



Network to and at CERN

Getting ready for LHC networking

Jean-Michel Jouanigot and Paolo Moroni
CERN/IT/CS

Summary



- ◆ **Current situation**
- ◆ **T0-T1 planning: LAN**
- ◆ **T0-T1 planning: WAN**



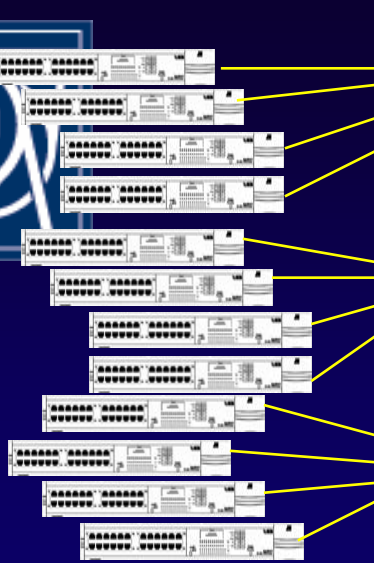
Current situation

- ◆ **General purpose network**
- ◆ **Technical network**
- ◆ **Experimental areas (pre-production)**
- ◆ **External network**
- ◆ **(firewall / HTAR)**

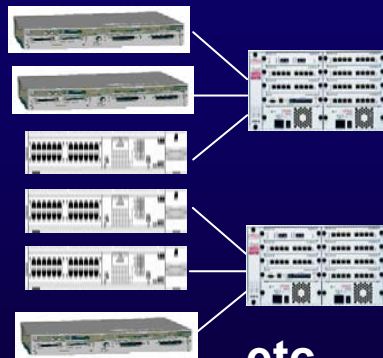
General-purpose network



Server Farms



..etc..



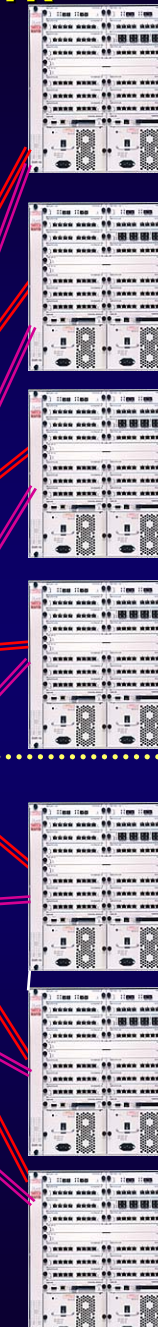
..etc..



