

# SDN con Openflow per il DAQ di KM3NeT

Lorenzo Chiarelli - Stefano Zani - Tommaso Chiarusi

Workshop GARR 2017 - Netvolution | Roma, 05-04-2017



# Sommario

- L'esperimento KM3NeT
- Elementi della rete di acquisizione
- Collegamento alla stazione di terra
  - Prime scelte tecnologiche per la costruzione di una rete «legacy»
- Problematiche e scelta di nuove tecnologie
  - Necessità di avere una rete programmabile
- Il testbed a Bologna (CNAF+KM3Net Bologna Common Infrastructure)
  - Definizione delle regole di forwarding dei pacchetti
- Deployment nella stazione di terra di Portopalo
- Impatto della nuova soluzione sull'esperimento
- Conclusioni e Sviluppi futuri

# L'esperimento KM3NeT

- Esperimento per lo studio dei neutrini

- Detector sottomarini

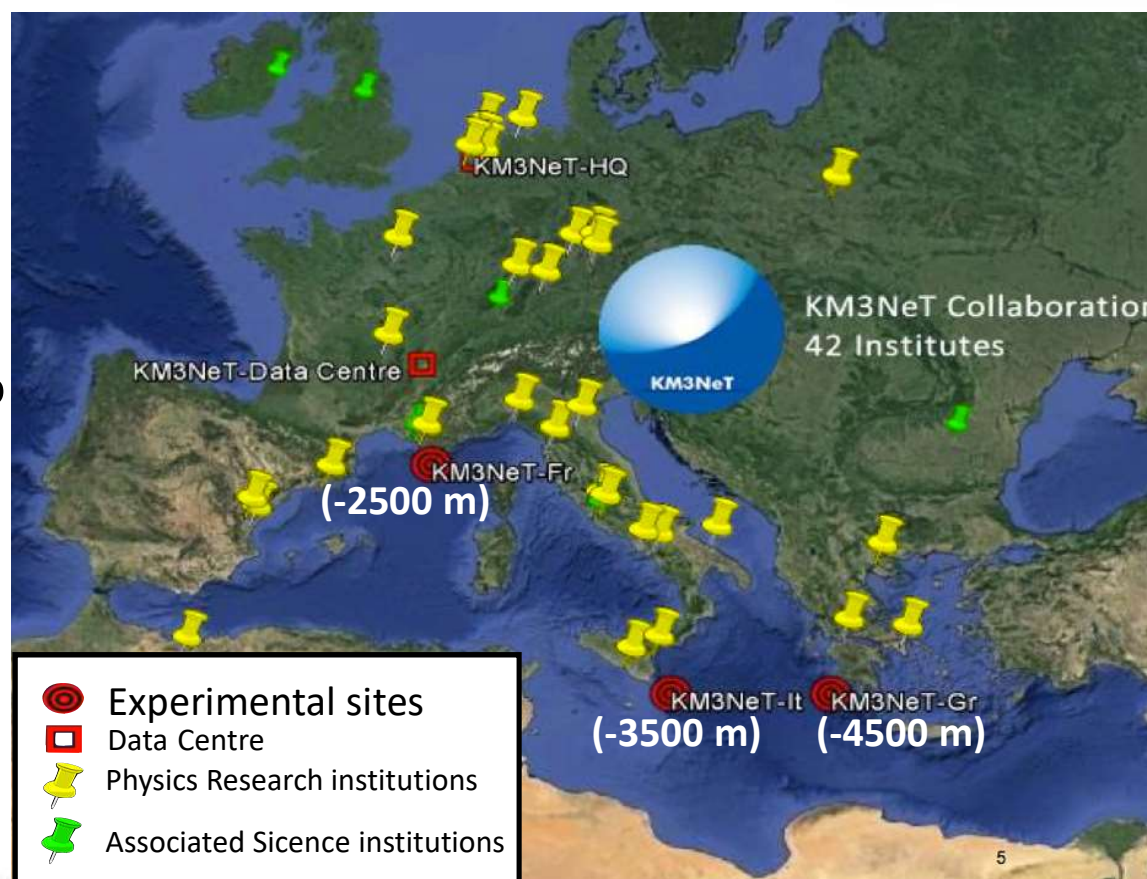
- Collegati online con una stazione di terra con cavi elettro-ottici lunghi 50-100 km

