is minimal data necessary to derive insight?
    - How do we integrate vis. into the Phylanx workflow?
  - We want feedback!
- Lunch!
- Indranil
  - Research background
- Bibek
  - LRA Updates
    - Current performance
  - Future work
    - Multiple LRA solvers
    - Distributed
    - Task inlining
- Shahrzad
  - ALS Updates
    - Performance
    - Improvements
      - Using solvers
      - Adding parallelization
      - Using sparse matrix or conjugate gradient
- Ali
  - IO in Phylanx
    - Support CSV, HDF5 (via HighFive)
  - Need to implement collective IO
- Parsa
  - Overview
    - Blaze
    - Primitives work
    - PhySL
    - Performance Counters
    - K-Means
    - Work on distributed
  - Research
    - Performance impaired distributed apps
    - Dynamic Data Migration
- Project Planning
- Steve's Demo
  - Chris- It might be nice to tell the Phylanx decorator to execute on the cloud
- End of the day