



# SPRACE Site Report

Guilherme Amadio

*SPRACE – UNESP*

# Computing resources

---

- 144 worker nodes
  - Physical CPUs: 288
  - Logical CPUs (cores): 1088
  - HEPSpec06: 13698
- 02 head nodes
  - CE: HTCondor-CE job gateway and HTCondor job scheduler
  - SE: dCache distributed storage system
- 04 auxiliary servers
  - Authentication: LDAP, GUMS
  - Shared filesystems: NFS, CVMFS
  - Proxy services (Frontier Squid)
  - Support services: monitoring, VM servers, DNS server, etc.

# Computing resources (cont.)

---

- 01 job submission server
  - [access.sprace.org.br](http://access.sprace.org.br) (Red Hat 6)
  - Work environment for:
    - ❑ Job submission (HTCondor stack)
    - ❑ File transfers (GridFTP clients, dCache clients, Globus clients)
    - ❑ C/C++, Fortran, Python
    - ❑ ROOT, CMSSW
- 12 storage servers
  - 1 PB (raw), 0.85 PB (effective)
  - 10 Gbps links (LAN, WAN)
  - Usage percentage: 81%

# SPRACE servers in the datacenter

---

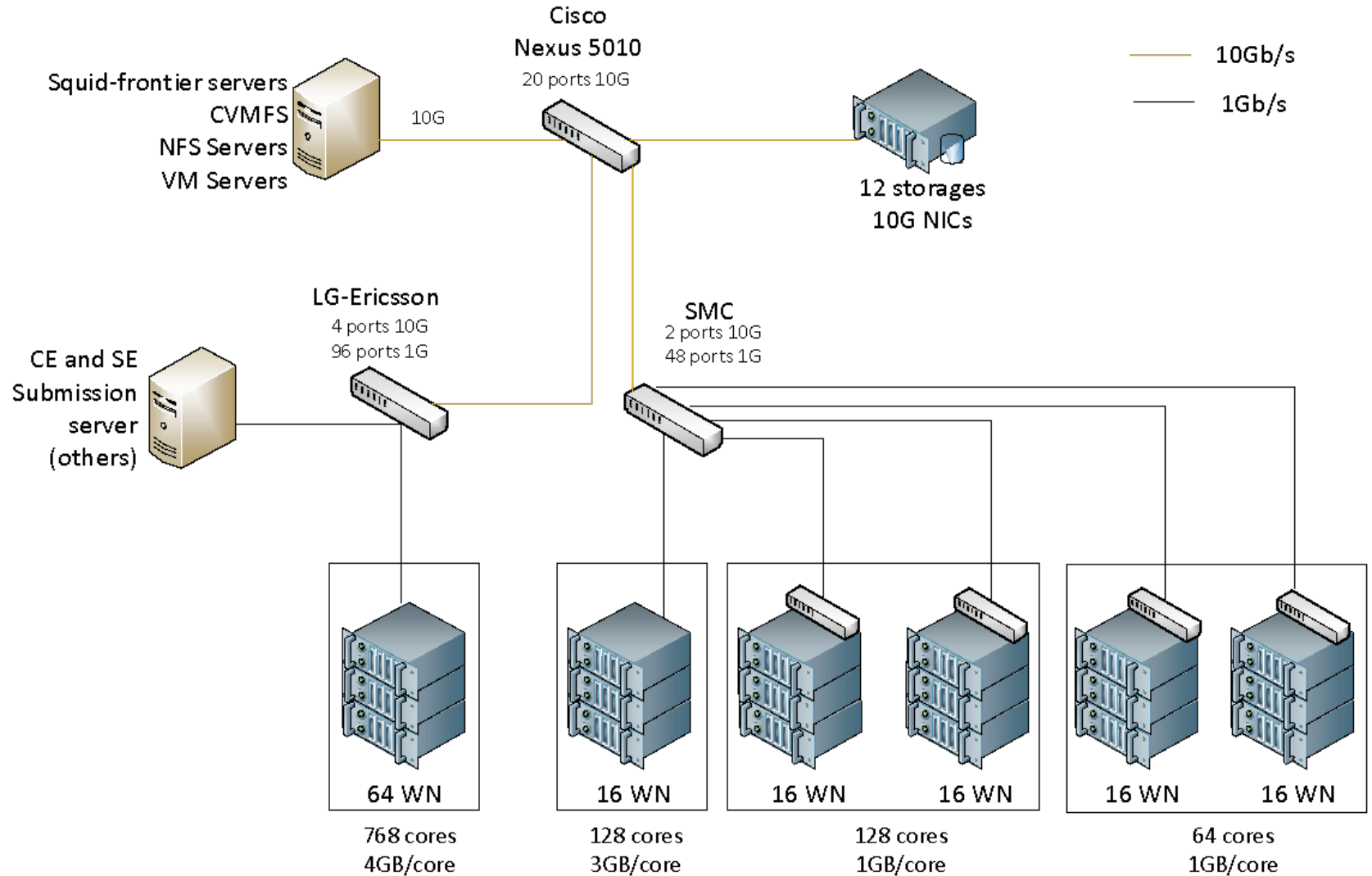


# Network Infrastructure

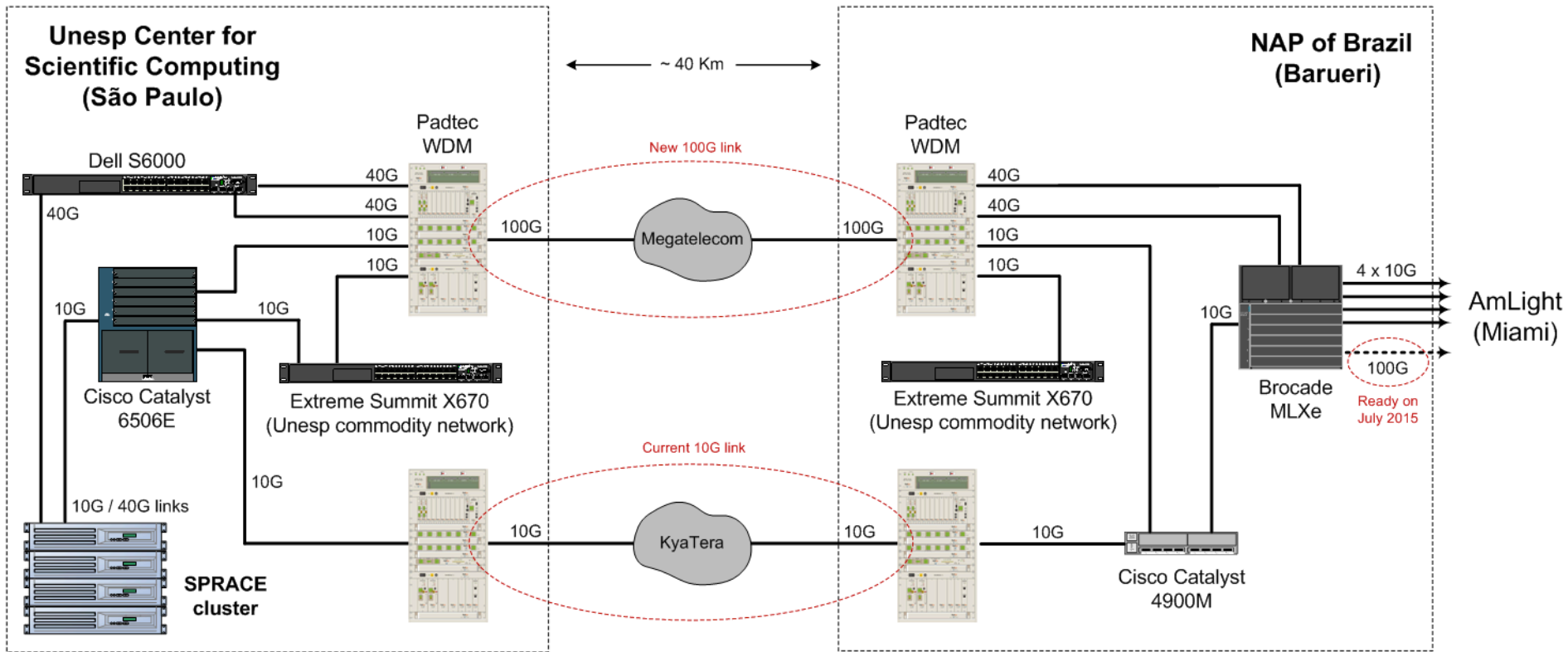
---

- LAN (cluster internal connections)
  - Worker nodes: 1 Gbps (to top-of-rack switches)
  - NFS, CVMFS, Frontier, VM servers, storage servers: 10 Gbps
- MAN (SPRACE to São Paulo Academic Network provider)
  - Nowadays: 10Gbps (2 redundant fiber links)
  - Upgrade:
    - ❑ 100 Gbps (2x 40G + 2x 10G links – deployed / operational in May 2015)
    - ❑ Links fully dedicated, independent from university commodity network)
- WAN (ANSP to AmLight in Miami)
  - Nowadays: 4x 10Gbps (2 Pacific side, 2 Atlantic side - ANSP & RNP)
  - Near future (July 2015):
    - ❑ 100Gbps (AmLight OpenWave project)