- Siti:

- Portopalo di Capo Passero IT (Astrofisica, ARCA)
- Tolone FR (Fisica di base, ORCA)
- Pilos GR (espansione)

**Consulenza del gruppo rete  
dell'INFN-CNAF per la rete di  
acquisizione dati**

Lorenzo Chiarelli - Stefano Zani - Tommaso Chiarusi



Workshop GARR 2017 - Netvolution | Roma, 5-04.2017

#3

Consortium  
GARR



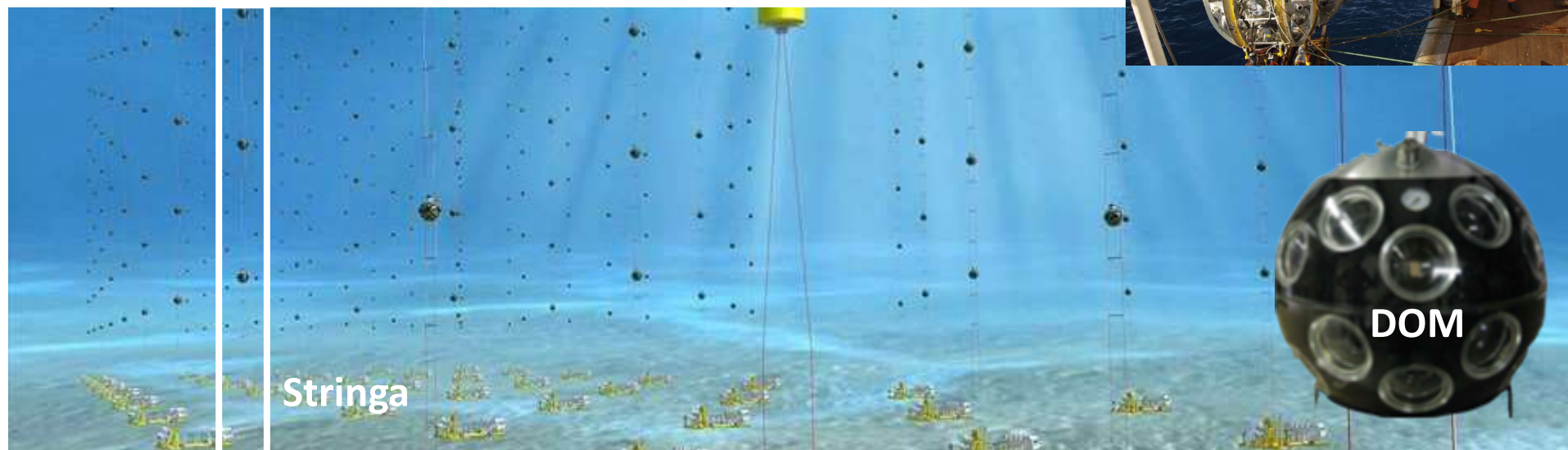
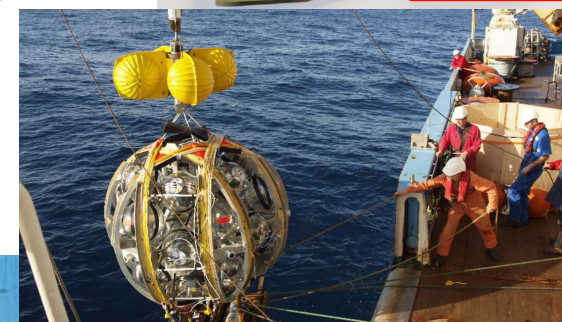
# Elementi della rete di acquisizione

- Digital Optical Module (DOM)
  - Elemento di rivelazione del detector
- Base
  - Altra sensoristica e controllo dei DOM

Organizzazione a  
Stringa  
1x BASE + nx DOM



- Ogni DOM e BASE hanno una Central Logic Board (CLB) che implementa lo standard **Ethernet (1G)**  
**Throughput MAX 200 Mbit/s**  
**MAC Addr 08:00:30:\***



# Elementi della rete di acquisizione

- White Rabbit Switch:

