

**GRNOG 14**  
**03/05/2023**

---

**Technical updates @GR-IX**

Marinos Dimolianis, GRNET/GR-IX



CONNECT. PEER. EXPAND.

POWERED BY **grnet**

# Outline

---

- 1. GR-IX Intro**
- 2. GR-IX Topology/Network Architecture Evolution**
  1. GR-IX v1
  2. GR-IX v2
  3. GR-IX v3
- 3. Enhanced Network Monitoring via Streaming Telemetry**
- 4. Automated Network Configuration Management**

# Outline

---

1. **GR-IX Intro**
2. GR-IX Topology/Network Architecture Evolution
  1. GR-IX v1
  2. GR-IX v2
  3. GR-IX v3
3. Enhanced Network Monitoring via Streaming Telemetry
4. Automated Network Configuration Management

# GR-IX Intro

---

## GRreek Internet EXchange (GR-IX) Member of **euro-IX**

- Connects in a common L2 domain key players of the Greek market
  - ISPs, Cloud Providers, CDNs, VoIP Providers, OTTs, etc.
- Two non-interconnected infrastructures (islands)
  - Athens (3 PoPs)
  - Thessaloniki (1 PoP)
- GR-IX today
  - 64 unique members
  - 3,931 Tbps of connected capacity
  - Peak traffic 700Gbps (x10 bandwidth growth in the last seven years)

# Outline

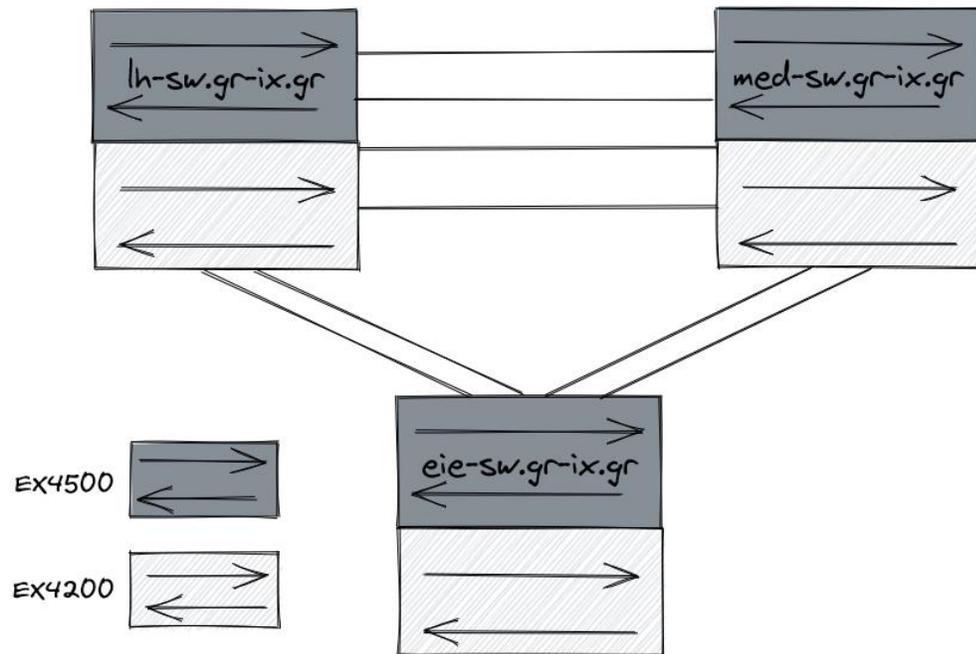
---

1. GR-IX Intro
2. **GR-IX Topology/Network Architecture Evolution**
  1. GR-IX v1
  2. GR-IX v2
  3. GR-IX v3
3. Enhanced Network Monitoring via Streaming Telemetry
4. Automated Network Configuration Management