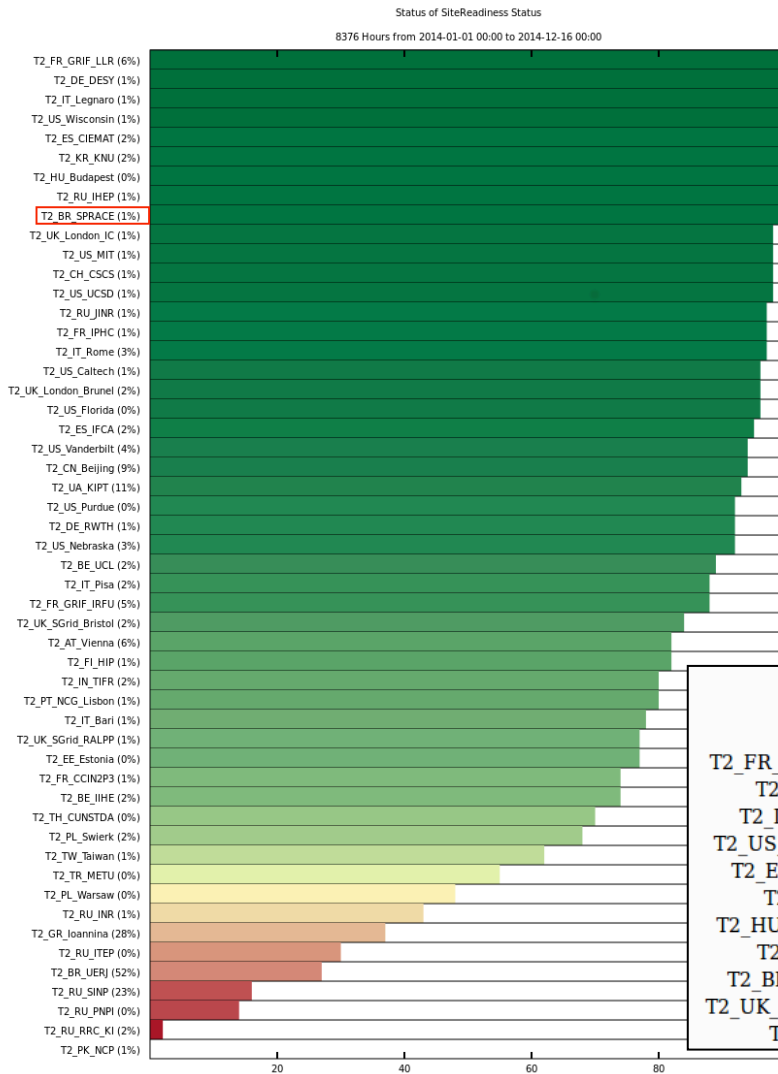
# SPRACE Internal Connectivity



# SPRACE External Connectivity

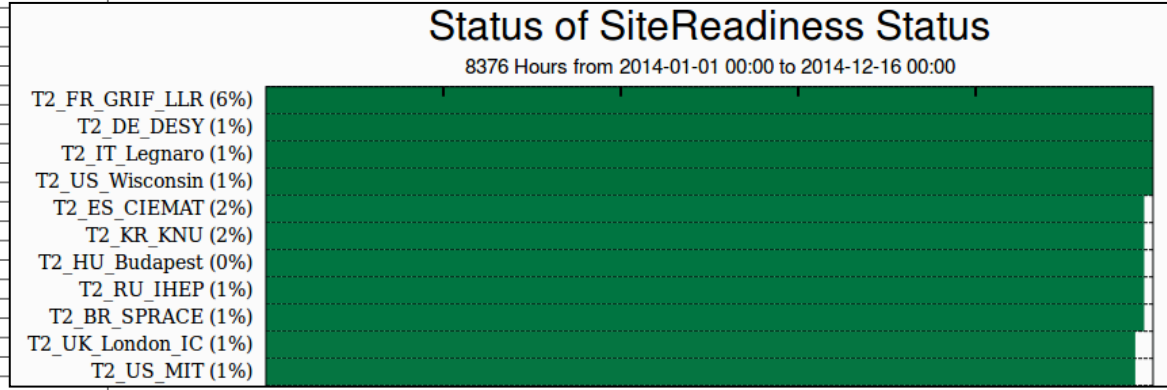


