

# LHCONE: LA RETE DEI TIER2

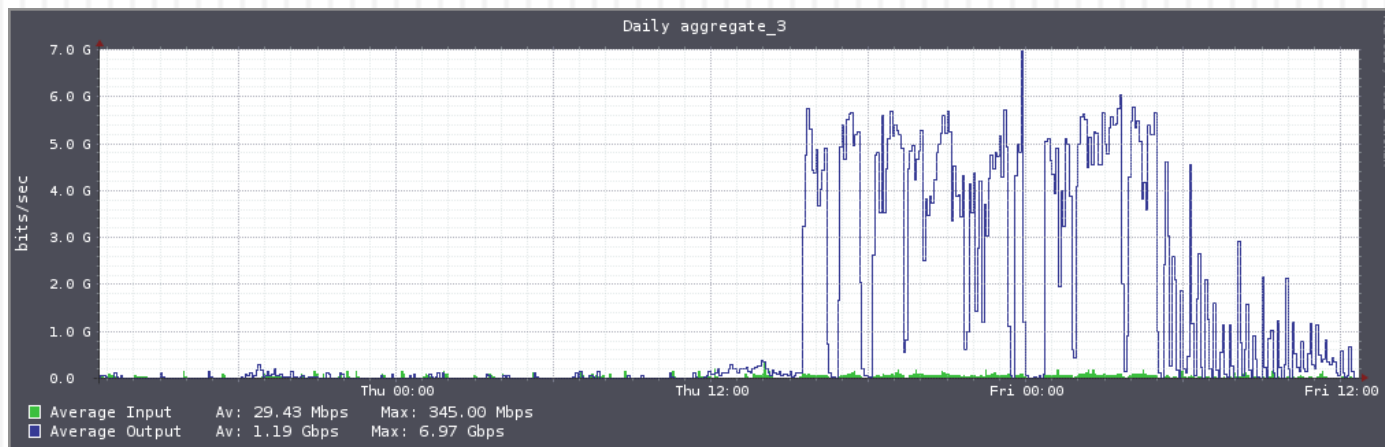
IL PUNTO DI VISTA DEL GARR

*16 - 20 maggio 2011  
Massimo Carboni*

# Le motivazioni

2

- Cambiamento del Data Model, da MONARC ad “ANARC”
- Incremento del traffico fra i siti LHC
  - su intervalli di tempo brevi
  - su direttrici non predicibili



- 6 Gbit/s di traffico CNAF -> MIT per ~18 ore
- Uno dei link transatlantici a 10Gbit/s è andato in saturazione
- Per anni il traffico CNAF-MIT era stato trascurabile

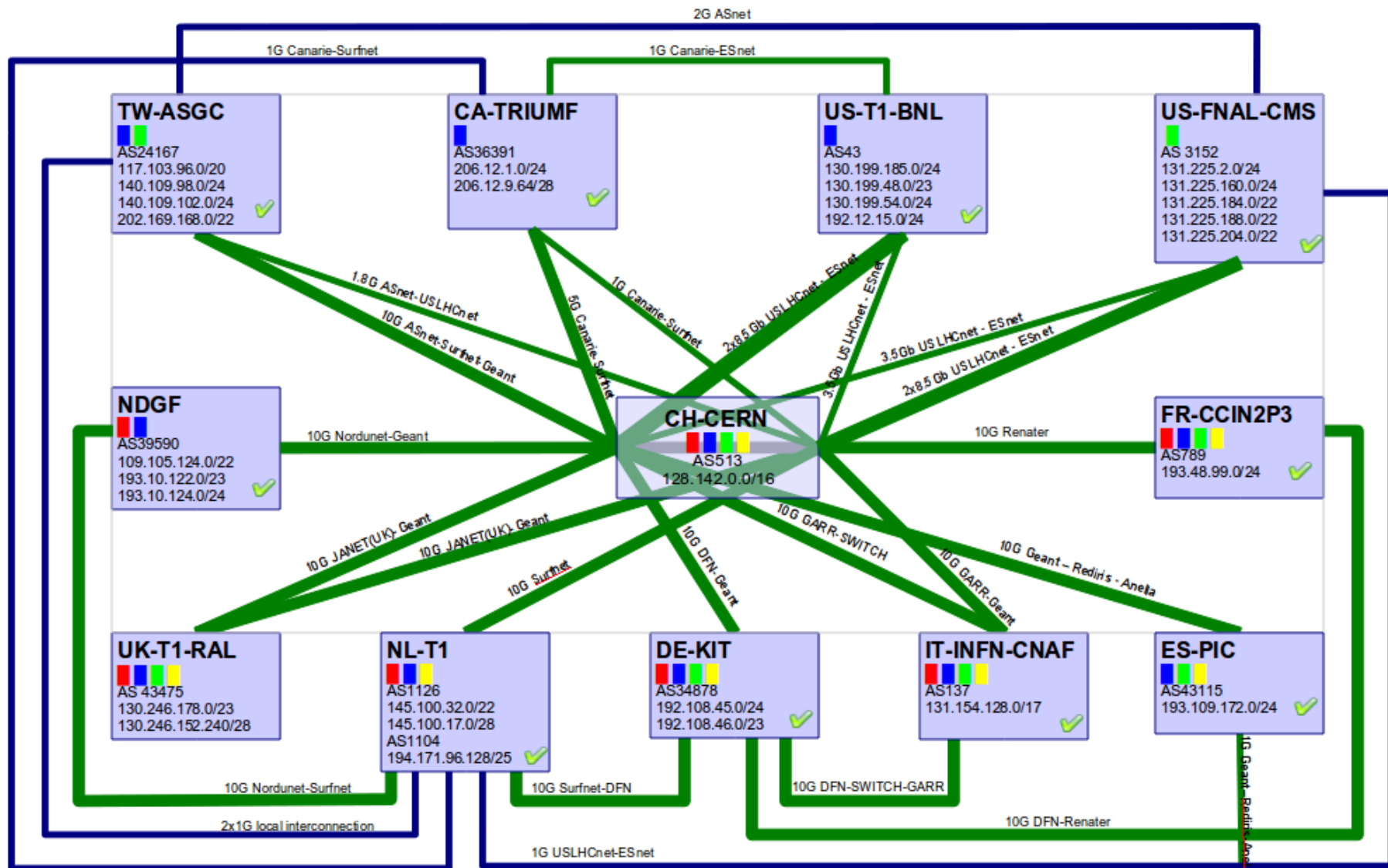
# Le conseguenze

3

- Le direttrici di traffico sono diventate imprevedibili a meno che i coordinatori del calcolo degli esperimenti non ce le dicano prima
- In queste condizioni il modello con connettività a circuiti (lightpath) non scala
  - Il numero di circuiti necessari diventa  $N*(N-1)/2$
- Bisogna andare verso un modello di tipo any to any (basato su IP)
  - E basato sull'overprovisioning
- LHCOPN, la rete costruita per lo scambio T0-T1 e T1-T1 non può veicolare questo tipo di traffico, in quanto ne sarebbe perturbata
- Bisogna creare una nuova rete *ad hoc* per lo scambio di dati fra T1/T2/T3

