CERN-CNAF DCI

Involved organizations

CERN, GEANT, GARR, INFN

Edoardo Martelli (CERN)
Lorenzo Chiarelli (GARR e INFN CNAF)
Donato De Girolamo (INFN CNAF)
S.Zani (INFN CNAF)

Guy Roberts (GEANT)
Gloria Vuagnin (GARR)
Paolo Bolletta (GARR)
Massimo Carboni (GARR)
Matteo Colantonio (GARR)

18-10-2023
LHCOPN-LHCONE meeting

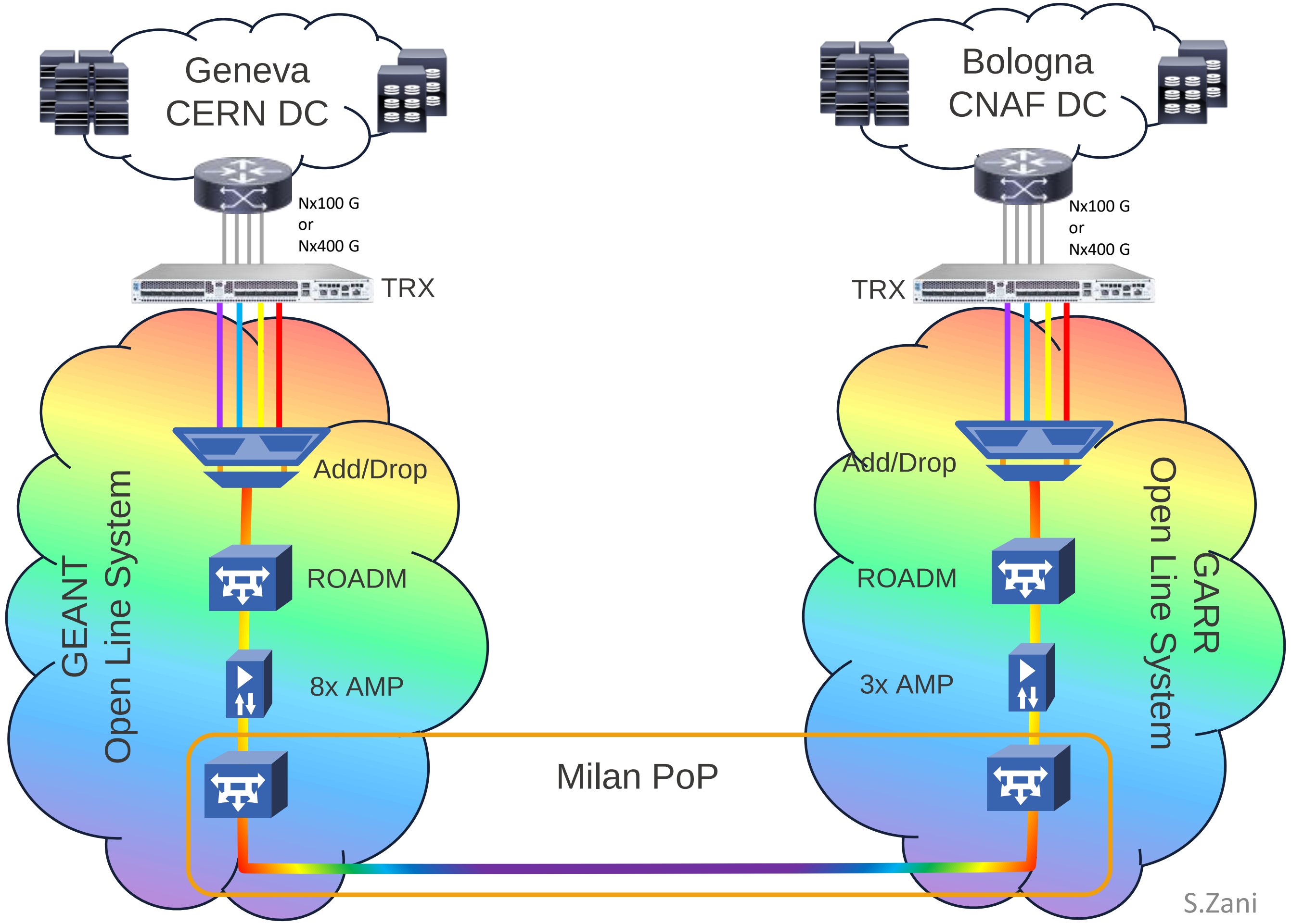
Main motivations to start a pilot project on an optical direct link between CERN and CNAF

