

Storage Performance Isolation: An Investigation of Contemporary I/O Schedulers

Sarala Arunagiri*

Seetharami Seelam†

Patricia J. Teller*

14 February 2007

1 Introduction

Consolidation of storage resources in data centers is becoming popular because it offers various types of benefits including the benefit of economy-of-scale. In these data centers, storage systems are shared by workloads of multiple customers. Each customer is provided with a Quality of Service (QoS) guarantee. In the absence of gross over-provisioning, in order to provide every customer with a level of I/O service conforming to the promised QoS, it is important to ensure performance isolation. This means that the performance experienced by the workload of a given customer must not suffer due to variations in I/O request stream characteristics of other customers. This is a huge challenge. In the last decade, several I/O schedulers have been designed that apportion shared storage resources among competing request streams. Most of these schedulers assume that the workloads presented to them can be serviced while conforming to the QoS promised to all customers. The feasibility of this is determined apriori using a provisioning tool.

2 Performance Isolation - Fine Distinctions

Some of the attributes that are commonly used to characterize I/O workloads are request rate, request size, burstiness, request type, and read/write ratio. These can be considered as high-level attributes of workload characteristics. This characterization does not capture the dynamic aspects of the workload imposed on a storage device. However, these dynamic aspects are characterized to some degree by the attributes of the streams that comprise the workload [1]. I/O stream attributes include, among others, run-count, jump-distance, and location-skew.

All of the schedulers considered in this study have performance isolation as one of their design goals. And, most of them provide performance isolation in the face of variations of most of the above mentioned high-level attributes. However, the question is: Do these schedulers provide performance isolation in the face of variations of any characteristic of I/O streams, for example, the runcount. If the answer to this question is yes, then we say that these schedulers provide absolute performance isolation. These schedulers must provide performance isolation both in terms of high-level attributes of workloads as well as the I/O stream attributes. In practice, there

may be reasons to relax some of these requirements. Motivated by an academic interest, we evaluated several contemporary I/O scheduling algorithms with respect to the property of absolute performance isolation.

3 Investigation of Schedulers

The focus of our study is a set of contemporary I/O schedulers which includes Cello, Facade, Triage, and the 2-Level Interposed Scheduler. Our analysis shows that the performance isolation provided by these schedulers is not absolute. For example, consider the 2-Level Interposed Scheduler, which employs a sophisticated algorithm, the design of which is based on statistics and queuing theory. In this case, let us say that one or more of the I/O request streams misbehave, i.e., they generate requests with behaviors different than those (implicitly) expected. As long as the misbehavior is in terms of arrival rates, there is a mechanism in the algorithm that reacts to it in order to provide performance isolation. However, if the misbehavior is in terms of run-count or jump-distance, then the algorithm cannot promise performance isolation. A summary of our study is provided in the table below.

Schedulers/Tools	Absolute Performance Isolation
Cello	
Disk Scheduling Algorithm	Yes
CFQ-CRR(P)	
Disk Scheduling Algorithm	Yes
YFQ	No
Facade	No
Interposed 2-Level Scheduler	No
Triage	No
SLEDS	Feasible
Chameleon	Feasible

4 Work In Progress

The results of this study motivated our work in progress: The design of an I/O scheduler that provides a high degree of performance isolation.

References

- [1] J. Wilkes. Traveling to Rome: QoS Specifications for Automated Storage System Management. In Proceedings of the International Workshop on Quality of Service. Springer-Verlag, June 2001. <http://citeseer.ist.psu.edu/wilkes01traveling.html>

*The University of Texas at El Paso

†IBM TJ Watson Research Center

0