# GR-IX v1

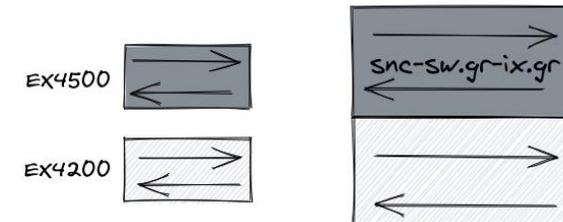
## Athens

- Virtual Chassis
- Pure L2 - Spanning Tree Protocol



## Thessaloniki

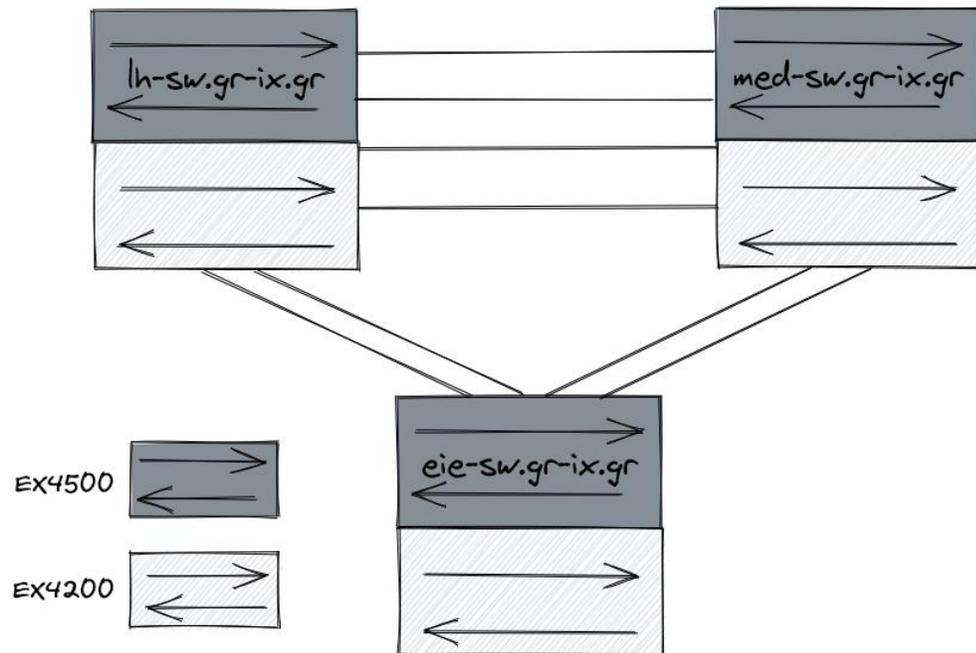
### Virtual Chassis



# GR-IX v1

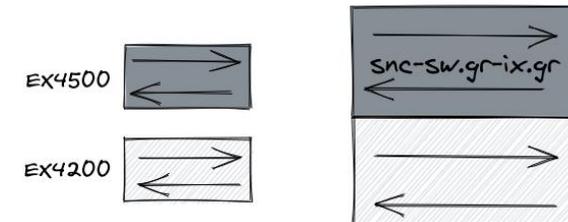
## Athens

- Virtual Chassis
- Pure L2 - Spanning Tree Protocol



## Thessaloniki

### Virtual Chassis

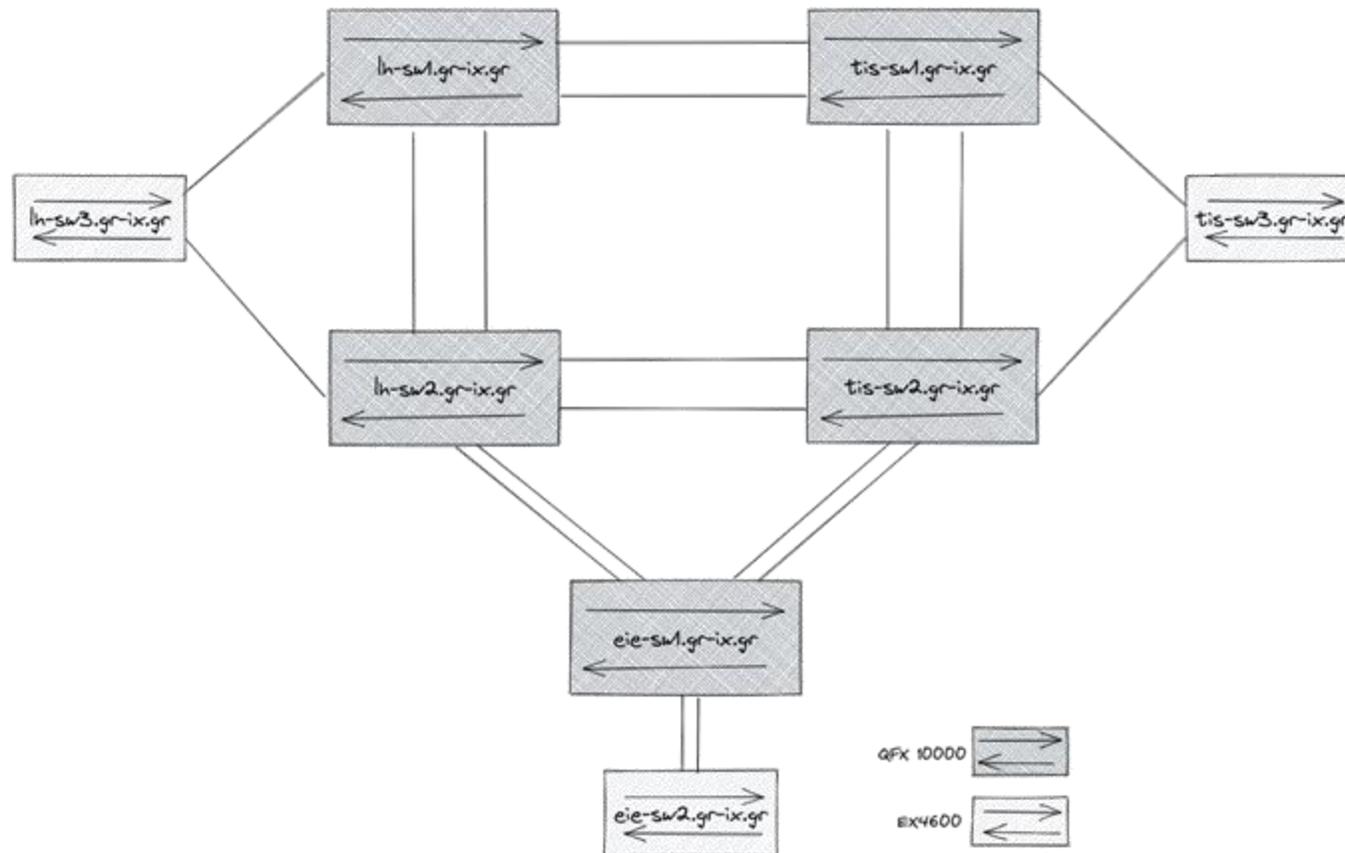


- 100G Ports?
- Scalability?

