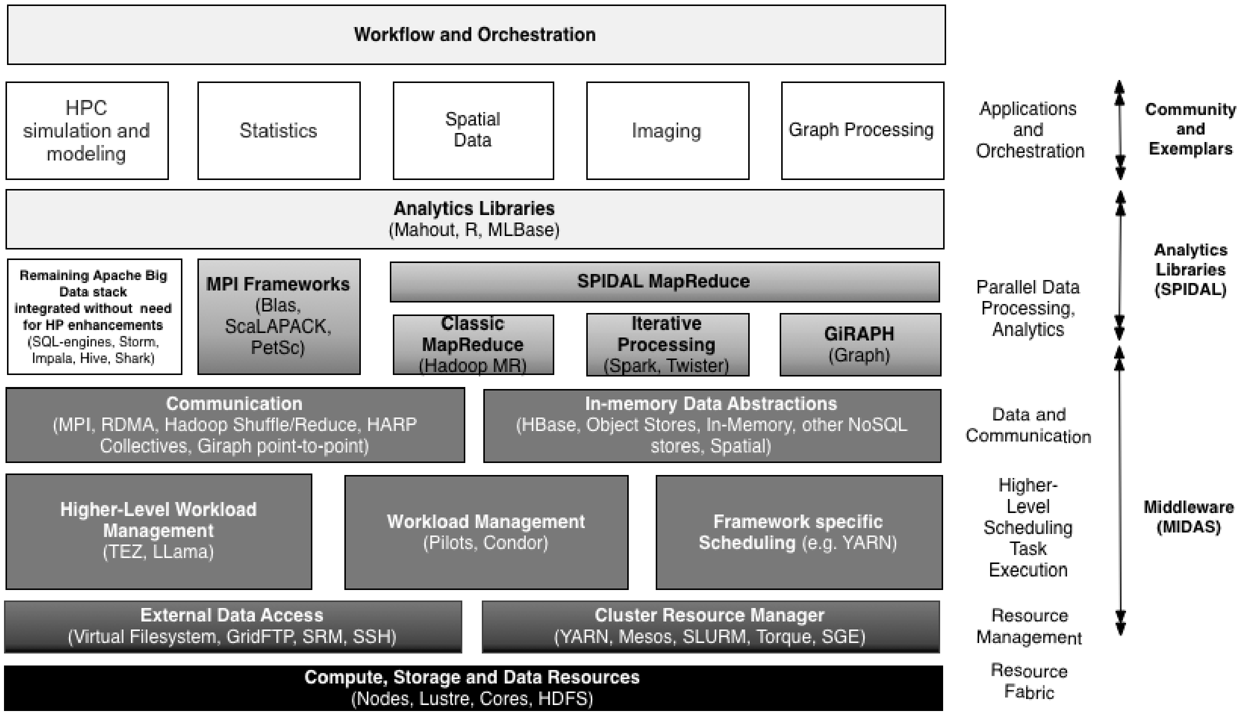
CIF21 DIBBs: Middleware and High Performance Scalable Analytics Libraries for Extreme Scale Data Intensive Science



HPBDS

The growing importance of data in many fields including physical and biological sciences, and the ability to derive insight and knowledge from increasing volumes of complex data, points to the importance of advanced analytics. Analytics needs to be able to utilize the full range of available infrastructure, however, the coupling between tools, analytic engines and infrastructure is often rigid, thus it is often difficult to employ existing solutions for contemporary environments that they were not natively or originally designed for. Further, many tools were developed at a time when parallelism was not essential. In addition, interoperability at multiple levels remains elusive, as well as difficult, and scalable yet general-purpose and broadly applicable solutions in the form of analytic libraries and abstractions are noticeable by their absence.

This project will build a **H**igh-**P**erformance **B**ig **D**ata **S**tack (**HPBDS**) for NSF production CI such as XSEDE. As part of HPBDS, we propose two fundamental building blocks: **M**iddleware for **D**ata-**I**ntensive **A**nalytics and **S**cience (MIDAS) and the **S**calable **P**arallel **I**nteroperable **D**ata **A**nalytics **L**ibrary (SPIDAL). SPIDAL will enable interoperable high performance data analytics and is based upon a careful analysis of architectures, tools and application characteristics/requirements. Although built from many existing components and capabilities, MIDAS is conceptualized and designed from first-principles with high-performance CI such as XSEDE as target architecture. MIDAS will support SPIDAL's interoperability and performance requirements.