Sincronizzazione temporale delle CLB

Protocollo White Rabbit PTP con HW dedicato (FPGA)

Master collegato al GPS



- Control Unit:

Macchina per il controllo del detector e del TriDAS<sup>1</sup>

E' anche il DHCP server



- Data Queue:

Macchina che riceve i dati dei DOM e li invia al resto del TriDAS per essere processati



- DOM Front End Switch (DFES) *Aggregazione*

- Star Center Switch Fabric (SCSF)

- Slow Control Base Data (SCBD)

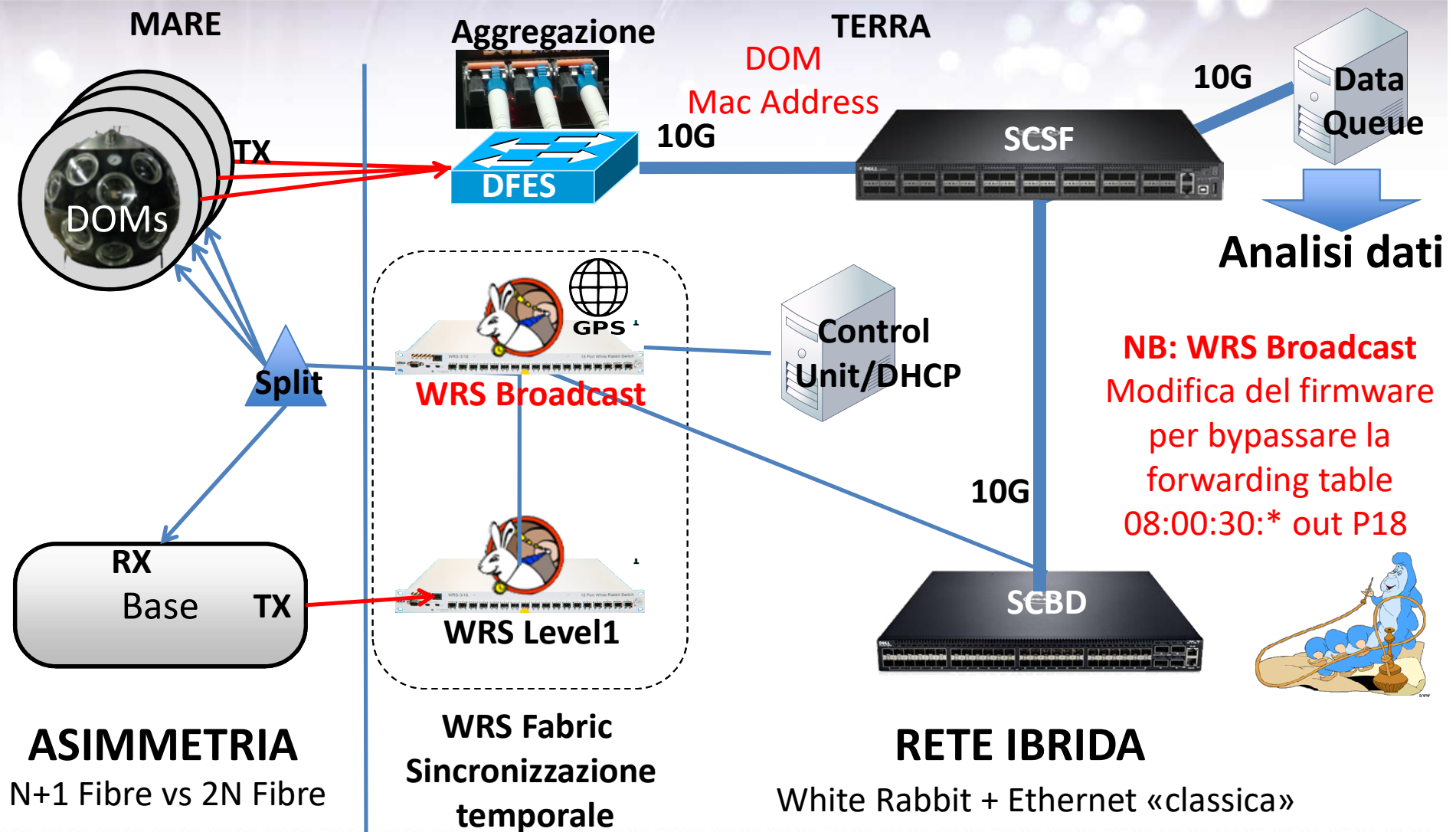
**Prima scelta tecnologica DELL:**

- Transceiver non proprietari
- Interesse al progetto KM3NeT
- Supporto tecnico

<sup>1</sup>Trigger and Data Acquisition System

# Collegamento alla stazione di terra

## Rete Layer 2





# Progettazione e scelta delle tecnologie

- WR Broadcast

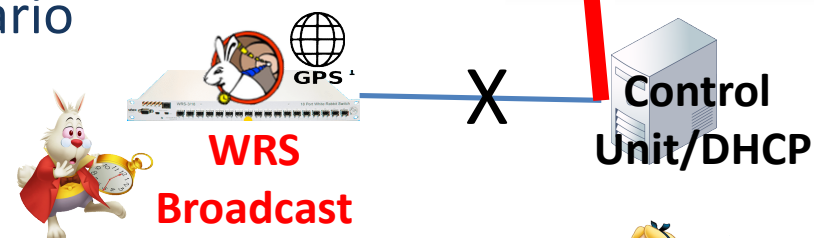
- Instabile (necessari continui riavvii) → Inaffidabile
- Non performante

→ Far passare solo il traffico necessario



- Spostare la Control Unit su SCSF

- CU collegata al «core»
- Possibilità di scalare in termini di banda (1G->10G)

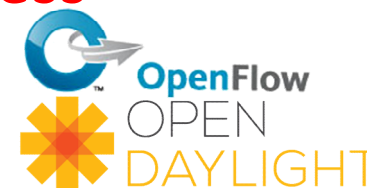


- Aggiungere un link dal White Rabbit L1 a SCBD (Loop)