Technical Network



Firewall
CIXP, Internet



COMPUTER CENTER

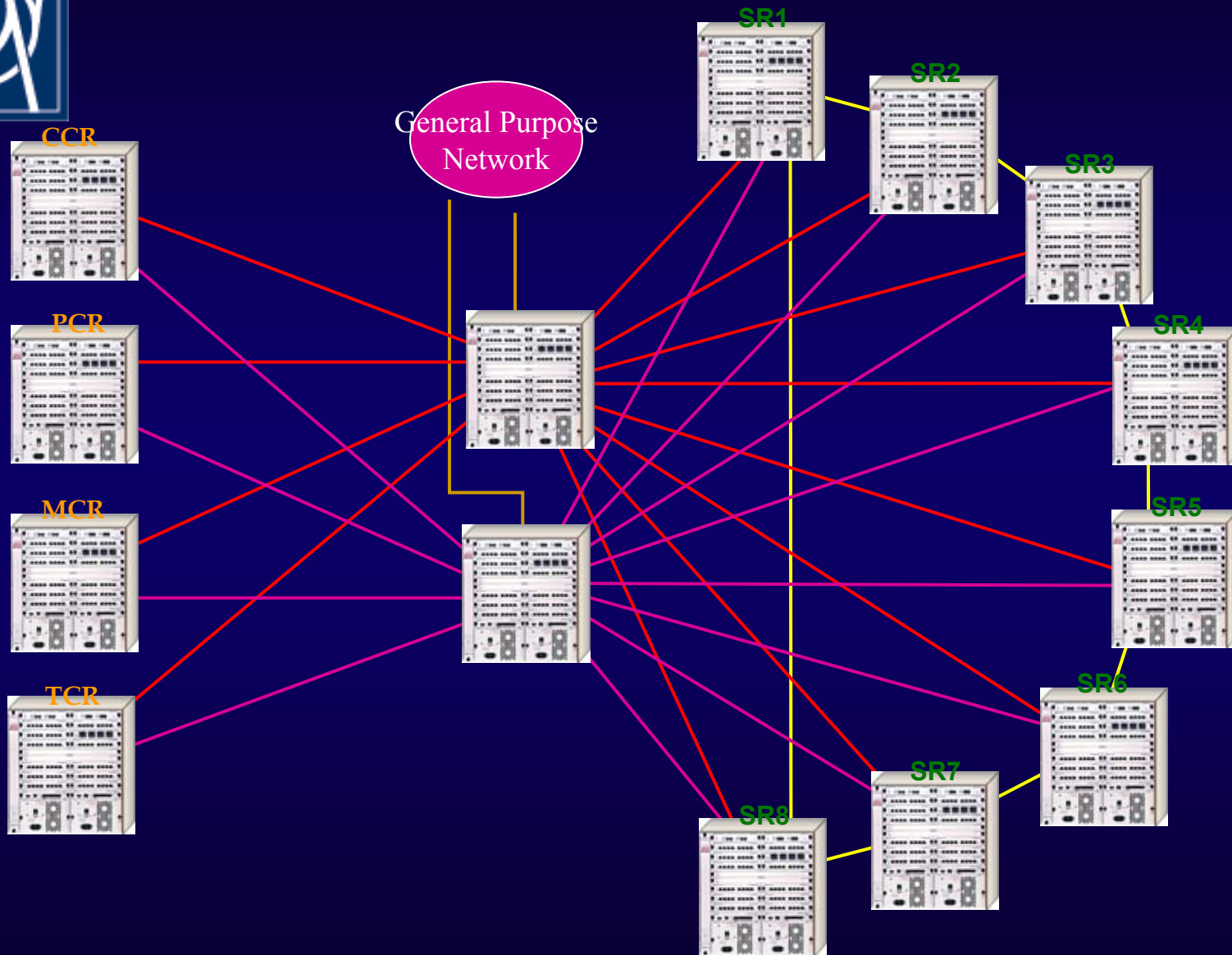
REMOTE MAJOR STARPOINTS

21 January 2005

T0/1 network meeting

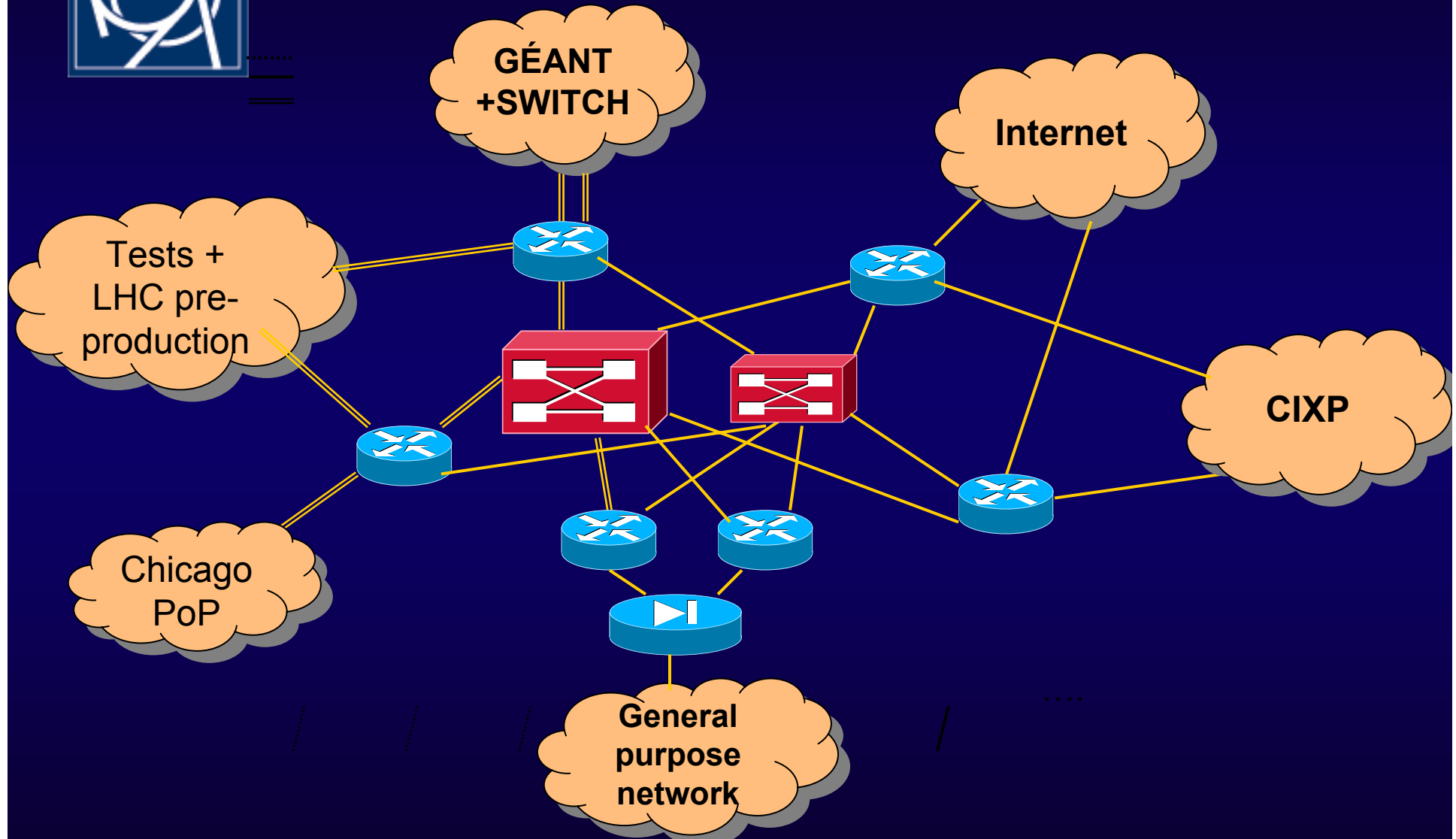


Technical network





External network





Firewall

This slide has been intentionally left blank



T0-T1 planning (LAN)

- ◆ New 2.4 Tb/s backbone to interconnect
 - ◆ LHC experiments (CERN Tier0)
 - ◆ general purpose network
 - ◆ CERN Tier1
 - ◆ T0-T1 WAN (regional Tier1's)
- ◆ Based on 10GE technology
- ◆ Layer 3 interconnections
- ◆ No central switch(es)
- ◆ Redundancy via multiple 10GE paths (OSPF)

