Question: We have concerns about data transfers becoming an issue:

- is USQCD storing the data in an optimal fashion?
- are there plans for facilitating data transfer between the labs? For example, plans for improving bandwidth? Any other plans for the future?

We take the sense of "optimal fashion" to mean organizing and storing the data in a manner that facilitates locating the data of interest, and places the data in locations that are suitable for high-speed inter-site transfers.

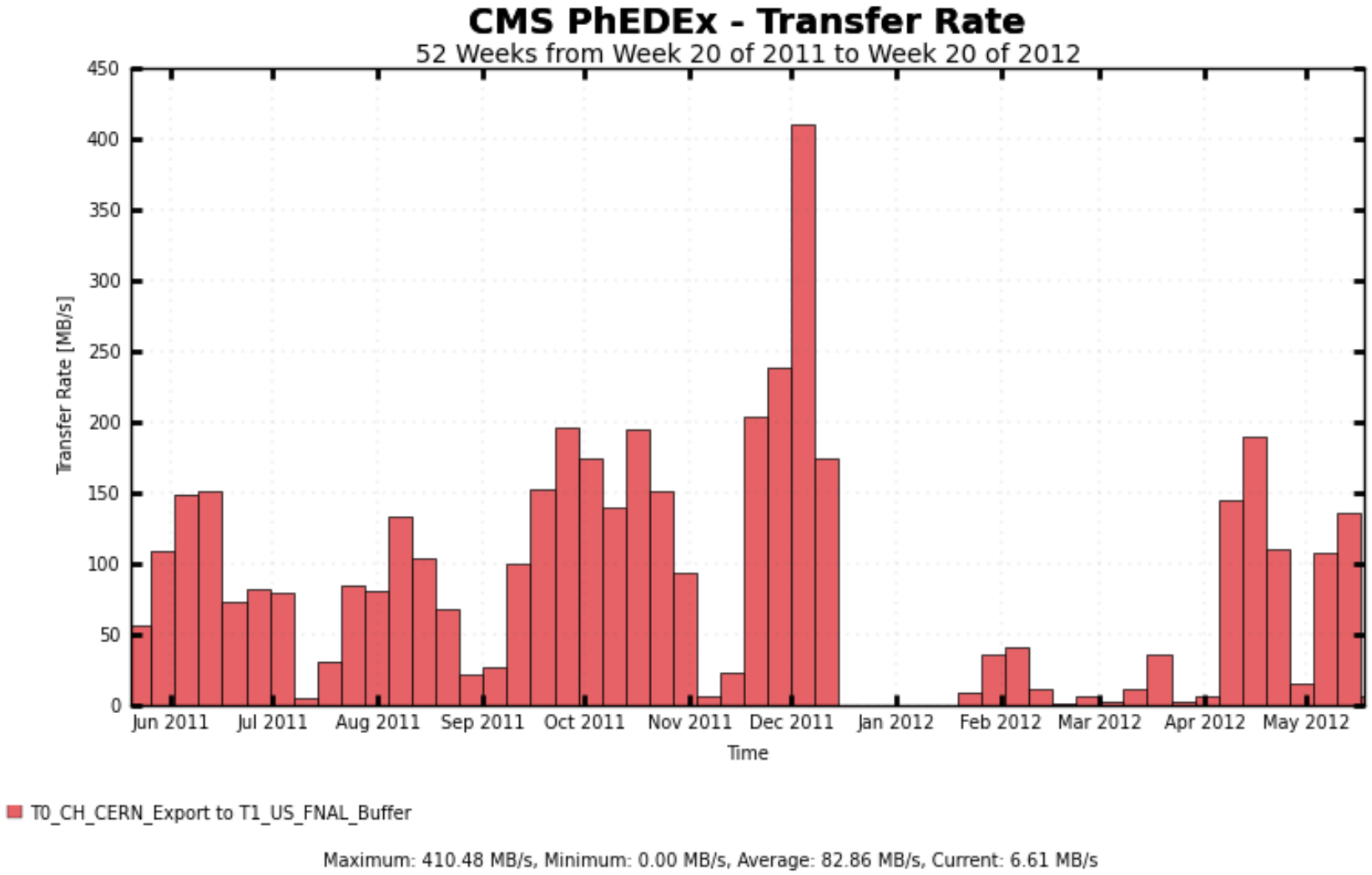
LQCD has a long history of storing (and sharing) our precious data – gauge configurations – at multiple sites. For example, copies of various MILC asqtad ensembles are kept in mass storage systems at Fermilab, NCSA, NERSC, UCSD. USQCD participates in the International Lattice Data Grid (ILDG), a data grid used for sharing gauge configurations, providing a standard for metadata markup and file layouts, searchable metadata catalogs, access controls (via Virtual Organizations), and tools for transferring data. Gauge configurations are also available from NERSC's "gauge connection".