- Forwarding dei pacchetti basandosi sui MAC Address ma non sugli FDB<sup>1</sup> degli switch . «Routing di MAC Address»

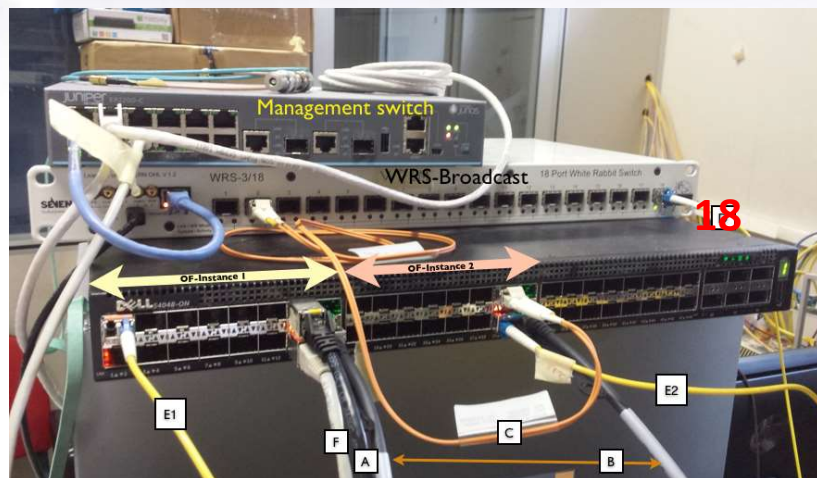
- **Possibile use case per una SDN! -> SW DELL +**



<sup>1</sup>Forwarding Data Base

# Testbed CNAF netlab + KM3Net Bologna Common Infrastructure

Ma la teoria funziona? Interazione con il controller: PUT di una regola



## FILE JSON

```
{
  "flow": [
    {
      "id": "1",
      "match": {
        "ethernet-match": {
          "ethernet-destination": {
            "address": "FF:FF:FF:FF:FF:FF"
          }
        }
      },
      "instructions": {
        "instruction": [
          {
            "order": "0",
            "apply-actions": {
              "action": [
                {
                  "output-action": {
                    "output-node-connector":
                      "openflow:303570285128704:85",
                    "max-length": "60"
                  }
                }
              ]
            }
          }
        ],
        "order": "0"
      },
      "flow-name": "SCSF_BCASTtoDHCP_CU",
      "installHw": "true",
      "idle-timeout": "0",
      "hard-timeout": "0",
      "cookie": "1",
      "table_id": "1"
    }
  ]
}
```

