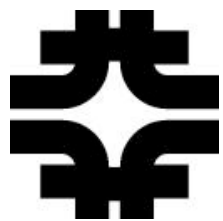
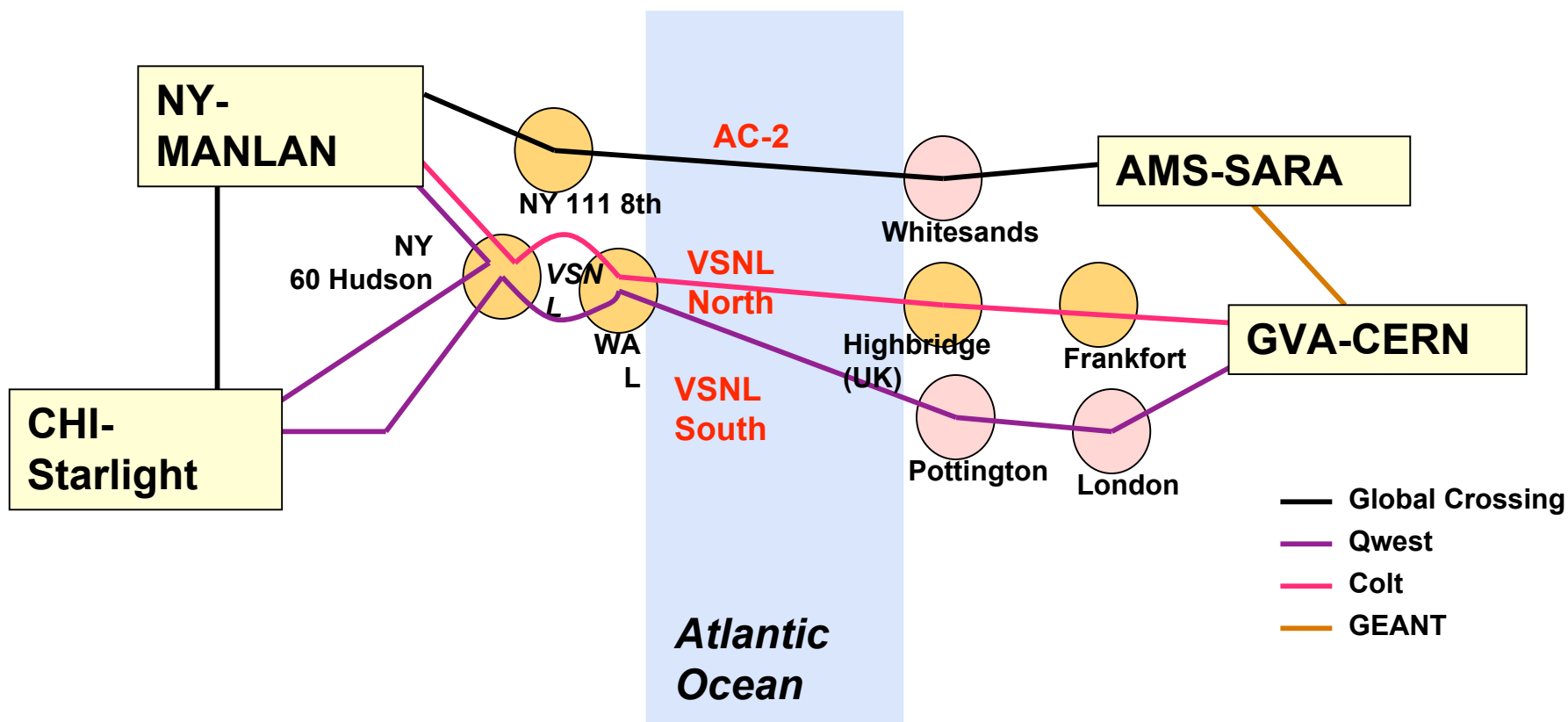


