Consortium GARR
THE ITALIAN EDUCATION & RESEARCH NETWORK

# GARR optical network evolution and DCI for Data Lake for Science

PAOLO BOLLETTA (GARR)

CERN, 16/01/2020

4th SIG-NGN Meeting

# Outline

- GARR Optical Network Update
- IDDLS project
- DCI Lab Trials
- Impact on network evolution
- Conclusions

# Status of the current infrastructure

## (2011) Huawei 1/10 G

- 10G/40Gbps channels not enough for the needs of the core
- Close to the end-of-life
- Maintenance (+2Y) ends 2020

## (2015) Infinera 10/40/100 G

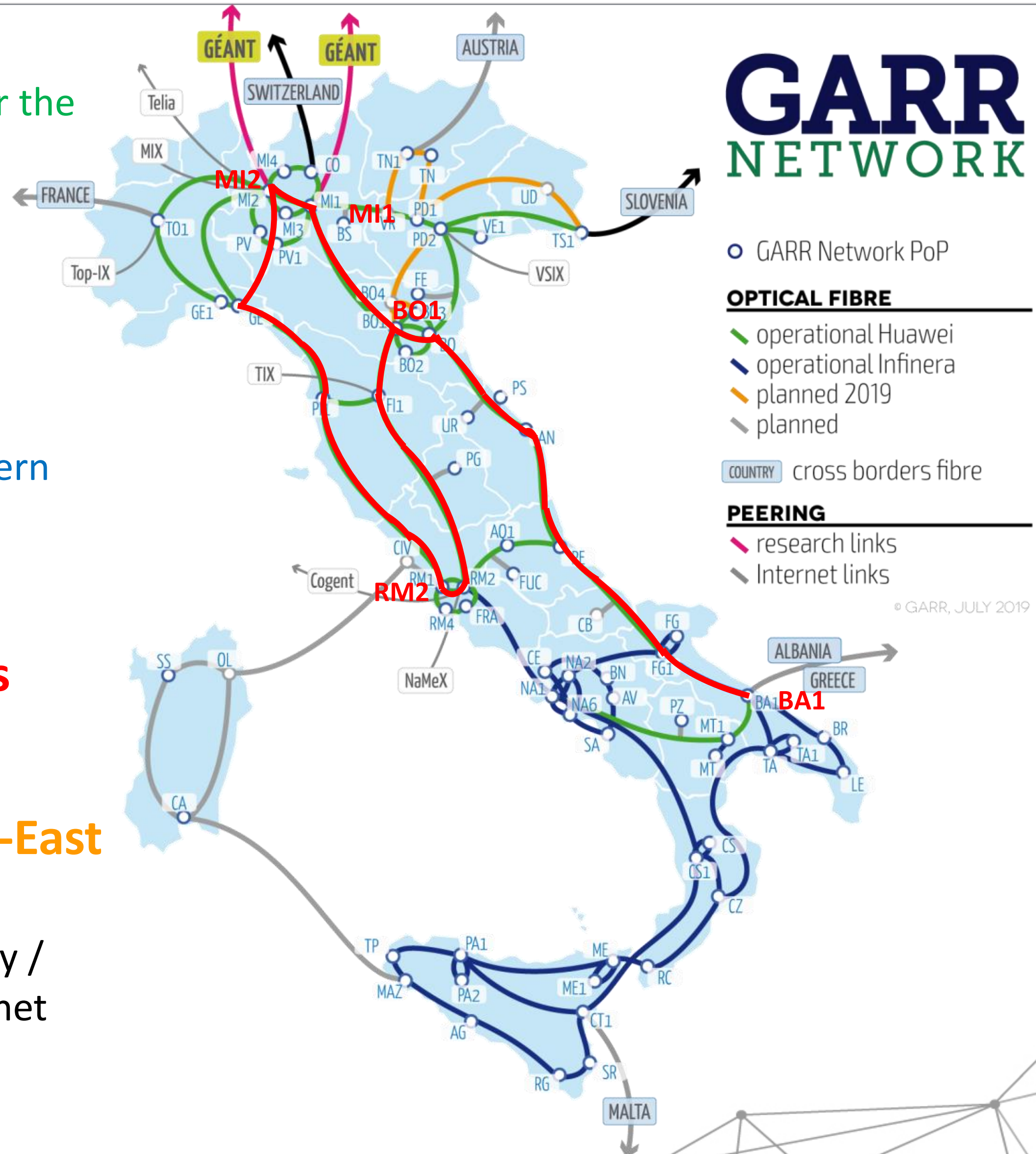
- 500Gbps superchannels
- Fine for the capacity needs of Southern Italy
- High power and space consumption

## (2017) AW 2x100G among Core PoPs

- Buffer solution – it works super fine

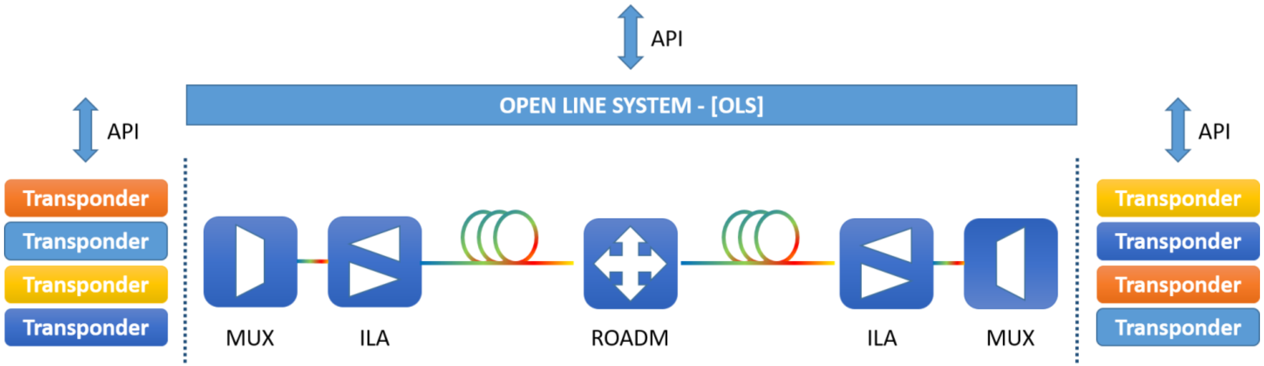
## (2019) New fiber footprint on North-East Area

