White Paper: Big Data, Simulations and HPC Convergence

Geoffrey Fox, Judy Qiu, Shantenu Jha, Saliya Ekanayake, Supun Kamburugamuve
School of Informatics and Computing ,Indiana University, Bloomington, IN 47408, USA
RADICAL, Rutgers University, Piscataway, NJ 08854, USA

In previous BDEC meetings we described two concepts "The Big Data Ogres" [1] and "The HPC-ABDS Software Stack" [2] that here we bring together and extend to provide an approach to the convergence of Big Data and HPC (simulations). Our approach [3] suggests a hardware architecture described later and it is built around an application analysis (the Ogres) and software model (HPC-ABDS). We now describe these two building blocks.

Converge Applications: The Big Data Ogres [4,5] built on a collection of 51 big data uses gathered by the NIST Public Working Group where 26 properties were gathered for each application. This information was combined with other studies including the Berkeley dwarfs [6], the NAS parallel benchmarks [7,8] and the Computational Giants of the NRC Massive Data Analysis Report [9]. This led to a set of 50 features divided into four views that could be used to categorize and distinguish between applications. The four views are Problem Architecture (Macro pattern); Execution Features (Micro patterns); Data Source and Style; and finally the Processing View or runtime features.

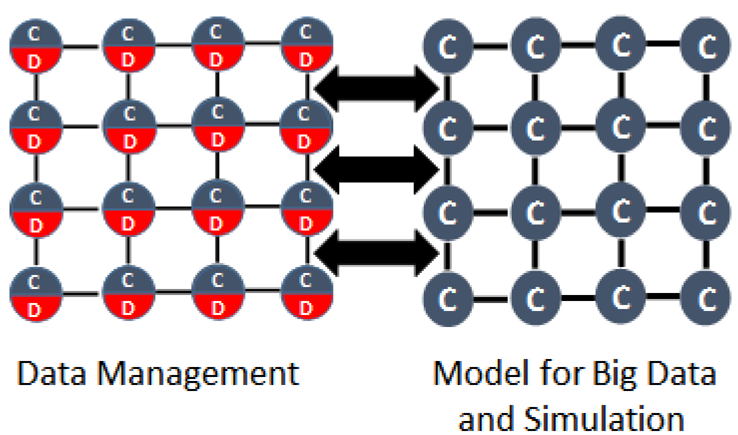
We generalized [3] this approach to integrate Big Data and Simulation applications into a single classification that we called convergence diamonds with the total facets growing to 64 in number and split between the same 4 views. A key idea was to separate for any application, data and model components which were merged together in the original Ogres. In Big Data problems, naturally the data size is large and this normal focus of work in this area. However, a model is essential to interpret data and this is of course the concern of the rapid advances in machine learning. Note the size of model can be much smaller than the data as in algorithms like clustering and dimension reduction. However, in applications like deep learning and topic modelling, the model can be huge. Parallelism has to be considered carefully both for data and model [10] and this leads to a new **convergence programming paradigm**. Turning to simulations where HPC has been most extensively explored, we again find that applications contain both data and model but typically it is now the model that is always large. For example, if one is solving particle dynamics or partial differential equations, then the model is the large numerical representation, while the data can be relatively small as in boundary conditions. On the other hand, there are examples like data assimilation for climate forecasting and data visualizations produced by simulations, where the data can be quite large.

We suggest that as long as one carefully compares apples with apples (e.g. big data model component with simulation model component), one can find many points of similarity between big data and simulations and see how one can support both with a common architecture that is separate in the handling of the different data and model components but NOT separated by the application type: Big Data and Simulation. These convergence diamonds are particularly useful in classifying benchmarks and we are pursuing this to design a "complete" (over the 64 facets) set of benchmarks which interestingly will link simulation and big data benchmark collections.

Software convergence builds on the HPC-ABDS High Performance Computing enhanced Apache Big Data Software Stack [2, 11], which arranges key HPC and ABDS software together in 21 layers showing where HPC and ABDS overlap. It for example, introduces a communication layer to allow ABDS runtime like Hadoop Storm Spark and Flink to use the richest high performance capabilities shared with MPI Generally it proposes how to use HPC and ABDS software together. Note that Apache is the largest but not only source of open source software.

The essential idea of our Big Data HPC convergence for software is to make use of ABDS software where possible as it offers richness in functionality, a compelling open-source community sustainability model and typically attractive user interfaces. Although ABDS has a good reputation for scaling, it often does not give good performance. To address this, we suggest augmenting ABDS with HPC ideas and we have illustrated this with Hadoop [12, 13], Storm (and its successor Heron) [14] Yarn/Mesos [15] and the basic Java environment [16]. We suggest using the resultant HPC-ABDS for both big data and big simulation applications. As one example we recommend using HPC enhanced MapReduce (Hadoop, Spark, Flink) for parallel programming for both simulations and the model (data analytics) of big data. .A byproduct of these studies is that classic HPC clusters make excellent data analytics engine.

Language Convergence requires one to understand how to make Java run as fast as C++ (Fortran) for computing and communication; the old Java Grande concept. It is surprising that so much Big Data work in industry uses the Java virtual machine but basic high performance Java methodology and tools are still incomplete. For example, Java OpenMP (threads) and OpenMPI (processes) need work [16]. The convergence Language Grande should support Python, Java (Scala), and C/C++ (Fortran)



Summarizing we recommend distinguishing data and model components but not to distinguish Big Data and Big Simulation applications -- each of which has data and model components. Our arguments suggest a system architecture sketched in figure, that supports the two components in any application: data management on the left side and model computation on the right. We recommend using HPC-ABDS software with basic principle of using the best practice ABDS software in BOTH components but adding where needed (as it nearly always is) HPC enhancements to improve performance.

