



FastForward I/O and Storage: ACG 7.4 Initial Data Analytics Pipeline Integration Demonstration

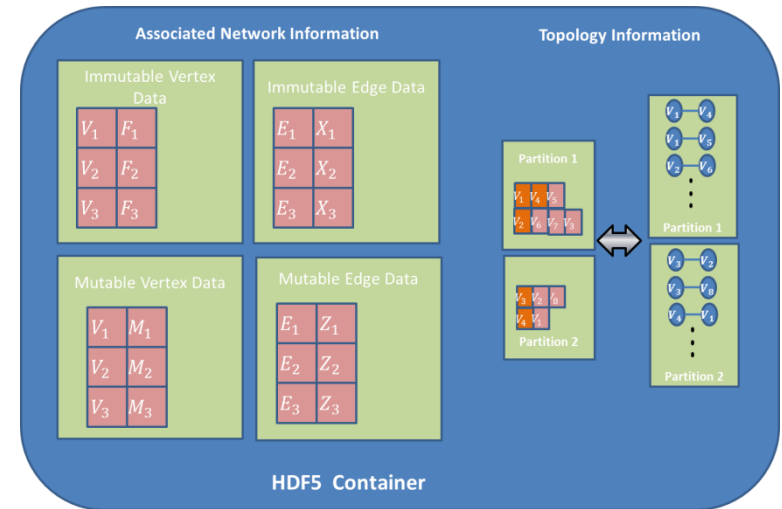
NOTICE: THIS MANUSCRIPT HAS BEEN AUTHORED BY INTEL UNDER THE SUBCONTRACT WITH LAWRENCE LIVERMORE NATIONAL SECURITY, LLC WHO IS THE OPERATOR AND MANAGER OF LAWRENCE LIVERMORE NATIONAL LABORATORY UNDER CONTRACT NO. DE-AC52-07NA27344 WITH THE U.S. DEPARTMENT OF ENERGY. THE UNITED STATES GOVERNMENT RETAINS AND THE PUBLISHER, BY ACCEPTING THE ARTICLE OF PUBLICATION, ACKNOWLEDGES THAT THE UNITED STATES GOVERNMENT RETAINS A NON-EXCLUSIVE, PAID-UP, IRREVOCABLE, WORLD-WIDE LICENSE TO PUBLISH OR REPRODUCE THE PUBLISHED FORM OF THIS MANUSCRIPT, OR ALLOW OTHERS TO DO SO, FOR UNITED STATES GOVERNMENT PURPOSES. THE VIEWS AND OPINIONS OF AUTHORS EXPRESSED HEREIN DO NOT NECESSARILY REFLECT THOSE OF THE UNITED STATES GOVERNMENT OR LAWRENCE LIVERMORE NATIONAL SECURITY, LLC.

Agenda

- Background and Demonstration objectives
- Demo Environment
- Page rank
- Conclusion and discussion on upcoming items

Milestones

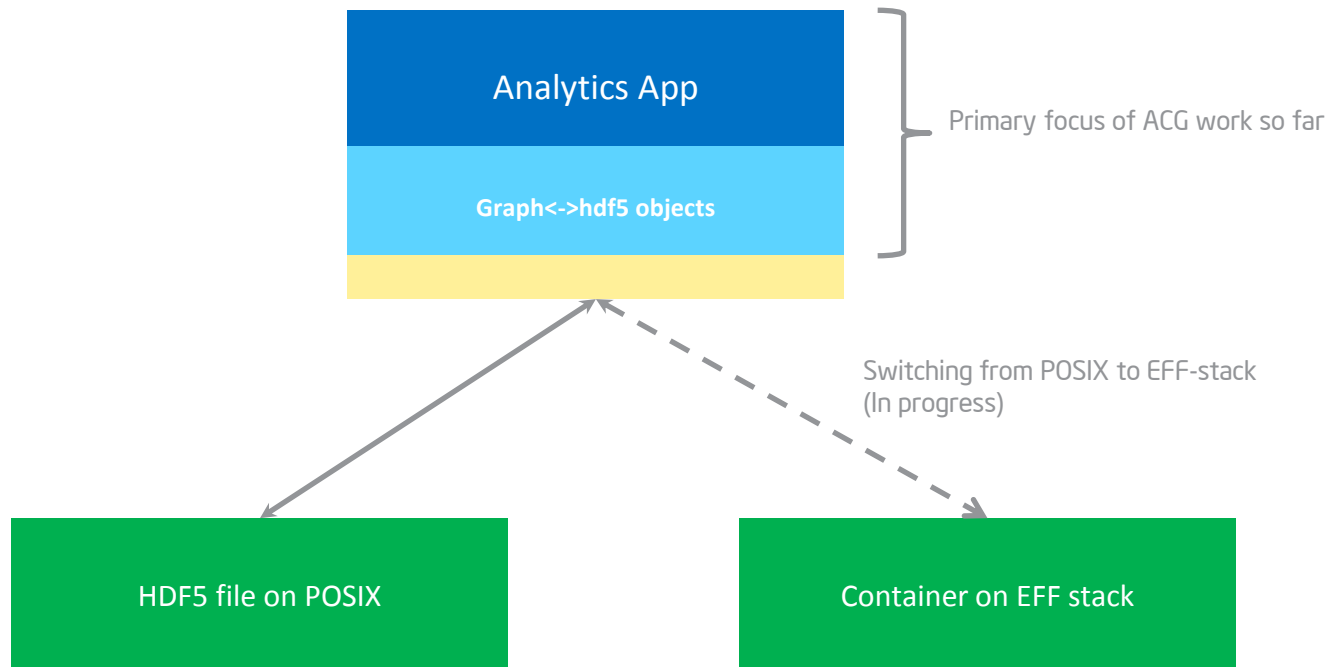
- Last quarter: we demonstrated a full graph computation – i.e., graph topology and network information.