- LHC Open Network Environment
  - ▣ Una rete dedicata agli scambi di dati
    - T1-T2
    - T2-T2
    - T1-T3 e T2-T3
- La nuova rete deve:
  - ▣ Avere scope globale
    - 5 continenti coinvolti
  - ▣ Scalare con l'elevato numero di partecipanti
    - 100+ siti coinvolti
  - ▣ Assicurare interoperabilità con la situazione attuale
    - Non deve precludere l'uso della rete attuale ai siti non ancora migrati o che non migreranno mai
  - ▣ Permettere una migrazione smooth
- In questo momento si sta predisponendo un prototipo di rete
  - ▣ Riutilizza risorse esistenti
  - ▣ Permette di cominciare a studiare le modalità di interconnessione

# LHCOPN



— T0-T1 and T1-T1 traffic  
 — T1-T1 traffic only  
 — Not deployed yet  
 (thick) >= 10Gbps  
 (thin) <10Gbps

■ = Alice ■ = Atlas  
 ■ = CMS ■ = LHCb  
 ✓ = internet backup available  
 p2p prefix: 192.16.166.0/24  
 edoardo.martelli@cern.ch 20110208



## □ Cosa offre:

### ■ Layer2 best effort connectivity

- VLAN any-to-any
- VLAN point-to-point

**SCENARIO  
PRIMARIO**

### ■ Point-to-point guaranteed bandwidth

- Dedicated circuit
- Bandwidth on demand

**SCENARIO  
DA VALUTARE**

## □ Cosa non offre:

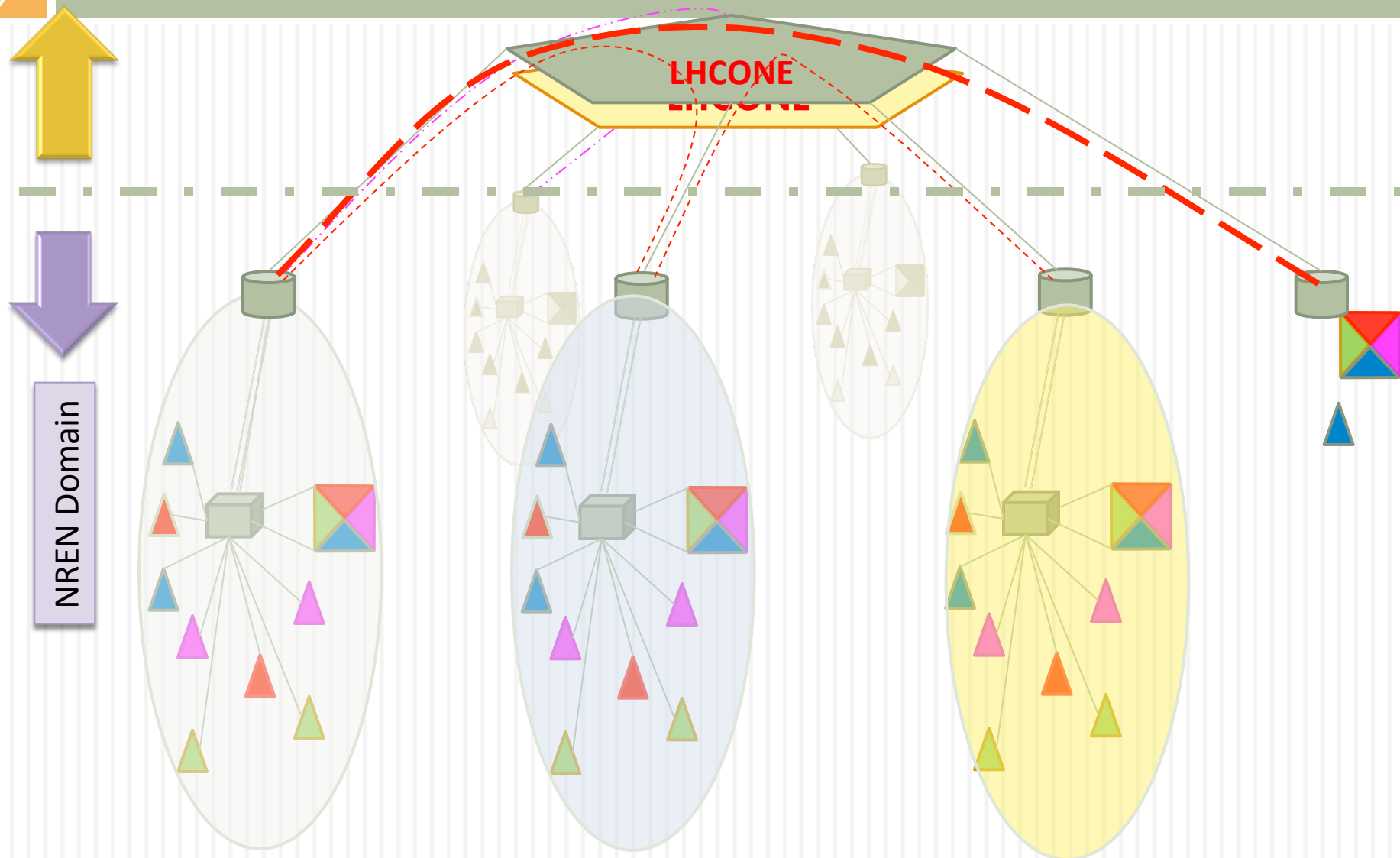
- Connettività verso il T0 (ai T2 non serve, ai T1 è assicurata da LHCOPN)
- Connettività tra T1 (è assicurata da LHCOPN)
- Connettività verso siti non connessi ad LHCONE (è assicurata dalla rete general purpose)