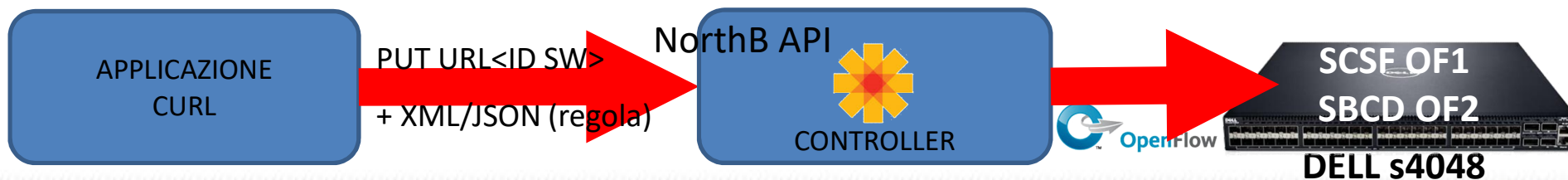
## «SHOW OPENFLOW INSTANCE 1»

Instance: 1, Table: acl, Flow: 3, Cookie: 0x1  
Priority: 32768, Internal Priority: 32768  
Up Time: 18d 22:09:46, Hard Timeout: 0  
seconds  
Idle Timeout: 0 seconds, Internal Idle  
Timeout: 0 seconds  
Packets: 73628, Bytes: 22238241  
Match Parameters:  
Valid Match: DMAC  
In Port : \* EType : \*  
SMAC : \*  
DMAC : ff:ff:ff:ff:ff:ff / ff:ff:ff:ff:ff:ff  
VLAN id : \* VLAN PCP : \*  
IP TOS : \* IP proto : \*  
Src IP : \* Dest IP : \*  
Src Port : \* Dest Port : \*

Actions:

Output: Te 1/11/3

Configurazione switch con Openflow  
Openflow 1.0 Helium (Ott 2014)  
Openflow 1.3 Beryllium (Feb 2016)





# Considerazioni su Openflow e sugli switch

**Si...funziona ma non allo stesso modo con tutti gli switch:**

- DELL serie S (FTOS 9)
  - Istanze multiple
  - Velocità delle porte configurabile (Speed, Autonegotiation)
  - Numerazione delle porte non immediata  
`openflow:303570285128704:85 -> 1/11/3`
- DELL serie N (DNOS)
  - Istanza singola
  - Porte forzatamente wire speed e con autonegoziazione
  - Corretta numerazione delle porte

**KM3NeT deve poter usare porte 1G (no auto) anche su switch 10G -> la scelta cade sulla serie S.**

# Regole di forwarding dei pacchetti

Rule	Source	Destination	Action	Description
SCSF-1	*	FF:FF:FF:FF:FF:FF	Out to CU (DHCP server)	Broadcast to CU
SCSF-2	08:00:30:00:00:00/ff:ff:ff:00:00:00	MAC[CU]	Out to CU	All DOMS to CU
SCSF-3	08:00:30:00:00:00/ff:ff:ff:00:00:00	MAC[DataQueue]	Out to DataQueue	All DOMS data to DQ
SCSF-4	MAC[CU]	08:00:30:00:00:00/ff:ff:ff:00:00:00	Out to SCBD	All controls to SBCD
SCBD-1	08:00:30:00:00:00/ff:ff:ff:00:00:00	*	out to uplink to SCSF	All Base data to SCSF
SCBD-2	Any from uplink to SCSF	08:00:30:00:00:00/ff:ff:ff:00:00:00	out to WRS-Broadcast	FROM SCSF to WRSB