Big data efforts within industry have made impressive gains in data-intensive computing, and have seemingly converged around the Apache stack, a distinctive feature is the existence of many implementations of the specific components of the Apache stack, providing sufficient richness in the trade-off between performance and capability. In contrast, within the scientific computing community, progress has been reliant either on long-term foundational advances or short-term hardware fixes. However, many industry solutions are not ready for “off the shelf” usage by the scientific community that uses XSEDE and high-end campus clusters. ***HPBDS will remedy this major gap and proffer an integrated solution that brings the best of recent advances to the service of extreme-scale science requirements on NSF’s current and future production platforms.*** HPBDS will utilize and expose the integrated relative technical strengths [1] of the two hitherto disjoint approaches and communities, yet it will focus on delivering these as production grade implementations that will bring the best-of-both to NSF’s shared-infrastructure -- such as XSEDE, OSG and other domain-specific infrastructure, as well as the software developments underway as part of the SI2 software program.

HPBDS will translate these applications characteristics, infrastructural requirements and existing capabilities into well-defined and implemented building blocks. HPBDS has three fundamental blocks: ***Parallel Analytics Libraries*** that capture system abstractions and expose application requirements, ***Middleware*** upon which to build such libraries that are interoperable yet high-performance, and ***Communities*** **and** ***Exemplars***.

Parallel Analytics Libraries: The high-performance community has prospered thanks to libraries like MPI, PETSc and SCALAPACK; SPIDAL brings this concept to data-intensive applications. SPIDAL will utilize capabilities of the underlying middleware that will be exposed via well-designed and engineered libraries.

Middleware: HPBDS will capture these common characteristics and requirements by identifying key abstractions. MIDAS is based on **abstractions** in the areas: a) Software defined System, b) Storage layer including a spatial access abstraction, c) Scheduling layer using advances in multi-level and application-level scheduling, d) Collective layer that permits Map Collective generalization of Iterative MapReduce e) Parallelism or Programming model which generalizes the hugely popular Hadoop and MPI SPMD models.

Community and Exemplars: Using the parallel analytic libraries that we will deliver, we will support multiple communities including but not limited to, users of graphs, clustering, statistics, imaging, spatial data, high-performance simulation & modeling. The graph algorithms come from the Network Science community in the CINET (VT) system; clustering libraries come from Mahout and published work from two partners (VT, IU); the statistics libraries are from existing IU MPI programs for dimension reduction, Mahout and R where we show how existing libraries can be refactored to HPBDS; the imaging libraries come from computer vision (IU) and pathology (Emory) applications; the spatial algorithms from Emory; the HPC simulation data visualization analytics comes from physics and chemistry.

Putting it together: HPBDS will enable multiple communities to use a core set of libraries SPIDAL -- graphs, statistics, imaging, spatial, HPC simulation and modeling -- on top of an abstractions-based (Job, data, communication, in-memory operations, parallelism) high performance middleware MIDAS. By integrating the two fundamental building blocks -- MIDAS and SPIDAL. HPBDS will provide new levels of scalability, application performance along multiple dimensions by combining application expertise, the broad Apache Big Data stack and best practice HPC. Specifically, HPBDS will, (i) implement resource management capabilities of MIDAS using job and data abstractions, as well as scalable and reusable fine-grained building blocks for HPC/high-end resources that realize communication and in-memory abstractions, (ii) define library interfaces (SPIDAL, with appropriate language bindings) to these fine-grained building blocks, (iii) provide scalable and parallel analytic libraries by integrating the aforementioned interfaces, the fine-grained building blocks with the resource management middleware capabilities of MIDAS, and (iv) refactor applications and exemplars to utilize SPIDAL application-facing libraries. SPIDAL libraries will be provided with a set of simple benchmark kernels. This proposal focuses on the data analytics libraries and needed abstractions built on a skeletal MIDAS middleware whose key features have already been demonstrated. We intend to improve MIDAS in performance and functionality outside this proposal

References:

[1] A Tale of Two Data-Intensive Paradigms: Applications, Abstractions, and Architectures, Shantenu Jha, Judy Qiu, Andre Luckow, Pradeep Mantha and Geoffrey Fox, Submitted to: IEEE BigData 2014 (under review), available at: http://arxiv.org/abs/1403.1528

[2] **High Performance High Functionality Big Data Software Stack**, G Fox, J Qiu and S Jha, in Big Data and Extreme-scale Computing (BDEC), 2014. Fukuoka, Japan. http://grids.ucs.indiana.edu/ptliupages/publications/HPCandApacheBigDataFinal.pdf