

High Performance High Functionality Big Data Software Stack

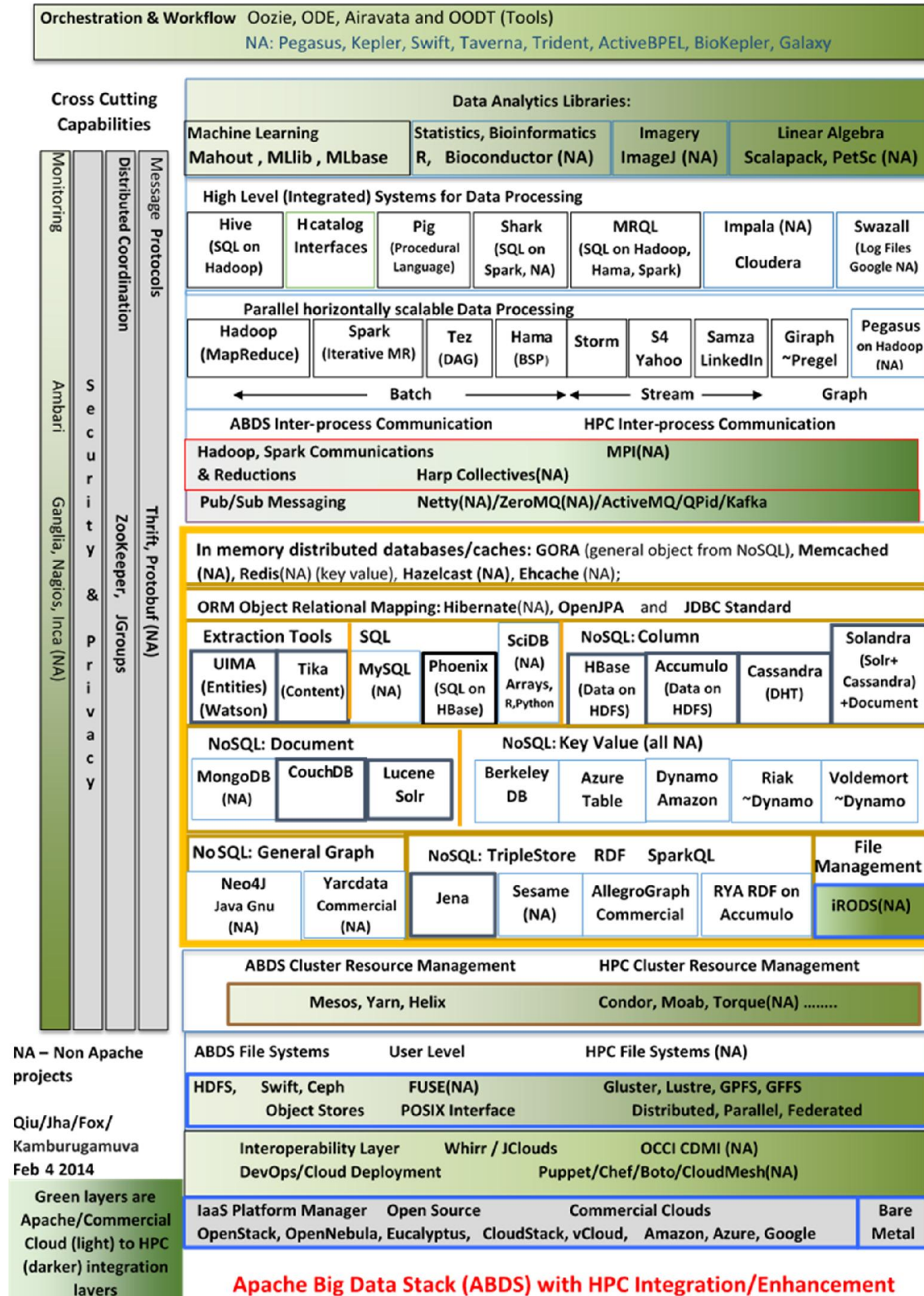
Geoffrey Fox, Judy Qiu, Shantenu Jha, Indiana and Rutgers University

Big data is important in all areas including research, government and commercial applications. The 51 use cases gathered in NIST [1] had 34, 8 and 9 use cases in these categories. Indeed the largest datasets are possibly those associated with commercial clouds including search and social media [2]. It is estimated that this year there are approximately 6 zettabytes of stored data [3] and CISCO estimates [4] almost a zettabyte of total IP traffic this year; with largest individual science applications like the LHC with “only” around 0.0001 zettabytes (100 petabytes). Other research applications in the NIST study were typically fractions of petabytes with much larger astronomy (LSST, SKA) projects underway and imagery (light sources, medical, surveillance, radar such as EISCAT-3D) large and growing in size with non-cardiac medical imagery around 70 petabytes a year in the USA [2]. This broad interest in Big data has spurred a frantic software activity much of it aimed at commercial cloud deployments. This was seen in NIST study in both architecture discussions and the details given in many use cases. Here we suggest that it is valuable to understand the large scale commercial approaches and understand how they can be made use of and as appropriate be integrated in the HPC and exascale data environments. This approach is helped by the broad use of Open source technology in the consensus commercial clouds, with in particular many Apache software projects contributing to what we can term ABDS or Apache Big Data Stack. This is depicted in a layered architecture on the figure on the following page.