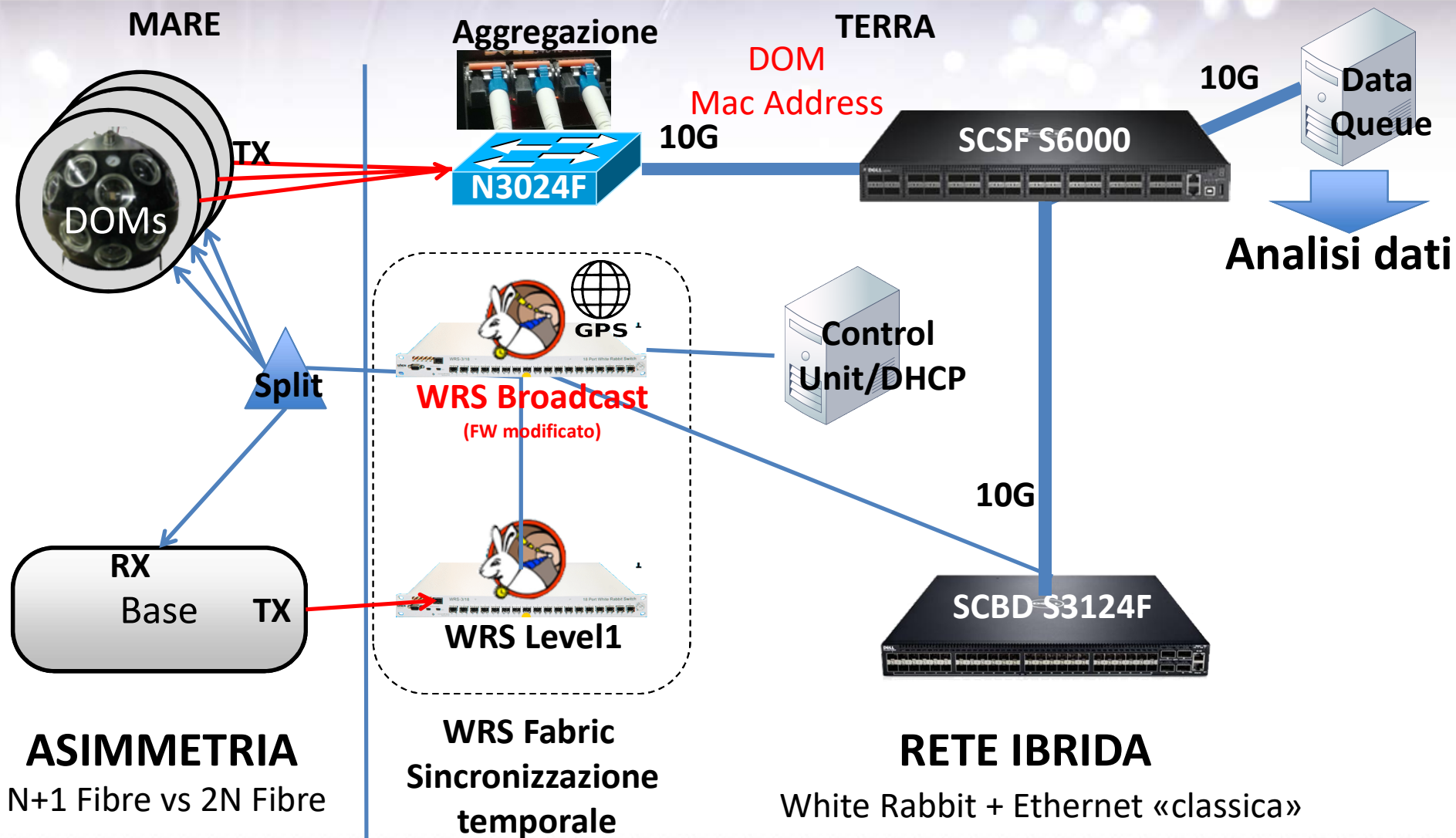
**HARD TIMEOUT = 0**

- Openflow 1.3 permette l'uso delle maschere di bit