# GR-IX v2

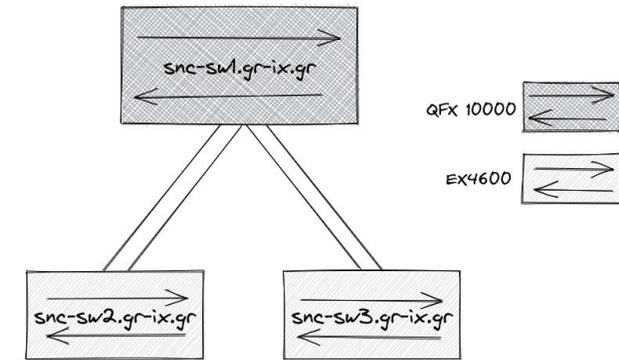
## Athens

### Collapsed spine fabric architecture



## Thessaloniki

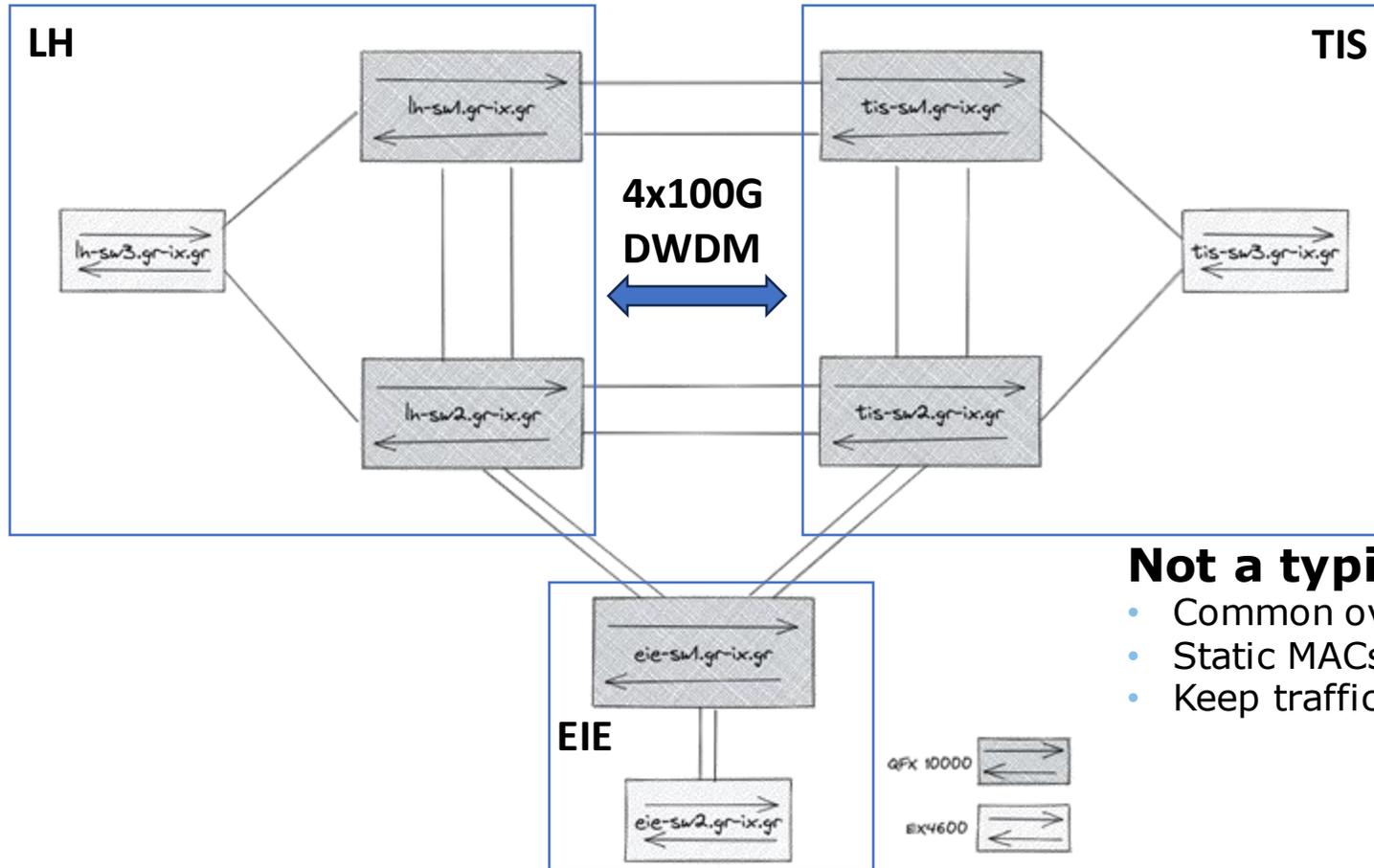
### Collapsed spine fabric architecture



# GR-IX v2

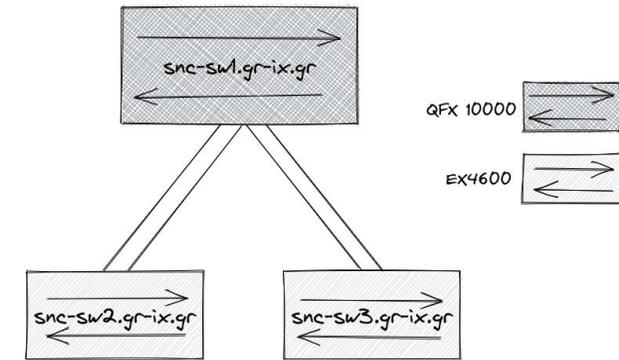
## Athens

### Collapsed spine fabric architecture



## Thessaloniki

### Collapsed spine fabric architecture



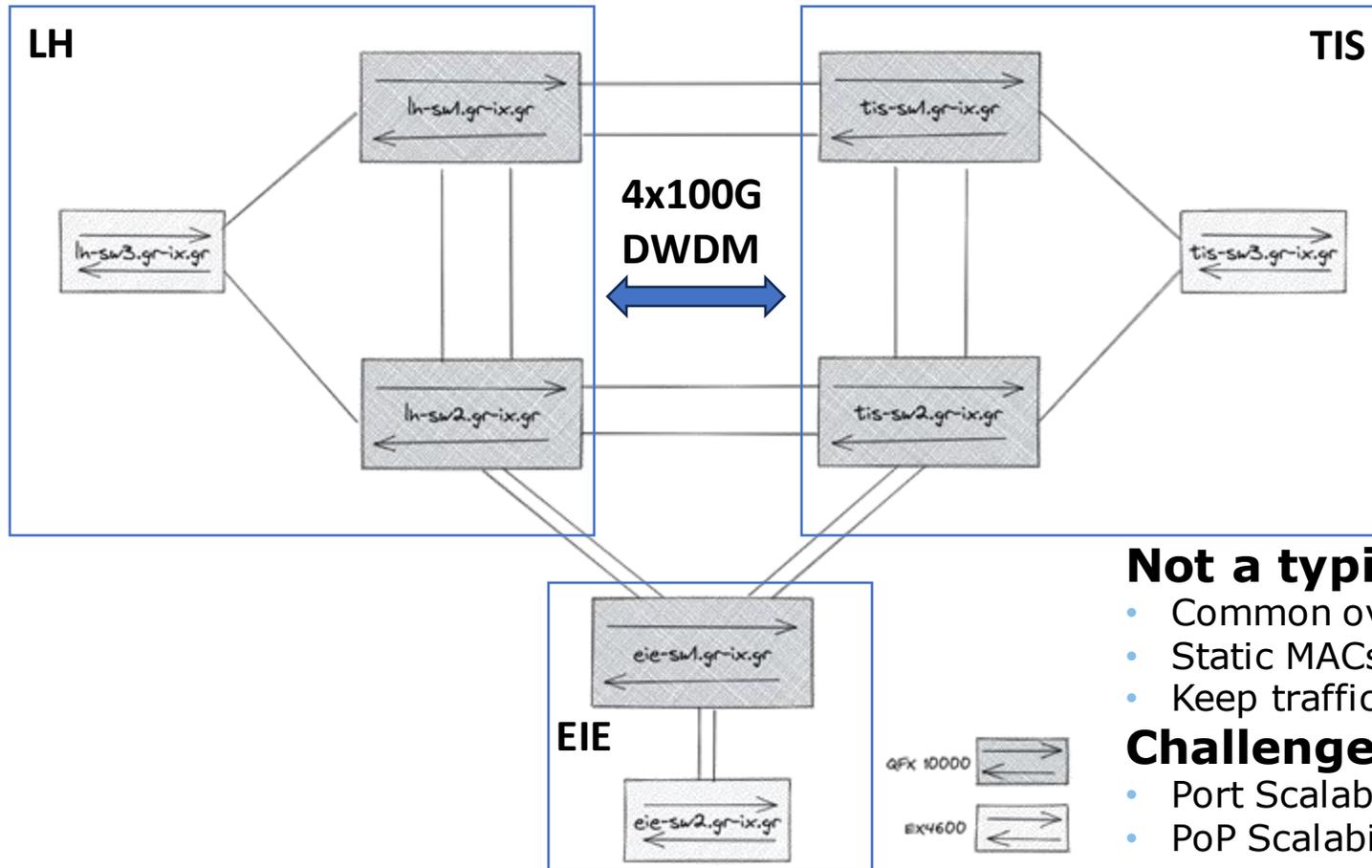
### Not a typical DC fabric

- Common overlay network between interconnected DCs
- Static MACs & L2 ACLs (no mac movement->prohibited)
- Keep traffic locally (switch->PoP)

# GR-IX v2

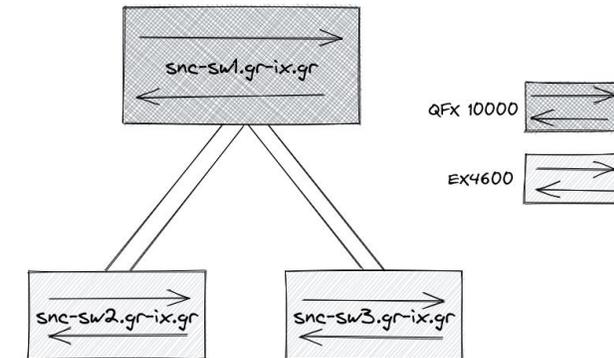
## Athens

### Collapsed spine fabric architecture



## Thessaloniki

### Collapsed spine fabric architecture



### Not a typical DC fabric

- Common overlay network between interconnected DCs
- Static MACs & L2 ACLs (no mac movement->prohibited)
- Keep traffic locally (switch->PoP)

### Challenges

- Port Scalability?
- PoP Scalability?