This figure show over 40 Apache Big Data projects [5] with a selection of other commercial and Open source technologies. We use ABDS to refer to all the projects and not just those in Apache with HPC-ABDS denoting the HPC integration. We also color key layers where interaction between HPC and ABDS seems either critical or timely. This diagram highlights the rich functionality of ABDS. It is not just Hadoop but has important contributions to dynamic deployment; virtualization; file systems; cluster management; object data models including consistency, table, key-value, document, graph and SQL; in memory caching; messaging middleware; programming model and run time including graph, streaming, iterative and classic MapReduce; high level (productivity) systems with well-known analytics and machine learning libraries. Workflow and orchestration is also well represented. Not all these projects are of the same value and quality but together they represent a highly capable and productive environment and we suggest that it will be very valuable to consider how to integrate HPC with ABDS. In our initial study [6], we have identified file systems, cluster resource management, file and object data management, inter process and thread communication, libraries, workflow and monitoring as areas where HPC-ABFS integration will be particularly valuable. Much of this has already started with studies for example of Lustre-Object Store-HDFS integration [7], comparison of Yarn with HPC schedulers, experimentation with NoSQL stores [8, 9], iterative MapReduce with a variety of communication mechanisms including those from MPI community [10]. Study of library performance shows that HPC solutions outrun current Mahout [11]; at the workflow level, probably grid and HPC approaches are very competitive/the best as they are in monitoring area. We envisage growing Mahout to include the core algorithms used in scientific research and by HPC integration deliver libraries of the quality seen in Petsc and Scalapack for HPC simulations. This way we will also develop data intensive kernels/patterns that are the equivalent of the Berkeley dwarfs [12] and NAS parallel benchmarks [13].

We suggest as well as working within these different layers, it is interesting to identify “different cross-layer pathways” through the ABDS making particular choices at each level and see how the approaches compare and integrate. We are looking at end-to-end applications in Image processing, social media analysis, exploratory simulations and deep learning [14] as such pathways. Key questions include. How much commercial/commodity technology can be re-used by science big data and by exascale systems?

Can we develop abstractions that allow HPC-ABDS to be independent of hardware details? Can we build a consensus big data software stack on HPC-ABDS? Can we use libraries extending Mahout to capture mini-apps?



References

1. NIST. *Big Data Initiative*. 2013 Report at http://bigdatawg.nist.gov/V1_output_docs.php Available from: <http://bigdatawg.nist.gov/home.php>.
2. Geoffrey Fox, Tony Hey, and Anne Trefethen, *Where does all the data come from?*, Chapter in *Data Intensive Science* Terence Critchlow and Kerstin Kleese Van Dam, Editors. 2011. <http://grids.ucs.indiana.edu/ptliupages/publications/Where%20does%20all%20the%20data%20come%20from%20v7.pdf>.
3. IDC iView (sponsored by EMC). *Extracting Value from Chaos*. 2011 June [accessed 2013 August 14]; Available from: <http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf>.
4. Cisco. *Visual Networking Index: Forecast and Methodology, 2012–2017*. 2013 May 29 [accessed 2013 August 14]; Available from: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_White_Paper.html.
5. Apache. *Bigtop project for the development of packaging and tests of the Apache Hadoop ecosystem*. [accessed 2014 February 9]; Available from: <http://bigtop.apache.org/>.
6. Shantenu Jha, Judy Qiu, Andre Luckow, Pradeep Mantha, and Geoffrey C.Fox, *A Tale of Two Data-Intensive Approaches: Applications, Architectures and Infrastructure*. 2014. To be submitted for publication.
7. Enterprise Tech (Software Edition). *Hadoop Speed Up From Lustre Will Attract Enterprises*. 2013 [accessed 2014 February 9]; Available from: <http://www.enterprisetech.com/2013/10/02/hadoop-speed-up-from-lustre-will-attract-enterprises/>.
8. Xiaoming Gao and Judy Qiu, *Supporting End-to-End Social Media Data Analysis with the IndexedHBase Platform*, in *Invited talk at 6th Workshop on Many-Task Computing on Clouds, Grids, and Supercomputers (MTAGS) SC13*. November 17, 2013. Denver, Co. http://grids.ucs.indiana.edu/ptliupages/publications/mtags_v7.pdf.
9. *KBASE: DOE Systems Biology Knowledgebase: Community-Driven Cyberinfrastructure for Sharing and Integrating Data and Analytical Tools to Accelerate Predictive Biolog.* [accessed 2012 August 28]; Available from: <http://genomicscience.energy.gov/compbio/>.
10. Bingjing Zhang, Judy Qiu, Stefan Lee, and David Crandall, *Large-Scale Image Classification using High Performance Clustering*. November 18, 2013. http://grids.ucs.indiana.edu/ptliupages/publications/CCGRID2014_TLarge-Scale%20Image%20Classification%20using%20High%20Performance%20Clustering.pdf.
11. Hui Li, Geoffrey Fox, Gregor von Laszewski, and Arun Chauhan, *Co-processing SPMD computation on CPUs and GPUs cluster*, in *Cluster Computing (CLUSTER), 2013 IEEE International Conference on*. 23-27 Sept. 2013, 2013. pages. 1-10. DOI: 10.1109/CLUSTER.2013.6702632.
12. Krste Asanovic, Rastislav Bodik, James Demmel, Tony Keaveny, Kurt Keutzer, John Kubiawicz, Nelson Morgan, David Patterson, Koushik Sen, John Wawrzynek, David Wessel, and Katherine Yelick, *A view of the parallel computing landscape*. Commun. ACM, 2009. 52(10): p. 56-67. DOI:10.1145/1562764.1562783
13. Rob F. Van der Wijngaart, Srinivas Sridharan, and Victor W. Lee, *Extending the BT NAS parallel benchmark to exascale computing*, in *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*. 2012, IEEE Computer Society Press. Salt Lake City, Utah. pages. 1-9.
14. Adam Coates, Brody Huval, Tao Wang, David Wu, Bryan Catanzaro, and Andrew Ng. *Deep learning with COTS HPC systems*. in *Proceedings of the 30th International Conference on Machine Learning (ICML-13)* 2013.