- North-Eastern region: closed topology / ARNES direct Interconnection / Lightnet interconnection
- ECMWF new site in Bologna



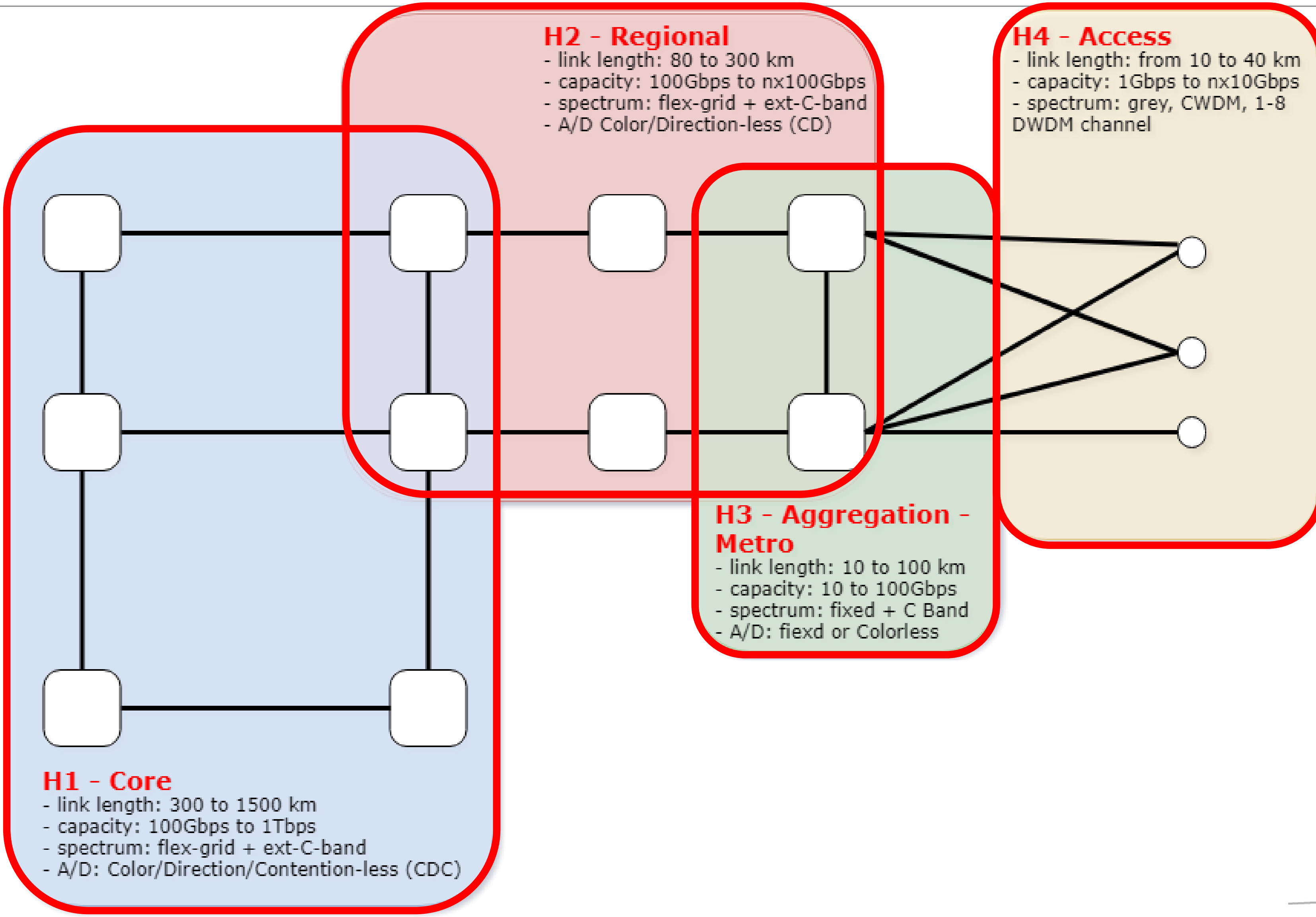
# How GARR sees its optical evolution

- Open Line System (partially disaggregated) to replace the oldest infrastructure