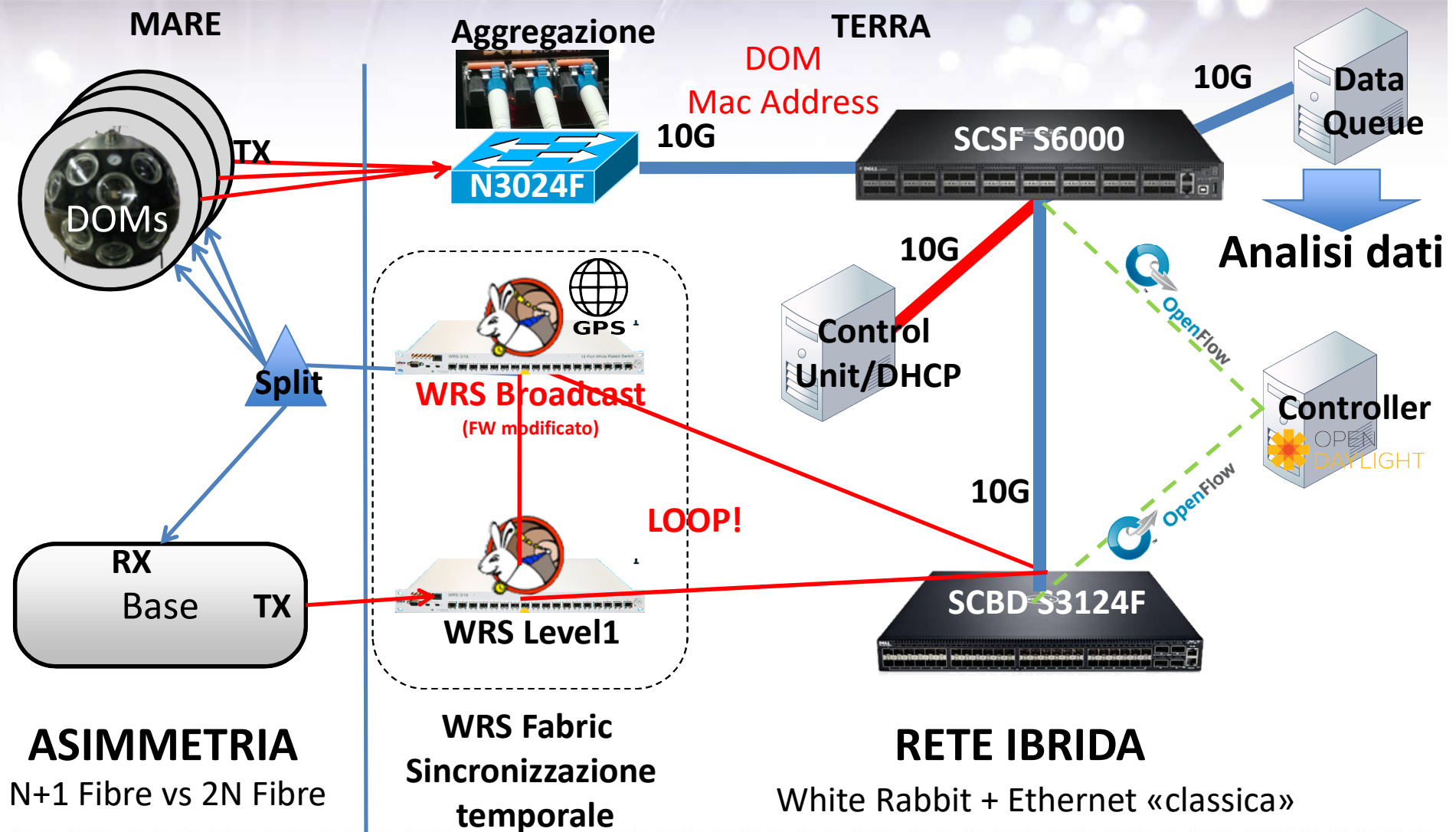
- Il numero delle regole non cresce al crescere dei DOM
- Routing di MAC Address

