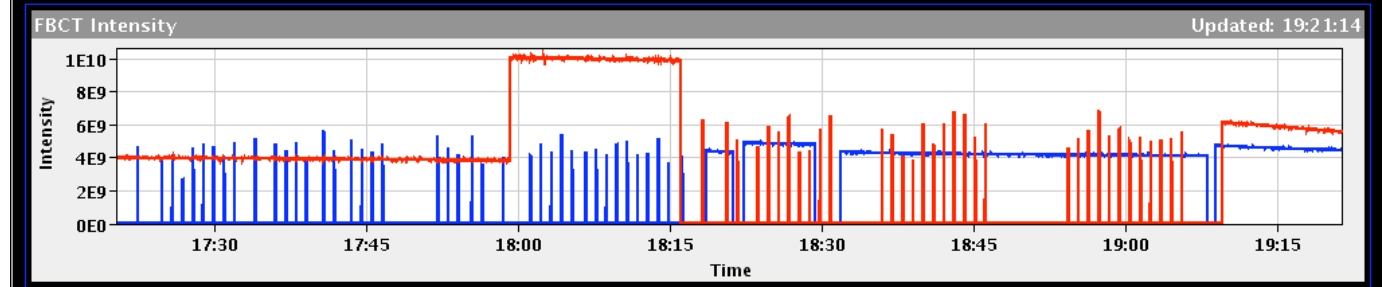# WLCG Future Network Requirements:

an overview of the WLCG Experiments use cases

*Disclaimer: for official experiment requirements please ask
to the experiments Computing Coordinators*