- Q7 & Q8!
 - *"The Subcontractor shall demonstrate an initial integrated data analytics pipeline using GraphLab/HAL for graph computation. The demonstration will show initial functional capability and load balancing."*
 - We will demonstrate integration with the stack, supporting a page rank graph computation (Q7), and a solving a complex text-analytics problem (Q8).

A very high-level summary update of this demo

- In-Scope Demonstration Elements
 - Simple data-analytics application
 - Graph computation kernel performing asynchronous vertex update
 - Application library runs over exascale-stack
- Out of Scope Demonstration Elements
 - Graph computation for full use-case applications



Demonstration Environment

Cluster Spec

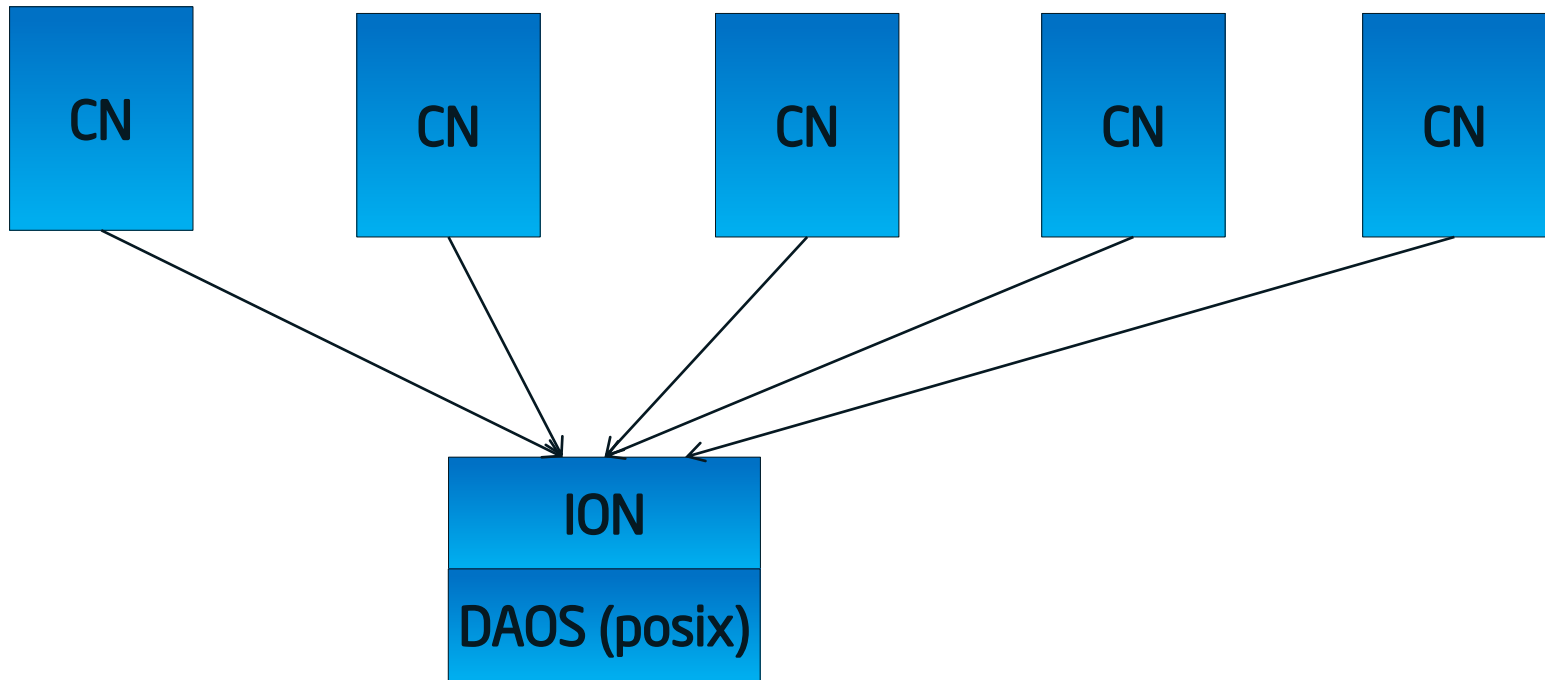
- 5 nodes
- 128GB Memory / node
- 8 quad-core / processor
- SATA 6Gb/s 7200RPM
16MB Seagate Barracuda ST500DM002



Note

- POSIX
- Small graph used for demonstration purposes

This Demo (simple pagerank)



Conclusions and Upcoming Work

Conclusions

- Covered demo criteria: Computational kernel loading non-trivial network-information along with graph-topology

Upcoming work

- ~~Computational kernel Interface with the exascale stack (through HDF5) (Q7)~~
- Detailed Benchmark and comparisons with other platforms (Q8)
- Full scale Graph Analytics on Exascale Stack (Q8)