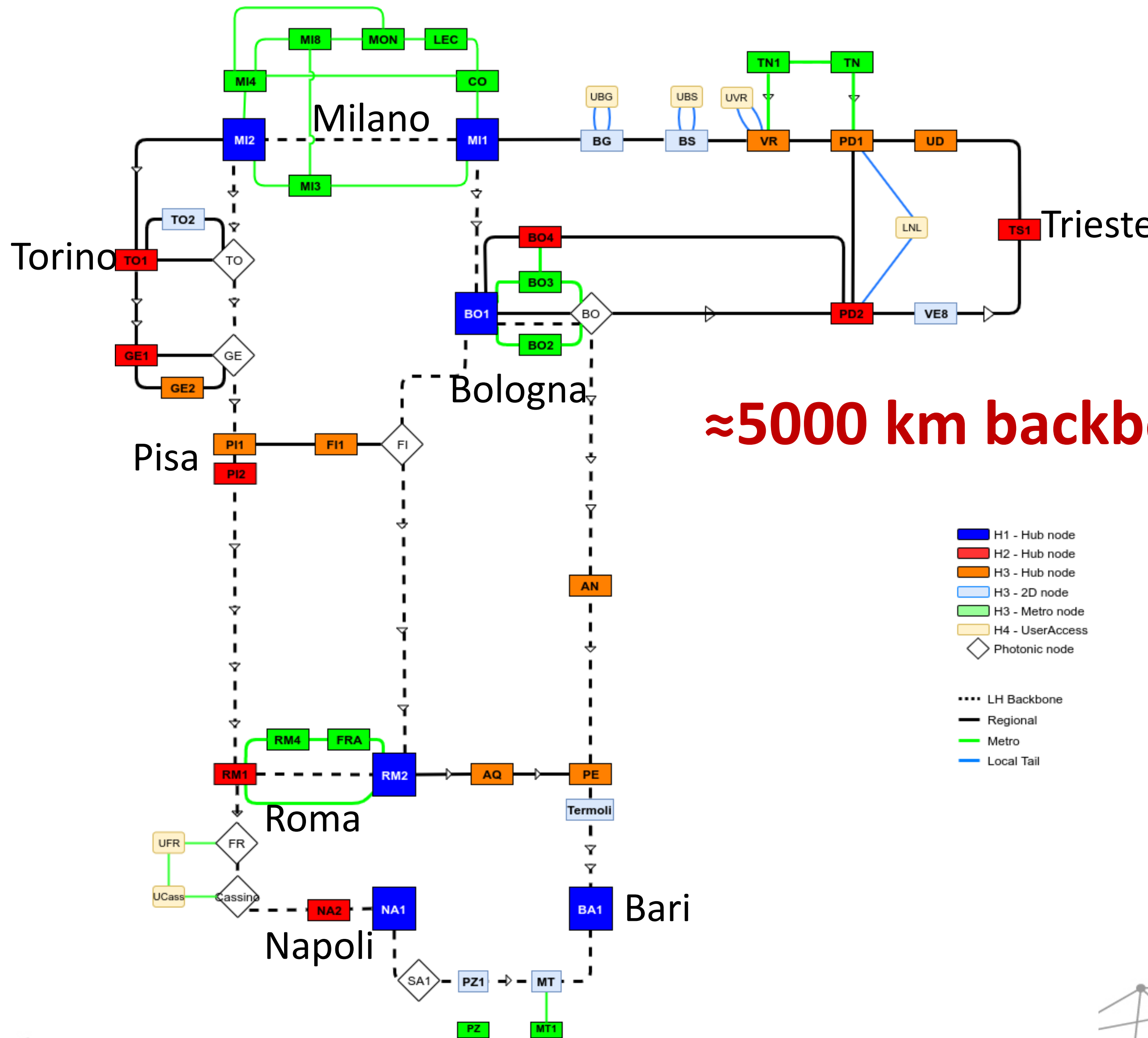



- April 2019: we issued a RFI on open line systems and optical transport network equipment (Juniper, ADVA, ECI, Huawei, Infinera)
- August: tender for DCI (Infinera G30) for an INFN-GARR joint project on Distributed Datalake for Science
- February 2020 – GARR-T Optical Transmission Network Tender
- September 2020 – start of the deployment phase

# Nodes and hierarchical optical network



# GARR-T optical fiber topology




**≈5000 km backbone fiber**

# Timeline

**Q2/Q3 2019**

- RFI with 5 vendors
- DCI equipment LAB trial
- Disaggregated platform LAB trial

**Q3/Q4 2019**

- DCI equipment procured for data lake project
- LAB trial Time Over WAN
- New network requirement definition and HLD

*February 2020*

***OLS e Transport Platform Tender***

**Q2 2020**

- Award of the Tender and start of the Field Trial phase

**Q3 2020**

- Field Trail outcome
- Phase 1 deplyment

*Q4 2020*

***Phase 1 migration***

**Deployment scheduled in different phases**

**2022 end of deployment and migration**

# USE CASE GARR – ELISA : High Capacity on demand

## Problem Statement

Complete separation of user and GARR control and management planes

**[MAIN GOAL]** Dinamically establish high capacity interconnections between user premises (DC)

Provide interconnections self-management tools to the users

DC/Transport Network integrated load and failure mangement

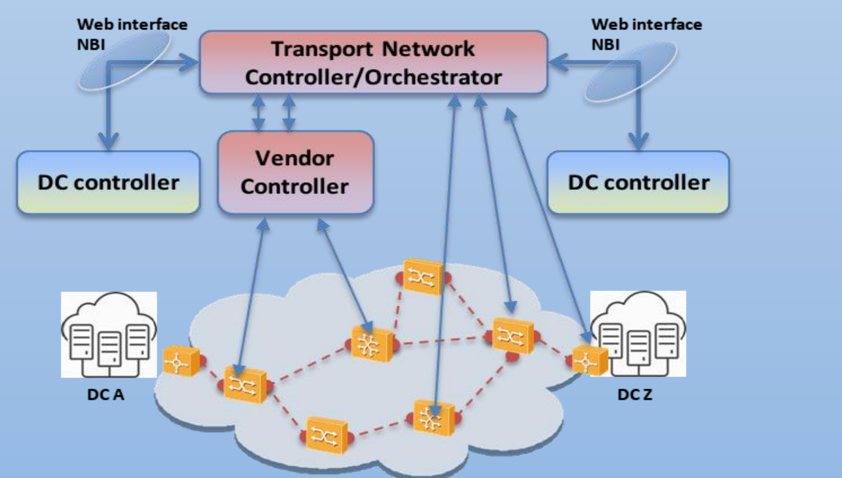
## Actors

- GARR
  - (transport) network controller/orchestrator
  - provisioning/monitoring/accounting performance/fault management tools
- User
  - (DC) controller
  - provisioning/monitoring/accounting performance/fault management tools
- Vendor controller