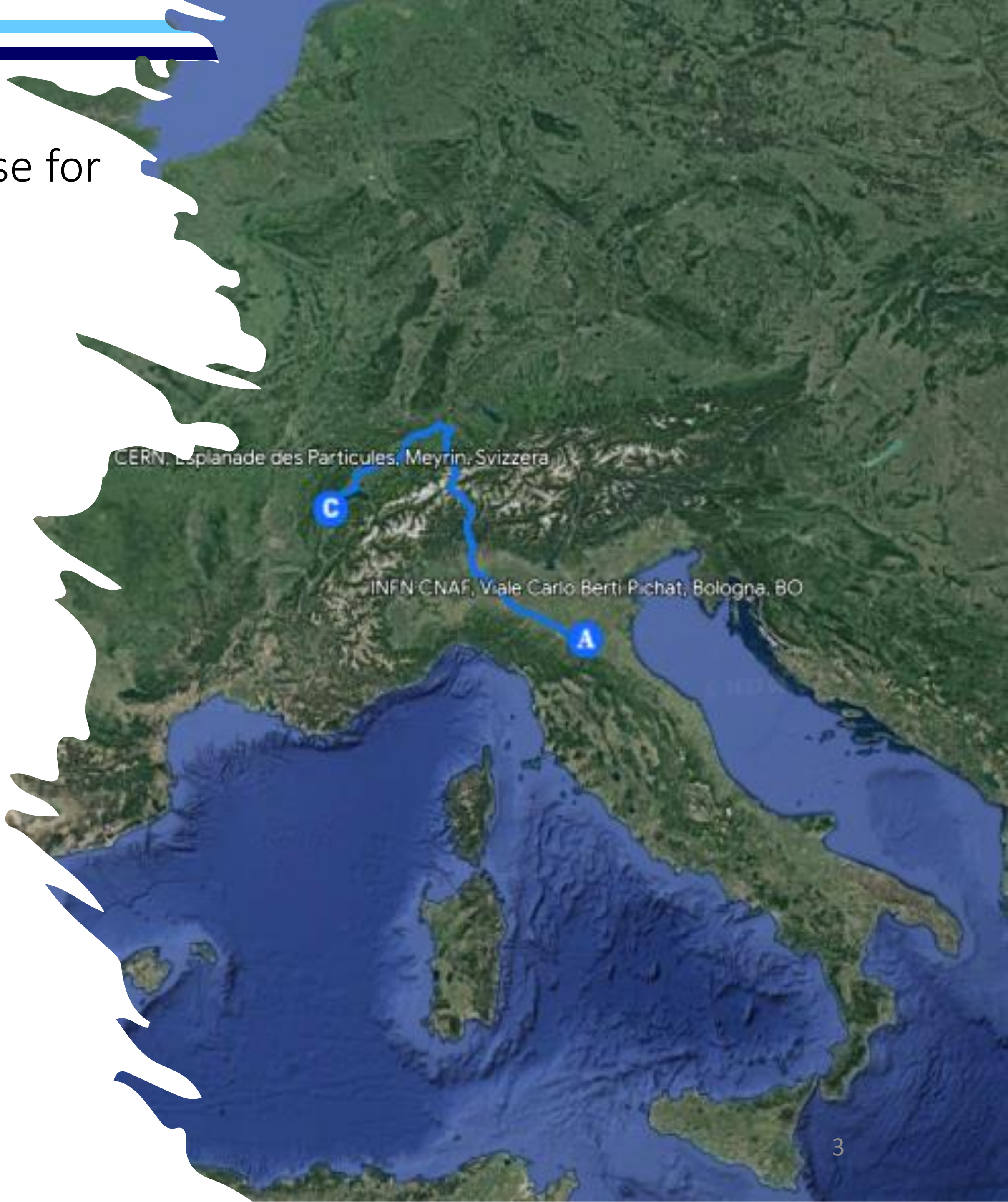
"HL-LHC connectivity requirements for TIER1s: More than 1Tbps in 2027" (S.Campana Presentation at LHCONE/OPN meeting 23/3/2021)

- A direct link with CERN without traversing too many **router interfaces** should be a **cheaper** solution for High Bandwidth link.
- A programmable transponder based interconnection infrastructure could be used to **dynamically resize the T0 –T1 Link** on the needs.
- A DCI interconnection could be used as the underlay network for a Data Lake spanning between CERN and CNAF
- High bandwidth end to end **optical channels for specific purposes** could be deployed in few minutes.
- More efficient use of the physical network infrastructure, sharing spectrum available on NREN backbones.

Shematic view of the DCI realized

Proposed in GEANT GN4-3 (WP7-T2) as a possible usecase for experimenting the multi domain **S**pectrum **C**onnection **S**ervice at about 1000 km of distance.