# GR-IX v3

## Athens

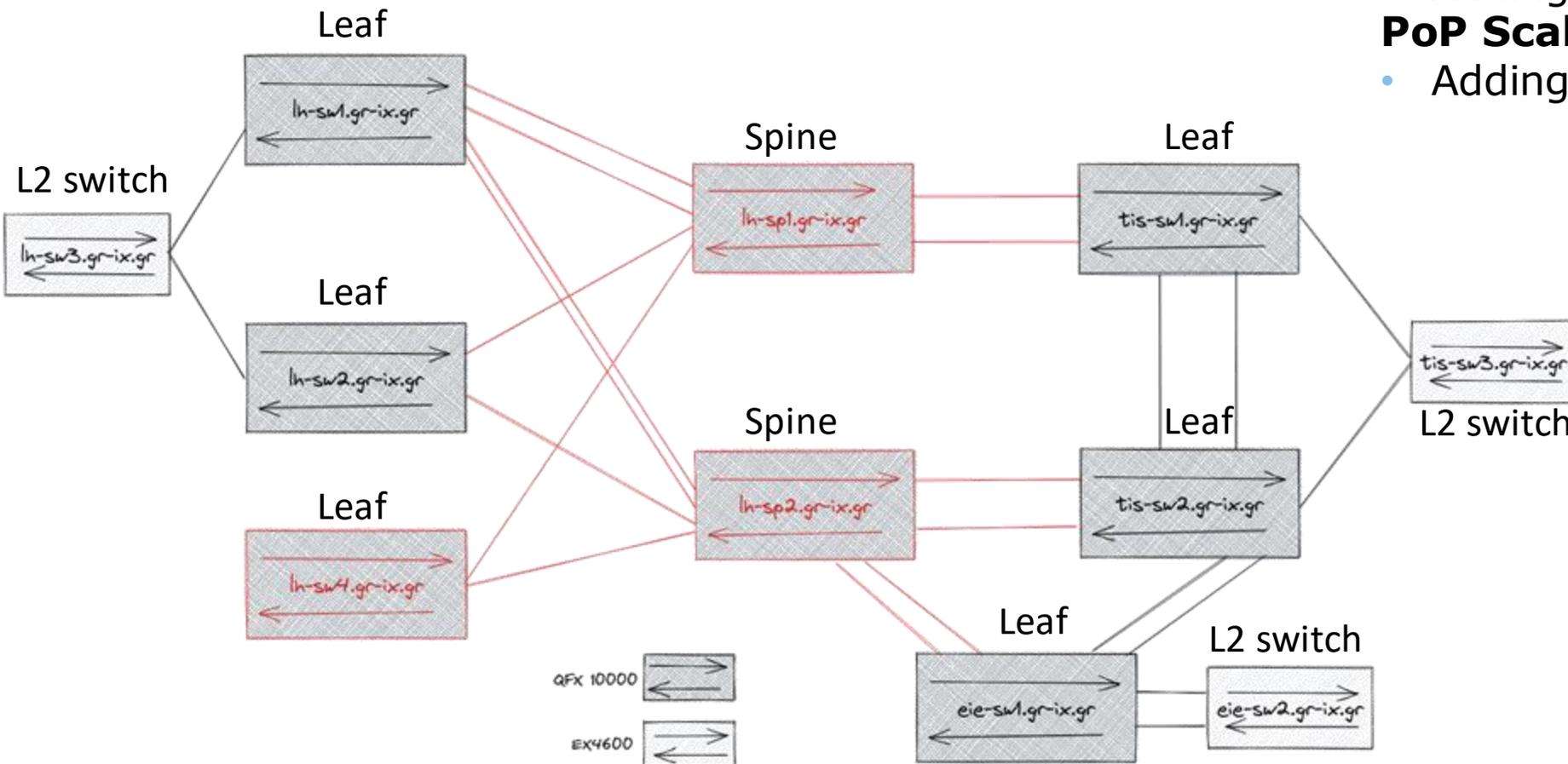
### Spine/Leaf fabric architecture

#### Port Scalability:

- Adding Leafs

#### PoP Scalability:

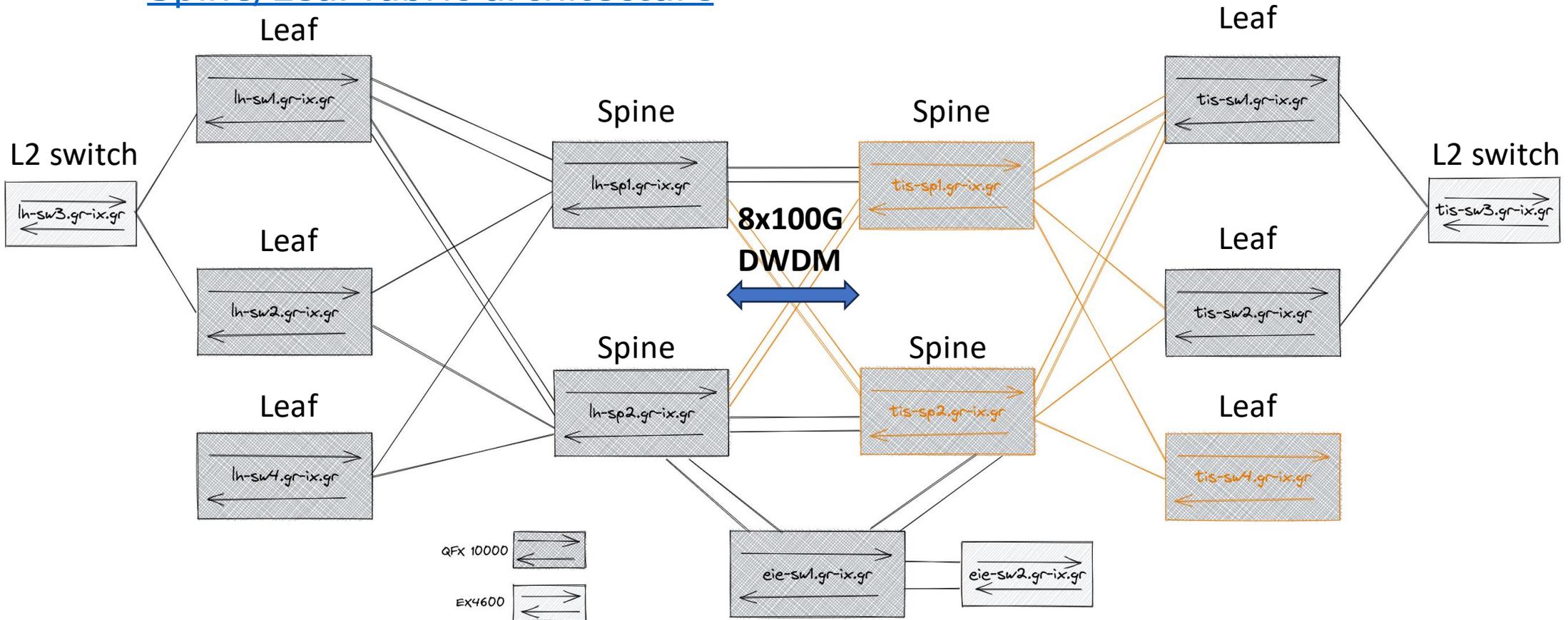
- Adding Leafs and/or Spines



# GR-IX v3

## Athens (Next Steps)

### Spine/Leaf fabric architecture



# Outline

---

1. GR-IX Intro
2. GR-IX Topology/Network Architecture Evolution
  1. GR-IX v1
  2. GR-IX v2
  3. GR-IX v3
3. **Enhanced Network Monitoring via Streaming Telemetry**
4. Automated Network Configuration Management

