LHCOPN and LHCONE: first experience gained in RRC-KI

Eygene Ryabinkin

National Research Centre “Kurchatov Institute”

Grid’2014, Dubna

Data transfer networks for LCG

Sometimes certain sites can run peering between them if this is justified. This is private peering and any pair of sites can do that independently.

Traffic between Tier-0 and Tier-1: raw data should be quickly transferred from CERN to Tier-1, so we need bandwidth and availability guarantees. This is an LHCOPN area.

Traffic between Tier-1 and Tier-2 sites usually goes via R&N networks provided by NRENs, but it is a lot better to get only layer-2 connectivity from them and optimize traffic with networking team that is closely coupled to the Grid sites. This is an LHCONE area.

R&N networks and public Internet cover the rest of the connectivity.

I'll try to share our experience gained during past year with LHCOPN and LHCONE.

НИЦ «Курчатовский институт»

LHCOPN: what is it

It is Tier-0 –Tier-1 and Tier-1 – Tier-1 network that uses dedicated channels.

Underlying technology is BGP-powered and LHCOPN is a closed routing cloud that extends just over Tier-0 and Tier-1 centres: currently 14 ASs.

Two policy points:

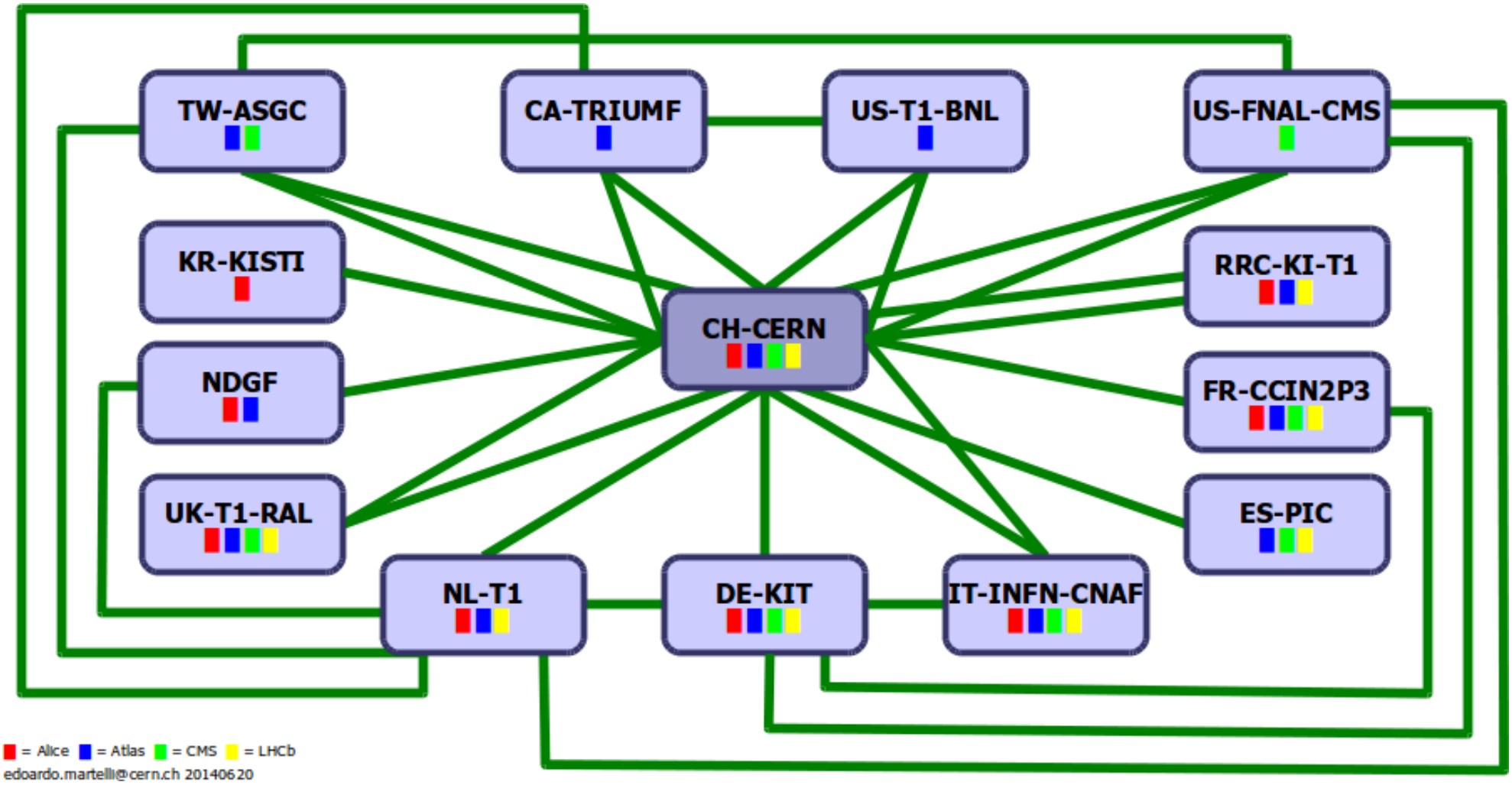
- only Grid nodes must be connected to this routing cloud and only LHC-related traffic should flow there;
- no central firewalling should be used on site-LHCOPN borders.