# Site Readiness



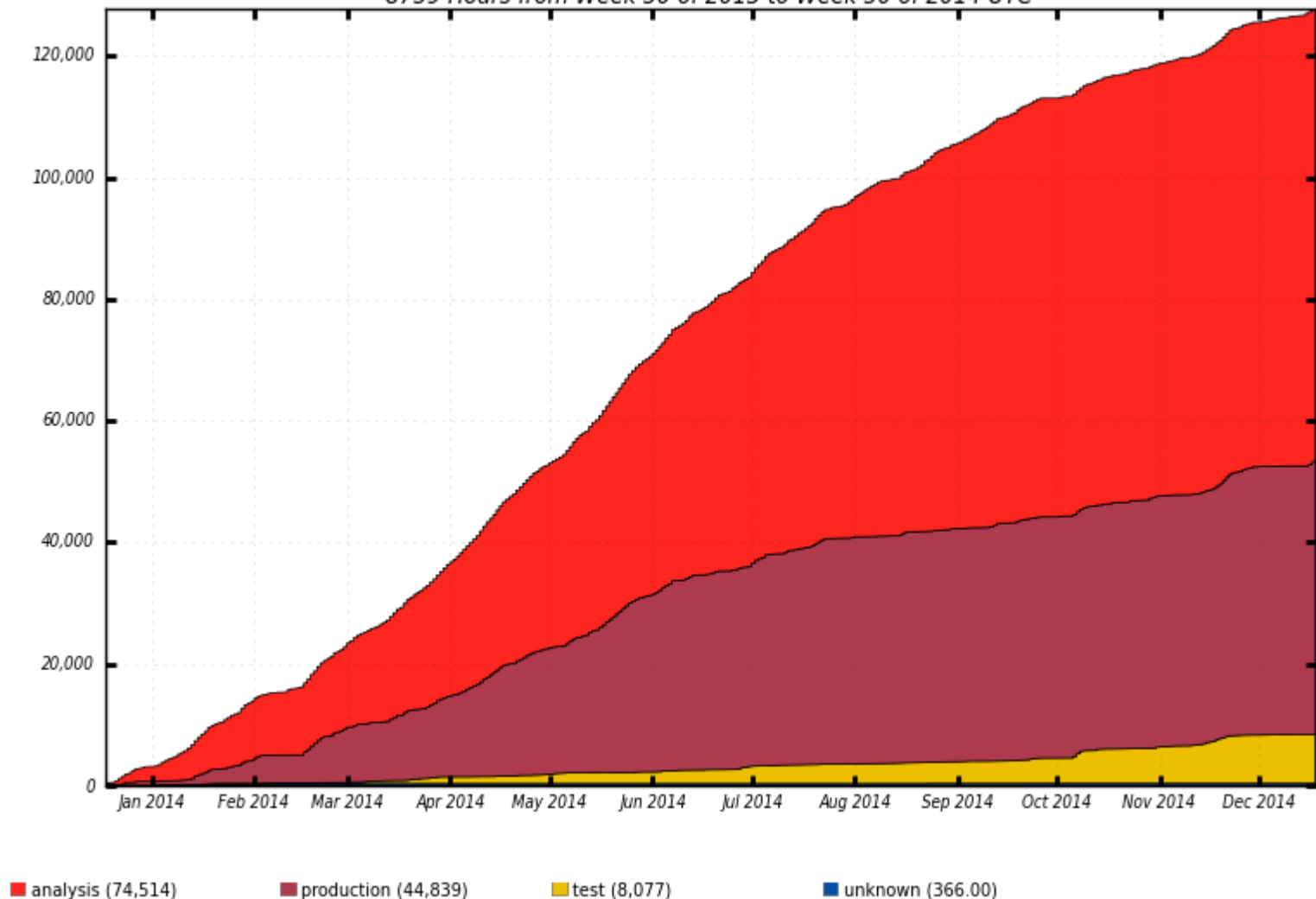
– Good results in availability and reliability