## □ Non impone constraint sul posizionamento dei dati

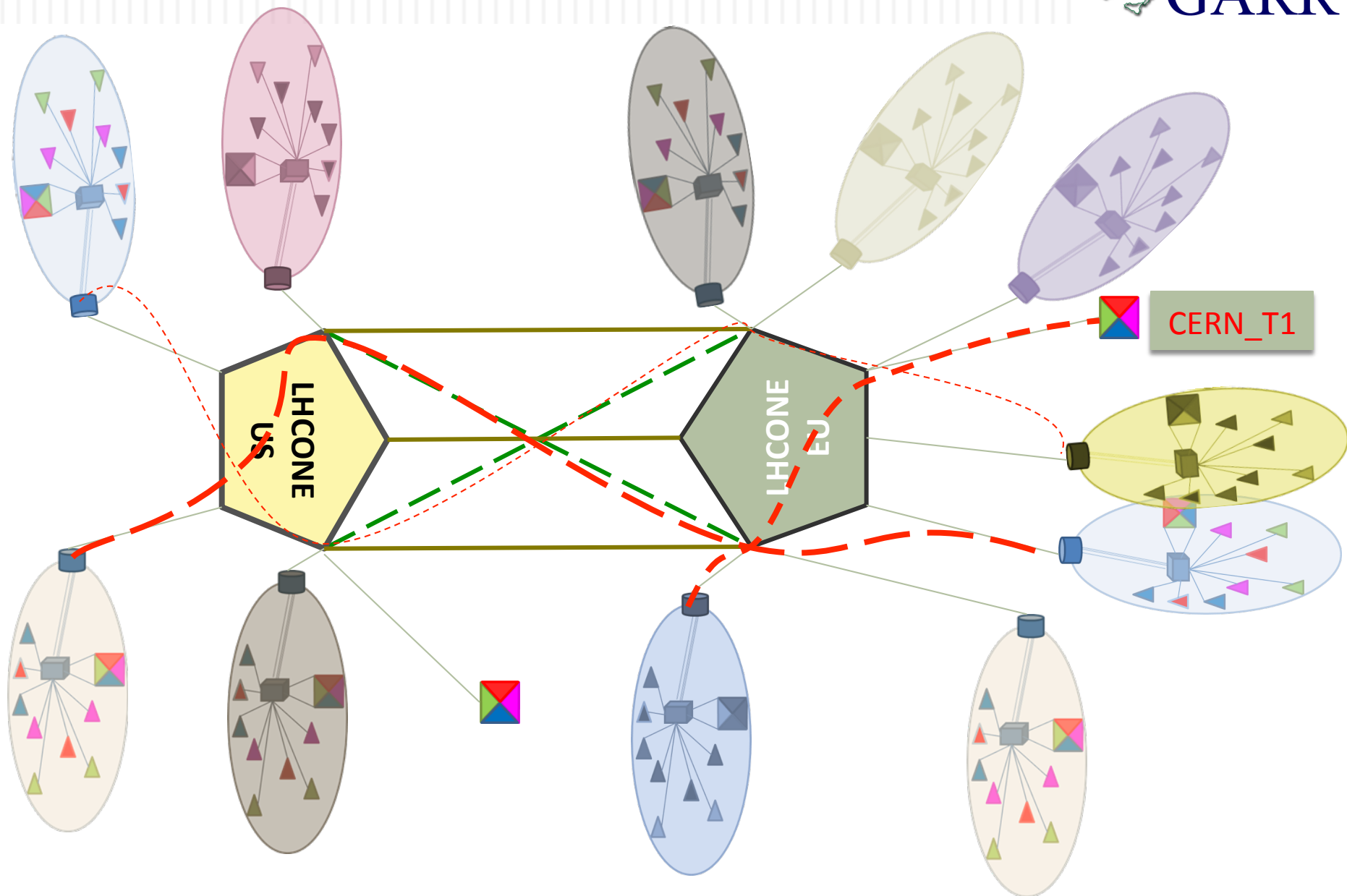
# Schema Logico LHCONE

Non entra nel merito del livello 3

7



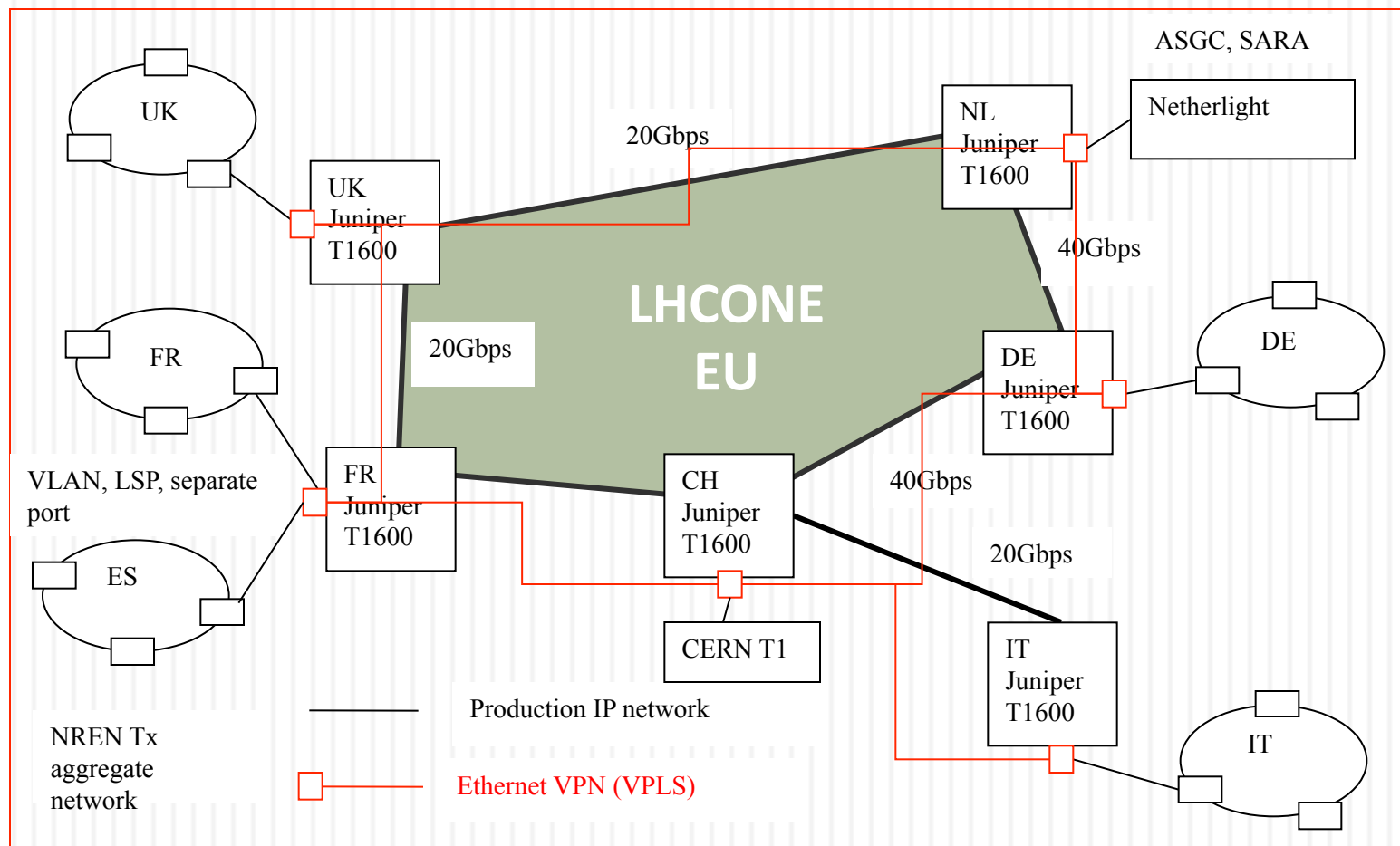
# LHCOne: US vs EU





# Il prototipo europeo

9



10

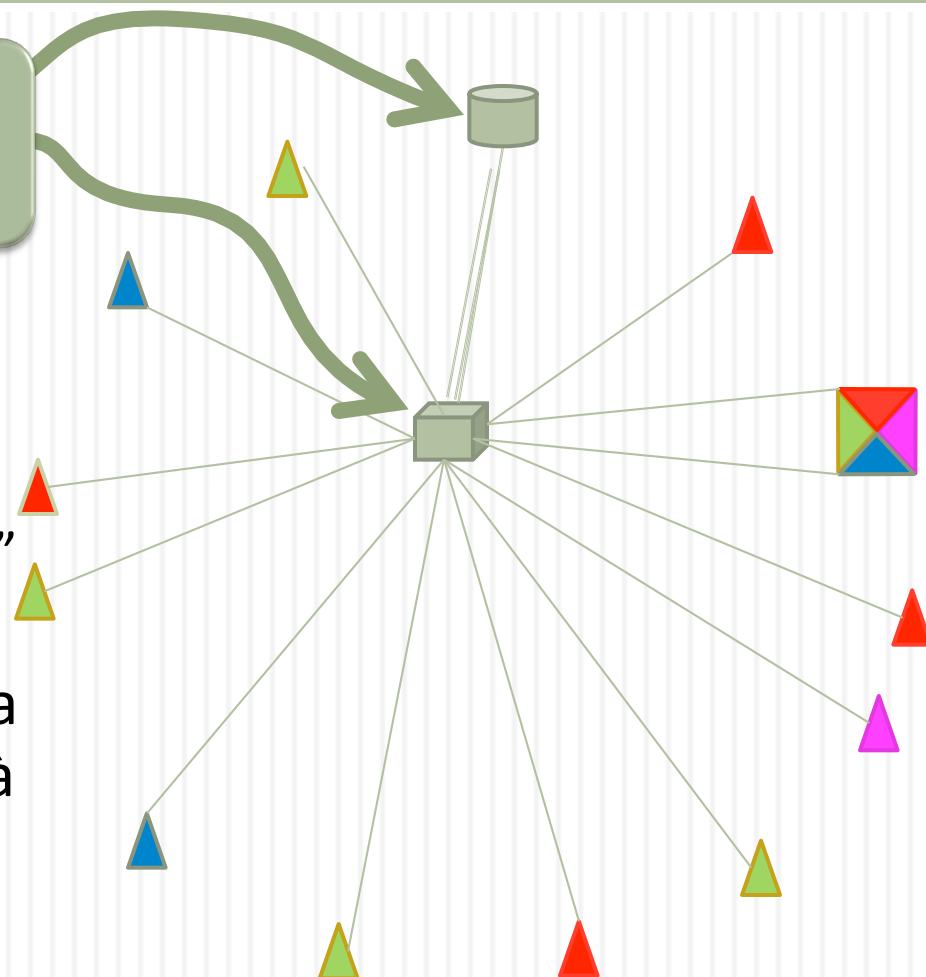
# La topologia interna

L'implementazione per INFN

# L'aggregazione nazionale

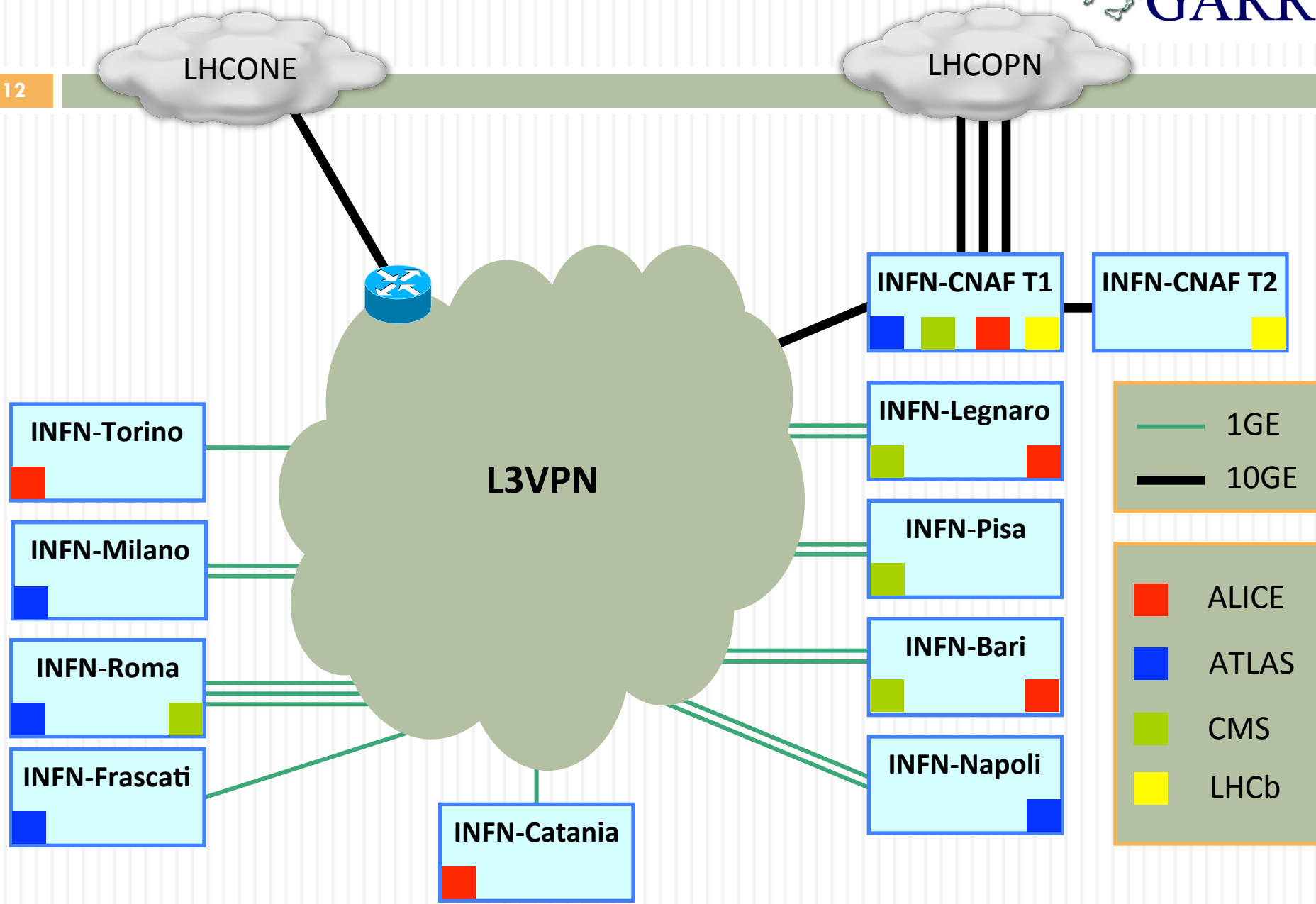
11

- Gestione nazionale del traffico L1/L2 e della connettività esterna di L3 alla frontiera
- Tier1: 20G-40G Iniziali
- Tier2: almeno 10G
  - ▣ Raddoppio in caso di necessità
  - ▣ Valutazione di modelli “dinamici” della capacità >10G
- Tier3: non compresi nella prima fase, solo connessi con capacità IP. Possibile integrazione in futuro anche con capacità 10G se con fibra GARR-X