You must have aggregated netblock for Tier-1 (typically /24 or around that, but really no hard restrictions).

At our speeds and number of connections, firewalls become bottlenecks (or rather pricey).

НИЦ «Курчатовский институт»

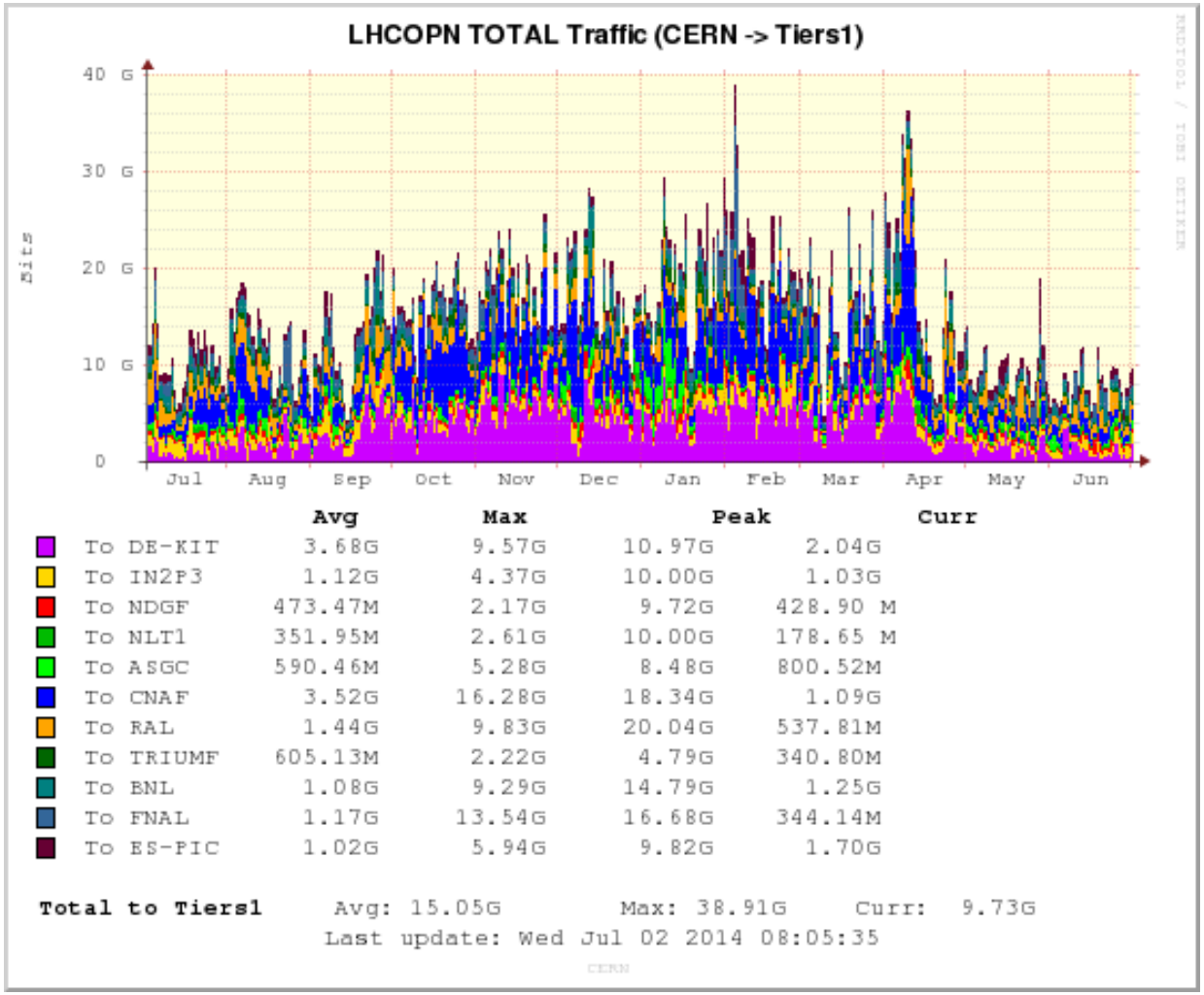
LHCOPN: topology



Color legend: red – ALICE, blue – ATLAS, green – CMS, yellow – LHCb.

НИЦ «Курчатовский институт»

LHCOPN: pushed data



Last year stats: average traffic – 15 Gbit/sec, peak traffic – 39 Gbit/sec.

НИЦ «Курчатовский институт»

LHCONE: your next best network