- ❑ Site admin hard work
- ❑ Datacenter (network, power, refrigeration) stability





days/day: CPU consumption All Jobs  
8759 Hours from Week 50 of 2013 to Week 50 of 2014 UTC



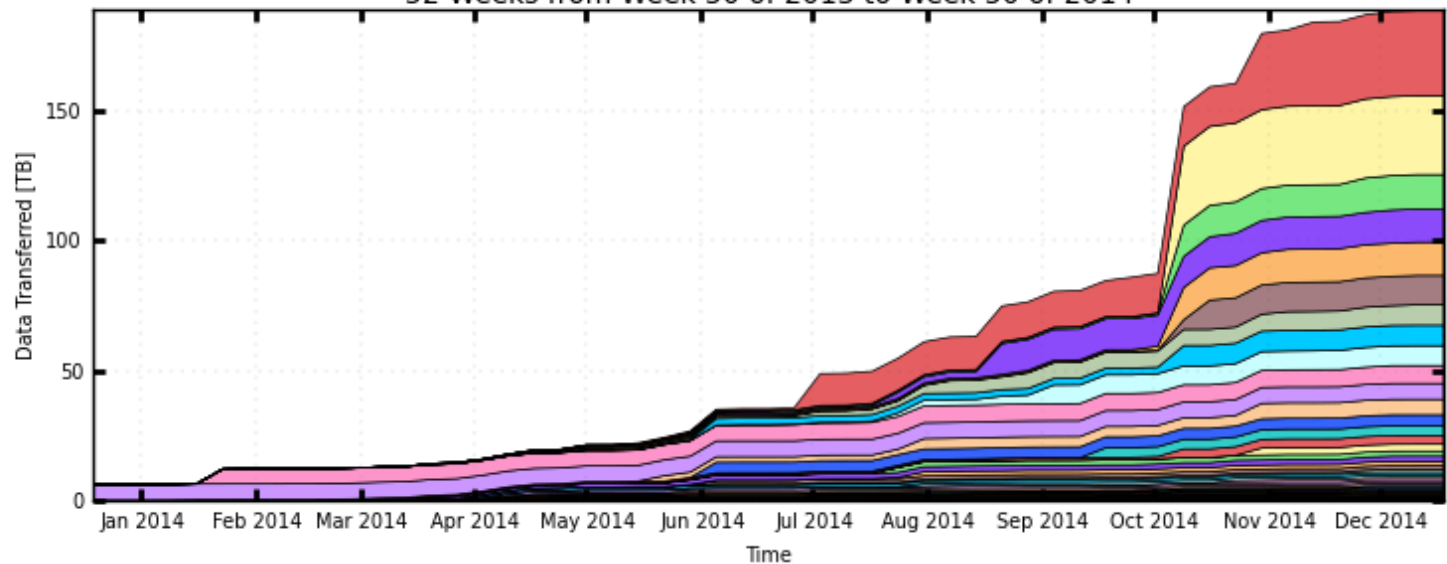
Total: 127,796 , Average Rate: 0.00 /s

CMS Jobs: CPU hours (days)  
December 2013 to December 2014

# Phedex Transfers

**188.7 TB** transferred **to** SPRACE during 2014

**CMS PhEDEx - Cumulative Transfer Volume**  
52 Weeks from Week 50 of 2013 to Week 50 of 2014



- |                       |                     |                   |                        |                     |
|-----------------------|---------------------|-------------------|------------------------|---------------------|
| ■ T1_US_FNAL_Disk     | ■ T2_US_Caltech     | ■ T2_CH_CERN      | ■ T1_ES_PIC_Disk       | ■ T2_IT_Legnaro     |
| ■ T2_US_Nebraska      | ■ T1_UK_RAL_Disk    | ■ T2_DE_DESY      | ■ T1_FR_CCIN2P3_Disk   | ■ T1_IT_CNAF_Disk   |
| ■ T3_US_Colorado      | ■ T2_US_FNAL_Buffer | ■ T2_US_Florida   | ■ T2_EE_Estonia        | ■ T2_US_Wisconsin   |
| ■ T2_KR_KNU           | ■ T2_IT_Rome        | ■ T1_RU_JINR_Disk | ■ T2_IT_Bari           | ■ T1_IT_CNAF_Buffer |
| ■ T1_DE_KIT_Disk      | ■ T2_US_Purdue      | ■ T2_RU_JINR      | ■ T2_UK_SGrid_RALPP    | ■ T2_US_Vanderbilt  |
| ■ T2_ES_CIEMAT        | ■ T1_UK_RAL_Buffer  | ■ T2_US_UCSD      | ■ T1_FR_CCIN2P3_Buffer | ■ T2_ES_IFCA        |
| ■ T2_FR_CCIN2P3       | ■ T2_CH_CSCS        | ■ T2_US_MIT       | ■ T2_FR_GRIF_LL        | ■ T1_ES_PIC_Buffer  |
| ■ T2_UK_London_IC     | ■ T2_BE_IHHE        | ■ T2_IT_Pisa      | ■ T1_DE_KIT_Buffer     | ■ T2_FI_HIP         |
| ■ T2_DE_RWTH          | ■ T2_RU_IHEP        | ■ T2_AT_Vienna    | ■ T2_FR_IPHC           | ■ T2_HU_Budapest    |
| ■ T2_UK_London_Brunel | ■ T2_BE_UCL         | ■ T2_FR_GRIF_IRFU | ■ T2_CN_Beijing        | ■ ... plus 2 more   |

