Large Scale Data Analytics on Clouds

Geoffrey Fox
School of Informatics and Computing
Indiana University
Bloomington IN 47408, USA
gcf@.indiana.edu

ABSTRACT

We summarize important overall issues affecting use of clouds to support Data Science. We describe the mapping of different applications to HPCC and Cloud systems and the architecture that support data analytics that is interoperable between these architectures.

Categories and Subject Descriptors

D.1.3 [Software Programming Techniques]: Concurrent Programming: Distributed programming; Parallel programming

General Terms

Performance, Design, Experimentation

Keywords

Clouds, Exascale, MapReduce, Iterative MapReduce, MPI, Data Science, HPCC, Programming Paradigms

1. CLOUDS+EXASCALE ECOSYSTEM

There are several important trend driving computing. We have the **Data Deluge** from Commercial (e.g. Amazon, e-commerce), Community (e.g. Facebook, Search), and Scientific applications (e.g. Analysis of LHC data, Genomics) with examples given just being representative of many others[1]. We have **light weight clients** from smartphones, tablets to sensors. The **multicore chip** architecture is reawakening parallel computing while it and GPGPU's (even more cores) are behind **Exascale** initiatives, which will continue drive to high end with a simulation orientation. **Clouds** with cheaper, greener, easier to use IT for (some) applications are growing in importance. They enable the lightweight clients by acting as a backend resource and answer the difficult question "what do we do all with all those cores on a chip". As that's not so easy to answer on a conventional client, this is one driver to lighter weight client (using smaller CPU chips) but on a server, each core can host a separate cloud service. These developments drive both research and education and will weave together as we look at data analysis in the clouds. **Curricula** based on the "Science of Clouds"[2, 3] and/or "Data Science" [4] are attractive as both area are predicted to generate several million jobs and not find the needed skills. Finally the need for data analytics links old (e.g. finance, retail) business and new (Web 2.0) business with science.

CloudDB'12, October 29, 2012, Maui, Hawaii, USA.

Clouds have many interesting characteristics including on-demand service, measured service, scalable elastic service, broad any-time any-where network access, pooling of resources leading to economies of scale in performance and electrical power (Green IT). These correspond to Infrastructure as a Service but there are also powerful new software models corresponding to Platform as a Service that are also important. We will see examples such as cloud support of sensors (lightweight clients) where IaaS with broad access drives cloud data analysis and others where novel MapReduce algorithms (i.e. PaaS) are most important. Areas like genomics are driven both by the need for the most effective computing combined with interest in new programming models like MapReduce[5]. The most visible and major data intensive area – analysis of LHC data from CERN – could use clouds (as can typical high throughput computing loads) but they have an effective operational grid solution.

Simulations have been explored on clouds but traditional super computers are typically required to get good performance on large highly parallel jobs. Clouds are currently only clearly get good performance on "bags of simulation tasks" with many small jobs that are not individually sensitive to synchronization costs. Synchronization costs are higher in clouds as virtualization leads to overheads both from software costs and difficulties in preserving locality. Thus we get classic HPC systems now moving inevitably to Exascale as likely to remain a critical part of the computing Cyberinfrastructure.

The above analysis suggests a "Clouds+Exascale" Cyberinfrastructure scenario and in next section we ask how data intensive applications map into this ecosystem.

2. EXAMPLE APPLICATIONS

Previously we have used the MapReduce paradigm to classify parallel applications into four major groups [6-9]..

Map-only applications are bags of independent tasks and clearly are suitable for clouds. This pleasingly parallel case includes not only LHC and similar science analysis but also support of the "Internet of Things" (IoT) [10] where each of the world's distributed devices (including smart phones) is backended by the cloud. The IoT is forecast to grow to 24 billion devices on the Internet by 2020. Robots are important sub-class of the IoT and cloud-backed robotics is very promising. The map-only case included "the long tail of science" (or indeed the "the long tail of most things") where one has parallelism over users each running smallish jobs that run effectively on clouds.

MapReduce jobs consist of independent maps and reducers with communication between tasks happening at the link between Map and Reduce. These of course cover many "Life-style Informatics" applications such as those used in the social media and search industries. Clouds can support this problem class well. There are some scientific applications of this class including for example basic statistical analysis (such as histogramming) common for example at final stage of LHC analysis.

Classic MPI jobs are those identified for supercomputers above and typically involve many small size point to point messages. This class is target of HPC systems and the domain of "Exascale" component of the computing ecosystem.