## Solution

- **A**: DC operators manually (?) ask for resources through a transport network interface (web/transport nms access)
- **B:** DC controller interacts with transport network through a network controller/orchestrator NBI
- **C:** DC App interacts with transport network through a network controller/orchestrator NBI

## Architecture/Diagram (e.g.)



# USE CASE: DCI over Spectrum Connection Service

Description:

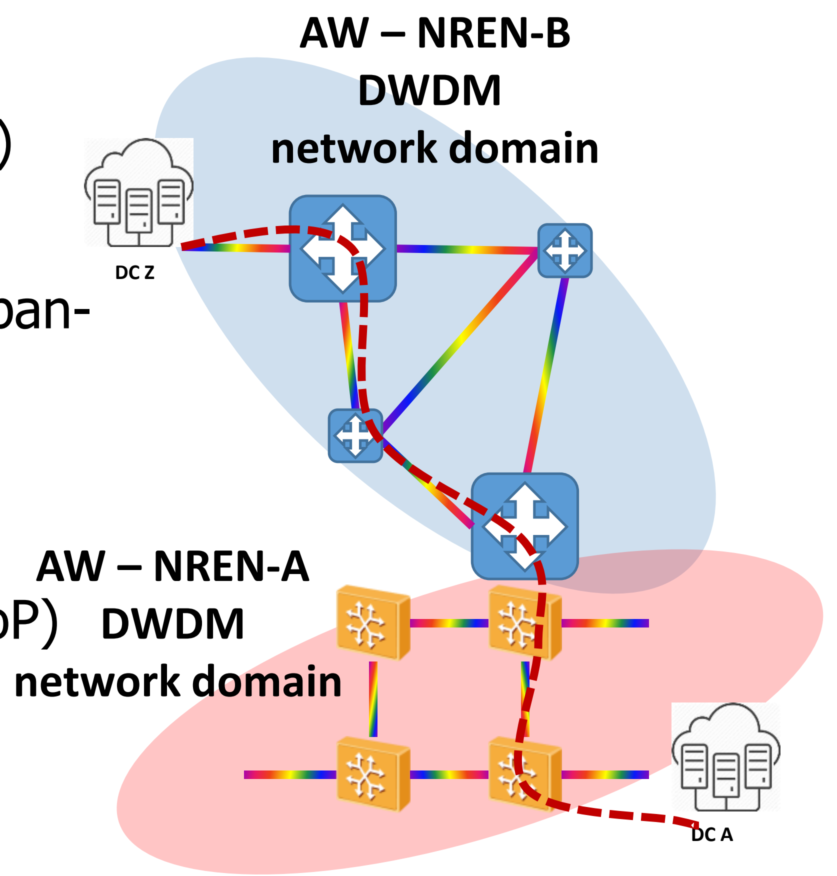
interconnection between 2 DC over multi-domain DWDM Networks (NREN + GEANT)

Prerequisites:

- Spectrum Connection Service over the pan-European network
- AW capable dwdm network domains
- DCI boxes
- Demarcation Points (e.g. MIL GEANT PoP)

Actors:

- Geant
- NRENs
- DC facilities



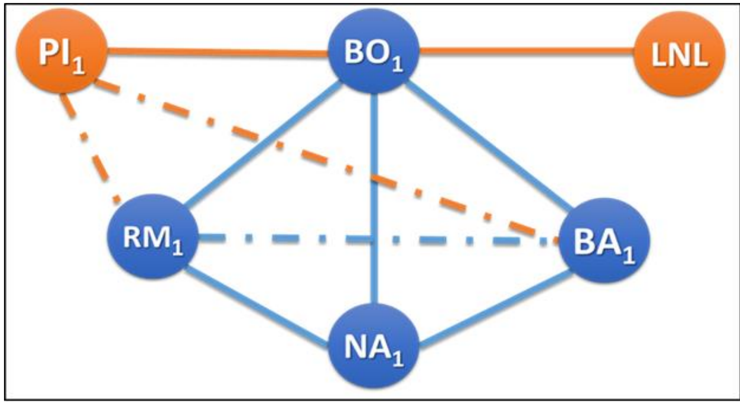
# IDDLS

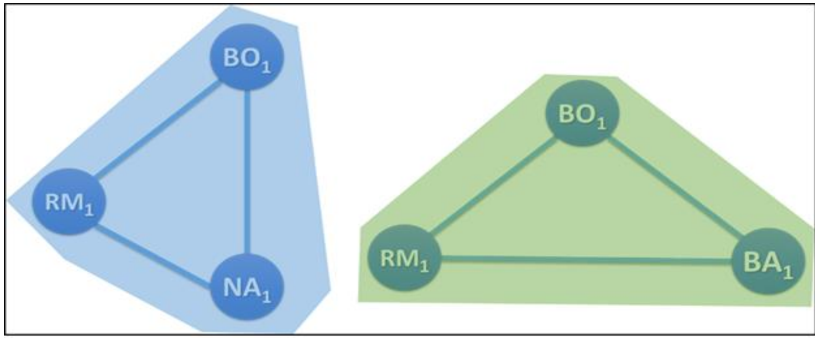
## Italian Distributed Data Lake for Science

# IDDLS: Italian Distributed Data Lake for Science

INFN-GARR collaboration to setup a prototype of an Italian DataLake exploiting:

- Last generation networking technologies provided by GARR
  - DCI (Data Center Interconnection) equipment
  - SDN (Software Defined Network) deployment
- Software for creating **scalable storage federations** provided by INFN
  - eXtreme-DataCloud (XDC) project
  - SCoRES project (INFN-NA)
- Real life use cases for testing
  - CMS
  - ATLAS
  - BELLE-II
  - Possibly involving LNGS experiments (XENON) and VIRGO