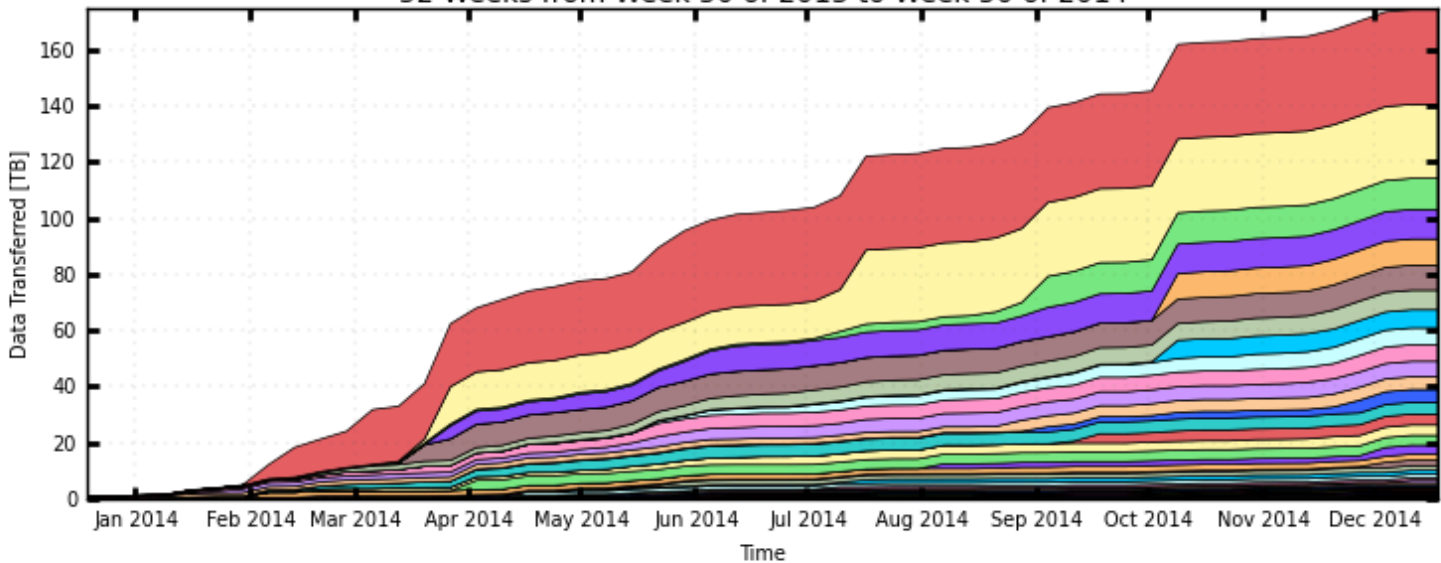
Total: 188.70 TB, Average Rate: 0.00 TB/s

