21st International Conference on Computing in High Energy and Nuclear Physics **CHEP2015** Okinawa Japan: April 13 - 17, 2015

Contribution ID: **194** Type: **poster presentation**

Understanding the CMS Tier-2 network traffic during Run-1

At the beginning of Run-1 CMS was operating it's facilities according to the MONARC model, where data-transfers were strictly hierarchical in nature. Direct transfers between Tier-2 nodes was excluded, being perceived as operationally intensive and risky in an era where the network was expected to be a major source of errors. By the end of Run-1 wide-area networks were more capable and stable than originally anticipated. The original data-placement model was largely superceded, and traffic was allowed between Tier-2 nodes.

Tier-2 to Tier-2 traffic in 2012 already exceeded the amount of Tier-2 to Tier-1 traffic, so it clearly has the potential to become important in the future. Moreover, while Tier-2 to Tier-1 traffic is mostly upload of monte-carlo data, the Tier-2 to Tier-2 traffic represents data moved in direct response to requests from the physics analysis community. As such, problems or delays there are more likely to have a direct impact on the user community.

Tier-2 to Tier-2 traffic may also traverse parts of the WAN that are at the 'edge' of our network, with limited network capacity or reliability compared to, say, the Tier-0 to Tier-1 traffic which goes the over LHCOPN network.

CMS is looking to exploit technologies that allow us to interact with the network fabric so that it can manage our traffic better for us, this we hope to achieve before the end of Run-2. Tier-2 to Tier-2 traffic would be the most interesting use-case for such traffic management, precisely because it is close to the users' analysis and far from the 'core' network infrastructure.

As such, a better understanding of our Tier-2 to Tier-2 traffic is important. Knowing the characteristics of our data-flows can help us place our data more intelligently. Knowing how widely the data moves can help us anticipate the requirements for network capacity, and inform the dynamic data placement algorithms we expect to have in place for Run-2.

This paper presents our analysis of the Tier-2 traffic during Run 1. We examine the geographical and temporal distribution of traffic, the transfer quality, and the correlation between data-movement and physics analysis groups. We conclude with recommendations for improving data-placement in Run-2 and suggestions for monitoring improvements to help us better understand the changes in this behaviour as our system evolves.

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

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

Track Classification: Track5: Computing activities and Computing models