Project Proposal

by

Aditya Sakhuja
(aditya.sakhuja@gatech.edu)

for

High Performance Parallel Computing

Prof. Richard Vuduc
(http://www.vuduc.org)

Title: HECURA: Configuration benchmarks for IO graphs

Goal: Benchmark various configurations of IOGraphs to determine the model for performance and algorithm(s) for deployment

Participants: Aditya Sakhuja (aditya.sakhuja@)

Sponsors: Jay Lofstead (lofstead@cc), Ada Gavrilovska (ada@cc)

Detailed Description:

Benchmark various configurations of IOGraphs to determine the model for performance and algorithm(s) for deployment.

The HECURA project on high performance I/O is concerned with how to efficiently move data in and out of large scale parallel machines, i.e., those used at the DOE National Labs. The idea of this project is to create rich methods for operating on data as it is being moved to/from storage. This is done with an asynchronous I/O interface termed DataTaps, with a data streaming overlay, termed I/O graphs, and with interesting storage backend (i.e., not just a file system, but potentially the LWFS (lightweight file system) object based store. The message transport uses the EVPath system to configure an overlay network with embedded computation.

GTC is a parallel fusion application running on a large number of "compute" nodes and performing periodic writes of large data sizes to a fairly smaller number of "I/O nodes". These are in turn responsible for further writing the data out to storage. The application is running on a local cluster with Infiniband interconnect.

The broad requirement for one of the extensions would be to consider the performance implications of using DataTaps and I/O graphs for large scale parallel machines.

- **Project specifics:**

  The details of the project build over the in place infrastructure are presented below.

The first stage would be to do a detailed performance analysis for determining the tradeoffs between data sizes, frequency of write operations, overheads of write operations, total throughput, etc. with the current policy.

The second stage would involve designing and implementing an algorithm to schedule the IO operation of individual "compute" nodes to improve the current behavior. Typically, this would mean to be able to deduce information such as "If I write out x amount of data, with frequency y, the system performance will suffer by n%".

Resources: WARP cluster, Might require working on a more dedicated HPC cluster (will need access to it)

Risks:

The infrastructure is in place and the system has been worked upon a lot. The probability of total failure and not able to benchmark the configurations is minimal if not eliminated all together.

Glossary:

- EVPath – Messaging mechanism
- DataTap – Asynchronous data staging mechanism
- GTC – Fusion code data source
- IO Graph – Concept of processing data in transit

References:

http://www.cc.gatech.edu/~smcmanus/6210/SpecialProjectIdeas.html