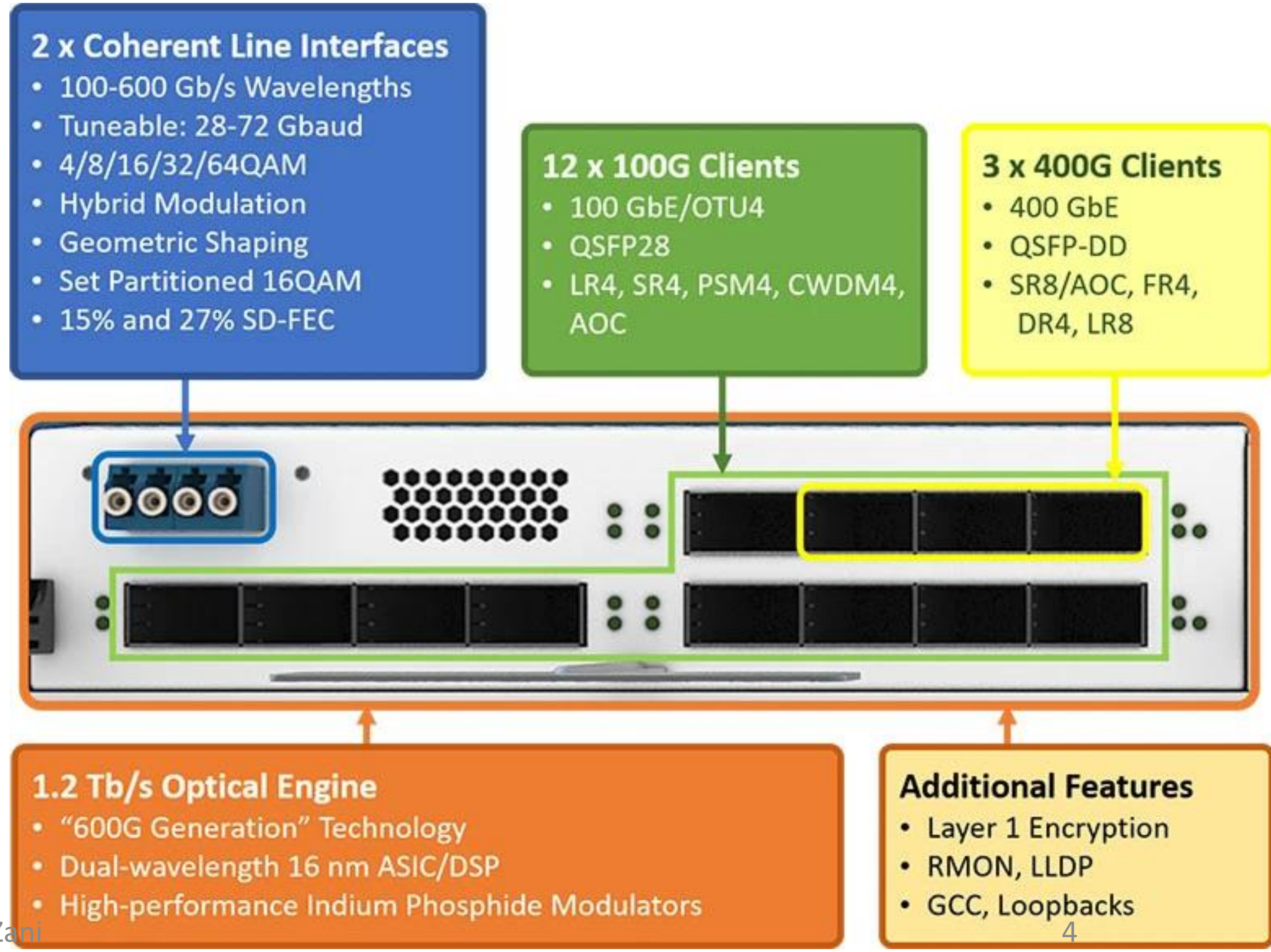
HW Installed for the experimental DCI CERN-CNAF

The transponders installed in test phase **are ready for the future production phase** (2 devices in each site).
Power consumption: < 700W
Budget per site: 80 K€

CERN
2 x Groove G30 (4 slot Transponder)
2 x CHM2T DCI Metro card (one for each G30)
2 x Plug QSFP28 (SR4) + 1 QSFP-DD 400G

CNAF
2 x Groove G30 (4 slot Transponder)
2 x CHM2T DCI Metro card (one for each G30)
2 x Plug QSFP28 (SR4) + 1 QSFP-DD 400G