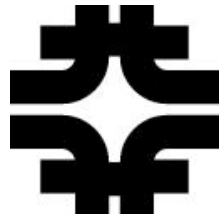
US CMS Computing Model and Networking

D. Petravick
March 29, 2007
Fermilab
ISGC 2007



USLHCNET

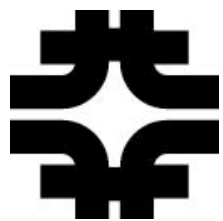




ESNet/I2-Newnet



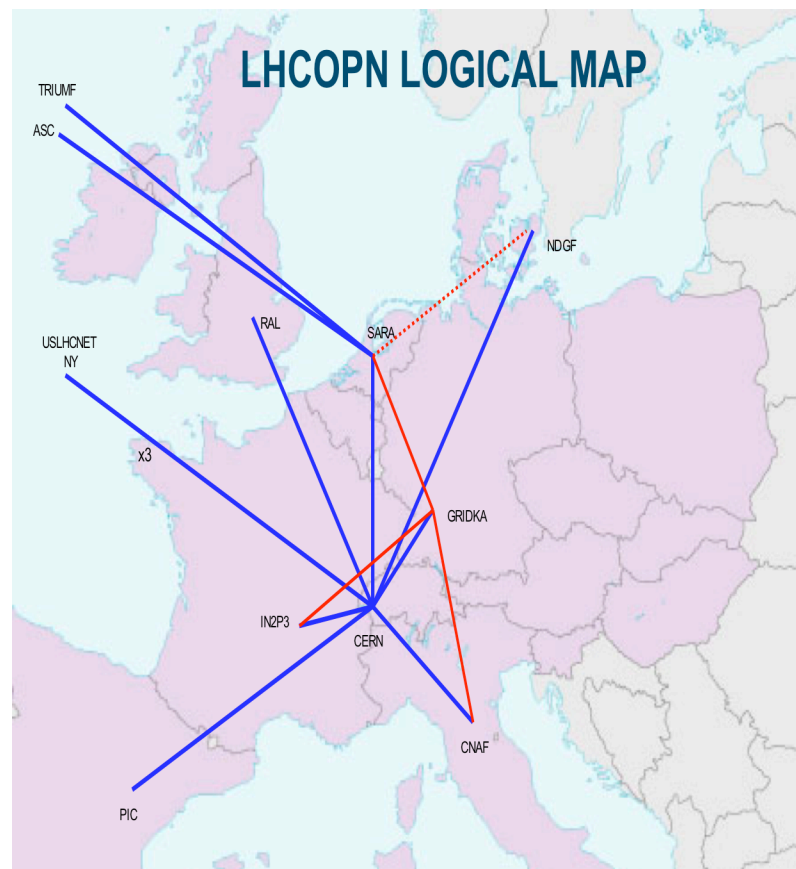
DLP -- ISGC 2007

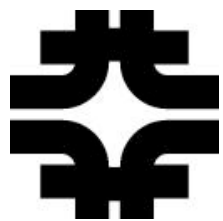


LHCOPN



- Overlay Network
- Mission:
 - Primary T0→T1
 - Secondary T1→T1.
 - Non-mission ,*T2
- Nominal Provisioning
 - 10Gb (Intra-Europe)
 - Variable (US)