# Timeline

## 3 years project

First year

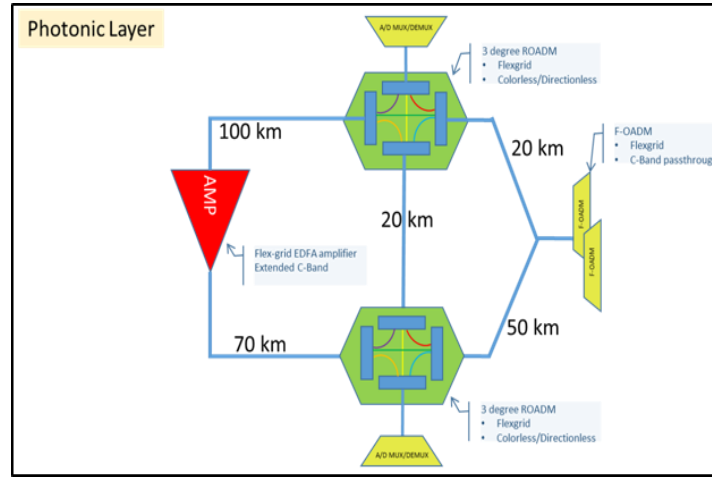
- Technology scouting for DCI equipment to be deployed by GARR
- Application (INFN) requirements analysis
- Network equipment acquisition (INFN and GARR) and Lab testing
- Deployment on mixed Lab+WAN environment of the networking equipment
- Creation of the DataLake on sites connected with standard networking and first prototype using DCI

Second year

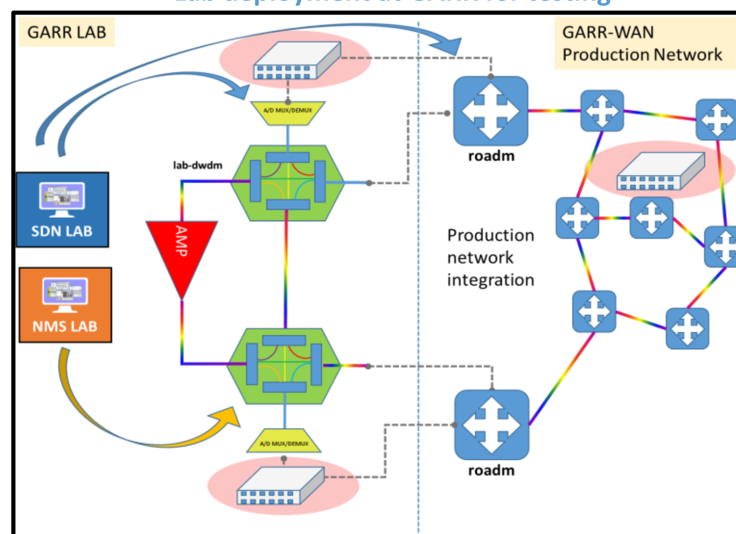
- Testing of the mixed (Lab+WAN) configuration
- Final creation of the DataLake on the 3 INFN sites with DCI systems
- Performance evaluation and comparison
- Possible acquisition of new equipment with increased performance

Third year

- Deployment only on WAN of the networking equipment
- Optimization of the DataLake
- Performance evaluation
- Final consideration




Lab deployment at GARR for testing




Mixed Lab+WAN deployment

# Milestones 2019 and status

**January 2019: Kickoff meeting @GARR**

- https://agenda.infn.it/event/17957/

**January – June 2019:** test transmission platforms in the GARR Lab

**End of June 2019: networking hardware chosen**

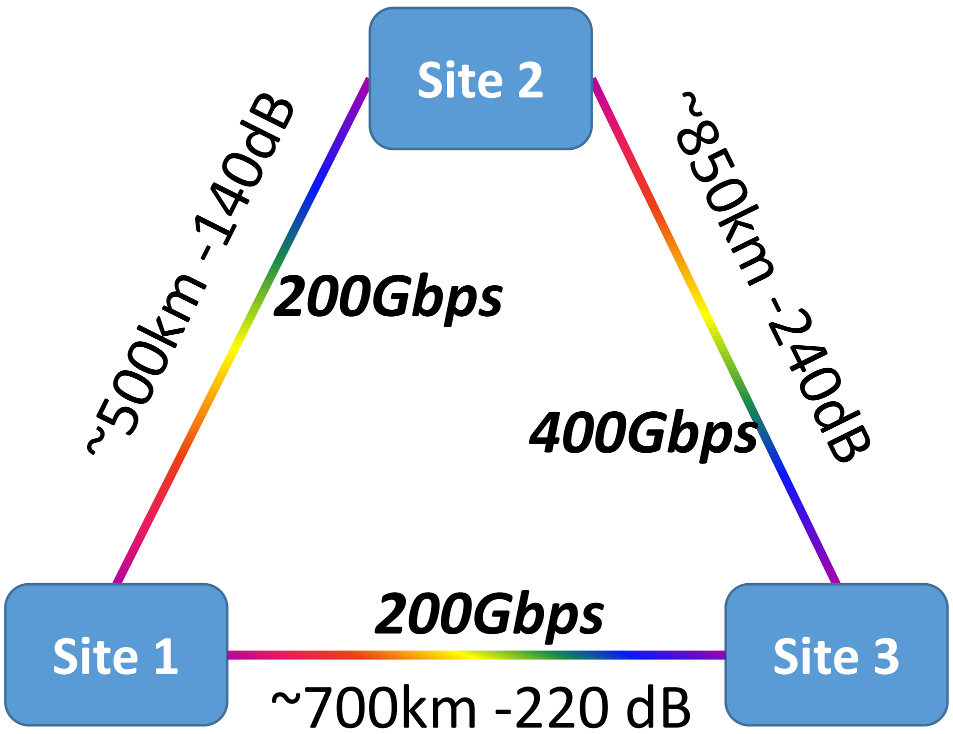
- **Transmission platfom and DCI scouting and lab testing**
  - Infinera Groove G30 Selected and under procurement
- **100Gb switches (acquired after tender)**
  - Arista 7050SX3 (32x1000Gbs)