Alessandro Di Girolamo

INFN-CNAF / CERN-IT

CERN IT Department
www.cern.ch/it

- LHC duty cycle ~ 14 / 24 hours
  - 50 K seconds / day

- Exps trigger rate (e.g. ATLAS &CMS) ~ 200 Hz

- Peaks of activity have to be taken into account

Alessandro Di Girolamo INFN-CNAF/CERN

Data Distribution: ATLAS

Prompt Reconstruction

320 MB/s

Tier0

CAF

Calibration & Alignment
Express Stream Analysis

CASTOR

1.5 GB/s

RAW Reprocessing
HITS Reconstruction

Tier1

Tier1

Tier1

Data Distribution: ATLAS

Prompt Reconstruction

320 MB/s

Tier0

CAF

CASTOR

Calibration & Alignment
Express Stream Analysis

1.5 GB/s

RAW Reprocessing
HITS Reconstruction

50-1000 MB/s

Tier1

Tier1

Tier1

Data Distribution: ATLAS

Prompt Reconstruction

320 MB/s

Tier0

CAF

CASTOR

1.5 GB/s

Calibration & Alignment
Express Stream Analysis

RAW Reprocessing
HITS Reconstruction

50-1000 MB/s

Tier1

Tier1

Tier1

Simulation
Analysis

50-500 MB/s

Tier2

Tier2

Tier2

hierarchical model

Data Distribution: CMS

Prompt Reconstruction

Tier0

CAF

Calibration Express
Stream Analysis

CASTOR

600 MB/s

Re-reconstruction
Skimming

50-1000 MB/s

Tier1

Tier1

Tier1

~20 MB/s

50-500 MB/s

Simulation
Analysis

Tier-2   Tier-2   Tier-2   Tier2   Tier-2   Tier-2   Tier-2   Tier2

mesh model

CERN IT Department
www.cern.ch/it
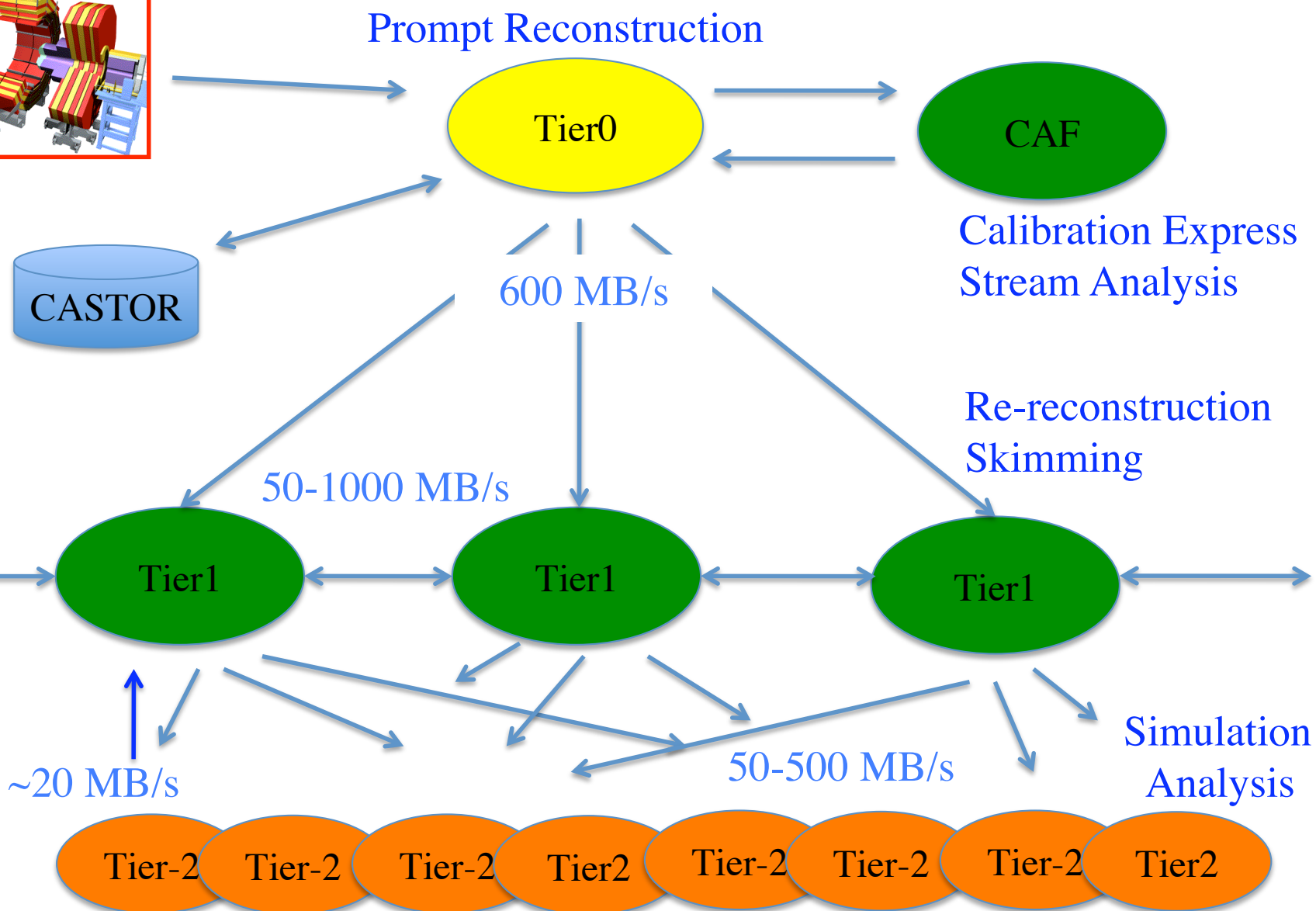
Data Distribution: LHCb

- ## Similar to ATLAS and CMS,
  - but data volume much smaller

- ## Tier2s computing model function: only simulation (not analysis)
  - Some big Tier2 might support analysis:
    - 40 MB/s IN, 10 MB/s OUT
    - Those Tier2s will be directly connected to the T0 (not to the Tier1 as in a per cloud model)