# Network Monitoring via Streaming Telemetry

---

- Streaming telemetry in QFX10K (22.3) and EX4600 (18.4)
- [Openconfig](#) (vendor-neutral) data model
  - Interface Status/Counters (polling period 3s)
    - Errors, Discards, Broadcasts
- Telemetry Stack (dial-in approach)
  - Collection/Processing (Telegraf 1.20.3)
  - Storage (InfluxDB 1.6.4)
  - Representation (Grafana 8.3.1)
- Telemetry data (fine-grained)
  - Continuously more & more are available by the vendor

**In production for more than a year**

# Streaming Telemetry to the rescue#1

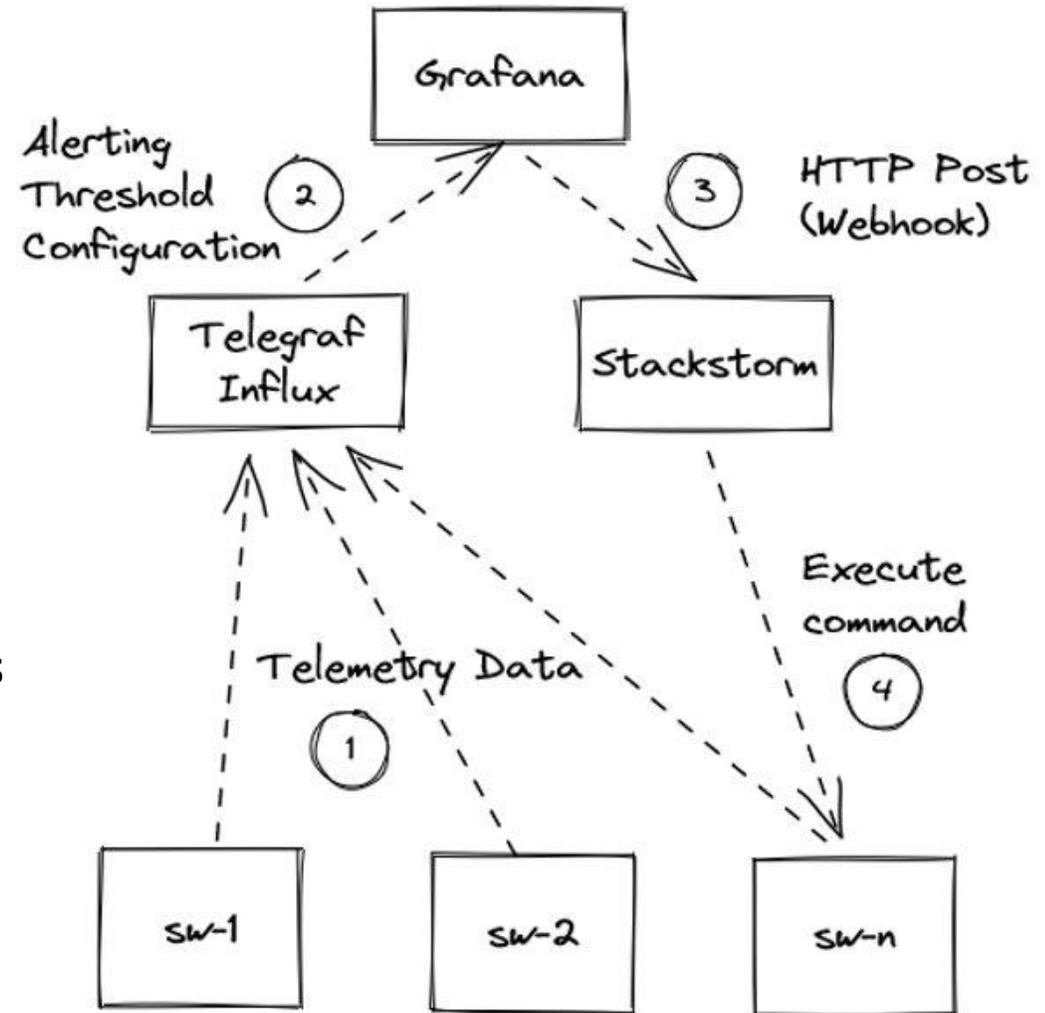
- Vendor bug identified by large broadcast packet rates



- Workaround -> command execution

## Procedure

1. Monitoring broadcast packet rate
2. Alerting based on predefined thresholds
3. Command execution on the switch



# Streaming Telemetry to the rescue#2

---

- BFD has been disabled at our Route Servers due to “*random*” flaps
- Reinitiated testing with our members
  - BFD flaps “strange” pattern

```
2023-03-03 09:12:00 <TRACE> rs_bfd: Session to XXX.XXX.XXX.XXX changed state from Down to Init
```

```
2023-03-03 09:12:01 <TRACE> rs_bfd: Session to XXX.XXX.XXX.XXX changed state from Down to Init
```

```
2023-03-03 19:13:11 <TRACE> rs_bfd: Session to XXX.XXX.XXX.XXX changed state from Down to Init
```

```
2023-03-03 19:13:11 <TRACE> rs_bfd: Session to XXX.XXX.XXX.XXX changed state from Down to Init
```

```
2023-03-05 05:12:48 <TRACE> rs_bfd: Session to XXX.XXX.XXX.XXX changed state from Down to Init
```

```
2023-03-05 05:12:48 <TRACE> rs_bfd: Session to XXX.XXX.XXX.XXX changed state from Down to Init
```

```
2023-03-05 19:13:05 <TRACE> rs_bfd: Session to XXX.XXX.XXX.XXX changed state from Down to Init
```