# Collegamento alla stazione di terra Portopalo di Capo Passero **Legacy**



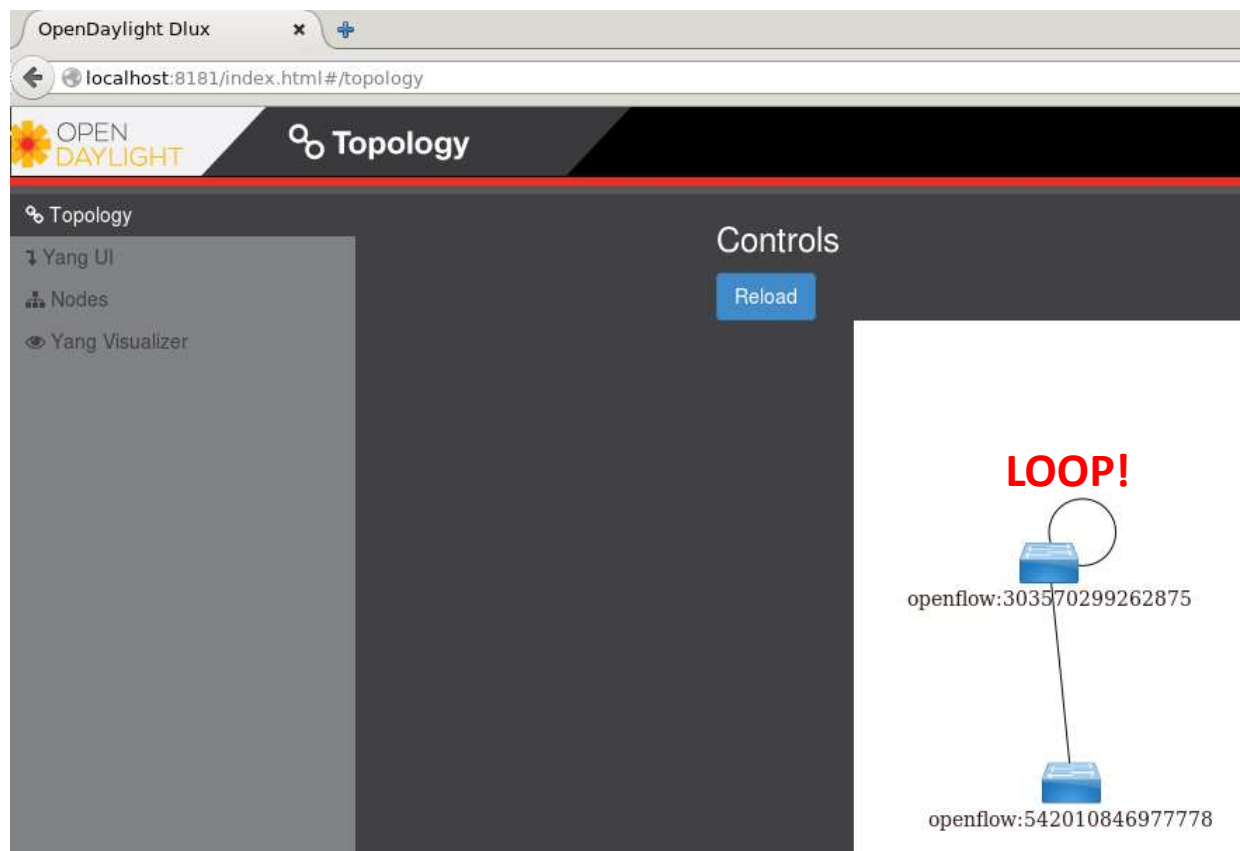


# Collegamento alla stazione di terra Portopalo di Capo Passero **SDN**



# Attenzione non è teoria ma pratica siamo in produzione!

- Dicembre 2016 SDN@Portopalo



# Impatto della nuova soluzione

## Traffico SCBD



- Stabilizzazione del traffico
- White Rabbit Switch Broadcast molto più affidabile.
- Uptime superiori (non si deve più riavviare)



# Considerazioni sulle tecnologie utilizzate

- Switch
  - Modelli diversi features e comportamenti diversi serie S vs N dipendenza dall'HW
- Topologia: Loop significa sempre Loop
  - BASE «impazzita»  
richieste ARP 127.0.0.1  
Casi particolari potrebbero portare a situazioni simili a loop classici  
Rete molto magliata attenzione a come si scrivono le regole
  - Il troubleshooting potrebbe risultare molto difficile
- Controller ODL
  - UI non User Friendly
  - Problemi di autenticazione



# Conclusioni e Sviluppi futuri

## Conclusioni

- Interazione con l'esperimento molto positiva che ha permesso di provare il paradigma SDN applicato ad un vero use case.
- Soddisfazione nel vedere che ha portato a buoni risultati!

## Sviluppi futuri

- Ridondare il controller
- Studio di regole per soppressione del traffico generato da anomalie degli elementi del detector.
- Costruzione di un'interfaccia/applicazione User Friendly



Grazie