Data Distribution: ALICE

Computing Model

Each ESD and AOD file produced at a T2 site has 4 replicas:
- Local replica
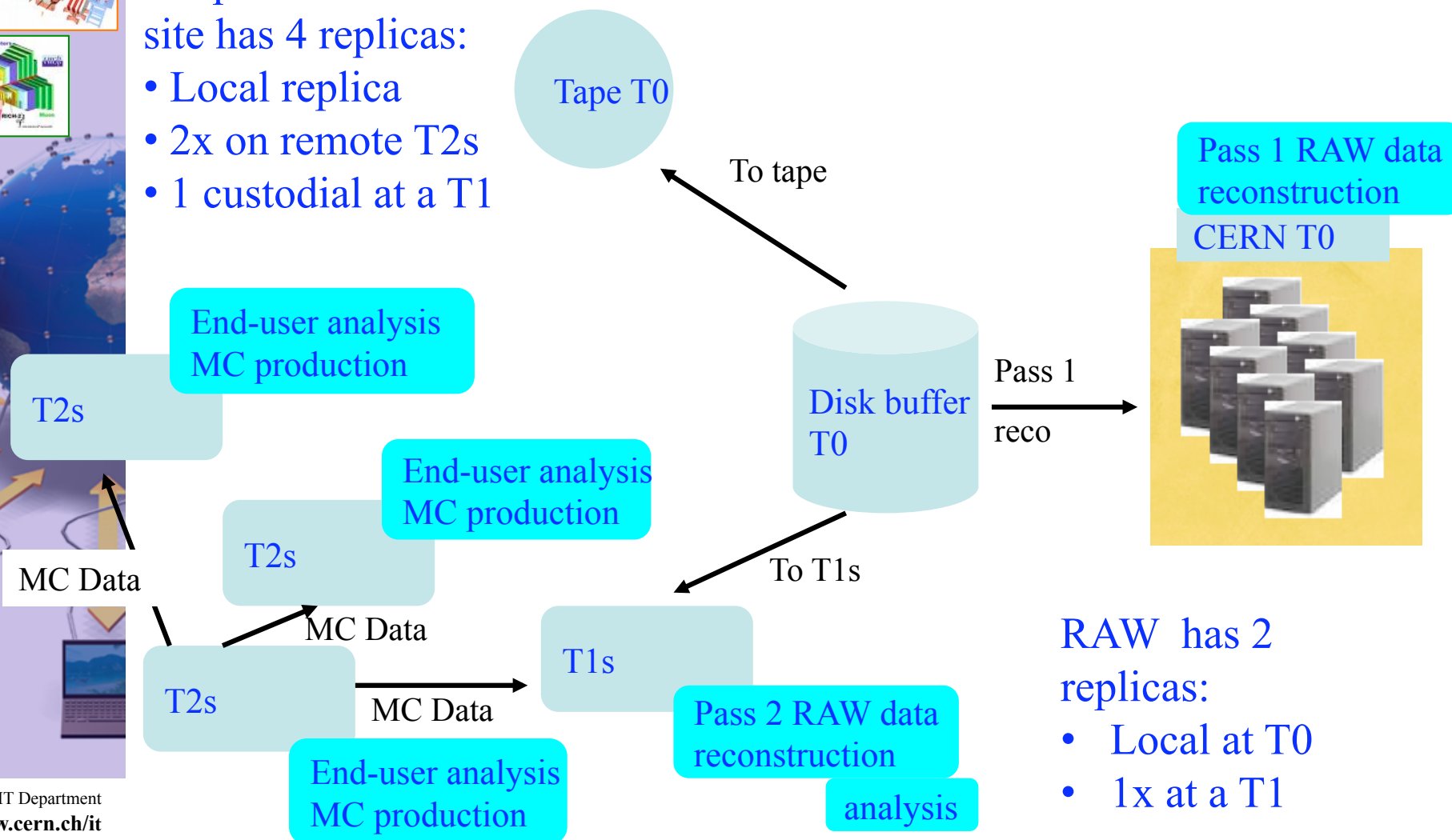- 2x on remote T2s
- 1 custodial at a T1

Tape T0

To tape

Pass 1 RAW data reconstruction CERN T0

Disk buffer T0

Pass 1 reco

End-user analysis MC production

T2s

End-user analysis MC production

T2s

MC Data

To T1s

MC Data

T1s

T2s

MC Data

Pass 2 RAW data reconstruction

End-user analysis MC production

analysis

RAW has 2 replicas:
- Local at T0
- 1x at a T1

ALICE: bandwidth for Tier2s

- Bandwidth is a f(#CPU) – 500 CPU ~100Mb/s

  – Incoming and outgoing traffic - replication of ESDs/AODs

  – Calculated from the (average) size of MC job output and job duration, valid also for ESDs/AODs from RAW data reconstruction