**Q1 2020: Networking systems deployment in a Lab+WAN configuration**

- First DataLake prototype

**Q2 2020: First evaluation runs**

# Market scouting / RFI




Reference Topology and Traffic Demands

AW transport over compensated DWDM network

2 approaches

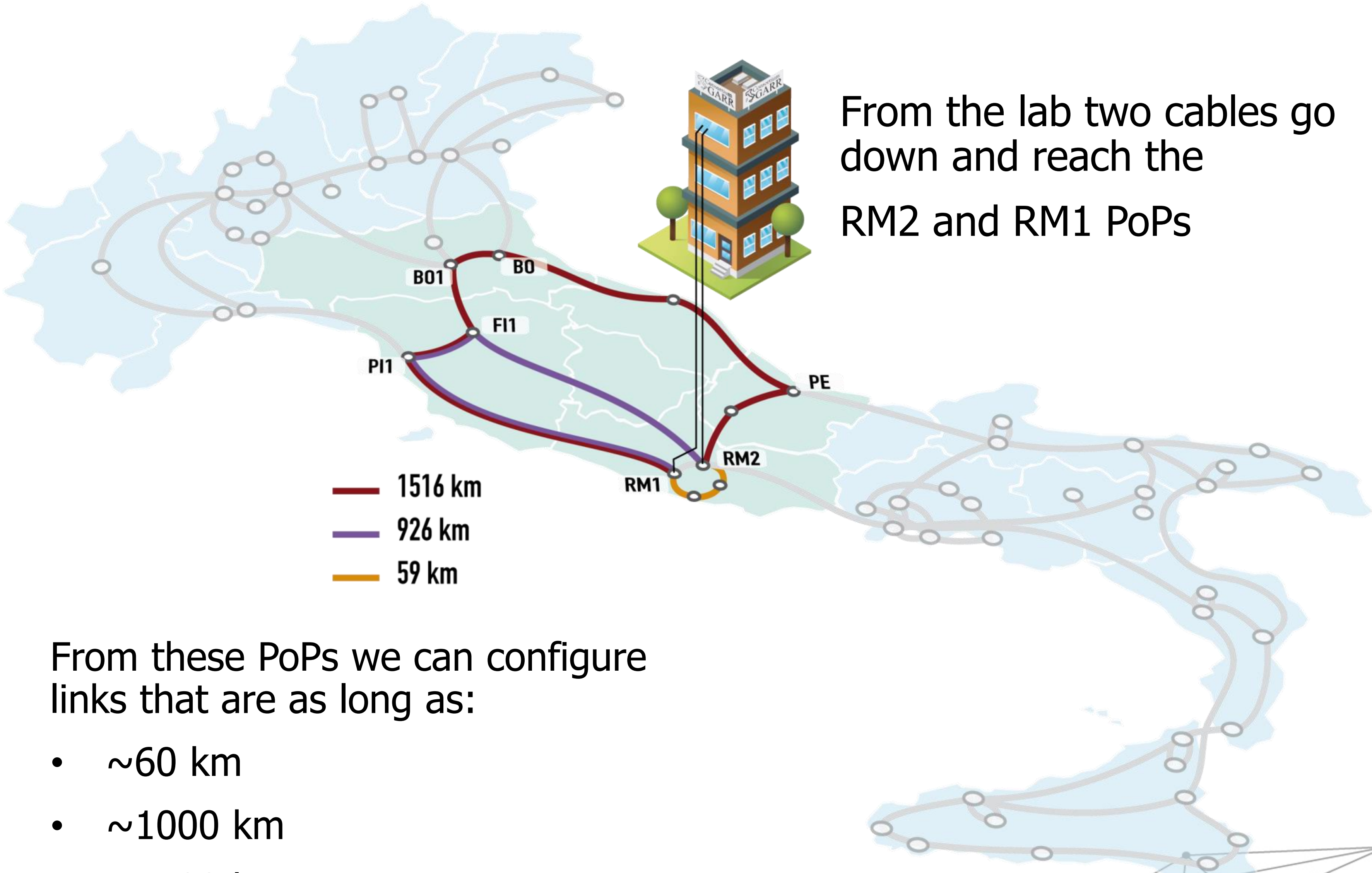
**Integrated :**
Packet+DWDM in a single BOX

Lacks in:
- Management and control
- Platform maturity
- Scalability

**Disaggregated:**
- DWDM DCI
- PKT layer Device for overlays and dynamic reconfiguration

Pros:
- Cheaper solutions
- Better scalability
- No dependencies in platform choice

# LAB and Field Trials




From the lab two cables go down and reach the

RM2 and RM1 PoPs

From these PoPs we can configure links that are as long as:

- ~60 km
- ~1000 km
- ~1500 km

# Paths on the production environment

| End Sites | km | dB | Ila | Raman | ROADM |
|---|---|---|---|---|---|
| RM1 – RMT | 22 | 10 | 0 | 0 | 2 |
| RMT -- FRA | 8 | 5 | 0 | 0 | 1 |
| FRA – RM2 | 29 | 12 | 0 | 0 | 1 |
| **Total** | **59** | **27** | **0** | **0** | **4** |

