



Contribution ID: 46

Type: **not specified**

Evaluating Infiniband Based Networking Solutions for HEP/NP Data Processing Applications

Tuesday, October 14, 2014 4:30 PM (30 minutes)

The Infiniband networking technology is a long established and rapidly developing technology which is currently dominating the field of low-latency, high-throughput interconnects for HPC systems in general and those included in the TOP-500 list in particular. Over the last 4 years a successful use of Infiniband networking technology combined with additional IP-over-IB protocol and Infiniband to Ethernet bridging layers was demonstrated well beyond the realm of HPC, covering various high throughput computing (HTC) systems, including data processing farms and private clouds devoted to HEP/NP data processing. With the recent advances of Mellanox VPI technology in 2013-2014 the 4X FDR IB now stands as the most versatile networking solution available for existing and future data centers that need to support both HTC and HPC oriented activities that can be seamlessly integrated into the existing Ethernet based infrastructure. Furthermore, it can be done completely transparently for the end users of these facilities, though certain modifications of the end user's activity-patterns are needed in order to utilize the full potential of the Infiniband based networking infrastructure. This contribution contains a detailed report on the series of tests and evaluation activities performed within the RACF over the last year in order to evaluate a Mellanox 4X FDR Infiniband based networking architecture (provided with an oversubscribed tree topology) as a potential alternative networking solution for both the RHIC and ATLAS data processing farms of the RACF, as well as the existing dCache and future Ceph based storage systems associated with them. Results of the price/performance comparison of such a networking system with a competing solution based on the 10 GbE technology (provided with non-blocking fabric topology) for a HEP/NP data processing farm consisting of 1500 compute nodes are presented. Job placement optimizations in Condor for the offline data processing farm of the PHENIX experiment were implemented in order to demonstrate a sample user activity-pattern that more optimally utilizes the Infiniband-based networking solution. The results of using those optimizations in production for the PHENIX experiment over the last 9 months are presented.

Primary author: Mr ZAYTSEV, Alexandr (Brookhaven National Laboratory (US))

Co-authors: Mr HOLLOWELL, Christopher (Brookhaven National Laboratory); Dr RIND, Ofer (Brookhaven National Laboratory); Dr WONG, Tony (Brookhaven National Laboratory); Mr STRECKER-KELLOGG, William (Brookhaven National Lab)

Presenter: Mr ZAYTSEV, Alexandr (Brookhaven National Laboratory (US))

Session Classification: Networking and Security

Track Classification: Security & Networking