



Contribution ID: 192

Type: oral presentation

Bandwidth-sharing in LHCONE, an analysis of the problem

Thursday, 16 April 2015 11:30 (15 minutes)

The LHC experiments have traditionally regarded the network as an unreliable resource, one which was expected to be a major source of errors and inefficiency at the time their original computing models were derived. Now, however, the network is seen as much more capable and reliable. Data are routinely transferred with high efficiency and low latency to wherever computing or storage resources are available to use or manage them.

Although there was sufficient network bandwidth for the experiments' needs during Run-1, they cannot rely on ever-increasing bandwidth as a solution to their data-transfer needs in the future. Sooner or later they need to consider the network as a finite resource that they interact with to manage their traffic, in much the same way as they manage their use of disk and CPU resources.

There are several possible ways for the experiments to integrate management of the network in their software stacks, such as the use of virtual circuits with hard bandwidth guarantees or soft real-time flow-control, with somewhat less firm guarantees. Abstractly, these can all be considered as the users (the experiments, or groups of users within the experiment) expressing a request for a given bandwidth between two points for a given duration of time. The network fabric then grants some allocation to each user, dependent on the sum of all requests and the sum of available resources, and attempts to ensure the requirements are met (either deterministically or statistically).

An unresolved question at this time is how to convert the users' requests into an allocation. Simply put, how do we decide what fraction of a network's bandwidth to allocate to each user when the sum of requests exceeds the available bandwidth? The usual problems of any resource-scheduling system arise here, namely how to ensure the resource is used efficiently and fairly, while still satisfying the needs of the users.

Simply fixing quotas on network paths for each user is likely to lead to inefficient use of the network. If one user cannot use their quota for some reason, that bandwidth is lost. Likewise, there is no incentive for the user to be efficient within their quota, they have nothing to gain by using less than their allocation.

As with CPU farms, some sort of dynamic allocation is more likely to be useful. A promising approach for sharing bandwidth at LHCONE is the 'Progressive Second-Price auction', where users are given a budget and are required to bid from that budget for the specific resources they want to reserve. The auction allows users to effectively determine among themselves the degree of sharing they are willing to accept based on the priorities of their traffic and their global share, as represented by their total budget. The network then implements those allocations using whatever mix of technologies is appropriate or available.

This paper describes how the Progressive Second-Price auction works and presents a prototype simulation of the auction. Results from the simulation are shown, addressing practical questions such as how are budgets set, what strategy should users use to manage their budget, how and how often should the auction be run, and how do we ensure that the goals of fairness and efficiency are met.

Primary author: Dr WILDISH, Tony (Princeton University (US))

Presenter: Dr WILDISH, Tony (Princeton University (US))

Session Classification: Track 6 Session

Track Classification: Track6: Facilities, Infrastructure, Network