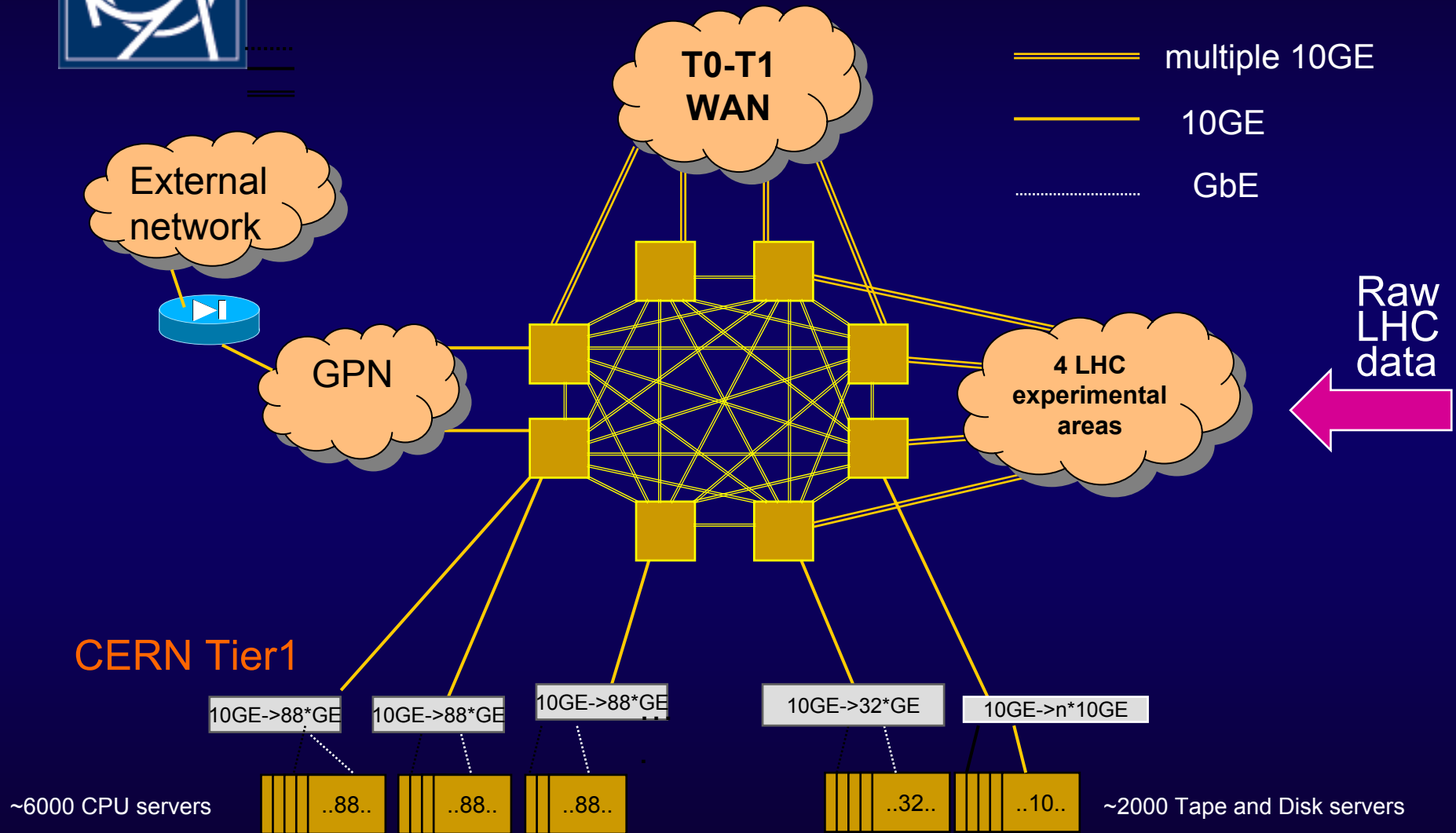


More about T0-T1 LAN

- ◆ Random paths through the backbone for load balancing (OSPF)
- ◆ IP addressing:
 - ◆ depends on the LHC WAN implementation,
 - ◆ RFC1918 addresses are likely for a lot of end systems
 - ◆ a data mover facility can help a lot (already successfully implemented for the BABAR experiment at IN2P3)
- ◆ Default route? Maybe not necessary
- ◆ Call for tender for the equipment being issued