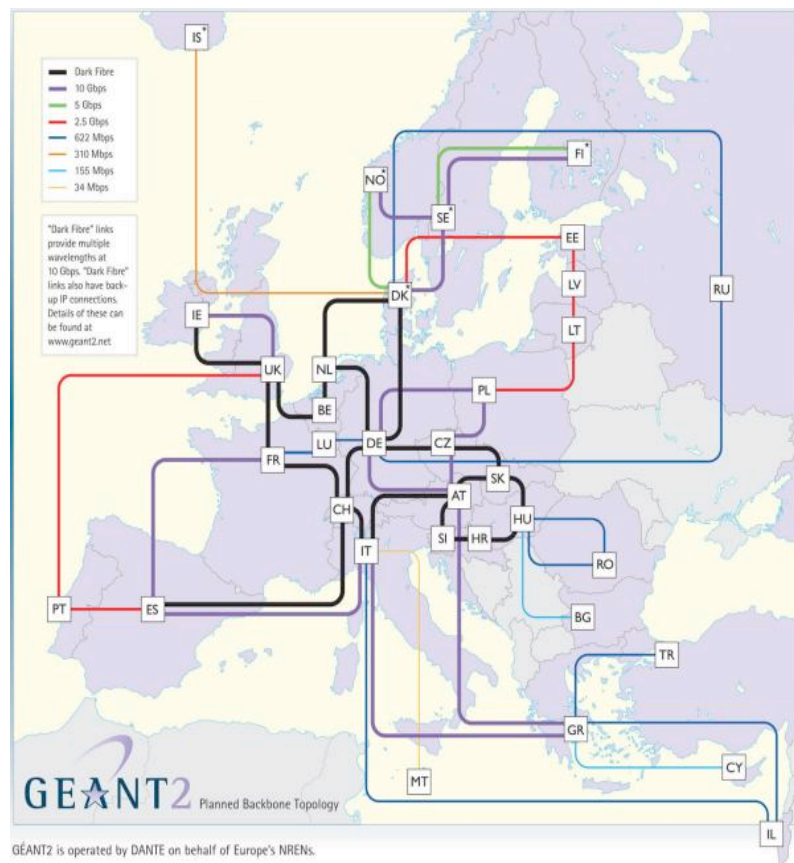


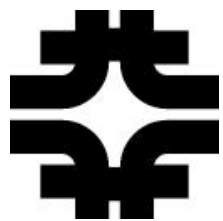


Geant2



- Links NRNENS, NRENS are the major stakeholders.
- Black lines are dark fiber
- T/A
 - Not shown are 3 T/A links.
 - Usual expectation is reciprocity.
 - ESNet does not have T/A links.

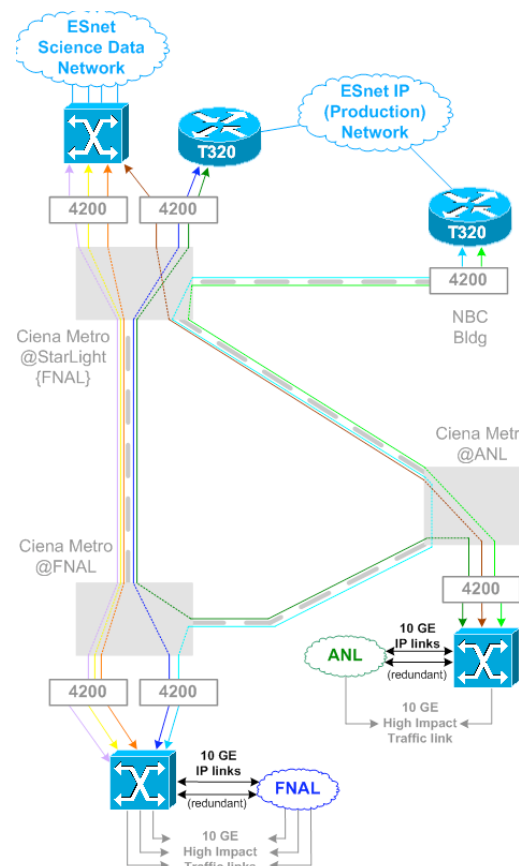


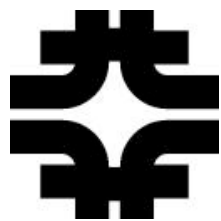


US overview



- CMS in the US:
 - T1 center at FNAL
 - T2 centers at Caltech, Florida, MIT, Nebraska, Purdue, UCSD, Wisconsin.
 - Other T3 centers.