At least for past 20 years Grid people want to control networks ;)

That's still hard, but success of LHCOPN made people to extend this model to T1 –T2 and T2 –T2 communications.

Key point of LHCOPN success is a small number of entities (couple of dozens): it makes operations rather smooth.

Now we $>$ 200 Tier-2 sites $\Rightarrow$ bare LHCOPN model is not an option.

Moreover, star-like or near full-mesh topologies isn't an option too.

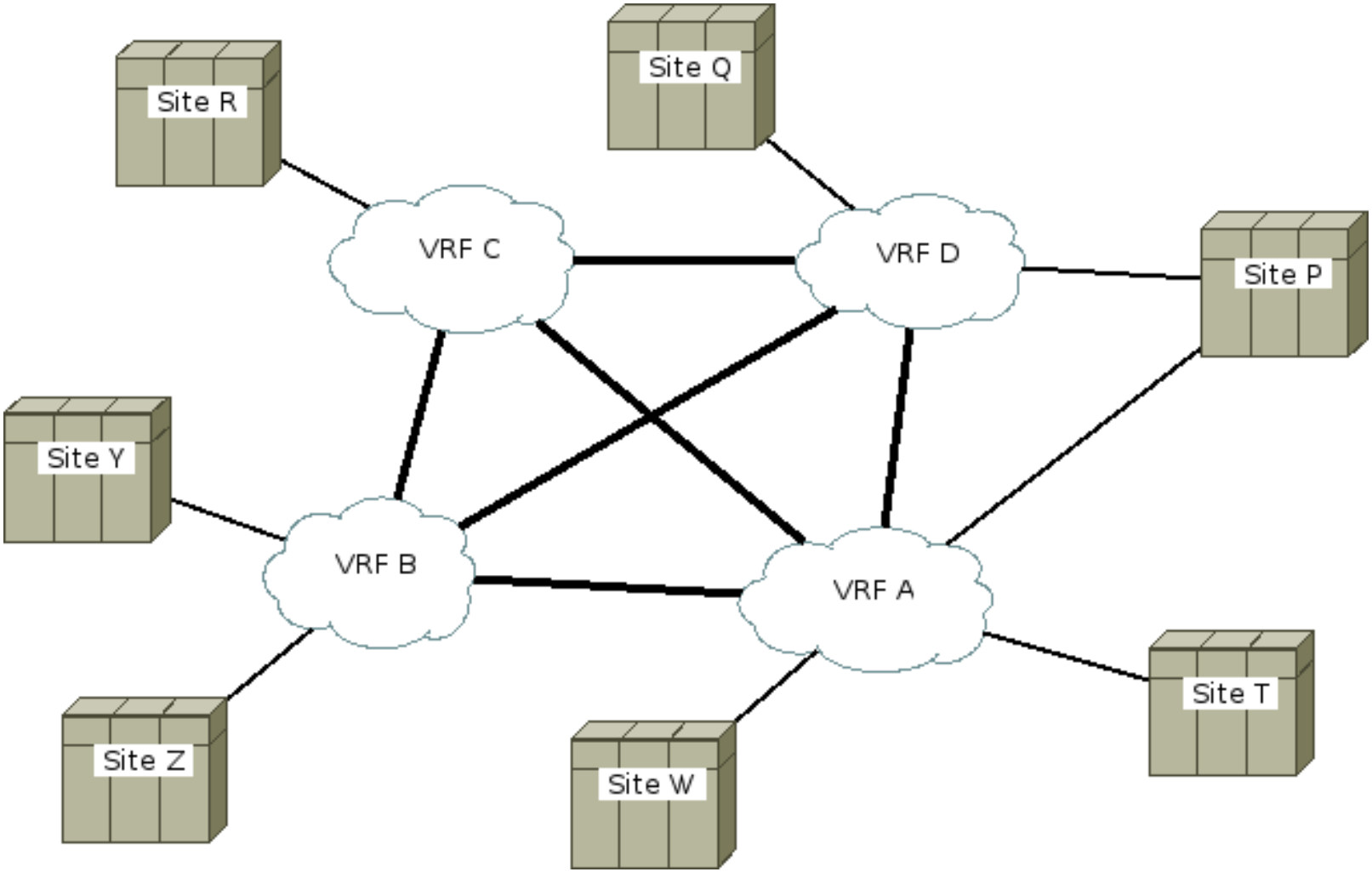
Really, LHCOPN shines because we can easily do traffic engineering and have tight control for policies and operations.