The final category has been called **Iterative MapReduce** [11-16] and is very clear in many data analysis applications. Many data analytics algorithm involve linear algebra at their core where the parallelism is well understood. These do not have the geometric parallelism of simulations but rather that of matrix rows, columns or blocks. Correspondingly we do not get many small messages but large reduction of broadcast (multicast) messages that are not as sensitive to latency overheads that are important for MPI structure of particle dynamics or partial differential equation solvers. Thus clouds are an interesting architecture and one can introduce a "Map Collective" programming abstraction that can be supported by either MPI or iterative versions of MapReduce.

Supporting the three categories suitable of clouds has important issues including especially the data architecture where one needs to move the computing to the data which is typically not easy in today's HPC or cloud environments. We discussed this in a previous note. In the last section we discuss a missing component that must be addressed.

3. DATA ANALYTICS LIBRARY

data analysis will need new robust algorithms to mimic the oft-quoted observation that HPC progress has benefited equally from Moore's Law-driven hardware improvements and from new algorithms. These observations motivate the introduction of SPIDAL, or the Scalable Parallel Interoperable Data Analytics Library, to address the analysis of big data. ***Figure 1*** shows the components of the project. We include communities with data intensive applications which need to identify what library members need to be built. Good existing examples are R [18] and Mahout [19] but these are not aimed at high performance needed for large scale applications. As shown in Figure 1, we identify six layers and also five broad abstraction areas [20] whose definition allows library members to be built in a way that is portable. One abstraction is **Jobs** where we can identify the Pilot job concept [21, 22] to obtain interoperably; **Communication** where we need both MPI and MapReduce patterns and will use iterative MapReduce to design a common abstraction; a **Data Layer** where one needs abstractions to support storage, access and transport (since SPIDAL algorithms will need to run interoperably with databases, NOSQL, wide area file systems and file systems like Hadoop's HDFS[23]). One also needs an **Application Level Data** abstraction between L2 and L3. Our final abstraction is the virtual machine or **Appliance** to deploy applications, where one could use a recently developed template approach [24-30] that can be realized on bare metal or commercial and private cloud VM managers. This supports both interoperability between different resources and preservation so that scientific results using SPIDAL will be reproducible.

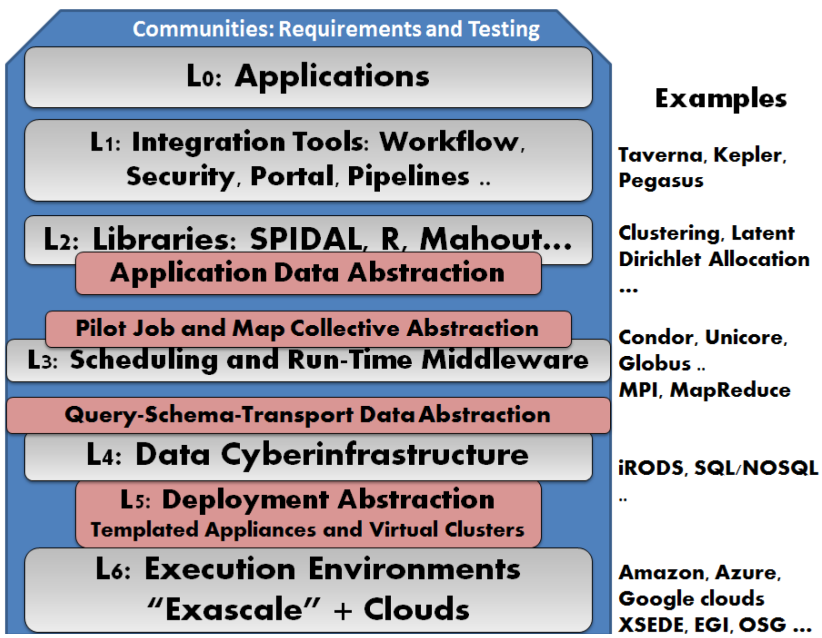

Figure 1: A Data Analytics Architecture with abstractions

Here we note that in the hugely successful but largely simulation-oriented HPCC activities starting around 1990, an important activity was the design and construction of core libraries such as PETSc, SCALAPACK (becoming PLASMA now [17]) and underlying technologies such as BLACS and MPI. Data intensive cloud applications require scalable parallel analysis routines and that these will cross many application areas just as the earlier HPCC libraries enable differential equation solvers and linear algebra across many disciplines. We further expect that reliable

4. ACKNOWLEDGMENTS

I would like to acknowledge Shantenu Jha, Madhav Marathe, Judy Qiu and Joel Saltz for discussions of SPIDAL architecture. This material is based upon work supported in part by the National Science Foundation under Grant No. 0910812 for "FutureGrid: An Experimental, High-Performance Grid Test-bed."