# Phedex Transfers

**174.1 TB** transferred **from** SPRACE during 2014

## CMS PhEDEx - Cumulative Transfer Volume

52 Weeks from Week 50 of 2013 to Week 50 of 2014



- |                      |                        |                    |                       |                    |
|----------------------|------------------------|--------------------|-----------------------|--------------------|
| ■ T1_US_FNAL_Disk    | ■ T3_US_FNALLPC        | ■ T2_CH_CERN       | ■ T2_DE_DESY          | ■ T2_US_UCSD       |
| ■ T1_FR_CCIN2P3_Disk | ■ T1_DE_KIT_Buffer     | ■ T2_US_MIT        | ■ T0_CH_CERN_Export   | ■ T1_UK_RAL_Buffer |
| ■ T1_IT_CNAF_Buffer  | ■ T1_FR_CCIN2P3_Buffer | ■ T2_BE_IHHE       | ■ T1_US_FNAL_Buffer   | ■ T2_AT_Vienna     |
| ■ T1_ES_PIC_Buffer   | ■ T2_US_Nebraska       | ■ T2_DE_RWTH       | ■ T2_IT_Pisa          | ■ T2_US_Florida    |
| ■ T1_IT_CNAF_Disk    | ■ T2_CN_Beijing        | ■ T2_US_Wisconsin  | ■ T2_FR_CCIN2P3       | ■ T2_BR_UERJ       |
| ■ T1_DE_KIT_Disk     | ■ T1_ES_PIC_Disk       | ■ T1_UK_RAL_Disk   | ■ T2_IT_Legnaro       | ■ T2_FR_IPHC       |
| ■ T2_FI_HIP          | ■ T2_US_Vanderbilt     | ■ T2_PT_NCG_Lisbon | ■ T2_UK_London_IC     | ■ T2_FR_GRIF_LL    |
| ■ T2_US_Purdue       | ■ T2_KR_KNU            | ■ T2_HU_Budapest   | ■ T2_US_Caltech       | ■ T2_IT_Bari       |
| ■ T2_CH_CSCS         | ■ T2_ES_IFCA           | ■ T2_TW_Taiwan     | ■ T2_IT_Rome          | ■ T2_ES_CIEMAT     |
| ■ T2_EE_Estonia      | ■ T3_US_Colorado       | ■ T2_RU_JINR       | ■ T2_UK_London_Brunel | ■ T1_RU_JINR_Disk  |

Total: 174.10 TB, Average Rate: 0.00 TB/s

# Site Planning for 2015

---

- **New storage server**
  - ~150 TB (usable) with 40G NIC card
  - SPRACE total effective storage space will reach 1 PB
  - Should be in production in April
- **16 new worker nodes**
  - Physical CPUs: 32 (Xeon E5-2600v3) 192 cores
  - Should be in production in July/August
  - Old worker nodes will be retired: 40 servers (96 cores)
- **New datacenter border switch**
  - Dell S6000: 32x 40G ports
  - To be connected to the 100G link to ANSP (2x 40G channels)
  - Deployed and tested during SC'14
  - Should be in production in May
- **SPRACE connection to LHCONE**
  - Under discussion - should be finished in May/June
  - ANSP (Brazil) → AmLight (Miami, USA) → ManLan through AtlanticWave