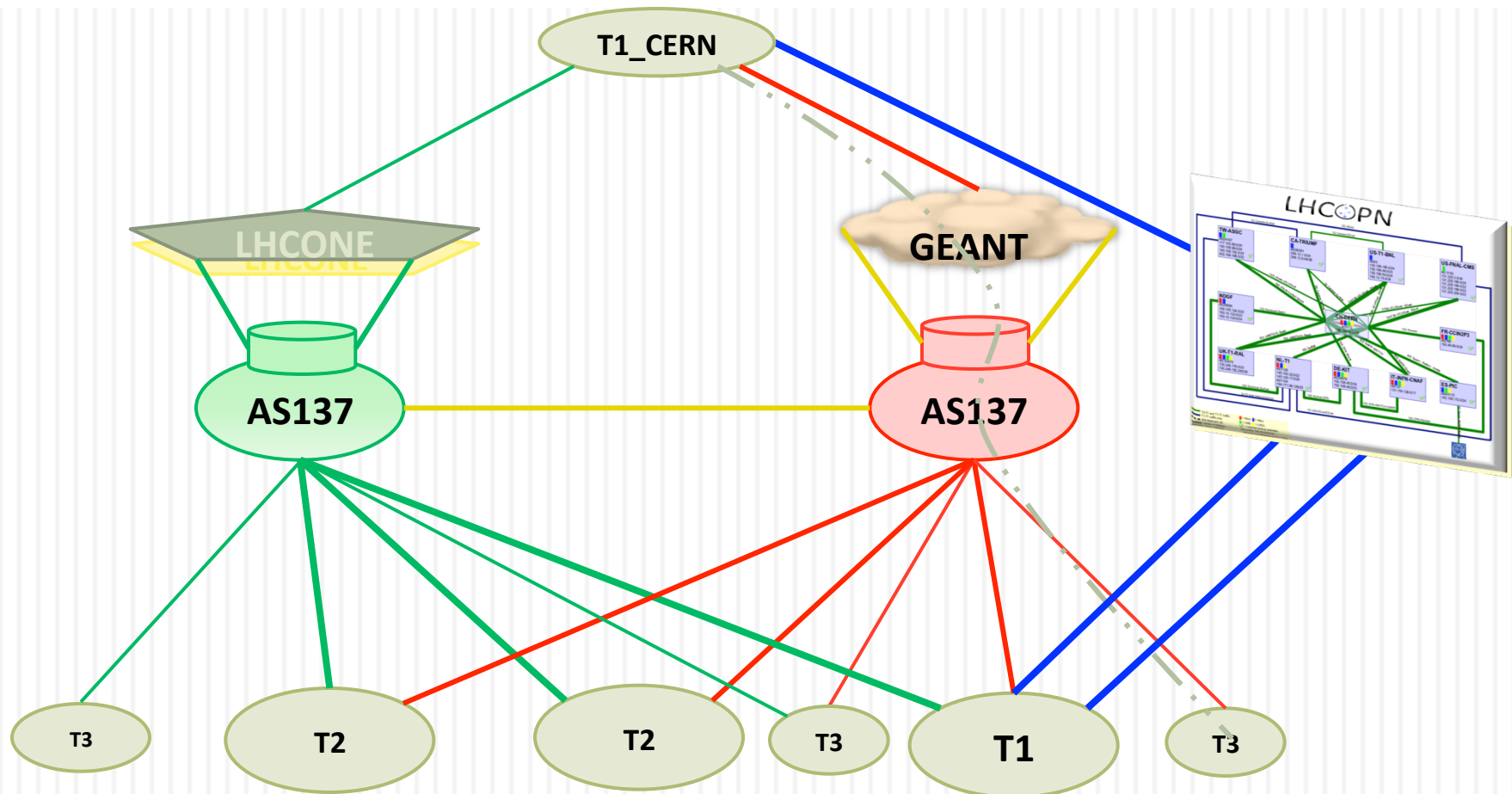
# Italian Tier2 prototype

12



# Schema di rete LHC T1/T2/T3

13

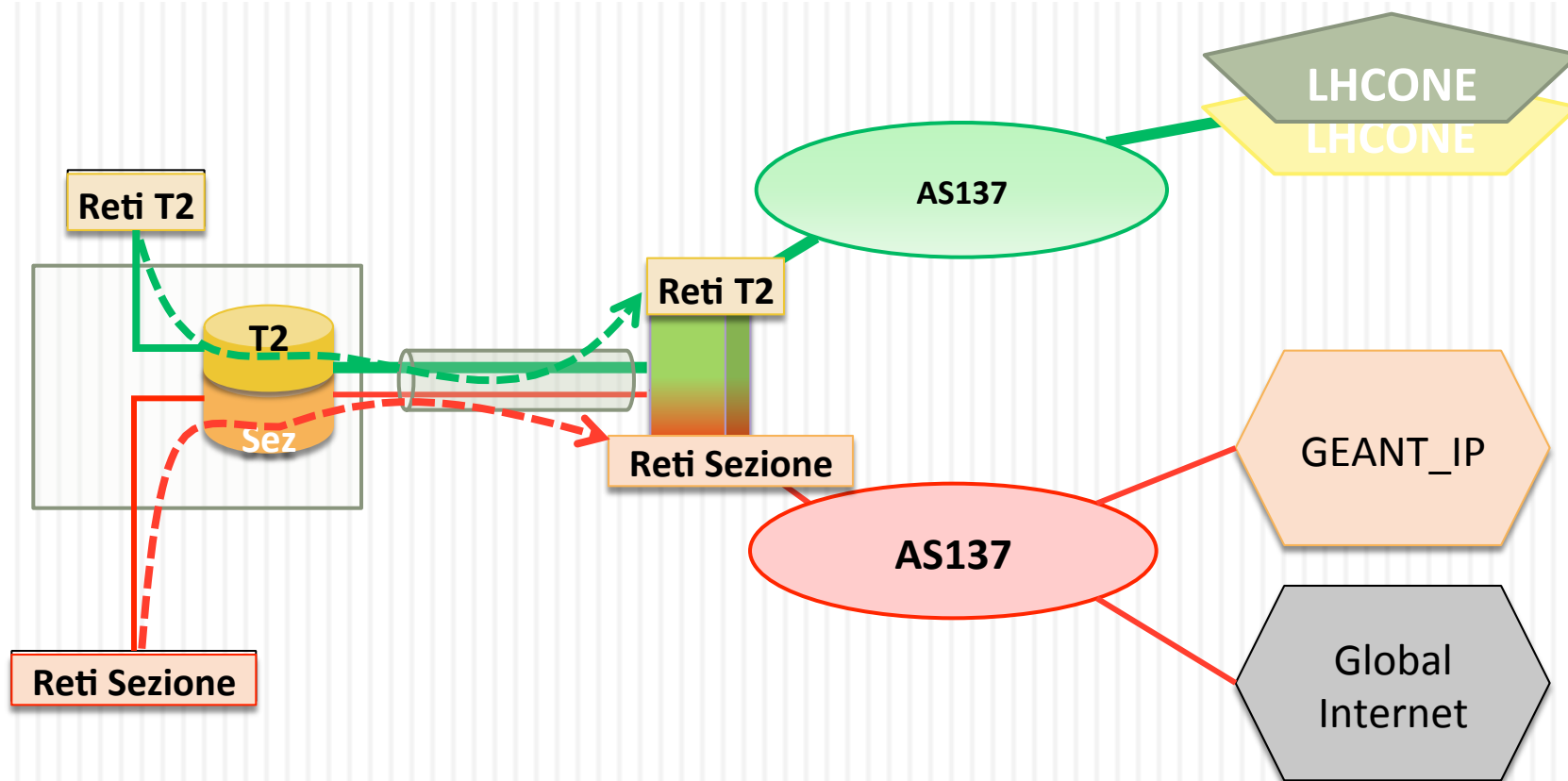


- La rete deve essere in grado di prendere decisioni diverse per le stesse destinazioni

LHCONE: la situazione INFN