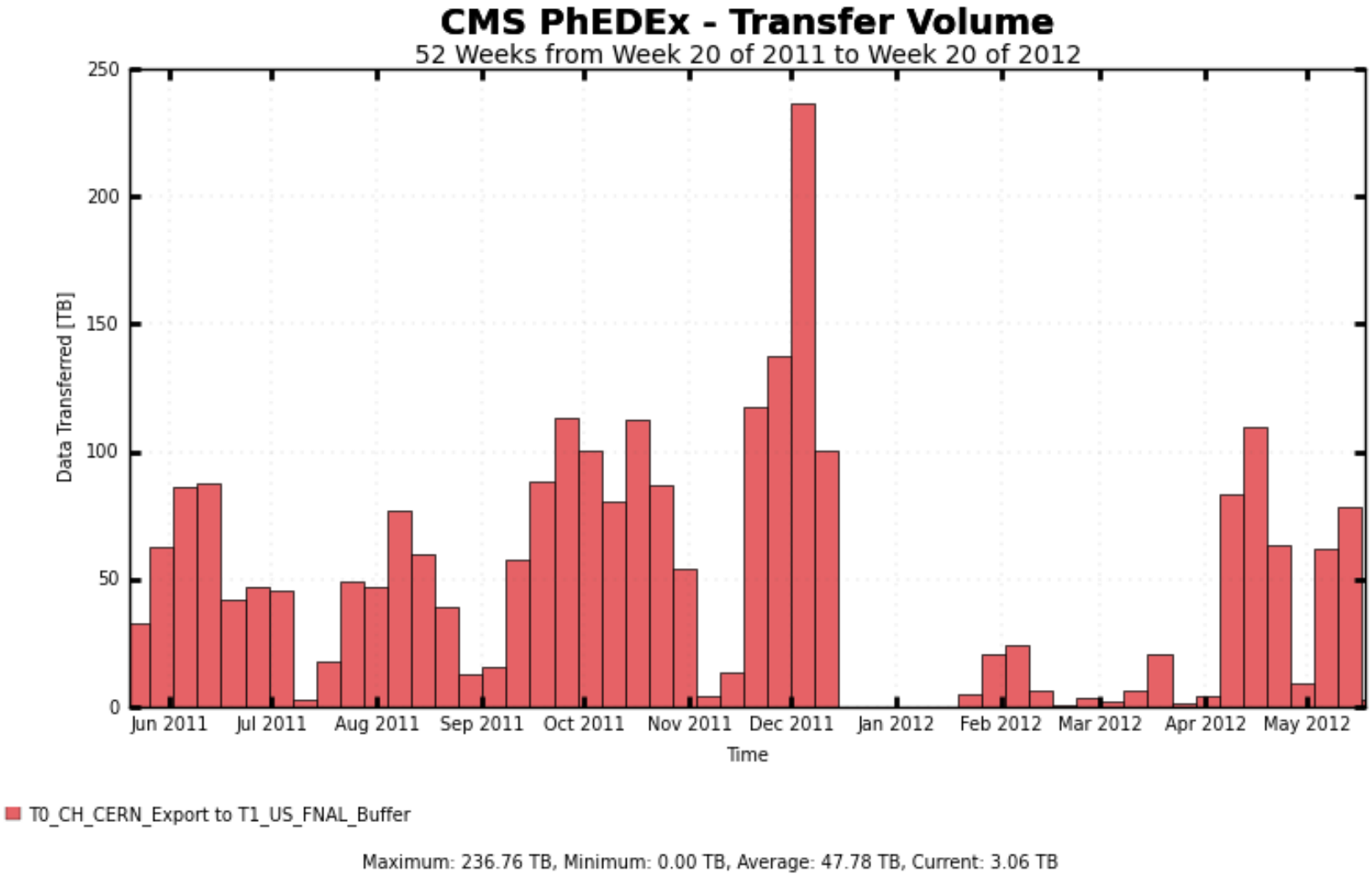
USQCD has strong experience with moving data generated at the LCF's to mass storage systems at Fermilab and JLab. For example, quark propagators generated by the "mixbk" project at the ALCF during the early years of the BG/P were transferred to Fermilab tape storage (volume was of order 100TB). Anisotropic clover action data have been generated for many years at ORNL and transferred to JLab.