Short Metro Ring

Intermediate Ring

| End Sites | km | dB | Ila | Raman | ROADM |
|---|---|---|---|---|---|
| RM1 – PI | 430 | 118 | 4 | 0 | 2 |
| PI – FI | 136 | 40 | 0 | 2 | 1 |
| FI – RM2 | 360 | 104 | 3 | 4 | 1 |
| **Total** | **926** | **262** | **7** | **6** | **4** |

| End Sites | km | dB | Ila | Raman | ROADM |
|---|---|---|---|---|---|
| RM1 – PI | 430 | 118 | 4 | 0 | 2 |
| PI–FI-BO1 | 259 | 74 | 0 | 4 | 2 |
| BO1-BO–PE | 466 | 144 | 4 | 0 | 3 |
| PE – RM2 | 361 | 96 | 2 | 0 | 2 |
| **Total** | **1516** | **432** | **10** | **4** | **9** |

Long Ring

# Test transponders 100G

1. Integrated Transponder
   Client: 100G Ethernet
   Line: CFP1 100G ACO - QPSK coherent

2. Pizza Box DCI - DCO – with pkt integration option
   Client: 100-Gbps (QSFP28) pluggable interfaces
   Line: 200-Gbps CFP2-DCO coherent DWDM pluggable interfaces, which support 100-Gbps QPSK and 200-Gbps 8QAM, and 16QAM modulation options

3. Pizza Box DCI - ACO – Disaggregated Transponder
   Client: up to 4x100GbE QSFP28
   Line: CHM1 sled – 400G 2xCFP2-ACO (100G QPSK, 150G 8QAM, 200G 16QAM)

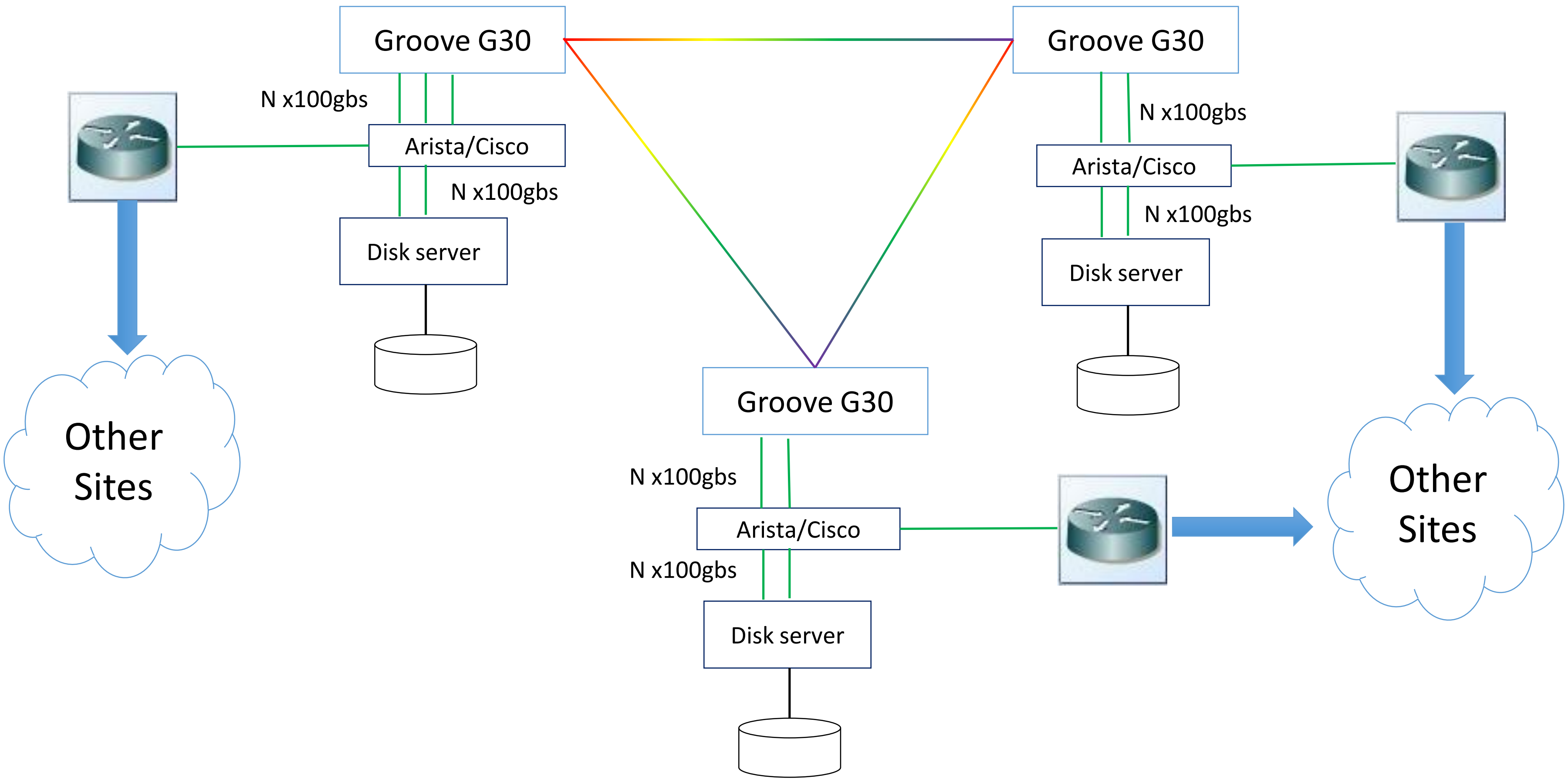
# Outcomes of the test sessions

**Pizza Box DCI - ACO – Disaggregated Transponder**
Its been only platform capable to address the reference topology in the current network
Infinera G30 Groove has been **selected** and **procured** for the project (**arrived yesterday!!**)