- Some T2 already showed some limitations

  – Not only bandwidth: also the case for site internal network

CMS nominal requirements for Tier2s

- Example for a 200 TB Tier2

  - 100 TB analysis

  - 100 TB MonteCarlo

  – 5 TB/day IN → 60 MB/s

  – 1 TB/day OUT → 12 MB/s

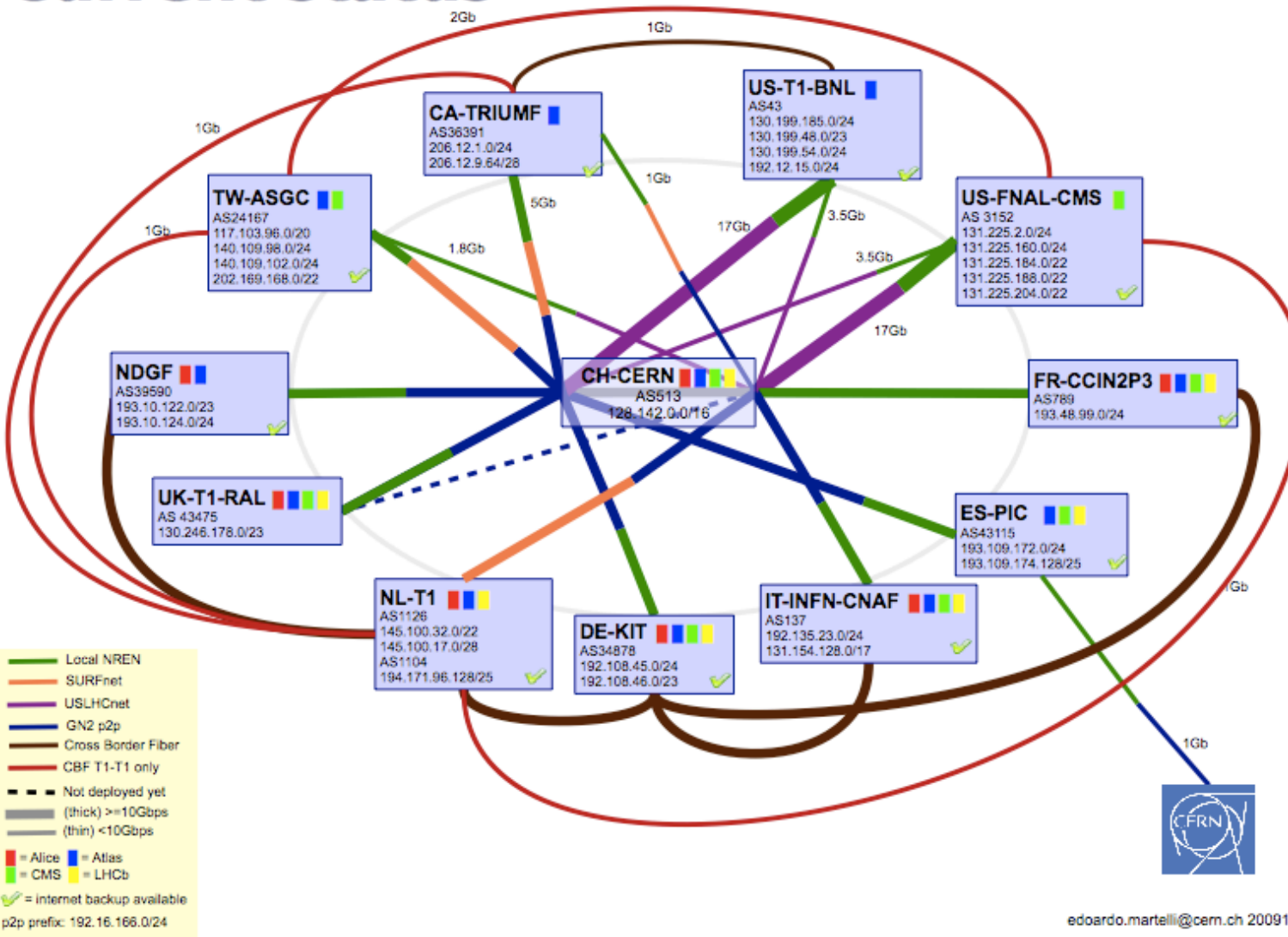 ! Data volume scale with Tier's size

Experiments data Reprocessing

- ## 3-5 times faster than normal datataking

  - Bandwidth needed: factor 3-5

  - Example for an ATLAS Tier2 (BIG: 100% share)

    - 40 MB/s times reprocessing speed + all the other activities:

      - 40 MB/s x 3

      - 40 MB/s normal data distribution
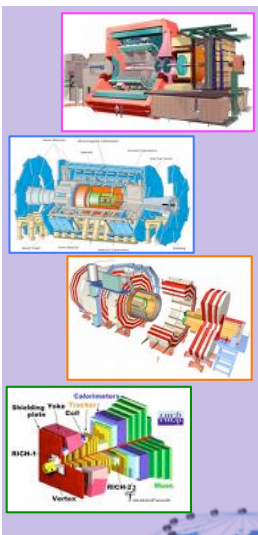
      - 40 MB/s MC data

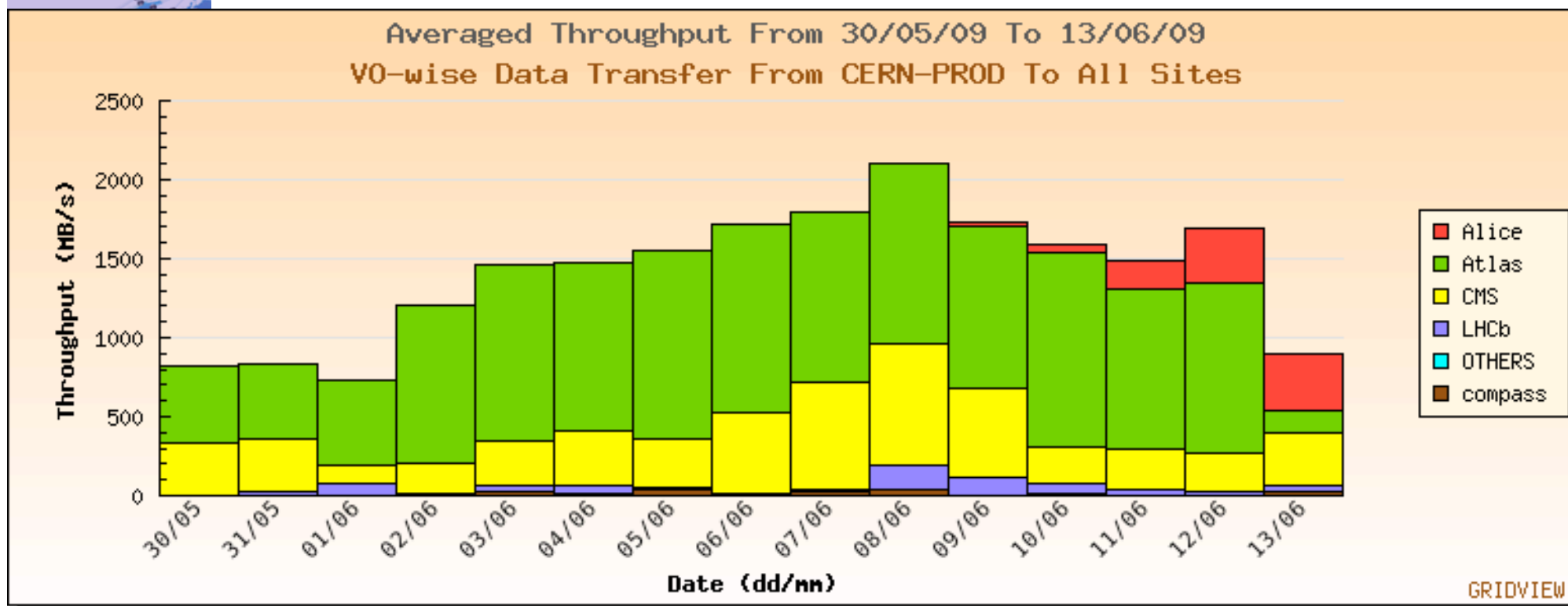# LHC OPN current status

Current status

Scale Testing for the Experiment Program '09

- Mainly 2 out of 4 exp together (ATLAS & CMS)

Averaged Throughput From 30/05/09 To 13/06/09
VO-wise Data Transfer From CERN-PROD To All Sites

Data Distribution from CERN to Tier1s ('2009-05-30 00:00' to '2009-06-13 23:30' UTC)