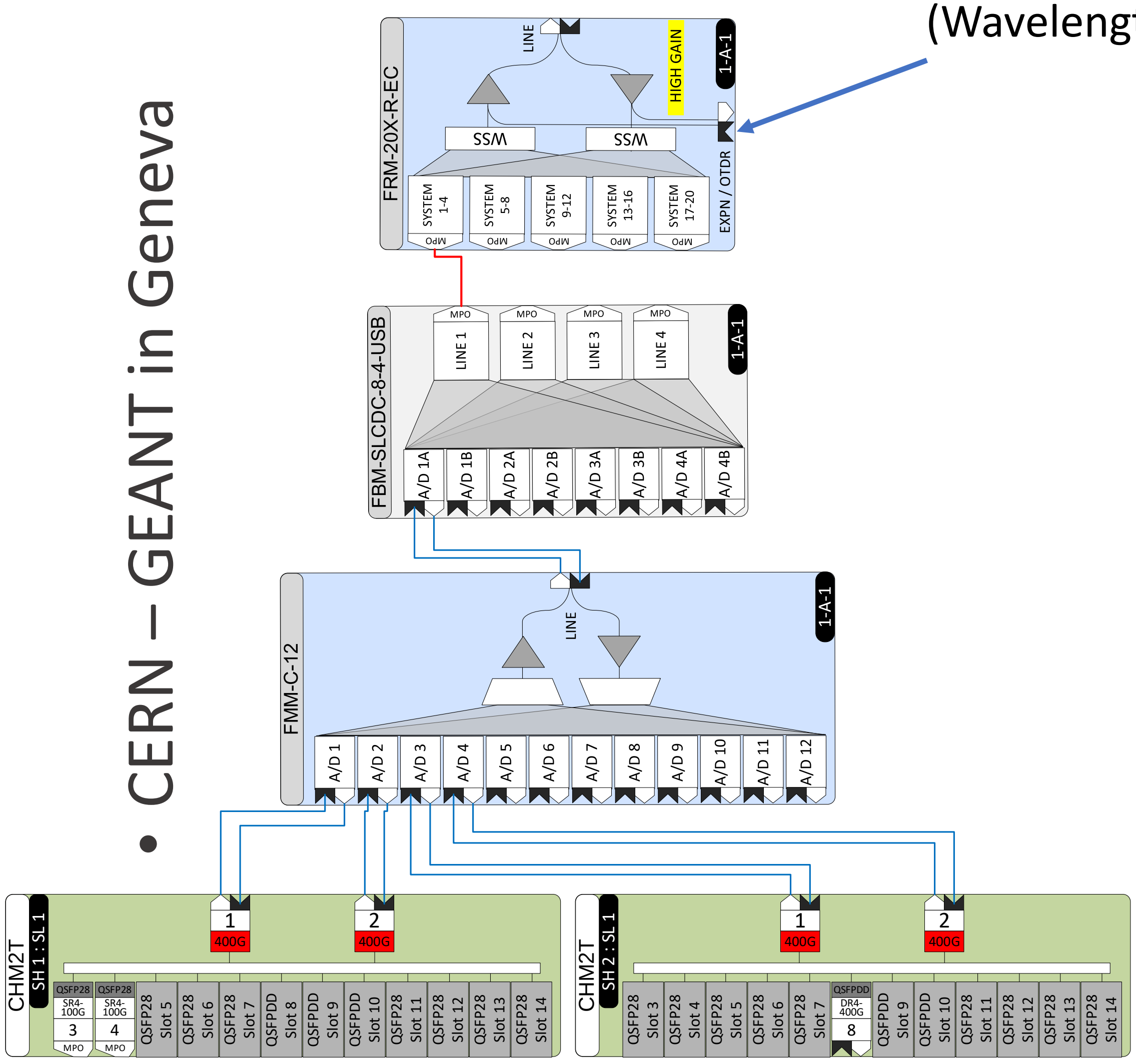
CHM2T