Acknowledgements
This work was partially supported by NSF CIF21 DIBBS 1443054, NSF XPS 1439007 and NSF OCI 1149432 CAREER. awards.

References

1) Geoffrey Fox, Shantenu Jha, Judy Qiu, Andre Luckow; “Ogres: A Systematic Approach to Big Data Benchmarks”, White paper at workshop on Big Data and Extreme-scale Computing (BDEC) Barcelona, January 29 to January 30, 2015, Barceló Sants Hotel http://www.exascale.org/bdec/sites/www.exascale.org.bdec/files/whitepapers/OgreFacets.pdf
2) Fox, G., Qiu, J., Jha, S.: “High Performance High Functionality Big Data Software Stack. Big Data and Extreme-scale Computing” BDEC Fukuoka, Japan February 26-28, 2014 http://www.exascale.org/bdec/sites/www.exascale.org.bdec/files/whitepapers/fox.pdf
3) Geoffrey Fox, Judy Qiu, Shantenu Jha, Saliya Ekanayake, and Supun Kamburugamuve, “Big Data, Simulations and HPC Convergence”, January 30 2016. DOI: 10.13140/RG.2.1.1858.8566
4) Fox, G., Jha, S., Qiu, J., Ekanayake, S., Luckow, A., “Towards a Comprehensive Set of Big Data Benchmarks”. Chapter in Big Data and High Performance Computing, Lucio Grandinetti and Gerhard Joubert, Editors. 2015, http://grids.ucs.indiana.edu/ptliupages/publications/OgreFacetsv9.pdf
5) Fox, G.C., Jha, S., Qiu, J., Luckow, A., “Towards an Understanding of Facets and Exemplars of Big Data Applications”, in 20 Years of Beowulf: Workshop to Honor Thomas Sterling’s 65th Birthday October 14, 2014. Annapolis http://dx.doi.org/10.1145/2737909.2737912
6) Asanovic, K., Bodik, R., Catanzaro, B.C., Gebis, J.J., Husbands, P., Keutzer, K., Patterson, D.A., Plishker, W.L., Shalf, J., Williams, S.W., et al., “The landscape of parallel computing research: A view from Berkeley. Tech. rep., UCB/EECS-2006-183, EECS Department, University of California, Berkeley (2006), Available from http://www.eecs.berkeley.edu/Pubs/TechRpts/2006/EECS-2006-183.html
7) Bailey, D.H., Barszcz, E., Barton, J.T., Browning, D.S., Carter, R.L., Dagum, L., Fatoohi, R.A., Frederickson, P.O., Lasinski, T.A., Schreiber, R.S., et al., “The NAS parallel benchmarks.” International Journal of High Performance Computing Applications 5(3), 63–73 (1991)
8) Van der Wijngaart, R.F., Sridharan, S., Lee, V.W.: “Extending the BT NAS parallel benchmark to exascale computing”. International Conference on High Performance Computing, Networking, Storage and Analysis. p. 94. IEEE Computer Society Press (2012)
9) National Research Council: “Frontiers in Massive Data Analysis”. The National Academies Press, Washington, DC (2013)
10) Bingjing Zhang, Peng Bo, Judy Qiu, "Model Data-Centric Computation Abstractions in Machine Learning Applications", 3rd Workshop on Algorithms and Systems for MapReduce and Beyond (BeyondMR2016), held in conjunction with SIGMOD/PODS2016, July 1, 2016
11) Jha, S., Qiu, J., Luckow, A., Mantha, P., Fox, G.C. “A tale of two data-intensive paradigms: Applications, abstractions, and architectures”. 2014 IEEE International Congress on Big Data (BigData Congress). pp. 645–652. IEEE (2014)
12) Bingjing Zhang, Bo Peng, Judy Qiu, “High Performance LDA through Collective Model Communication Optimization”, International Conference on Computational Science (ICCS2016) conference, June 6-8, 2016, San Diego, California.
13) Zhang, B., Ruan, Y., Qiu, J.: “Harp: Collective communication on Hadoop.” IEEE International Conference on Cloud Engineering (IC2E) conference (2014)
14) Supun Kamburugamuve, Saliya Ekanayake, Milinda Pathirage, Geoffrey Fox, “Towards High Performance Processing of Streaming Data in Large Data Centers” HPBDC 2016 IEEE International Workshop on High-Performance Big Data Computing in conjunction with The 30th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2016), Chicago, Illinois USA, Friday, May 27th, 2016
15) Andre Luckow, Ioannis Paraskevakos, George Chantzialexiou, Shantenu Jha, ”Hadoop on HPC: Integrating Hadoop and Pilot-based Dynamic Resource Management” HPBDC 2016 IEEE International Workshop on High-Performance Big Data Computing in conjunction with The 30th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2016), Chicago, Illinois USA, Friday, May 27th, 2016
16) Saliya Ekanayake, Supun Kamburugamuve and Geoffrey Fox, “SPIDAL: High Performance Data Analytics with Java and MPI on Large Multicore HPC Clusters”, Proceedings of 24th High Performance Computing Symposium (HPC 2016), April 3-6, 2016, Pasadena, CA, USA as part of the SCS Spring Simulation Multi-Conference (SpringSim'16)