Moreover, R&E network operators are now skilled in Grid-related operations.

So, cloud-like model was laid out and adopted.

НИЦ «Курчатовский институт»

LHCONE: the architecture



НИЦ «Курчатовский институт»

LHCONE: key points

Again, it is (another) private BGP cloud.

VRF: it is a backbone object that peer with all other VRFs. The name is very unfortunate, but typically it is an R&E network or Tier-1 center.

VRF's functions:

- they enforce policies: again, only LHC-related traffic must flow across LHCONE, end sites should not be transit entities;
- they form data backbone and they are liable for keeping it alive and kicking;
- they perform traffic engineering to enable good link utilization.

Currently VRFs are CERN, ESnet, Internet2, GEANT, NORDUnet and CANARIE.

End-sites: Tier-1 and Tier-2 centers.

Again, no centralized firewalling at site-LHCONE border.

НИЦ «Курчатовский институт»

LHCONE and LHCOPN: the interplay

Tier-1 participate both in LHCONE and LHCOPN, so how it is handled?

In the past there was a strict separation that was kept at least operationally: if you can go via LHCOPN, you must, the rest goes via LHCONE.

In LHCOPN's "mandatory star" and "optional T1-T1" link structure more than 50% of traffic will go via CERN. So, 2 European T1 sites will have larger than geographical RTT.

With the current success of LHCONE and its services, Tier-1 sites are allowed to go via LHCONE even for LHCOPN-type traffic.

Good for GEANT: it connects at least 4 Tier-1 sites.