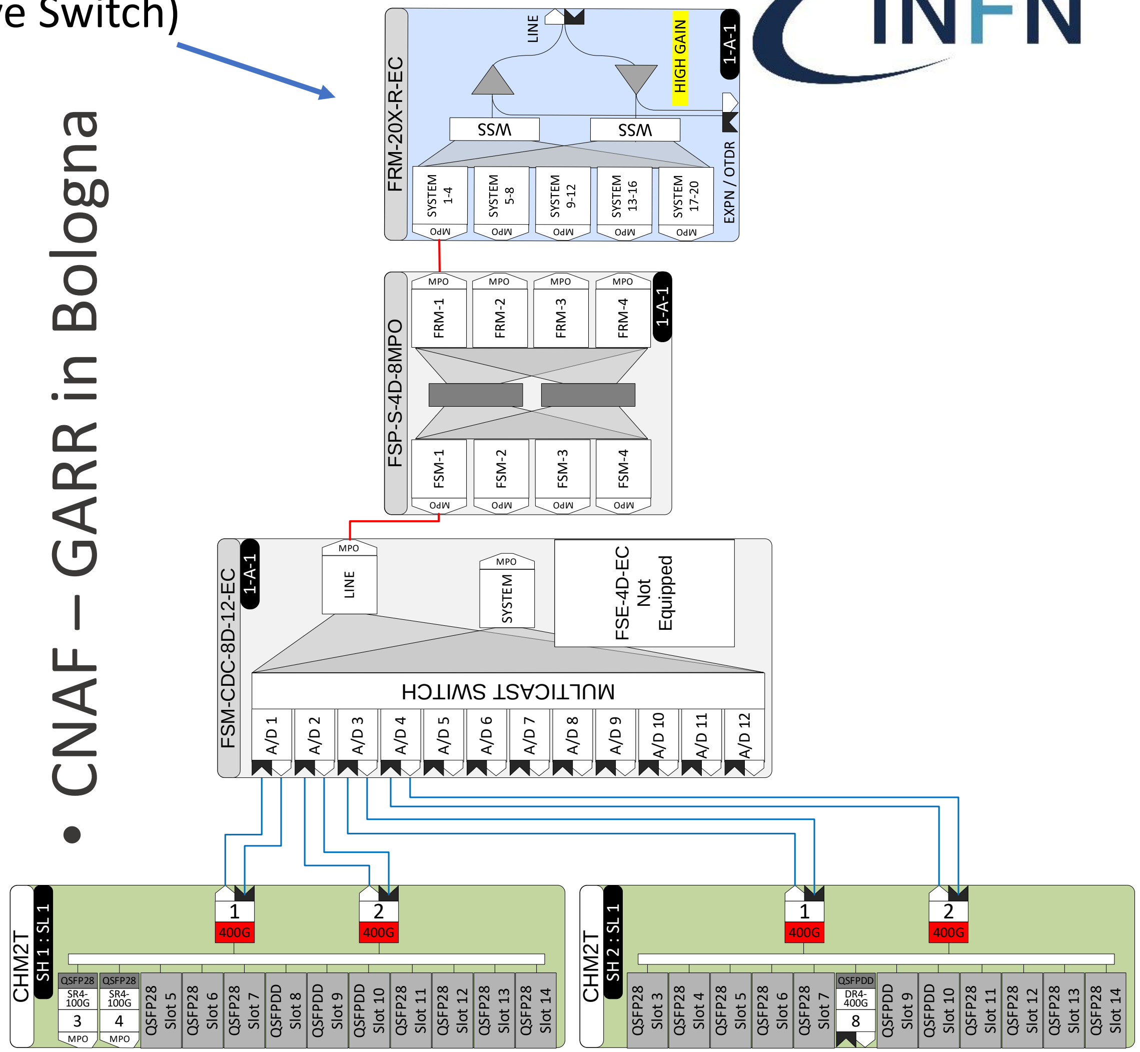
UNIs

ROADM WSS
(Wavelength Selective Switch)