5. REFERENCES

1. 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.
2. IDC. *Cloud Computing's Role in Job Creation*. 2012 [accessed 2012 March 6]; Sponsored by Microsoft Available from: http://www.microsoft.com/presspass/download/features/2012/IDC_Cloud_jobs_White_Paper.pdf.
3. *Cloud Computing to Bring 2.4 Million New Jobs in Europe by 2015*. 2011 [accessed 2011 March 6]; Available from: http://www.eweek.com/c/a/Cloud-Computing/Cloud-Computing-to-Bring-24-Million-New-Jobs-in-Europe-by-2015-108084/.
4. James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and A.H. Byers. *Big data: The next frontier for innovation, competition, and productivity*. 2011 [accessed 2012 August 23]; McKinsey Global Institute Available from: http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation.
5. Jeffrey Dean and Sanjay Ghemawat, *MapReduce: simplified data processing on large clusters.* Commun. ACM, 2008. 51(1): p. 107-113. DOI:http://doi.acm.org/10.1145/1327452.1327492
6. Fox, G.C., R.D. Williams, and P.C. Messina, *Parallel computing works!* 1994: Morgan Kaufmann Publishers, Inc. http://www.old-npac.org/copywrite/pcw/node278.html#SECTION001440000000000000000
7. Geoffrey C. Fox, *Data intensive applications on clouds*, in *Proceedings of the second international workshop on Data intensive computing in the clouds*. 2011, ACM. Seattle, Washington, USA. pages. 1-2. DOI: 10.1145/2087522.2087524.
8. Jaliya Ekanayake, Thilina Gunarathne, Judy Qiu, Geoffrey Fox, Scott Beason, Jong Youl Choi, Yang Ruan, Seung-Hee Bae, and Hui Li, *Applicability of DryadLINQ to Scientific Applications*. January 30, 2010, Community Grids Laboratory, Indiana University. http://grids.ucs.indiana.edu/ptliupages/publications/DryadReport.pdf.
9. Judy Qiu, Jaliya Ekanayake, Thilina Gunarathne, Jong Youl Choi, Seung-Hee Bae, Yang Ruan, Saliya Ekanayake, Stephen Wu, Scott Beason, Geoffrey Fox, Mina Rho, and H. Tang, *Data Intensive Computing for Bioinformatics*. December 29, 2009. http://grids.ucs.indiana.edu/ptliupages/publications/DataIntensiveComputing_BookChapter.pdf.
10. Kai Hwang, Geoffrey Fox, and Jack Dongarra, *Distributed and Cloud Computing : from Parallel Processing to The Internet of Things*. 2011: Morgan Kaufmann Publishers
11. Thilina Gunarathne, Bingjing Zhang, Tak-Lon Wu, and Judy Qiu, *Scalable Parallel Computing on Clouds Using Twister4Azure Iterative MapReduce* Future Generation Computer Systems 2012. To be published. http://grids.ucs.indiana.edu/ptliupages/publications/Scalable_Parallel_Computing_on_Clouds_Using_Twister4Azure_Iterative_MapReduce_cr_submit.pdf
12. Judy Qiu, Thilina Gunarathne, and Geoffrey Fox, *Classical and Iterative MapReduce on Azure*, in *Cloud Futures 2011 workshop*. June 2-3, 2011. Microsoft Conference Center Building 33 Redmond, Washington United States. http://grids.ucs.indiana.edu/ptliupages/presentations/Twister4azure_June2-2011.pptx.
13. Yingyi Bu, Bill Howe, Magdalena Balazinska, and Michael D. Ernst, *HaLoop: Efficient Iterative Data Processing on Large Clusters*, in *The 36th International Conference on Very Large Data Bases*. September 13-17, 2010, VLDB Endowment: Vol. 3. Singapore. http://www.ics.uci.edu/~yingyib/papers/HaLoop_camera_ready.pdf.
14. SALSA Group. *Iterative MapReduce*. 2010 [accessed 2010 November 7]; Twister Home Page Available from: http://www.iterativemapreduce.org/.
15. J.Ekanayake, H.Li, B.Zhang, T.Gunarathne, S.Bae, J.Qiu, and G.Fox, *Twister: A Runtime for iterative MapReduce*, in *Proceedings of the First International Workshop on MapReduce and its Applications of ACM HPDC 2010 conference June 20-25, 2010*. 2010, ACM. Chicago, Illinois. http://grids.ucs.indiana.edu/ptliupages/publications/hpdc-camera-ready-submission.pdf.
16. Matei Zaharia, Mosharaf Chowdhury, Michael J. Franklin, Scott Shenker, and Ion Stoica, *Spark: Cluster Computing with Working Sets*, in *2nd USENIX Workshop on Hot Topics in Cloud Computing (HotCloud '10)*. June 22, 2010. Boston. http://www.cs.berkeley.edu/~franklin/Papers/hotcloud.pdf.
17. *Parallel Linear Algebra for Scalable Multi-core Architectures (PLASMA) project*. [accessed 2012 September 6]; Available from: http://icl.cs.utk.edu/plasma/index.html.
18. *The Comprehensive R Archive Network*. [accessed 2012 August 22]; Available from: http://cran.r-project.org/.
19. *Apache Mahout Scalable machine learning and data mining* [accessed 2012 August 22]; Available from: http://mahout.apache.org/.
20. Shantenu Jha, Murray Cole, Daniel S. Katz, Manish Parashar, Omer Rana, and J. Weissman, *Distributed Computing Practice for Large-Scale Science & Engineering Applications* Concurrency and Computation: Practice and Experience (in press), 2012.
21. Andre Luckow, Mark Santcroos, Ole Weidner, Andre Merzky, Pradeep Mantha, and Shantenu Jha, *P*: A Model of Pilot-Abstractions*, in *8th IEEE International Conference on e-Science*. 2012.
22. Pradeep Kumar Mantha, Andre Luckow, and S. Jha, *Pilot-MapReduce: an extensible and flexible MapReduce implementation for distributed data*, in *Third international workshop on MapReduce and its Applications*. 2012.
23. Apache. *HDFS Overview*. 2010 [accessed 2010 November 6]; Available from: http://hadoop.apache.org/hdfs/.