# Accesso: 1 Router che non parla BGP

15

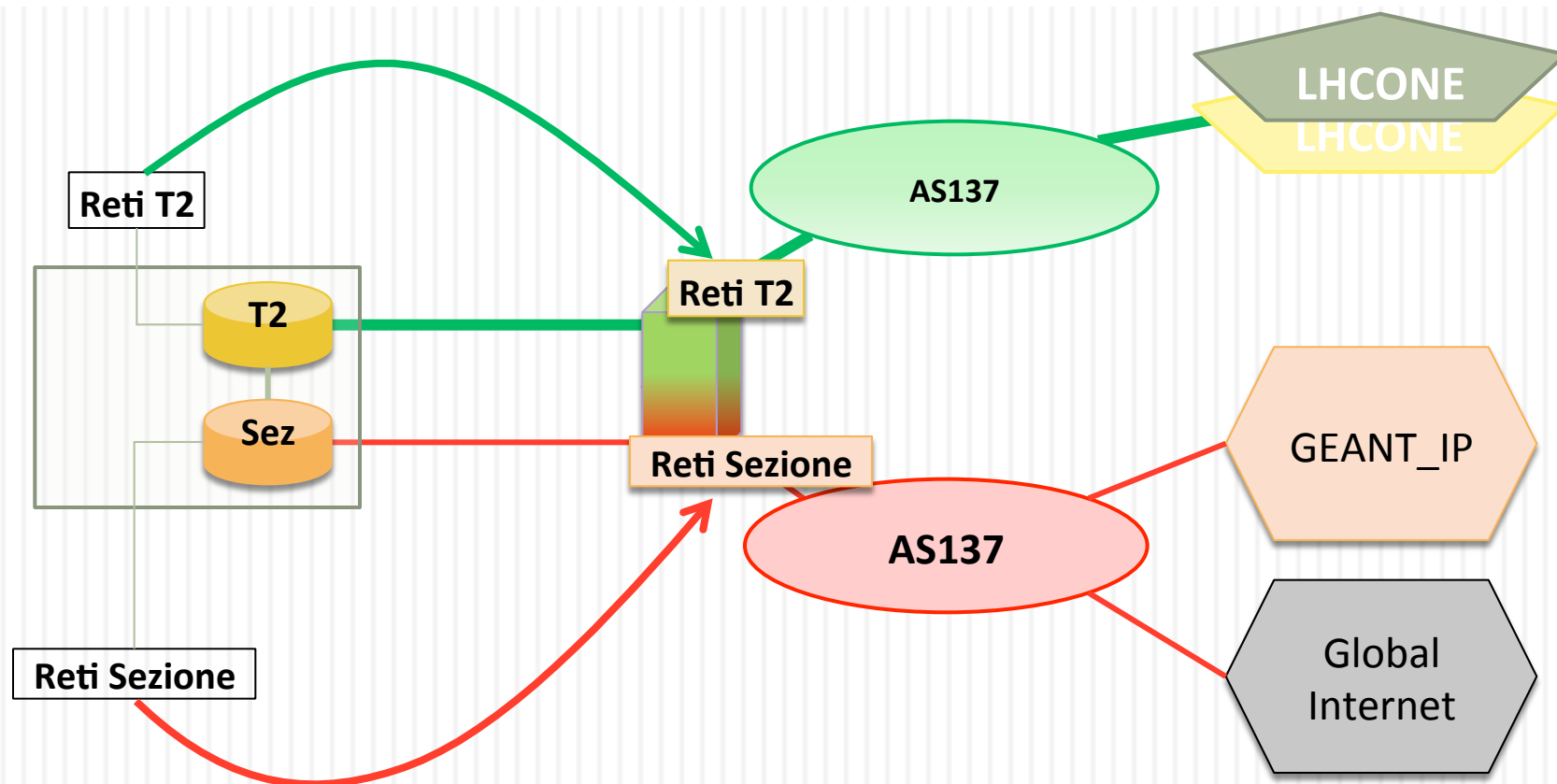


- Bisogna creare 2 collegamenti logici sul singolo collegamento fisico
- Il router unico deve utilizzare il Policy Based Routing, ovvero deve instradare verso i 2 collegamenti logici sulla base della network sorgente

