Domain Science At-Scale

Submitted to the "NSFCloud Workshop on Experimental Support for Cloud Computing"

Paul Ruth, Anirban Mandal, Ilya Baldin
RENCI/UNC Chapel Hill

Today's science is completely reliant on the cyberinfrastructure to carry on its functions - from traditional enterprise applications to sophisticated distributed systems that tie together heterogeneous cyber resources in support of computationally-intensive research or data analytics. The demands of modern scientific collaborations drive many changes to the way the infrastructure is constructed, attempting to extract maximum utility from limited infrastructure investments, while addressing the needs of individual science domains. The introduction of cloud technologies and cloud providers, like Amazon, Microsoft and Google showed one approach to gain efficiencies by virtualizing most of the cyber resources and shifting parts of the infrastructure into the public cloud.

At the same time, in addition to the traditional institutional and grid resources, the emergence of new software technologies (OpenStack, VMWare, Eucalyptus, Hadoop, Mesos, xCAT and so on) allowed some organizations to create and to take advantage of a private cloud infrastructure geared towards internal needs. They however lack the scale of the public clouds and do not permit the flexible merging of heterogeneous infrastructures belonging to different organizations for the purposes of collaborative research activities. They are vertically integrated, allowing interoperation only with alike resources. They also lack integration with transit provider networking technologies to support frictionless data movement between elements of these clouds - a critical requirement of modern collaborative science.

As part of GENI we have constructed ExoGENI testbed [2] that represents a new kind of a federated cloud infrastructure with a deep integration of networking functions. ExoGENI, in addition to traditional GENI needs, focuses on running domain science applications on a widely distributed federated cloud infrastructure [3,4,5]. Thus far, GENI has focused on designing and building a suite of infrastructure to support research in network science and engineering. In order to truly transform next generation internets, GENI must consider the requirements of the users of next generation internets. One class of user that has applications that can utilize current GENI resources "at-scale" are domain scientists who currently use existing high-throughput and high-performance computing facilities. These scientists have experience scaling applications well beyond the current capabilities of GENI. Leveraging this experience can inform future directions of GENI. Further, these scientists are quickly running into the "big-data" problem, a solution to which will certainly include next-generation "big-network" technologies.

One example application that we have deployed on ExoGENI is the ADCIRC storm surge model [1] for the North Carolina Forecast System (NCFS) [6,7]. ADCIRC is a tightly coupled finite element model that is known to scale to 10000+ MPI processes on TACC's Ranger and relies on its large number of compute nodes and specialized networks, such as InfiniBand, to achieve very low latencies and high bandwidths. This scale far exceeds the capability of GENI limiting our ability to provide real-time data during a weather event.

One of the main challenges in forecasting hurricane events is the need for massive amounts of dedicated compute resources for the duration of a weather event. This requires the coordinated orchestration of data movement into and out of slices (2-30 GB per ensemble member). Ensembles computing high-resolution grids are currently limited by the ability or inability to acquire sufficient compute resources. We will experiment with deploying slices of resources that can satisfy the real-time constraints to compute 20-100 ensemble members for a given hurricane forecast advisory. Further, the purpose of these simulations is often to provide government agencies with real-time decision support information. The workload needs to be mobile and utilize resources that are themselves not at risk from the weather event that is being modeled.

The developing NSFCloud systems present a much larger pool of resources compared to what is available in GENI and are thus complimentary to the existing GENI infrastructure and better answer the needs of large science applications. The resources available in NSFCloud can help scale our experiments with ADCIRC and other scientific codes to help tune these HPC applications to the dynamic cloud environment, and, conversely, help shape the evolution of the infrastructure to support the future needs of these applications.

In order to support our experiments we will expect the following features from NSF Cloud infrastructure:

- Fast internal interconnects to support data-intensive HPC workloads
- Fast dynamically-allocatable storage to allow for movement of data in and out of NSF Cloud infrastructure
- Large scale of available resources (10K cores) across sites, available on short notice to support bursty demands of the exemplar applications
- Dynamically programmable WAN connections to emerging programmable high-speed networks provided by Internet2 (AL2S) and ESnet to allow for moving data between NSF Cloud sites and other resources
- Open programmable interfaces to compute, storage and network resources that applications and workflows can invoke in order to provision their infrastructure
- Programmable linkage to existing GENI resources. We will want to allow applications to transition their executions from a widely distributed cloud environment like ExoGENI into more concentrated pools of resources like NSF Cloud and vice versa, depending on their needs.

References

[1] Ruth, P.; Mandal, A.; (2014), '*Toward Evaluating GENI for Domain Science Applications',* in 'International Workshop on Computer and Networking Experimental Research using Testbeds (CNERT 2014)', Oct 2014, to appear.

[2] Baldine, I.; Xin, Y.; Mandal, A.; Ruth, P.; Heerman, C. & Chase, J. (2012), *ExoGENI: A multi-domain infrastructure-as-a-service testbed* 'Testbeds and Research Infrastructure. Development of Networks and Communities', Springer Berlin Heidelberg, pp. 97--113.

[3] Mandal, A.; Ruth, P.; Baldin, I.; Xin, Y.; Castillo, C.; Rynge, M. & Deelman, E. (2014), *Leveraging and Adapting ExoGENI Infrastructure for Data-Driven Domain Science Workflows*, *in* 'Research and Educational Experiment Workshop (GREE), 2014 Third GENI', pp. 57--60.

[4] Mandal, A.; Ruth, P.; Baldin, I.; Xin, Y.; Castillo, C.; Rynge, M. & Deelman, E. (2013), *Evaluating I/O aware network management for scientific workflows on networked clouds*, *in* 'Proceedings of the Third International Workshop on Network-Aware Data Management', pp. 2.

[5] Ruth, P.; Mandal, A.; Xin, Y.; Baldine, I.; Heerman, C. & Chase, J. (2012), *Dynamic network provisioning for data intensive applications in the cloud*, *in* 'E-Science (e-Science), 2012 IEEE 8th International Conference on', pp. 1--2.

[6] Dietrich, J.; Zijlema M.; Westerink, J.; Holthuijsen, L.; Dawson, C.; R. L. Jr., R. Jensen.; Smith, J.; Stelling, G.; and Stone G. (2011), *'Modeling hurricane waves and storm surge using integrally-coupled, scalable computations,'* Coastal Engineering, vol. 58, no. 1, pp. 45 – 65, 2011.

[7] Atkinson, J.; Wamsley, T.; Westerink, J.; Cialone, M.; Dietrich, C.; Dresback, K.; Kolar, R.; Resio, D.; Bender, C.; Blanton, B. et al., *'Hurricane storm surge and wave modeling in southern Louisiana: A brief overview,'* in Estuarine and Coastal Modeling X, M. Spaulding, Ed. ASCE, 2008.