

QoS Provisioning Framework for an OSD-based Storage System*

Yingping Lu
University of Minnesota
lu@cs.umn.edu

David H. C. Du
University of Minnesota
du@cs.umn.edu

Tom Ruwart
IOPerformance Inc.
tmruwart@iopformance.com

Abstract

Quality of Service (QoS) is crucial for certain applications such as multimedia. As the Object-based Storage Device (OSD) protocol emerges as the next generation storage technology, QoS provisioning for OSD-based systems has also received a great deal of attention. In this paper, we propose a QoS framework for OSD-based storage system that integrates both the network QoS and storage QoS. We examine the existing OSD specification and analyze the QoS requirements for applications on OSD clients. Based on the QoS requirements analysis, we propose a three-level QoS specification. We further elaborate on extensions to the existing OSD and iSCSI protocol to support our QoS specification. These extensions are then incorporated with the current OSD reference implementation. Finally, we discuss both the implementation structure and issues encountered as part of this study.

1. Introduction

With the ubiquity of TCP/IP networks and the increased popularity of network applications, Quality of Service (QoS) has been extensively studied. QoS is crucial for real-time applications running on a network, such as streaming video and voice over IP, which demand certain guarantees of network bandwidth and delay variance. Generally speaking, QoS is an end-to-end issue, i.e. from application to application [8]. The QoS enforcement involves transport, network and end system. There has been plethora of QoS research in the TCP/IP network transmission domain [9-14], as well as disk scheduling in end-systems [15-25] to satisfy different QoS requirements in storage data transfers. However, these two aspects of QoS (network QoS and storage QoS) are often studied separately [8]. For example, the storage scheduling schemes rarely consider the effect of network's condition.

Recent advances in storage technology have brought about the Object-based Storage Device (OSD) protocol [1][2]. The Storage Networking Industry Association (SNIA) OSD Standard [1] and the Lustre Project [3] represent two OSD-centric efforts. The first deals with the development of a standard OSD protocol that is part of the SCSI standard and implies interoperability with any standard SCSI disk drive that implements the OSD SCSI command extensions. On the other hand, Lustre, is focused on the development of a file system that communicates with object devices rather than traditional block devices over an OSD-like protocol. This is an important aspect of the evolution of OSD as a protocol as Lustre represents an "application" that can fully utilize an OSD hence making it a good platform for testing and validating OSD concepts such as QoS.

Compared to traditional *block-level* data access, OSD offloads storage management functions (such as space allocation) from a traditional file system to the object device, and consequently offers *object-level* data access to its clients. As a result, the OSD protocol imposes new challenges and opportunities for QoS provisioning to its clients.

In most of the previous QoS studies it was assumed that the server had a *dedicated* storage system with a direct I/O connection. In this study, we assume the OSD device is *shared* by a number of clients connected through network rather than directly attached to a single server. Furthermore, not only will each server have *different QoS requirements* for their object access, but each server and/or object access may also have quite *different network access characteristics*. For example, each server and/or object device connection may have a different data path, thus implying different available network bandwidth, variable delay characteristics, ... etc. The *intelligence* embedded on object devices can be used to assess the QoS (and other) requirements of its data objects, monitor the current condition and capacity of storage devices, and to measure the existing networking conditions on its interface to its clients. The focus of our study is to explore the *integration of both the storage QoS and network QoS* for QoS provisioning and guarantees in this new environment.

* This work is supported by StorageTek, Veritas, Engenio, and Sun Microsystems through their memberships in the University of Minnesota Digital Technology Center Intelligent Storage Consortium (DISC).