ACCROCCO

# Accesso: 2 Router che non parlano BGP

16



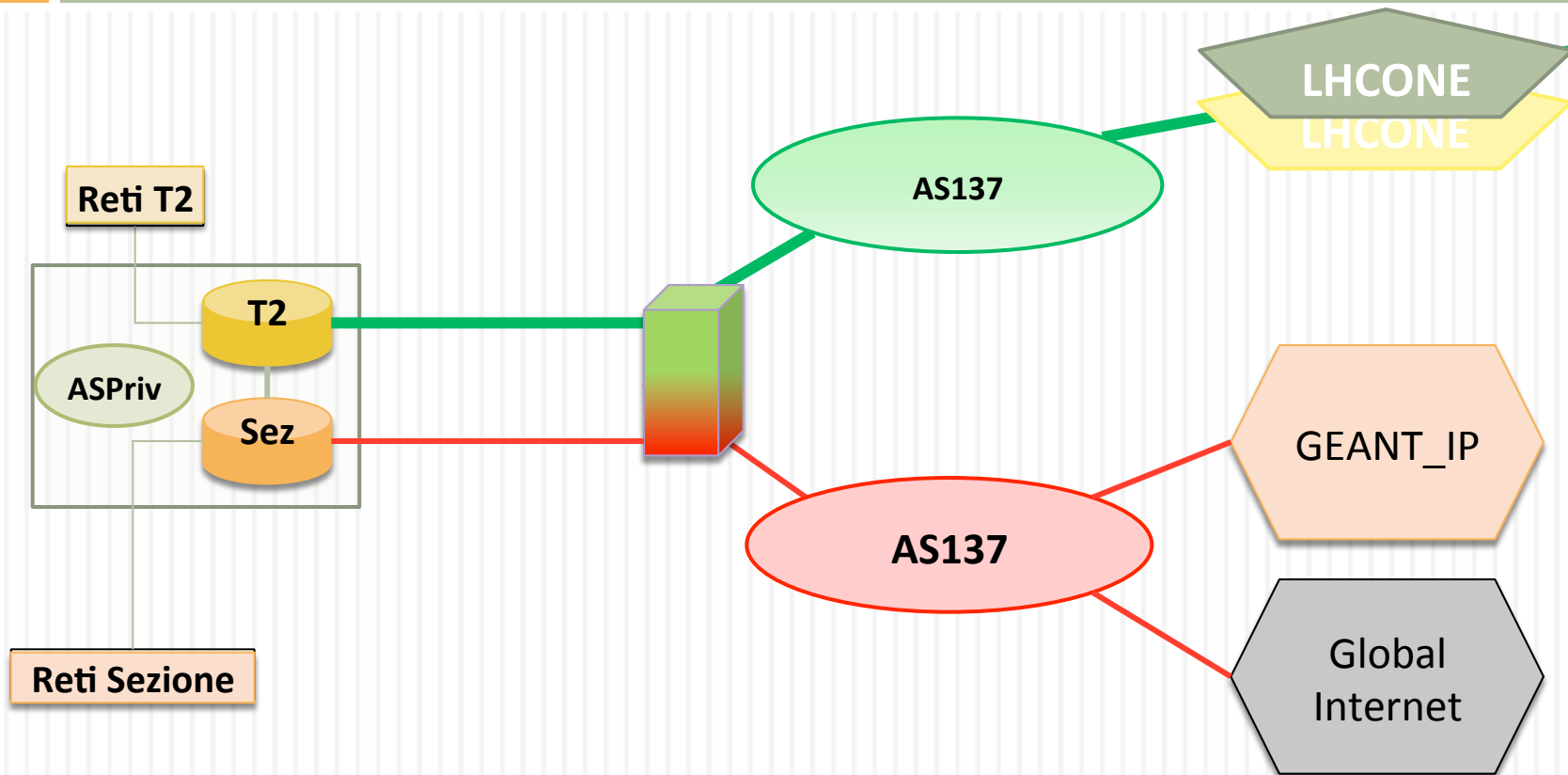
- I collegamenti sono distinti anche dal punto di vista fisico
- Le rotte vengono installate staticamente sui router GARR

TRANSITORIA



# Accesso: 2 Router che parlano BGP

17



- I collegamenti sono distinti anche dal punto di vista fisico
- Le rotte vengono installate sui router GARR dinamicamente
  - È possibile gestire il backup automatico della connettività T2 attraverso il collegamento della sezione

PREFERITA

- Siti che hanno un collegamento fisico dedicato ad LHC e separato da quello di laboratorio/sezione
  - Bari: 2GE (dedicato, fisicamente separato)
  - Catania: 1GE (dedicato, fisicamente separato)
  - Milano: 2GE (dedicato, fisicamente separato)
  - Roma1: 3GE (dedicato, fisicamente separato)
- Siti che hanno un collegamento condiviso con laboratorio/sezione
  - LNF: 1GE (condiviso con il lab, no VLAN)
  - LNL: 2GE (condiviso con il lab, no VLAN)
  - Napoli: 2GE (logicamente separato dalla sezione mediante VLAN)
  - Pisa: 2GE (logicamente separato dalla sezione mediante VLAN)
  - Torino: 1GE (condiviso con la sezione , no VLAN)
  - CNAF: 10GE (condiviso con il lab, no VLAN)

# In conclusione

- ❑ GARR sta predisponendo un prototipo mediante la realizzazione di una L3VPN sulla rete IP attuale
- ❑ Il prototipo potrà includere dal primo momento tutti i siti LHC italiani
- ❑ GARR assicurerà il collegamento del prototipo alla infrastruttura europea e gestirà le politiche di routing internazionali
- ❑ I siti INFN dovranno preferibilmente:
  - ❑ presentare un collegamento logico o fisico separato
  - ❑ presentare subnet T2 ben distinte da quelle della sezione/laboratorio
  - ❑ gestire il routing via BGP
- ❑ In ogni caso GARR ha la possibilità di assicurare l'interconnessione anche ai siti che non hanno:
  - ❑ un collegamento logico o fisico separato
  - ❑ un doppio router
  - ❑ almeno un router che parli BGP