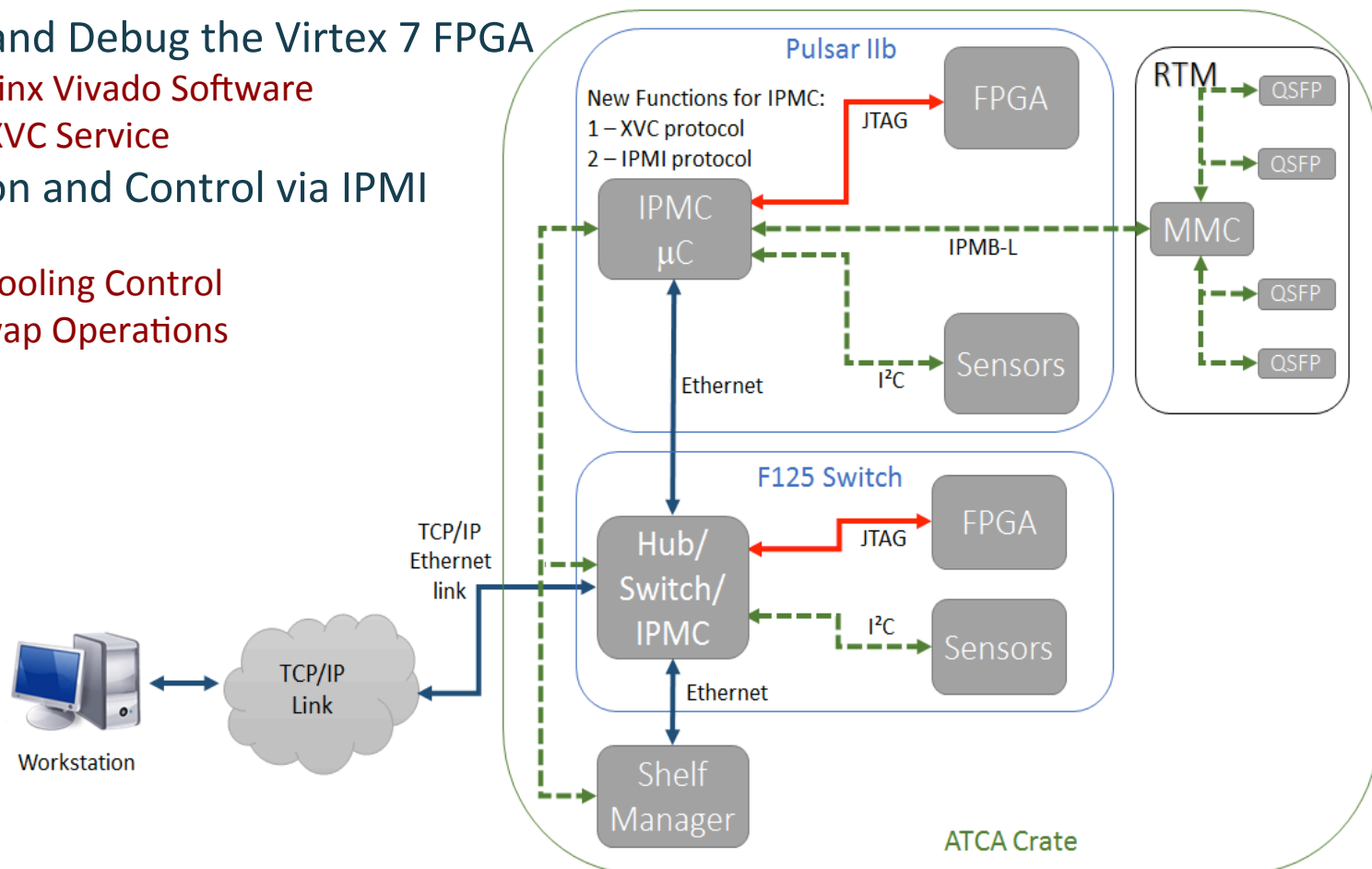
# Site Planning for 2016

---

- **New storage servers**
  - SPRACE total effective storage space will approach 1.5 PB
  - Storage servers with 40G NIC cards
- **32 new worker nodes**
  - Physical / logical CPUs: investigating
- **Old worker nodes will be retired**
  - 24 servers (96 cores)
  - Retirement should happen as soon as new worker nodes enter production phase
- **New control servers**
  - CE, SE will be upgraded
  - New servers will have 40G NIC cards (LAN, WAN)
  - Old CE, SE will be used to upgrade auxiliary servers
- **New datacenter border switch**
  - A second Dell S6000 should be bought for redundancy purpose

# SPRACE CMS L1TT Collaboration

- New Remote Control Functions for ATCA Carrier Pulsar IIb IPMC for the L1TT Operator:
  - Program and Debug the Virtex 7 FPGA
    - ❑ Use Xilinx Vivado Software
    - ❑ IPMC XVC Service
  - Supervision and Control via IPMI protocol
    - ❑ Auto Cooling Control
    - ❑ Hot Swap Operations



# IPCC: GeantV Development

---

- IPCC — Intel Parallel Computing Center
  - Parallelization and Vectorization of GeantV
- Geant: Geometry and Tracking
  - Framework to simulate passage of particles through matter
- Geant4MT: Multi-threaded version of Geant4
  - Event-level parallelism
- GeantV: Geant Vector Prototype
  - Multi-threaded and vectorized
  - Take advantage of new hardware
    - ❑ GPUs, Xeon Phi
  - New programming models:
    - ❑ Nvidia's CUDA, MIC architecture

# Challenges

---

- Datacenter power and cooling issues
  - New SPRACE servers will require power and cooling capacity
  - A 150KVA no-break system is ready to be installed (costly)
  - Cooling issue: there is no room for new CRAC units
  - Alternatives:
    - ❑ Hot-air collectors
    - ❑ Server redistribution (thermal management)
    - ❑ Retirement of old servers
    - ❑ Hot/cold aisle containment
- Network issues
  - 4 X 40G module for ANSP Brocade switch should be bought
  - We still do not have a plan for IPv6 deployment
- Manpower
  - Short on dedicated manpower to take care of SPRACE farm