- CERN – GEANT in Geneva



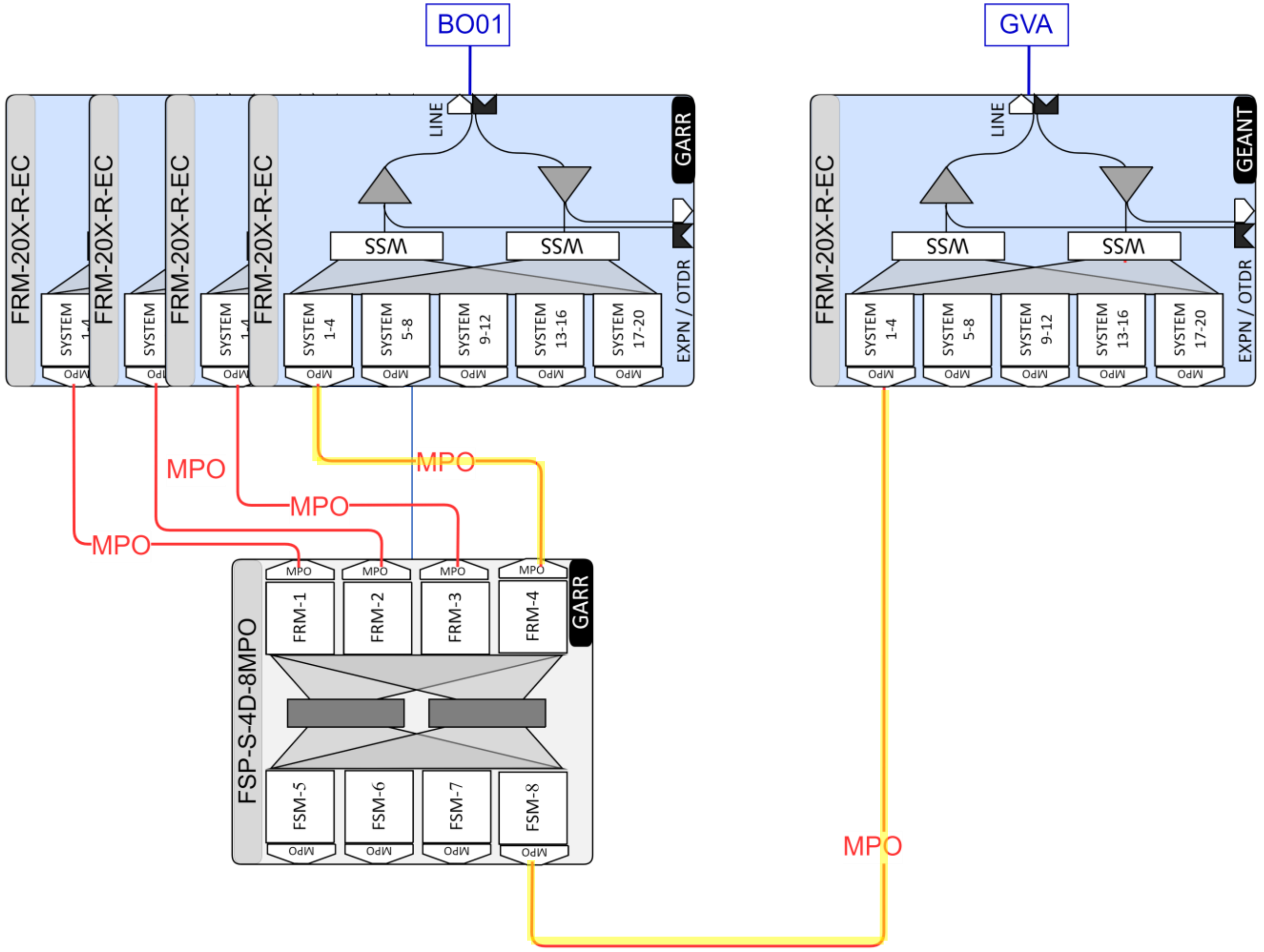
- CNAF – GARR in Bologna



Configuration of the "Line system" CERN-GEANT side and CNAF-GARR side

NNI (POP GEANT-GARR POP located in Milan)

- WSS directly interconnected
  - **Privacy:** Each circuit is present only on a fiber (no broadcasts)
  - **Spectrum Policing:** circuits are interrupted if the optical power is out of the allowed range.
  - **Spectrum Shielding:** The optical signal is not affected by overlapping with possible spurious signals on the same frequency