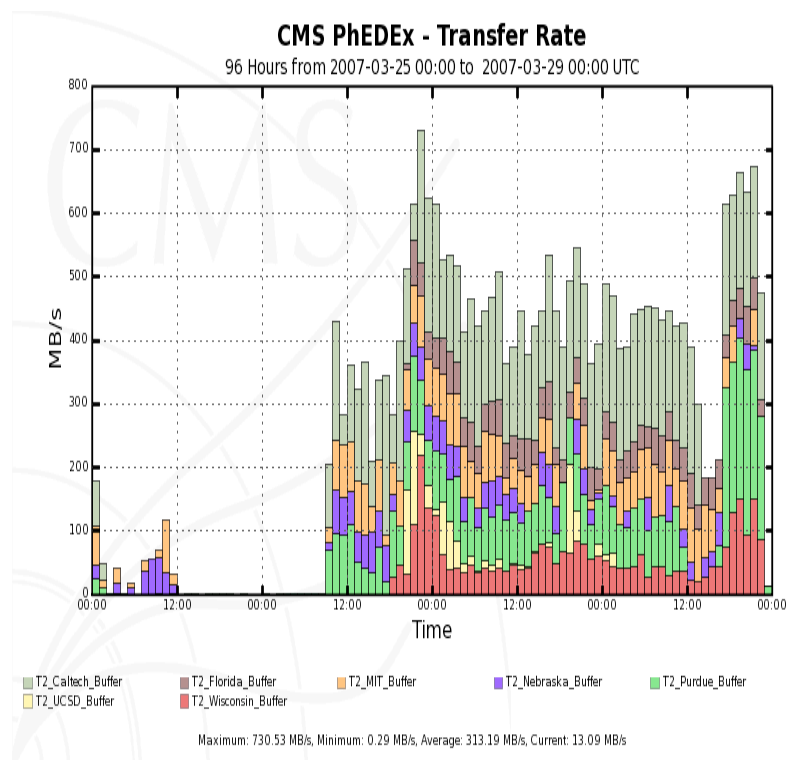


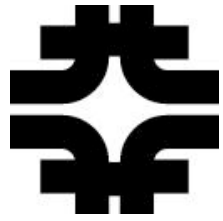


US Status



- All T2 sites having/on track to have 10 Gig Tail Ckts.
- Chicago Man is being commissioned, FNAL: \geq
 - Eight 10 Gig ckts of various types.
- Throughputs of a few 100 MB/sec to/From FNAL are routine.
- Evaluation is in context of production system.

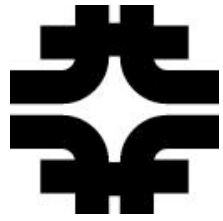




Current Integration Focus



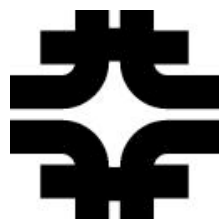
- Commissioning the data model, compared to the end sites.
- Applying perfsonar to Circuits
- Removing/Mitigating end-host bugs and problems.
- Investigating issues with managed bandwidth networks.
- Tier 3
 - Organized under umbrella of OSG
 - Networking workshops organized by Internet 2.



Computing Model Significant Xfers



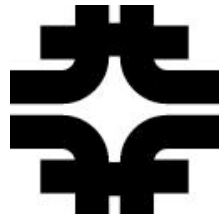
Transfer	Intra-US	Trans-Oceanic
T0 -> us T1	N/A	USLHCNET/ LHCOPN (Nren/Geant/ESNet tertiary)
Us T1->T1	N/A	USLHCNET/LHCOPN Nren/Geant/Esnet
US T1 -> ASGC	N/A	Apan/Twaren
FNAL>T2 Europe FNAL > asia	Various	Nren/Geant/I2 Esnet/Geant/Nren Esnet/Apan/Twaren



Newer Work



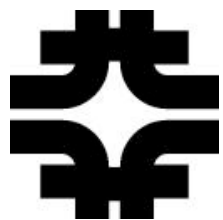
- Appreciated: The computing model implies
 - FNAL giving good service to offshore T2.
 - FNAL is a huge T1! Most T2 pledges are offshore!
 - Offshore T1 giving good service to US T2s.
- Interim goal:
 - 80% of rate compared to distant intra-EU xfers.
 - Viewed from the deployed data system.
- New role! Need to work with I2/G2/ESNET
 - The Reach of USLHCNET is CERN + partial T1's
 - Seems to imply reciprocal provisioning.