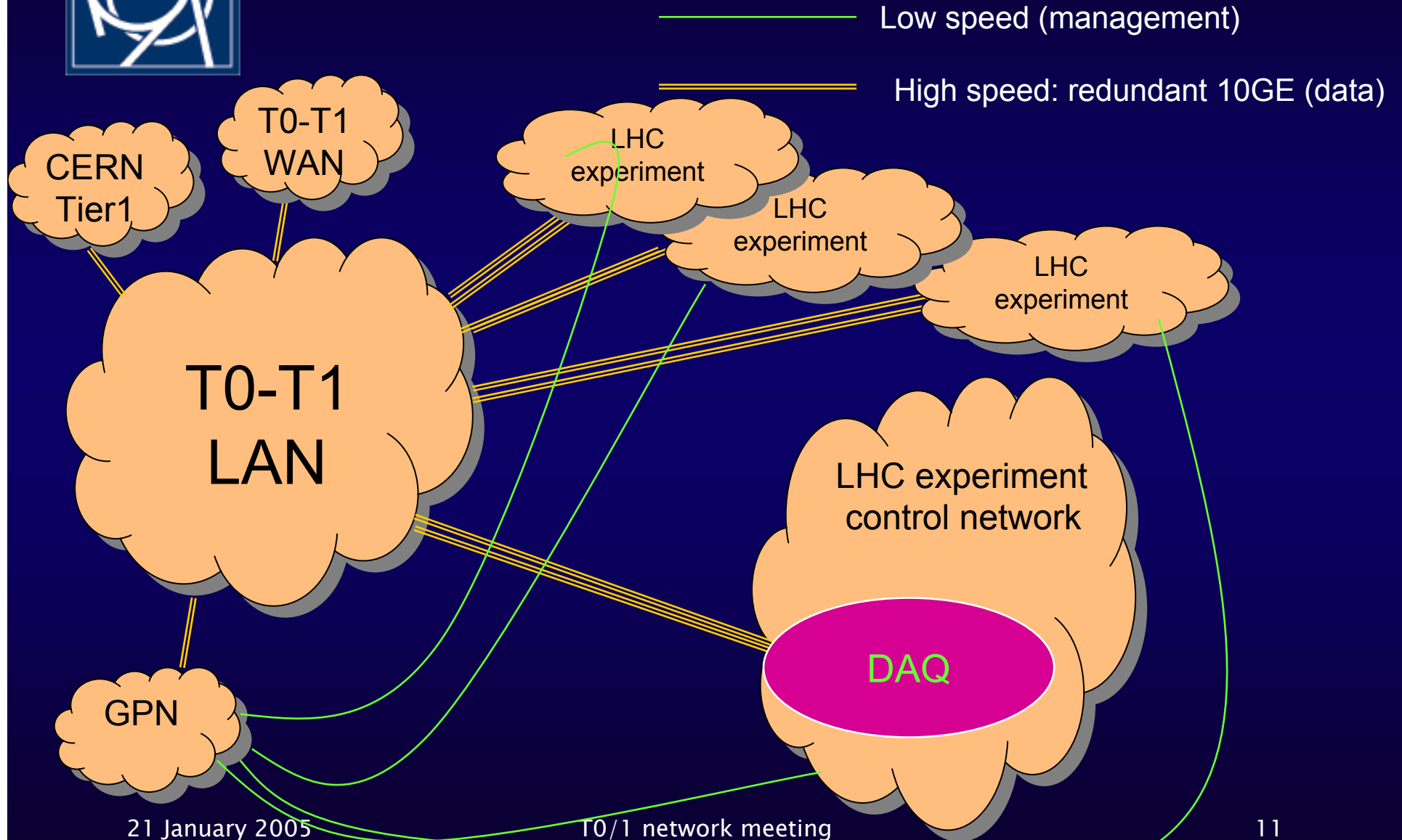


T0-T1 network at CERN (LAN)





Tier0 network (LHC experimental areas)





T0-T1 WAN: progress

- ◆ A lot of progress has been made:
 - ◆ 10 Gb/s equipment is commonly available (although not yet cheap): STM-64 (10GE WAN PHY), 10GE LAN
 - ◆ 10 Gb/s capacity (SDH, wavelength, WDM over dark fibre) is affordable
 - ◆ long-distance, high-speed TCP is feasible, although with special Linux tuning