The host laboratories have strong offsite network connectivity. BNL and Fermilab, tier 1 sites respectively for Atlas and CMS, currently have 10+ Gbit offsite network connectivities. Fermilab's connectivity will be upgraded to 100 Gbit via a dedicated 100G link and several 10G links to the new 100G ESnet5 later this year. JLab increased their connectivity to 10G in 2009. At all three host sites, LQCD access to these very large offsite pipes requires provisioning sufficient WAN capacity to/from the LQCD systems. At Fermilab, our current 1G connection to the site WAN will be upgraded this year to multiple 1G connections to a 10G backbone.

LQCD data in the mass storage systems (robotic tape systems) at FNAL and JLab are readily accessible to authorized collaboration members via the various well-known high performance transfer protocols (SRM, GridFTP, etc..). Two weeks ago, the Fermilab Globus Online endpoint was used to transfer MILC HISQC gauge configurations between NCSA and Fermilab mass storage systems. In response to a request at the All Hands Meeting, we plan to provision next month a G.O. endpoint within the Fermilab LQCD Lustre filesystem to be used in transferring eigenvector data from BNL (order 100 TB). Should G.O. prove, as we anticipate, to be a convenient mechanism for our users to perform high volume transfers, we plan to provision endpoints within the 3 sites' parallel file systems.

Individual LQCD projects in general are cited at just one of the host sites, and tend not to move more often than once, or in very rare cases, twice a year (although there are exceptions, with some project receiving allocations at both cluster sites). Large scale data movement between sites for most projects therefore is infrequent. Should more frequent transfers be needed, all three host labs have sufficient network capabilities that could be employed in part by LQCD.


CMS PhEDEx – Transfer Rate
52 Weeks from Week 20 of 2011 to Week 20 of 2012
Transfer Rate [MB/s]
450
400
350
300
250
200
150
100
50
0
Jun 2011
Jul 2011
Aug 2011
Sep 2011
Oct 2011
Nov 2011
Dec 2011
Jan 2012
Feb 2012
Mar 2012
Apr 2012
May 2012
Time
T0_CH_CERN_Export to T1_US_FNAL_Buffer
Maximum: 410.48 MB/s, Minimum: 0.00 MB/s, Average: 82.86 MB/s, Current: 6.61 MB/s


CMS PhEDEx - Transfer Volume
52 Weeks from Week 20 of 2011 to Week 20 of 2012
Data Transferred [TB]
250
200
150
100
50
0
Jun 2011
Jul 2011
Aug 2011
Sep 2011
Oct 2011
Nov 2011
Dec 2011
Jan 2012
Feb 2012
Mar 2012
Apr 2012
May 2012
Time
T0_CH_CERN_Export to T1_US_FNAL_Buffer
Maximum: 236.76 TB, Minimum: 0.00 TB, Average: 47.78 TB, Current: 3.06 TB