Components of NSF 144305

- Design and build SPIDAL – a Scalable Parallel Interoperable Data Analytics Library
- Domain specific libraries – mainly from project
- Core Machine Learning Libraries
- High Performance for Java, Libraries and Implementations: HPC and Clouds with Data Converged Classification
  - 64 Features in 4 views for HPC

Building Blocks of Proposal

64 Features in 4 views for HPC-Big Data Converged Classification

HPC-ABDS Apache Big Data Stack

**Applications in action**

MIDAS and Biomolecular Simulations

- Previous Work (HPC-ABDS): Cloud-HPC Interoperable software with performance of HPC and rich functionality of commodity Apache Stack
- Core Machine Learning Libraries
- Domain specific libraries – mainly from project
- High Performance on Java, Libraries and Implementations: HPC and Clouds with Data Converged Classification
  - 64 Features in 4 views for HPC

**Applications in action**

MIDAS and Biomolecular Simulations

- Parallel Interoperable Data Analytics Library
- Deriving 50 Ogres and 64 Convergence Features of data intensive Applications
- BIG Data Application Analysis
- MIDAS and Pathology Informatics.
- Vision, Spatial Geographical Information Network Science, Epidemiology, Computer Science
- DevOps functionality of commodity Apache Stack software with performance of HPC and rich -HPC-ABDS Apache Big Data Stack

**Applications in action**

**HPC-ABDS**

MIDAS and SPIDAL Java

High Performance Middleware and Language


http://spidal.org

**Applications in action**

Multi-Scale Imaging and Spatial Data Analytics

<table>
<thead>
<tr>
<th>Multi-Scale Imaging and Spatial Data Analytics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark of Trajectory File Formats</td>
</tr>
<tr>
<td>2D and 3D image analysis</td>
</tr>
</tbody>
</table>

Biomolecular Simulations: Parallel analysis in the MIDAS Analysis Library

- Performance strongly depends on the format
- XTC file size plays a key role in determining performance

**Applications in action**

WebPlotViz – Browser Visualization of High Dimensional Data

WebPlotViz is a 2D/3D data point browser that can visualize very large volumes of 2D or 3D data, as points in a virtual space and enable users to explore the virtual space interactively. WebPlotViz also includes support for Time Series Data plots.

**Applications in action**

**Polar Remote Sensing Algorithms**

**Counting Triangles in Massive Networks**

Many applications in data mining, network analysis, social science, and database systems

**Applications in action**