Distributed Data and Software Defined Systems

Geoffrey Fox, Indiana University

Recently several tools have emerged to dynamically build computing environments at the level of individual nodes. These environments can be based on Bare Metal or Virtual machines and the software build includes IaaS (Infrastructure as a Service e.g. operating system) or that plus PaaS (Platform as a Service e.g. MPI and data libraries) and SaaS (Software as a Service e.g. full data analytics). These tools include templated image libraries [1], authentication and authorization, accounting and metrics [2], user interfaces (dashboards like Nimbus Phantom or OpenStack Horizon), DevOps (e.g. Salt, Chef, Puppet), dynamic provisioning [2] and higher level tools at scheduling level and above. These can be combined with tools like Rocks to manage clusters and support software defined (virtual) clusters. Further at the NaaS (Network as a Service) layer [3], technologies like OpenFlow and projects like GENI support software defined networks that allow higher performance interconnections; combining these ideas, we find software defined systems. The GENI Network Stitching Architecture would perhaps be used [4]. As well as building general distributed compute grids, we consider here the special case of a software defined system that includes localized computing (e.g. a cluster, cloud or even exascale machine) linked using software defined networks to distributed data with software defined networks.

Consider the architecture of a data repository which often in the past focused on storage and access to the data. However many researchers now need systems that manage both the (big) data and provide computing (data analytics) on the data. Here we are often told to bring the computing to the data to avoid overheads of data transport. Indeed recent commercial systems such as those at Google, Facebook, Amazon, Twitter are architected as clouds supporting analytics with co-located storage and satisfy the principle of bringing computing to the data. However, this is not universally applicable for the cyberinfrastructure supporting data-intensive science, which features a diversity of distributed data analyzed by a distributed research team. We have good examples in the LHC data analysis system with data distributed and analyzed across the world. The primary data analysis of LHC uses compute-data affinity in a grid to produce event parameters, but also needs to access data across the globe, making it more efficient (as only parts of distributed files are needed) than transferring files. However, this model is not obviously sufficient in other disciplines where one traditionally uses repositories like GENBank (Biology), NSIDC (Polar data) and EOSDIS (Earth Satellite data), which do not always have enough attached computing to support science analysis. We can imagine adding a cluster (cloud) in front of each discipline repository but how do we determine the right size and scalability for it? How can it be elastic if it is single use? Further suppose we want to do environmental genomics; we need genomics as well as environmental data, which are typically not in same repository as they are gathered by multidisciplinary studies. One solution would be to locate all data next to the same giant compute environment – Amazon, Azure or Google or an exascale supercomputer. I expect this to be used in some fields but it does not seem to be a general solution. It’s hard to get all data co-located!

Note that in genomics, data is now gathered by a multitude of distributed low cost “individual” sequencers such as the Illumina MiSeq which has an instrument cost of ~$100K and can produce many gigabytes of data a day. This analysis highlights a scenario different from the single large data source like the LHC, where data is stored in multiple places across continents. However in many analyses one needs this distributed data to be presented to the analytics as a co-located data system perhaps set up as virtual Hadoop file system.

I suggest support of a suite of experiments that compare this streaming virtualized data model with the traditional model where files are copied from distant to local storage. One should consider multiple data architectures including databases, HDFS (data parallel storage with addons like Hbase), Lustre (wide area file systems) and the object stores exemplified by Amazon S3 and OpenStack Swift. The analytics needed include some pleasing parallel as in sensor data analysis; some with weak coupling as in Genomic sequence comparison as with Blast; others with interactive browsing as with Geographic Information Systems in Earth Science and closely coupled clustering seen in Genomics and Network Science. These analytics span performance issues including I/O dominated and communication intensive applications. Other follow-on projects could include software defined networks supporting (cloud) bursting to cope with temporary overloads.

Acknowledgements

These ideas have been developed with Ilia Baldine, Jose Fortes, Gregor von Laszewski, and Martin Swany

References

1. 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
2. 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.
3. E. Kissel and M. Swany, *Evaluating High Performance Data Transfer with RDMA-Based Protocols in Wide-Area Networks*, in *14th International Conference on High Performance Computing and Communications (HPCC-2012)*. June 25-27, 2012. Liverpool UK. http://damsl.cs.indiana.edu/projects/phoebus/sdn_xsp.pdf.
4. *GENI Network Stitching Architecture - Example*. [accessed 2013 April 21]; Available from: http://geni.maxgigapop.net/twiki/pub/GENI/NetworkStitching/geni-network-stitching-architecture-example.pdf.