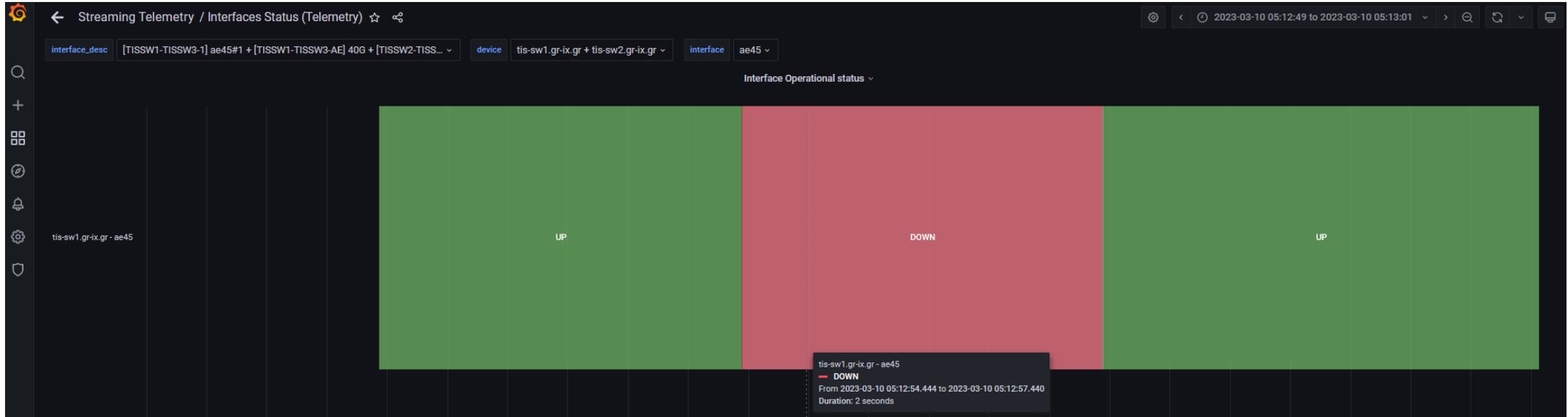
```
2023-03-05 19:13:05 <TRACE> rs_bfd: Session to XXX.XXX.XXX.XXX changed state from Down to Init
```

# Streaming Telemetry to the rescue#2



- Logs->LACP flap due to high CPU
- BFD session flap

# Streaming Telemetry to the rescue#2



- Logs->LACP flap due to high CPU
- BFD session flap
- **Current situation: Steady BFD sessions with three different members (three different vendors)**
- **BFD to be enabled at our Route Servers in the near future**

# Outline

---

1. GR-IX Intro
2. GR-IX Topology/Network Architecture Evolution
  1. GR-IX v1
  2. GR-IX v2
  3. GR-IX v3
3. Enhanced Network Monitoring via Streaming Telemetry
4. **Automated Network Configuration Management**

# Automated Network Configuration Management

---

- **IXPManager (Source of Truth)**
  - custom API -> exposing business and technical data
    - Switches
    - VLANs
    - Members ports
    - Backbone ports/interfaces

# Automated Network Configuration Management

---

- **IXPManager (Source of Truth)**
  - custom API -> exposing business and technical data
    - Switches
    - VLANs
    - Members ports
    - Backbone ports/interfaces
- **Ansible**
  - Configuration (Jinja2) templates
    - Members ports (e.g., static MAC)
    - Servers ports
    - Backbone ports
    - BGP underlay
    - BGP overlay
    - Streaming Telemetry
    - VLANs

# Next Steps

---

- **GR-IXv3 (finalization)**
- **BFD**
- **Blackholing**
- **Marketplace**
- **400G?**
- **Interconnection with other IXes?**

# Do not forget to be a good netizen

---

**DROP RPKI  
INVALIDS**

**AGGREGATE**

# Thank you!



CONNECT. PEER. EXPAND.

POWERED BY grnet

Marinos Dimolianis, GRNET/GR-IX

Questions/Thoughts?



Follow us and Stay Connected



[gr-ix.gr](http://gr-ix.gr)



[@GRIXsocial](https://twitter.com/GRIXsocial)



[@GR-IX::The Greek Internet Exchange](https://www.linkedin.com/company/GR-IX::The%20Greek%20Internet%20Exchange)



[info@gr-ix.gr](mailto:info@gr-ix.gr)