STEP'09

- Tier0 → Tier1, T1 ↔ T1, T1→ T2
  Data Distribution for one experiment
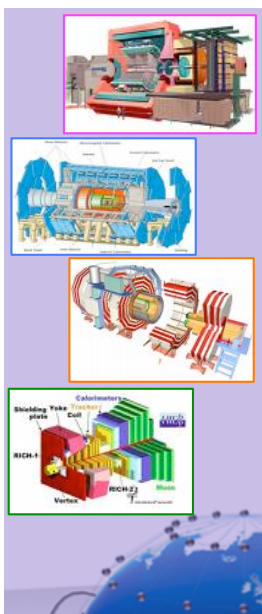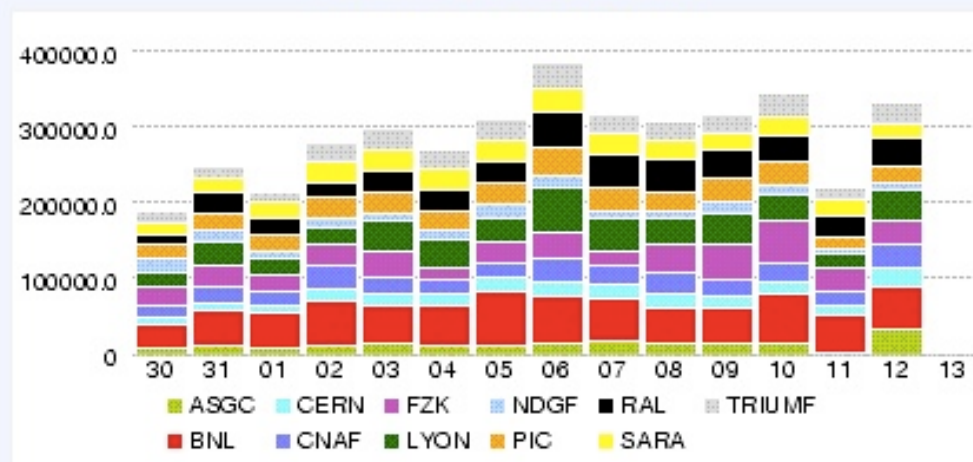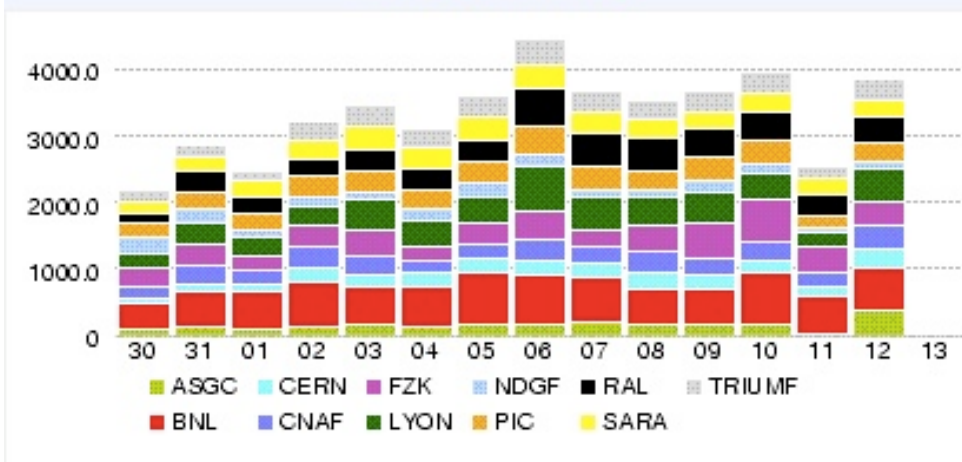
'2009-05-30 00:00' to '2009-06-13 23:30' UTC
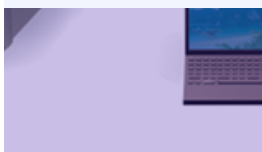
✓ ~ 3 GB/s with peaks of 5.5 GB/s