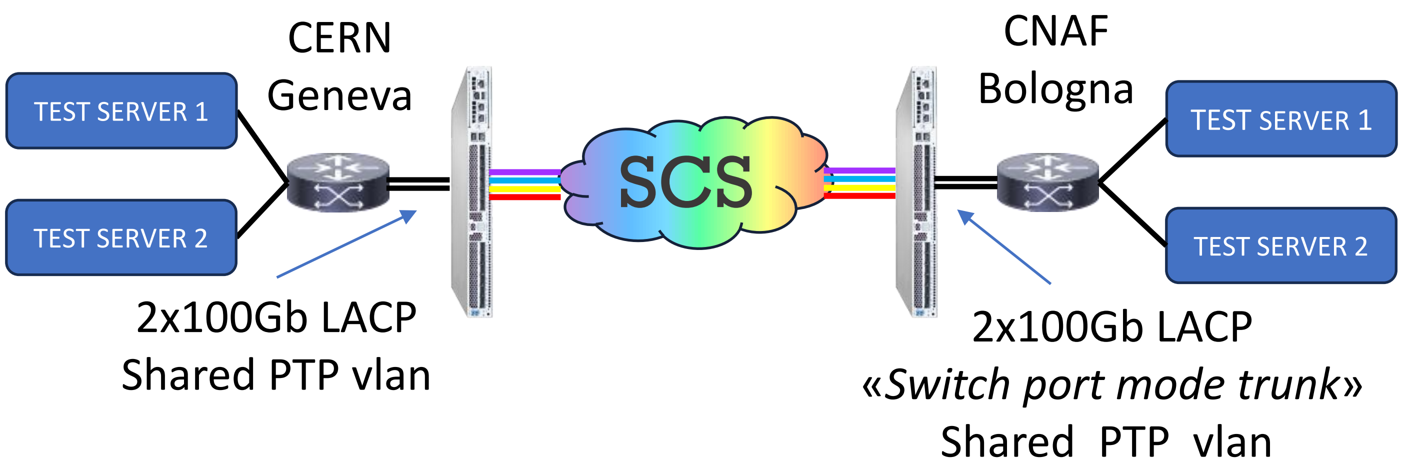


FRM = 1x20 twin WSS w/ amps

FSP-S = shuffles MPO

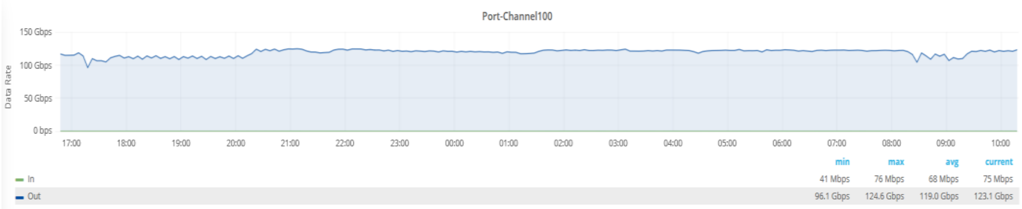
Current state of DCI (CERN-CNAF)
1.6 Tbps available end to end

On 1000 km distance using **69 Gbaud DP-16 QAM*** modulation, it is possible to reach **400Gbps per carrier**
Having 4 line ports on transpoders, it is possible to reach 1.6 Tbps on this «Circuit» that could be used as up to 4x400Gb Ethernet or 16x100Gb Ethernet .
The spectrum occupation for each carrier is 100Ghz, so 400Ghz (10% of the C band) are suffcient to transport **1.6 Tbps with 9.5 ms of Round Trip Time (Standard routed path is about 13.5ms)**



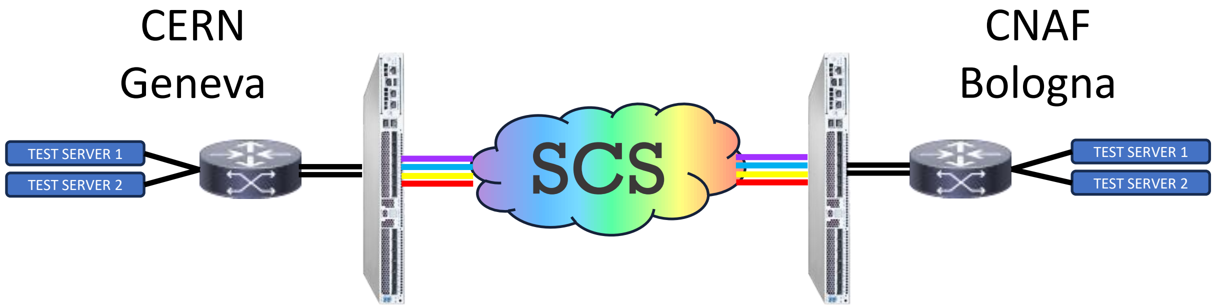
First end to end IP traffic injection test:

Connected 2 servers in each site to inject about 100Gbps **sustainted** traffic (iperf) for days in order to test the **stability** of the circuit and to give the opportunity to tune up the monitoring tools on line systems.



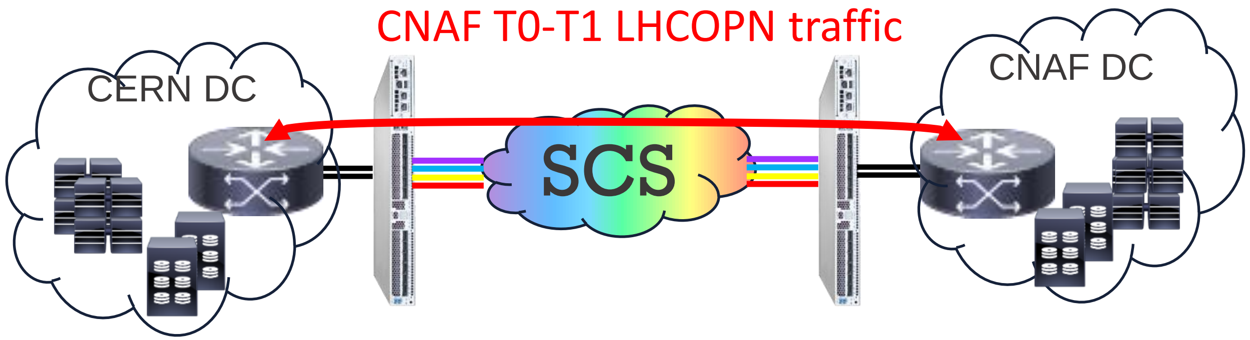
*Dual-Polarization 16 Quadrature Amplitude Modulation

DCI pilot state (CERN-CNAF)
Next steps



Grow in terms of capacity adding 100G Ethernet ports and activating 400G Ethernet interfaces.
Test different transceivers (also third party transceivers on Transpoders)