24. Jonathan Klinginsmith, M. Mahoui, and Y. M. Wu, *Towards Reproducible eScience in the Cloud.*, in *Third International Conference on Cloud Computing Technology and Science (CloudCom)*. November 29 - December 1, 2011. DOI: 10.1109/CloudCom.2011.89.
25. Jonathan Klinginsmith and Judy Qiu, *Using Cloud Computing for Scalable, Reproducible Experimentation*. August, 2012.
26. Gregor von Laszewski, Hyungro Lee, Javier Diaz, Fugang Wang, Koji Tanaka, Shubhada Karavinkoppa, Geoffrey C. Fox, and Tom Furlani, *Design of an Accounting and Metric-based Cloud-shifting and Cloud-seeding framework for Federated Clouds and Bare-metal Environments*, in *Workshop on Cloud Services, Federation, and the 8th Open Cirrus Summit*. September 21, 2012. San Jose, CA (USA). http://grids.ucs.indiana.edu/ptliupages/publications/p25-vonLaszewski.pdf.
27. Geoffrey C. Fox, Gregor von Laszewski, Javier Diaz, Kate Keahey, Jose Fortes, Renato Figueiredo, Shava Smallen, Warren Smith, and Andrew Grimshaw, *FutureGrid - a reconfigurable testbed for Cloud, HPC and Grid Computing*, Chapter in *On the Road to Exascale Computing: Contemporary Architectures in High Performance Computing*, Jeff Vetter, Editor. 2012, Chapman & Hall/CRC Press http://grids.ucs.indiana.edu/ptliupages/publications/sitka-chapter.pdf
28. Javier Diaz, Gregor von Laszewski, Fugang Wang, and Geoffrey Fox, *Abstract Image Management and Universal Image Registration for Cloud and HPC Infrastructures*, in *IEEE CLOUD 2012 5th International Conference on Cloud Computing* June 24-29 2012. Hyatt Regency Waikiki Resort and Spa, Honolulu, Hawaii, USA http://grids.ucs.indiana.edu/ptliupages/publications/jdiaz-IEEECloud2012_id-4656.pdf
29. J. Diaz, A. J. Younge, G. von Laszewski, F. Wang, and G. C. Fox, *Grappling cloud infrastructure services with a generic image repository*, in *CCA11: Cloud Computing and Its Applications*. April 12-13, 2011. Argonne National Laboratory, USA. http://grids.ucs.indiana.edu/ptliupages/publications/11-imagerepo-cca.pdf.
30. Javier Diaz, Gregor von Laszewski, Fugang Wang, Andrew J. Younge, and Geoffrey Fox, *FutureGrid Image Repository: A Generic Catalog and Storage System for Heterogeneous Virtual Machine Images*, in *3rd IEEE International Conference CloudCom on Cloud Computing Technology and Science*. November 29 - December 1 2011. Athens Greece. http://grids.ucs.indiana.edu/ptliupages/publications/jdiazCloudCom2011.pdf