20

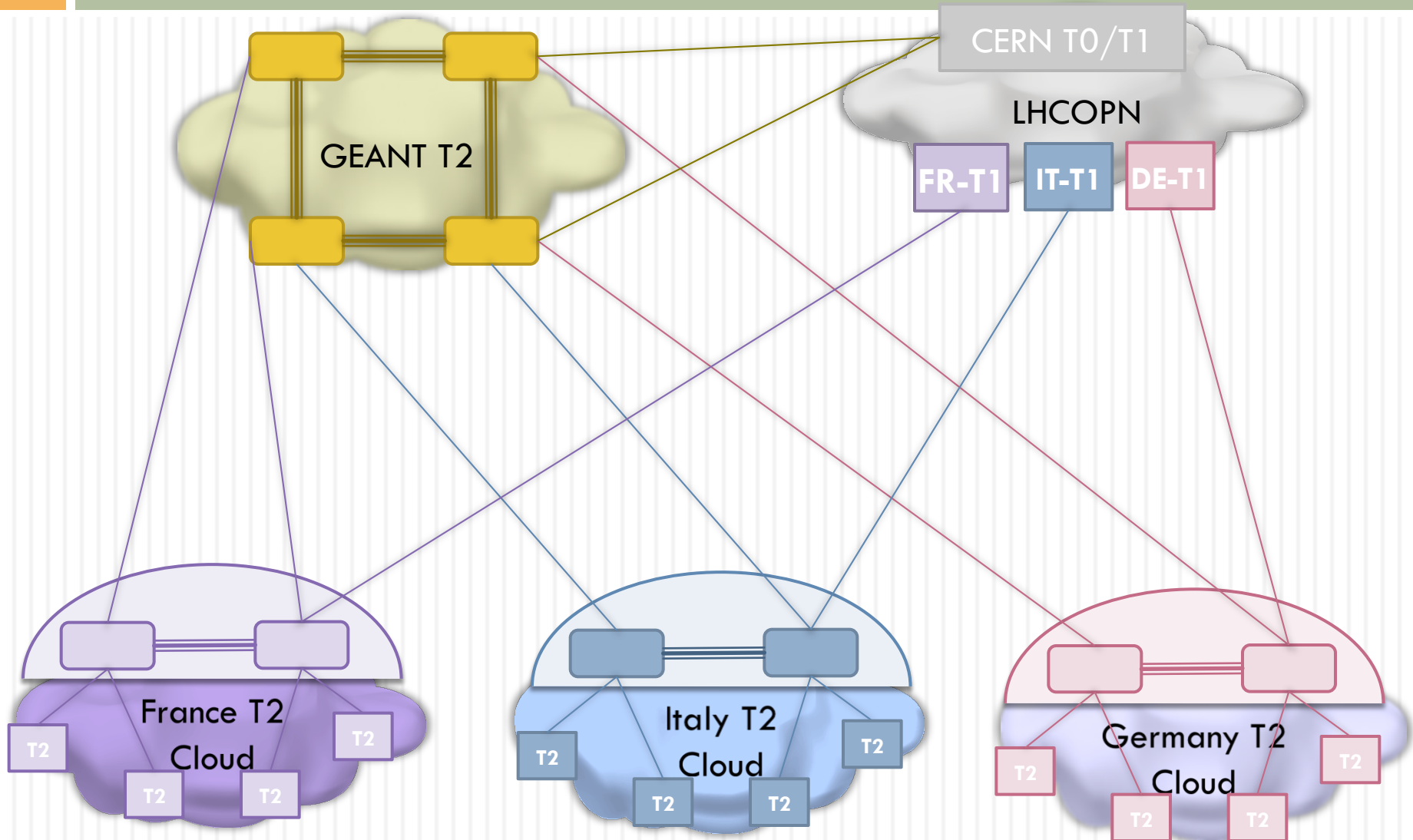
Grazie

21

# Backup slides

# LHCT2s prototype interconnection

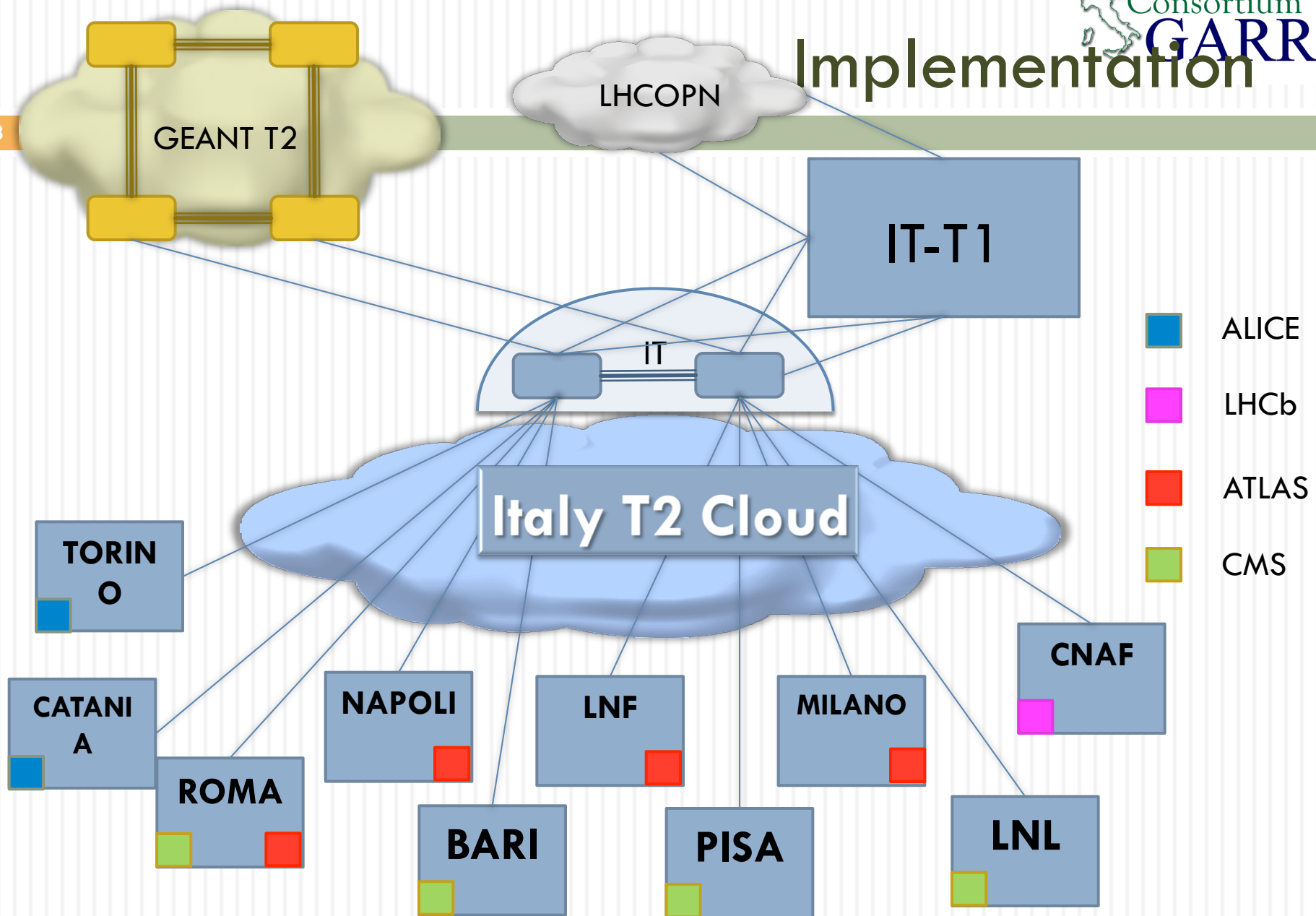
22



# IT-T2 Cloud Implementation

Consortium  
GARR

23



# Schema di rete per i T1/T2 in Italia

24

