Contribution ID: **23**   Type: **Parallel**

A strategy for load balancing in distributed storage systems

Tuesday 22 May 2012 13:55 (25 minutes)

Distributed storage systems are critical to the operation of the WLCG. These systems are not limited to fulfilling the long term storage requirements. They also serve data for computational analysis and other computational jobs. Distributed storage systems provide the ability to aggregate the storage and IO capacity of disks and tapes, but at the end of the day IO rate is still bound by the capabilities of the hardware, in particular the hard drives. Throughput of hard drives has increased dramatically over the decades, however for computational analysis IOPS is typically the limiting factor. To maximize return of investment, balancing IO load over available hardware is crucial. The task is made complicated by the common use of heterogeneous hardware and software environments that results from combining new and old hardware into a single storage system.

This paper describes recent advances made in load balancing in the dCache distributed storage system. We describe a set of common requirements for load balancing policies. These requirements include considerations about temporal clustering, resistance to disk pool divisioning, evolution of the hardware portfolio as data centers get extended and upgraded, the non-linearity of disk performance, garbage collection, age of replicas, stability of control decisions, and stability of tuning parameters. We argue that the existing load balancing policy in dCache fails to satisfy most of these requirements.

An alternative policy is proposed, the weighted available space selection policy. The policy incorporates ideas we have been working on for years while observing dCache in production at NDGF and at other sites. At its core the policy uses weighted random selection, but it incorporates many different signals into the weight. We argue that although the algorithm is technically more complicated, it is in our experience easier to predict the effect of parameter changes and thus the parameters are easier to tune than in the previous policy. It many cases it may not even require manual tuning, although we need more empirical data to conclude that.

The new policy has been integrated into dCache 2.0 using a new load balancing plugin system developed by NDGF. It has been used in production at NDGF since end of August 2011. We will report on our experiences. A qualitative and quantitative analysis of the policy will be presented augmented by simulations and empirical data.

Although our algorithm has been developed and is used in the context of dCache, the ideas are universal and could be applied to many storage systems.

Author: BEHRMANN, Gerd (NDGF)

Co-author: WADENSTEIN, Erik Mattias (Unknown)

Presenter: WADENSTEIN, Erik Mattias (Unknown)

Session Classification: Computer Facilities, Production Grids and Networking

Track Classification: Computer Facilities, Production Grids and Networking (track 4)