T0-T1 WAN: progress (continued)

- ◆ More progress being made:
 - ◆ GN2 is coming in Europe with new services and research activities
 - ◆ Several interesting initiatives in North America and in Europe (dark fibre-based networks, etc.)
 - ◆ Several interesting monitoring tools exist or are being developed
 - ◆ Pre-production simulation (robust data challenge): a useful ongoing experience
 - ◆ Firewall with HTAR works for non-LHC traffic and for some pre-production



T0-T1 WAN: issues

- ◆ Still several open questions:
 - ◆ how will Tier1's connect to Tier0 (directly, one upstream, layered upstreams, ...)?
 - ◆ backup routing ?
 - ◆ non-homogeneous Tier1 requirements?
 - ◆ any Tier1-Tier1 traffic via Tier0?
 - ◆ IP addressing: routable or RFC1918 ?
 - ◆ does every Tier1 have enough routable addresses?
- ◆ and ...



T0-T1 WAN: more issues

- ◆ ...what about
 - ◆ security ?
 - ◆ Tier2's ?
 - ◆ compatibility between GRID middleware and network design?
 - ◆ special tuning for WAN data transfers?
 - ◆ compatibility between high speed flows and some network devices (Juniper M160)?
 - ◆ management, monitoring, troubleshooting?
- ◆ Anything else?



Recommendations (I)

- ◆ Allow for diverse regional requirements, but standardise NOW on the T0-T1 physical interface:
 - ◆ 10GE LAN PHY (LR/SR ?)
 - ◆ STM-64/OC192
 - ◆ 10GE WAN PHY (?)
 - ◆ Other interfaces also possible in the pre-production phase (GbE, multiple GbE, STM-16)
- ◆ Take advantage of useful experience (robust data challenge)
- ◆ Define clearly the operational responsibilities across multiple administrative domains



Recommendations (II)

- ◆ Select equipment which is expected to work reliably for some years
- ◆ A data mover facility (spooling system) helps with several issues:
 - ◆ IP addressing needs
 - ◆ security
 - ◆ WAN data transfer optimisation
- ◆ Select proven and stable technology: smooth network operations and easy troubleshooting are essential



Recommendations (III)

- ◆ Security is essential
- ◆ Monitoring is essential
- ◆ Allocate suitable (routable) subnets, dedicated to LHC production purposes
- ◆ If not enough routable IP addresses, ask RIPE-NCC (or ARIN, or APNIC) for more and do so **NOW**



Recommendations (IV)

- ◆ Never mind if the network is just a boring production tool: being at the bleeding edge is not essential in this situation
- ◆ LHC physics is the research target, not LHC networking



LHC WAN: a possible design

- ◆ Assumptions: if ...
 - ◆ Tier1's connect at layer 3
 - ◆ backup routing is a requirement and it is acceptable via research IP networks (not more than two-three Tier1's down at the same time)
 - ◆ Tier1-Tier1 traffic is allowed via Tier0 (*although this would not be Tier0's preference...*)
 - ◆ Tier1 and Tier0 addresses are publicly routable and every Tier1 has allocated a **SMALL** number of subnets for inter-Tier0/1 traffic
 - ◆ BGP routing using the "natural" ASN and routable prefixes
 - ◆ no default route (or no default route towards T0): is it possible?