Results Oct Workshop



- Global T2 transfers recognized as important element of CMS data model.
- Incumbent Routed Networks have the capability and mission to support this, though no add'l resources
- Idea: Pursue USLHCnet provisioning of bandwidth for ESN Net G2 peering.
- Subsequently -- Achieved some exemplars of comparable rate in context of computing model



Status:

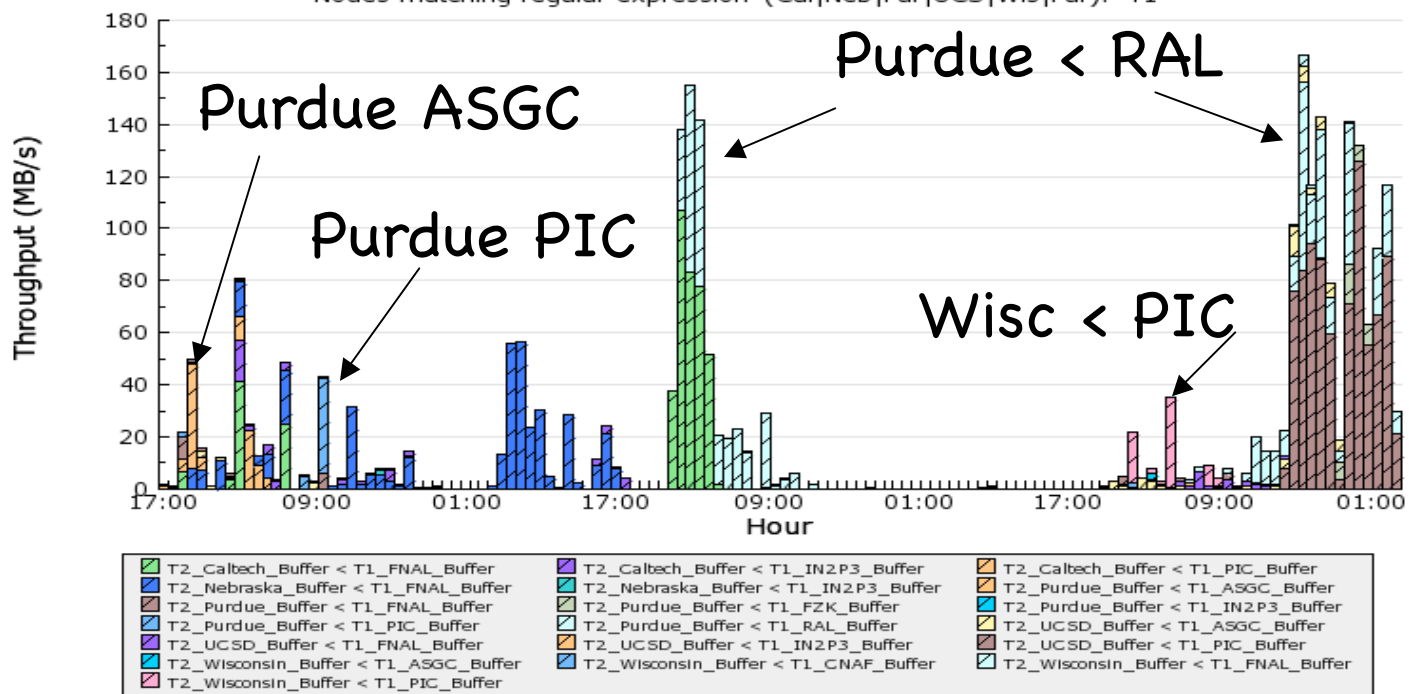
T1 -> T2 CSA06

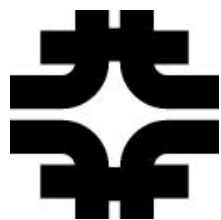
Peak T/A >70 MB/sec



PhEDEx Prod Data Transfers By Link

132 Hours from 2007-01-12 17:00 to 2007-01-18 04:00 GMT
Nodes matching regular expression '(Cal|Neb|Pur|UCS|Wis|Pur).*T1'