LHC OPN during STEP'09

CERN IT Department
www.cern.ch/it
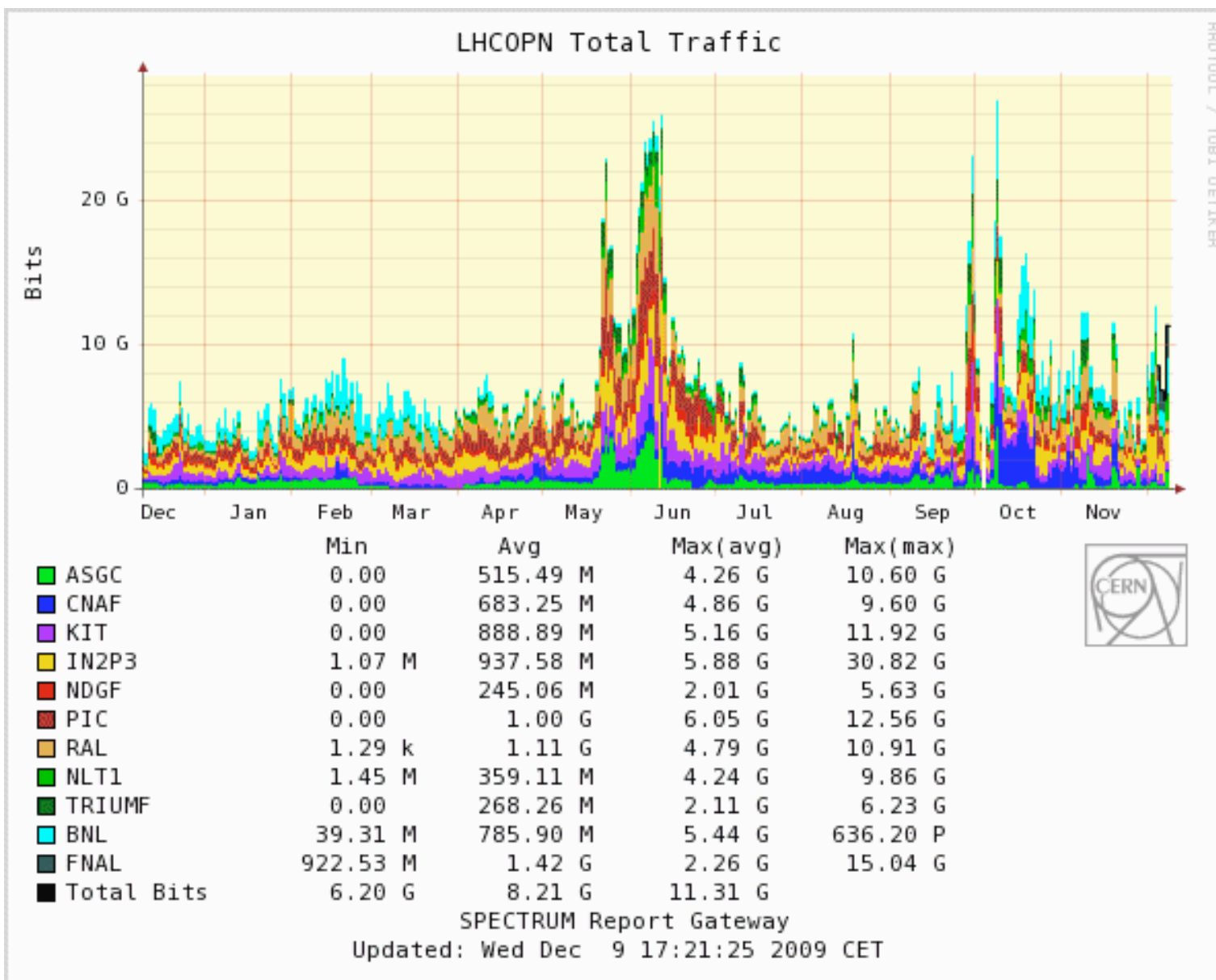
Main points

- ## RAW data to T1 tapes:

  - This makes the T0 - T1 links different from all other.
    - ✓ T0-T1 backup paths is important
  - ! Any additional features on the OPN should not endanger this prime functionality!

- ## T1-T1 data volume:

  - 'first' design: T1-T1 smaller than T0-T1
  - ✓ Now: same order of magnitude of T0-T1

- ## Tier2s: their importance is increasing

  - Very different sizes: some T2 are similar to T1
  - ! List of "golden T2s" may vary with time:
    - Some T2s may improve significantly in a matter of months

(some possible) Conclusions

✓ Tier0 → Tier1s & T1 ↔ T1: OK

- Bandwidth for (some big) Tier2s may become an issue:

  ? Will some huge Tier2 have enough bandwidth in few years from now?

  ? Is there enough network expertise at the Tier2s?

  ! … not only external bandwidth is challenging …

CERN IT Department
www.cern.ch/it