- 100G LR (10km)
- 100G FR (2km)
- 400G FR4 (2km)
- 400G LR4 (10km)



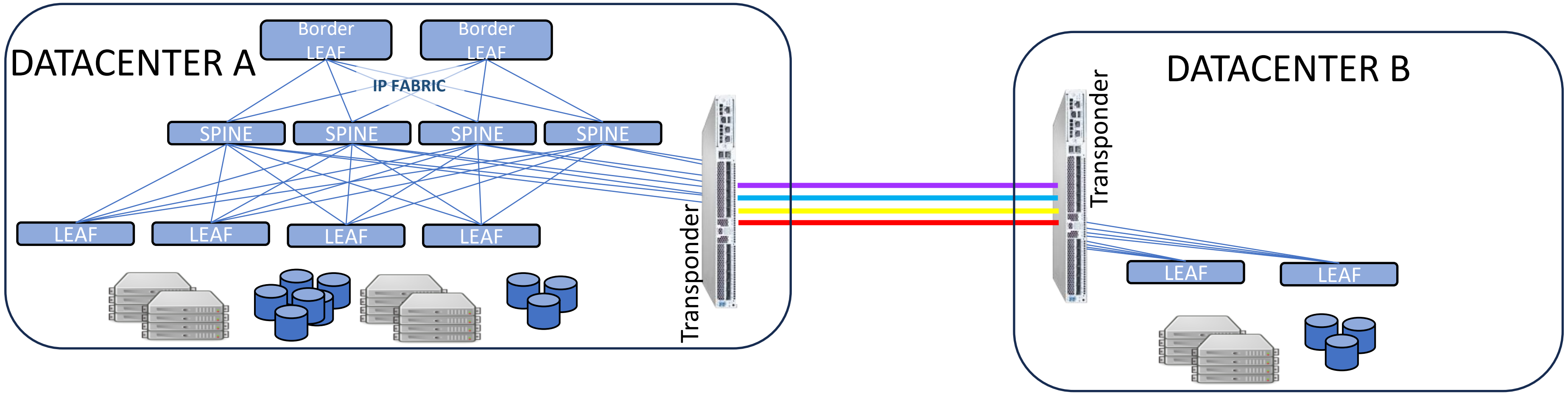
Connect DCI directly to Production CORE Routers at CNAF and to production Routers ad CERN
Move OPN (T0-T1) Traffic on DCI circuit **to test the link with real traffic** setting up the backup path on currently routed lhcopn connectivity.

New opportunities

Considering that, this kind of DCI gives the possibility to connect up to 16 x 100G Ethernet or 4x400G Ethernet interfaces to whatever device both sides having only 9.5 ms RTT penalty the possible applications could be many.

Some examples:

- A «Classic» BGP peering with a dynamic number of ports contributing to the total capacity - **Easy growing in capacity (Shure interest for future production OPN link CNAF)**
- L2 Stretching of specific networks or IP Fabric stretching between the DCs having for example Leaves located in different Datacenters using the DCI links between spines and Leaves



That’s it

Backup Slides

- https://indico.cern.ch/event/983436/contributions/4226012/attachments/2213578/3746895/LHCOPN-LHCONE-Mar2021.pdf

Network needs for HL-LHC

T1	%ATLAS	%CMS	% Alice	% LHCb	ATLAS+CMS Network Needs (Gbps) Minimal Scenario in 2027	Alice Network Needs (Gbps) Minimal Scenario in 2027	LHCb Network Needs (Gbps) Minimal Scenario in 2027	LHC Network Needs (Gbps) Minimal Scenario in 2027	LHC Network Needs (Gbps) Flexible Scenario in 2027
CA-TRIUMF	10	0	0	0	200	0	0	200	400
DE-KIT	12	10	21	17	450	80	70	600	1200
ES-PIC	4	5	0	4	180	0	20	200	400
FR-CCIN2P3	13	10	14	15	450	60	60	570	1140
IT-INFN-CNAF	9	15	26	24	480	110	100	690	1380
KR-KISTI-GSDC	0	0	12	0	0	50	0	50	100
NDGF	6	0	8	0	110	30	0	140	280
NL-T1	7	0	3	8	140	10	30	180	360
NRC-KI-T1	3	0	13	5	50	50	20	120	240
UK-T1-RAL	15	10	3	27	490	10	110	610	1220
RU-JINR-T1	0	10	0	0	200	0	0	200	400
US-T1-BNL	23	0	0	0	450	0	0	450	900
US-FNAL-CMS	0	40	0	0	800	0	0	800	1600
(atlantic link)					1250	0	0	1250	2500
Sum	100	100	100	100	4000	400	410	4810	9620