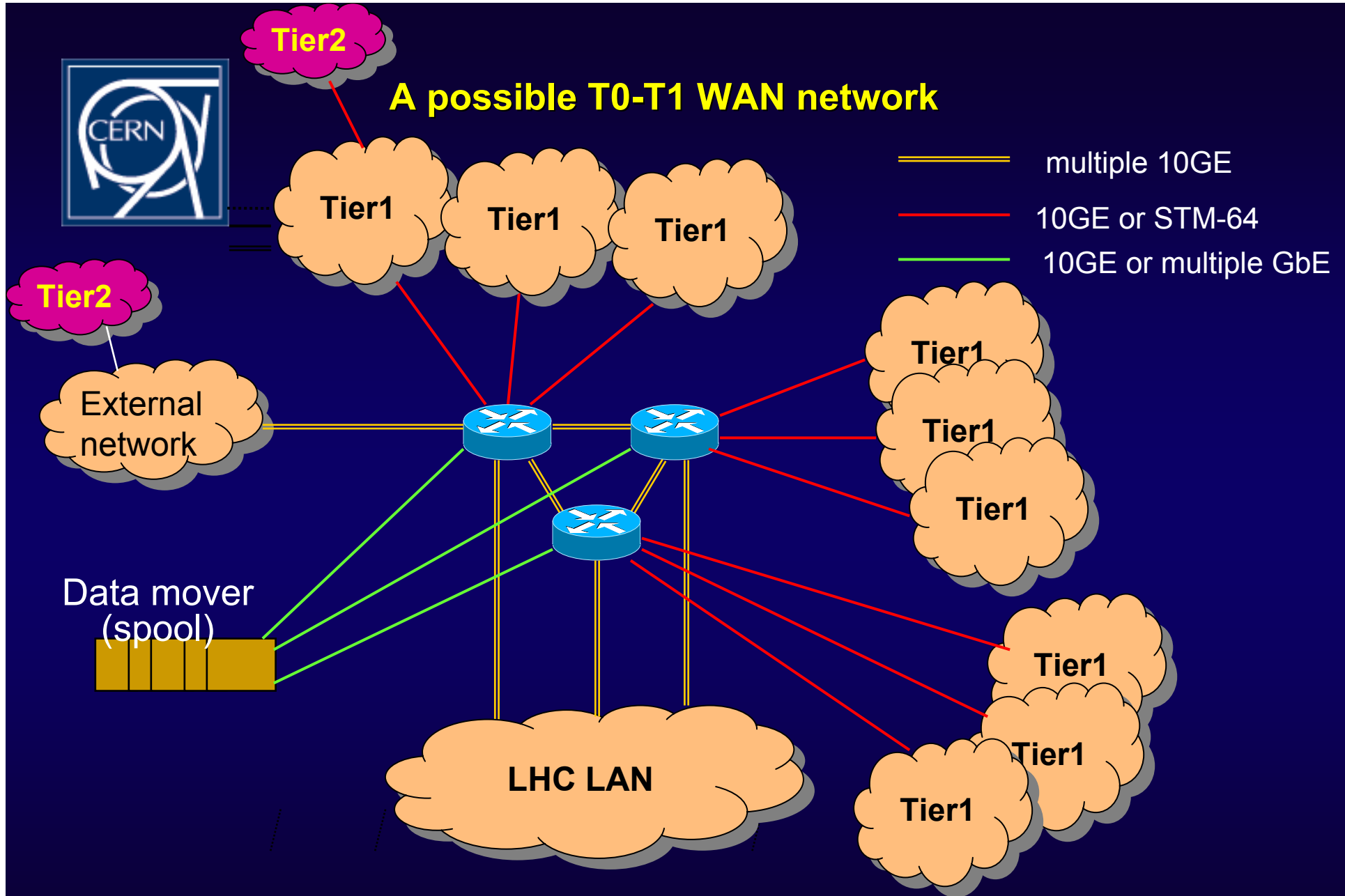


A possible design (continued)

- ◆ ...and if ...
 - ◆ basic security is provided via layer 3 ACLs (allowed subnets and, if possible, port numbers)
 - ◆ Tier1's may have some non-homogeneous requirements
 - ◆ no Tier2 directly connected to Tier0, but some may be allowed to exchange traffic at less than 10 Gb/s
 - ◆ alternatively, some T0-T2 traffic may transit via an intermediate T1
 - ◆ a spooling system (data mover) is used as buffer between sites to optimise long-distance data transfer and reduce public IP addresses needs
- ◆ ... then ...



A possible T0-T1 WAN network





Thank you

Questions?