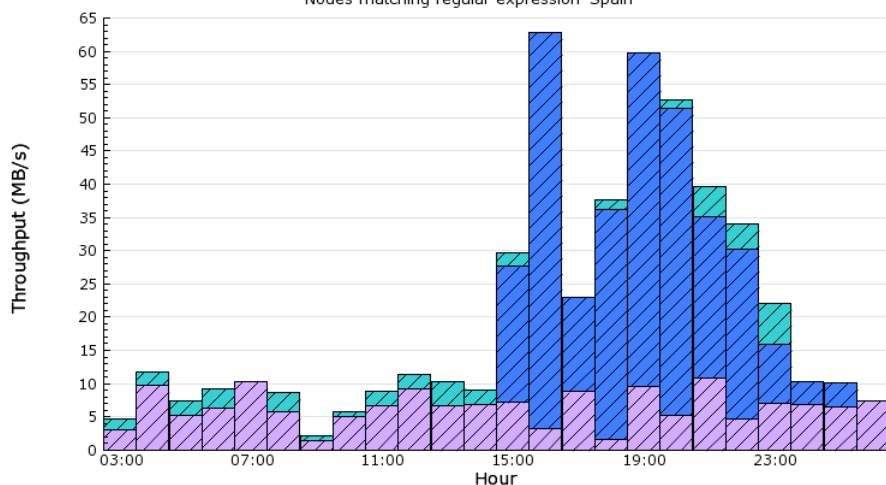


Status : FNAL->T2 (CSA06)



PhEDEx Prod Data Transfers By Link

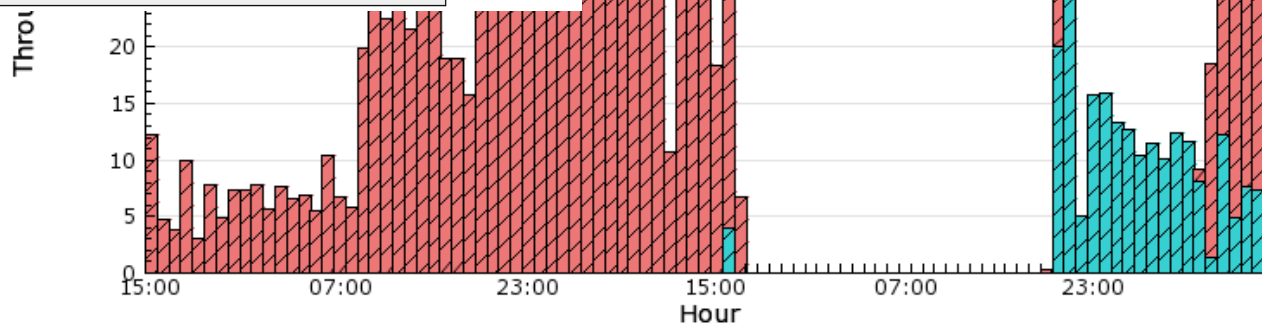
24 Hours from 2006-10-17 03:00 to 2006-10-18 02:00 GMT
Nodes matching regular expression 'Spain'



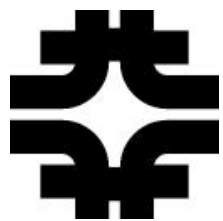
Data Transfers By Link

2006-10-15 15:00 to 2006-10-19 14:00 GMT
Nodes matching regular expression 'Estonia'

Legend for PhEDEx Prod Data Transfers By Link:
T2_Spain_Buffer < T1_PIC_Buffer (purple)
T2_Spain_IPCA < T1_FNAL_Buffer (blue)
T2_Spain_IPCA < T1_PIC_Buffer (green)



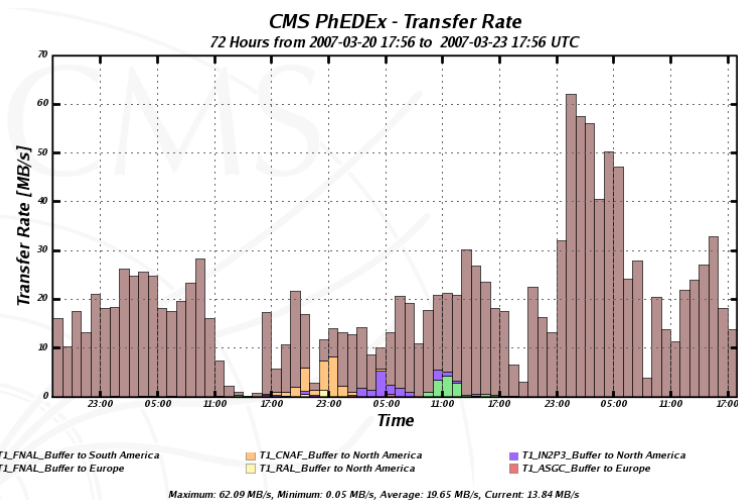
Legend for Data Transfers By Link:
T2_Estonia_Buffer < T1_FNAL_Buffer (cyan)
T2_Estonia_Buffer < T1_RAL_Buffer (red)

