Beyond the Internet: Convergence of Networking and Storage Becomes A Must

by David Hung-Chang Du
University of Minnesota, Minneapolis, MN 55455

The Internet today has grown to an enormously large scale. Devices large and small are connected globally from anywhere on the earth via either wireless or wired connections. With the rapid advancement of technology, we now also have cheap and small devices with high computing power and large storage capacity. These devices are designed to improve our daily life by monitoring our environment, collecting critical data, and executing special instructions. These devices have gradually become a dominant part of our Internet. Many imaging, audio and video data are converted from analog to digital and digital data are generated at an alarming rate. As a result, unprecedented amount of data are available. How to manage and look for the desired information becomes a great challenge. How to preserve these data becomes a crisis.

Both network and storage technologies have made tremendous improvement over the last few decades. However, it is still true that our networking research focuses on how to efficiently deliver data end-to-end to satisfy certain quality of services (QoS) including security and privacy concerns. The semantic of the data and how data to be stored, organized and accessed are not a concern of networking research.

On the other hand, storage research has focused on designing and developing cheaper and higher capability storage devices, various storage systems architectures are proposed to support efficient accesses of both structured and unstructured data. Network is considered as a black-box and a means to enable remote storage accessing by storage researchers. Typically, for satisfying certain QoS requirement of a storage system, over provision of storage devices is the only way to go. However, this approach is no longer working since more data are stored in remote sites and has to be accessed via Internet. Cloud storage and iSCSI (SCSI commands over Internet) become more and more popular.

We argue in the future software defined infrastructure, both networking and storage will have to be fully integrated. From the networking side, this integration will have to be beyond the current thinking of software-defined network and named data network since both of them have not considered the management issues of storage including where data should be located and migrated, when to duplicate a data or reduce the copies of a data in response to the increasing or decreasing of demand of the data, how to guarantee the reliability or availability of data considering both network and storage failures. From storage side, cloud storage has the advantages of elastic, scale up, and cheaper (reduced TCO). However, its performance cannot be guaranteed without the assistance from network side. Currently, the users have decided which data should be stored locally or in cloud storage. It will be great if we can seamlessly integrated local and cloud storage into one system that can ensure all accessed data in local storage and no longer needed data in cloud storage. The data migration between local and cloud can fully consider the network situation and can shorten the migration latency. Currently storage industry is pushing for object-based storage devices which require meta-data servers to handle all the management issues. If these management functions can be fully integrated into the proposed NFV, it will create a much more efficient software defined infrastructure.