Chapter 11
Data Transfer and Storage in Cloud Computing

Yushi Shen
Microsoft Corporation, USA

Ling Wu
EMC² Corporation, USA

Yale Li
Microsoft Corporation, USA

Shaofeng Liu
Microsoft Corporation, USA

Qian Wen
Endronic Corp, USA

ABSTRACT

When network bandwidth is no longer a bottleneck for Internet applications, we still face the challenges of how to utilize the bandwidth in full when transferring data. This might sound straightforward, but it is far from being easy or straightforward. Lots of work has been done to improve the utilization of abundant network bandwidth, while maintaining the same reliability, efficiency, and fairness of slow protocols, like TCP. In this chapter, the authors introduce the background of emerging bandwidth-intensive applications, related works that are in place to solve these issues, and their limitations to make the challenges clear to the readers. Similar challenges also present for large-scale data preservation.

INTRODUCTION

The Internet can never be fast enough. As users of the Internet, we have all witnessed the Internet speed changing from 56Kbps via a modem to 10/100Mbps via ADSL or cable networks, and the new networks are being installed in hundreds of millions of homes around the world. Our utilization of the Internet has also evolved: Besides browsing simple web pages, more and more are using the Internet to watch high-resolution videos, do real-time sharing and video chatting and so on. There

DOI: 10.4018/978-1-4666-4801-2.ch011

is a clear trend on how the Internet is being used, as new network applications appear. The development of network technologies just barely meet the requirements of these new network applications, and in many cases, applications must succumb to the available network bandwidth. Although new applications have leveraged the continuous development of the Internet infrastructure, at the same time, they put great pressure on the future Internet technologies for further development. This is perhaps why we never feel that the Internet can be sufficiently fast for us. These are today's challenges, as we want to bring everything into the cloud.

As time goes on, many applications that used to appear in science fiction stories become our reality due to increasingly fast networks. Gradually, the lower-level communication modules responsible for data transfer are becoming increasingly application critical. Demands are high for improving the performance of those modules. We cannot change the trend that future applications are required to have higher network speeds and latency. We should rather think whether the potential of the current network resources has been fully utilized. More specifically, the current communication modules need to be improved, such that data can be transferred more efficiently. We believe these network communication modules should be adaptive to network applications, instead of the other way around. Since different applications may have different characteristics, communication modules need to take these differences into consideration. Communication protocols may vary between different types of network applications. The research in this section has been initially motivated by one specific type of application that emerged in the recent years: large scale digital content transfer, streaming and archiving. This is especially true in Cloud storage, where people can store and retrieve massive data whenever they need it.

There has been massive amounts of research done on optical networks in the recent years. As one example, CineGrid has led a trend of using high bandwidth networks to transfer very high quality digital content, for real-time movie showings and digital preservation. Although many successful demonstrations of 4K video streaming have been performed over the fiber networks, we foresee the challenges in designing new protocols, which scale to applications demanding higher bandwidth and higher resolution digital content in the future. Today, ultra-high-resolution displays have become standard infrastructure in scientific research. These displays are typically built by tiling an array of standard LCD displays onto a display wall, driven by a PC cluster. There are some different types of display walls:

- The HIPerSpace uses sixty 30' LCD displays, to form a display wall with more than 200 million pixels;
- The StarCAVE uses 34 high-definition projectors to construct a 3D virtual space, where people can work with 3D virtual reality objects.

Meanwhile, high-speed research optical networks make it possible for scientists to use these ultra-high resolution displays over long distances, for such scientific applications as very-high-definition video streaming or conferencing, real-time data visualization generated by remote scientific instruments and so on. As a perfect example of the combination of display walls and a high-speed network, the OptIPuter research project, funded by the US National Science Foundation, constructed an 1Gbps-10Gbps optical network infrastructure, as well as the middleware to facilitate the interactive access to the remote gigabytes to terabytes of visualization data objects, and bring them to a visual interface, the OptIPortal.

However, scaling up visualization devices, from a single PC with a single display to a cluster of PCs with a cluster of displays, has brought up challenges on how to feed data to these display devices. These challenges are due to the limitations of traditional single data source communication