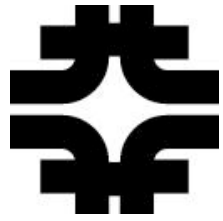


Recently



- 5-week testing cycle established in CMS.
- This week's emphasis on T[01]
- Next week's testing cycle will include T2's showing recent good throughput





Commissioning activities



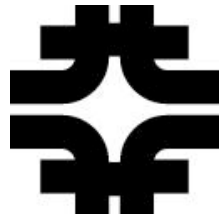
- Beginning to organize to systematically commission the global T1→T2 aspects of the computing model.
- Relying on I2 to organize T3 networking in the US.
 - Initial assumption: T3's will use access mechanisms like more like T2's than not.
- Hope to use Wide Area working group meeting to resolve networking level issues.



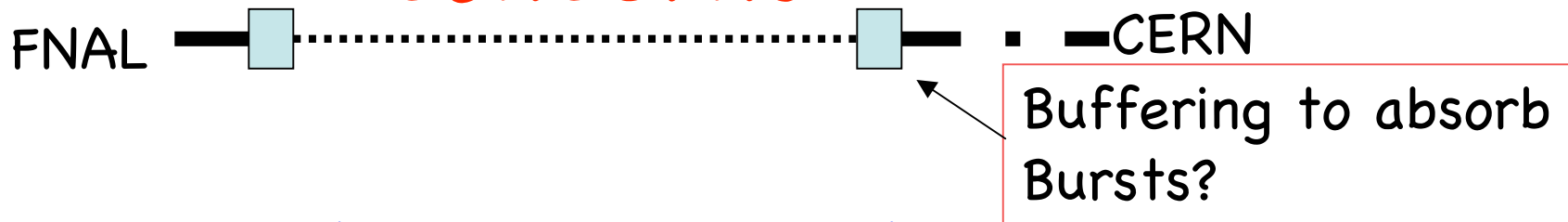
Circuits



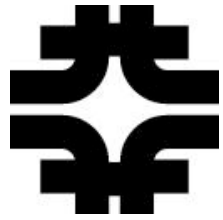
- FNAL has less-than-layer-3-circuits to
 - OPN, IN2P3, Purdue, Caltech, Florida, Wisconsin, UCSD.
 - Circuits are significant, but not ubiquitous.
- Circuits are made of segments provided by many networks.
 - No routing? No diagnostics.
- We are
 - early deployers, collaborating with PerfSonar framework. see this: http://cnmdev.lrz-muenchen.de/e2e/lhc/G2_E2E_view_e2elink_FERMI-IN2P3-IGTMD-001.html



Managed bandwidth concerns.



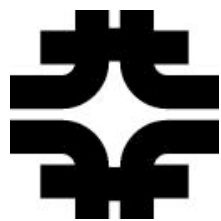
- Newnet and USLHCNET will deploy Cienna Core Directors to supply managed bandwidth ckts.
 - New! Thin Pipe Between Thick Pipe
 - Investigate:
 - Will there will be pile-up,
 - congestion at beginning of thin pipe.
 - USLHCNET: “dynamic control” What does this mean?



End Host work



- June: Significant crashing of end hosts, some variable rate.
- Linux Kernel investigations, results:
 - Tuning guide for 320bit intel servers
 - Basically eliminated crashes at the T1
 - Priority inversion in Linux Scheduler
 - Local service -> interactive
 - Staging looked like batch
 - Patch in Scientific Linux.
 - "Kernel buzz"



Summary



- Intra-US provisioning on track.
- T/A work:
 - Transfers can approach order 1 GB/sec.
 - Two sources of T/A provisioning T0 - USLHC, T2 incumbents.
 - Attempting to use USLHCNET as underlying provisioning for both.
- Original Investigations needed to mitigate/remove bugs in Linux Kernel.
 - Separate work to get long term fix.
- Proposed New features in network
 - Must work to understand impact.