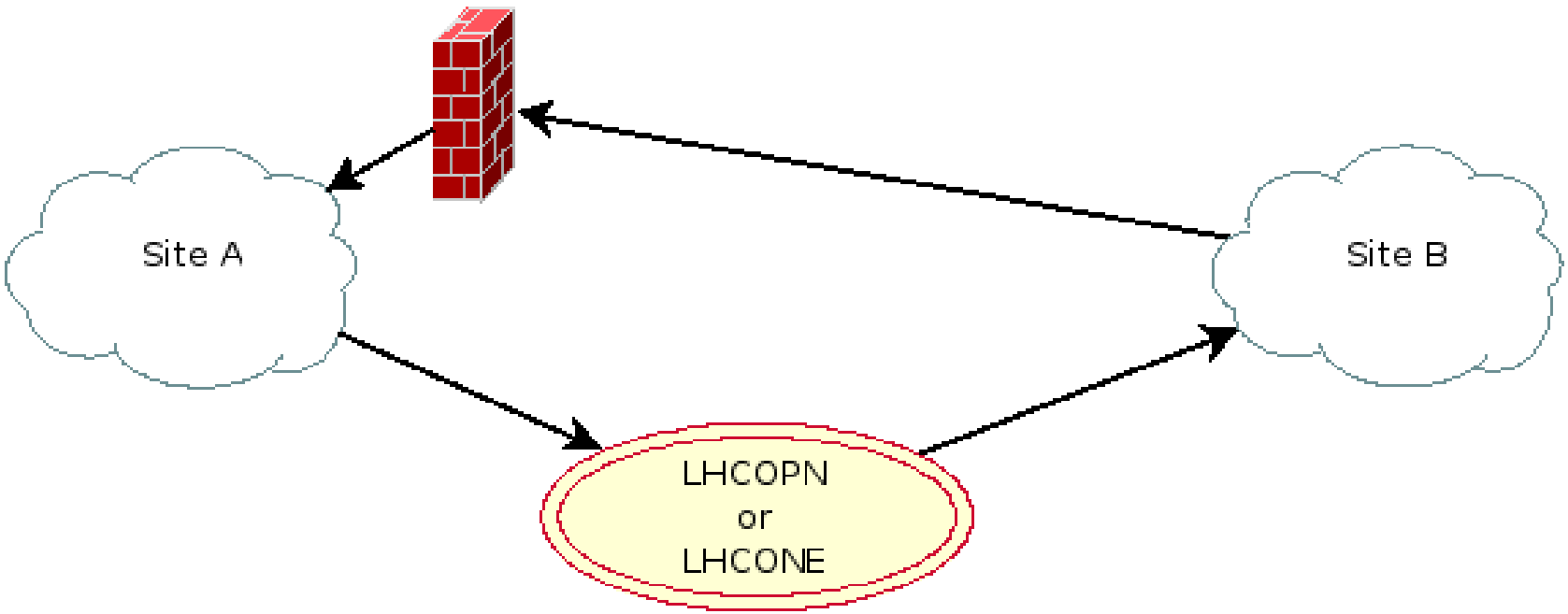
There was another idea to allow big Tier-2's to go into LHCOPN, but it wasn't accepted (yet).

НИЦ «Курчатовский институт»

LHCONE and LHCOPN: need of symmetric routing

Obvious reason: on-purpose network is usually better.

Less obvious reason: site-wide firewalling.



TCP sessions just won't get established in this case.

НИЦ «Курчатовский институт»

LHCONE: case for RRC-KI and RRC-KI-T1

- From 1/2 to 3/4 of traffic moved from GEANT to LHCONE: that's very good.
- Many big sites (that generate most traffic) are already in LHCONE.
- We use VRF (real ones) to do BGP: it is a lot simpler to run than policy-based routing.
- No need for separate AS: BGP cloud is private, so LHCONE announces won't leave it.
- Currently we have transit from NORDUnet and getting transit from CERN, so we a kind of multi-homed LHCONE end-site.
- The plan is to create Russian VRF (in LHCONE sense) for all reginal Tier-2 sites: most likely it will be run by RRC-KI.

So, for both our Tier-1 and Tier-2 the answer
is that LHCONE really helps and provide good service.

НИЦ «Курчатовский институт»

LHCONE: fine technical points

You should have your IP space for LHC-related machines to be aggregated: plan for that or start transition now.

Currently can put only SEs into LHCONE, but with federated storage WNs should be put there too.

Run VRFs, avoid policy-based routing: FZK does PBR and it is a real pain for both them and other parties.

Lay our your routing priorities:

- site-to-site peerings;
- LHCOPN (if any);
- LHCONE;
- R&E networks and general Internet (adjust to taste).

Don't fear private BGP clouds: not another full view, 155 announced networks are now in LHCONE.

НИЦ «Курчатовский институт»