Amazon S3 for Science Grids: a Viable Solution?

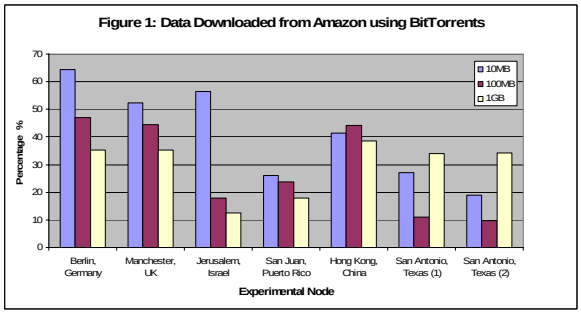
Mayur Palankar[1] Ayodele Onibokun[1] Adriana Iamnitchi[1] Matei Ripeanu[2]
[1]University of South Florida
{mpalanka, aoniboku, anda}@cse.usf.edu
[2]University of British Columbia
matei@ece.ubc.edu

The volume of data produced and shared by data-intensive scientific collaborations is rapidly growing. For example, modern high-energy physics experiments, such as DZero, typically acquire more than one TB of data per day and may move up to ten times as much [1]. As a result, significant resources, both human and equipment, are allocated to support the data-intensive operations of these communities leading to high storage and management costs. Recently, amazon.com has introduced the *Simple Storage Service* (S3), a novel storage utility. S3 aims to provide data storage as a service in a cheap, highly available, reliable way, on a pay-as-you-go basis. Additionally, S3 offers tools for developers to integrate S3 with data-intensive applications from personal desktop remote backup solutions to multi-tiered Internet applications that may use S3 to support their data-storage tier. With tens of thousands of computer systems around the world [2], S3 aims to provide a highly-scalable data infrastructure that has low-latency, is reliable, and provides 99.99% data availability and low server-side latency [3].

These characteristics make S3 a good candidate to offload storage support for data-intensive scientific collaborations. We aim to evaluate whether offloading data storage from in-house maintained mass storage hardware to S3 is a feasible and cost-effective alternative for today's scientific collaborations like DZero [4], CERN, and SLAC. To this end we have characterized S3' availability and client observed download/upload performance using a collection of our own nodes and nodes located in the PlanetLab experimental platform [5]. We use this characterization in conjunction with real traces from scientific communities: more than two years of traces from the DZero Experiment, a high-energy physics collaboration that spans 18 countries and has more than 500 active users. Our objective is to estimate the feasibility, observed performance, and costs of a hypothetical S3-supported DZero collaboration.

We find that this is possible in terms of feasibility and performance but the cost of using S3 would be relatively high. Costs can, however, be reduced by using BitTorrent [6] along with S3. We have verified this assumption by deploying BitTorrent on a set of PlanetLab nodes (Figure 1). Additionally, we propose design guidelines that exploit data usage patterns to improve performance, and, more importantly, to minimize the data access costs of accessing S3-stored data by introducing user-managed collaborative caching in the system. In effect, our proposed design is driven by S3 billing structure: we use S3 for the costly tasks of providing data availability and durability (for these two costs are driven up by specialized hardware and nontrivial engineering effort) and employ caching at the edges of the system to reduce the access volume when the usage patterns allow. It turns out that the design not only reduces our S3 bill but also significantly improves performance due to a cacheable workload specific to these collaborations.

Finally, motivated by our scientific collaboration scenario, we identify requirements that are not satisfied by the current S3 design. While S3 successfully supports relatively simple scenarios (e.g., personal data backup) or it can be easily integrated in the storage tier of a multi-tiered Web application (e.g., in Microsoft Software Development Network to serve its Direct Student downloads), we found that the existing S3 security functionality is strikingly inadequate to support complex, collaborative environments like the ones in today's scientific collaborations. More precisely, S3 lacks in terms of access control, support for delegation and auditing, and makes implicit trust assumptions between S3 and its clients. This lack of functionality is even more troubling when direct financial loss is at stake (rather than simpler to deal with data-loss risks or system reduced availability).



References:

[1]. A. Iamnitchi *et al.* "Filecules in High-Energy Physics: Characteristics and Impact on Resource Management," *HPDC 2006, Paris, June 2006.*
[2]. S. Garfinkel, "Commodity Grid Computing with Amazon's S3 and EC2", USENIX ;login, February 2007.
[3]. Amazon Web Services, http:/s3.amazonaws.com
[4]. The DZero Experiment, http://www-d0.fnal.gov
[5]. PlanetLab Consortium, http://planet-lab.org
[6]. BitTorrent, http://www.bittorrent.com