**Developed and teste monitoring tools for disaggregated DCI:**
**SNMP**
**INFLUXDB**
**GRAFANA**


**Details presented @ CEF 2019:**
**https://www.cesnet.cz/wp-content/uploads/2019/09/gloria.pdf**

# IDDLS Pilot Diagram

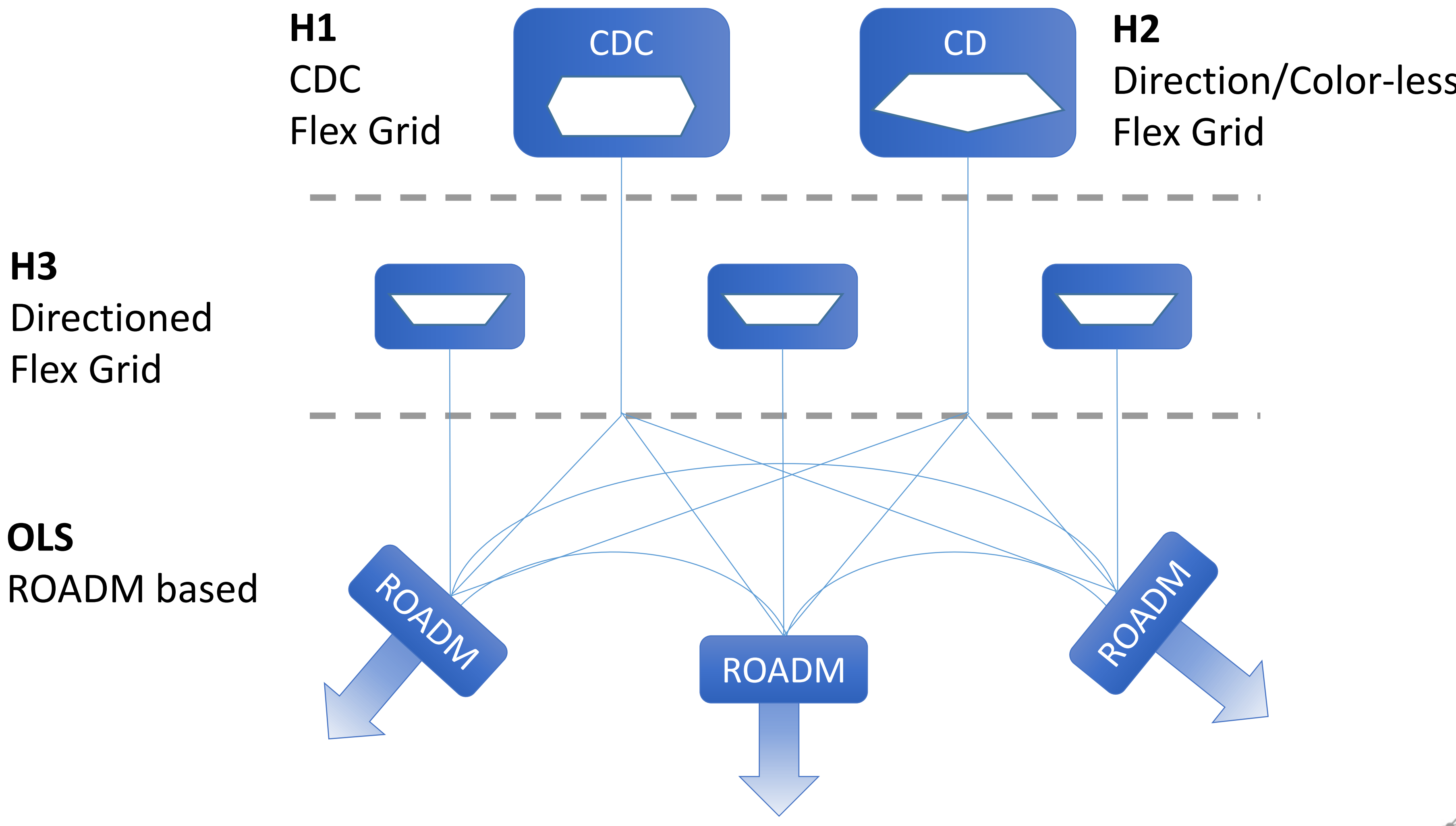


# Next Steps

- Use Case Demonstration
- Provisioning and service configuration via APIs
- Topologies and services reconfiguration automated with Ansible (DCI + Packet Layer)
- 3rd party controller management and control (ONOS)
- 3rd party planning and design tool (TIP GNpy)

# Impacts on network evolution

# OLS node A/D features scalability



# Spectrum Arrangement

**Extended C-BAND 4.8 THz (96 Ch)**

| | | | |
|---|---|---|---|
| **25% Backbone** | H1 ←→ H1 | 100/200Gbps+ SuperChannel | Flex Grid |
| **25% Regional** | H2 ←→ H1<br>H3 ←→ H2 | 100/150/200 Gbps | Fixed Grid 50GHz |
| **25% Spectrum Sharing** | e2e | Spectrum Slice | Flex Grid |

**25% future expansions**

NB 25% = 24 channels @50GHz

# Open approach to API/NBI

APIs and NBIs are mandatory for integration and net operations in a multivendor environment

*Network and equipment are logically ‘aggregated’*



**APIs:**

- REST
- NETCONF
- RESTCONF

**Data Model (YANG) equipment:**

- ONF Core model
- OPEN ROADM
- OpenConfig

**Data Model (YANG) network :**

- ONF Transport API
- IETF TEAS TE Topology

*Scope and target of these data models are not overlapped and span between different applications*



Consortium GARR THE ITALIAN EDUCATION & RESEARCH NETWORK

# Conclusions

- We are along the way …
- A lot of work done but also to be done

… since we are talking about lakes, seems to be easy to find birds of a feather and flock together

# Q&A