UNIVERSITAT POLITÈCNICA DE CATALUNYA

Optical Communications Group

Optical Network Automation

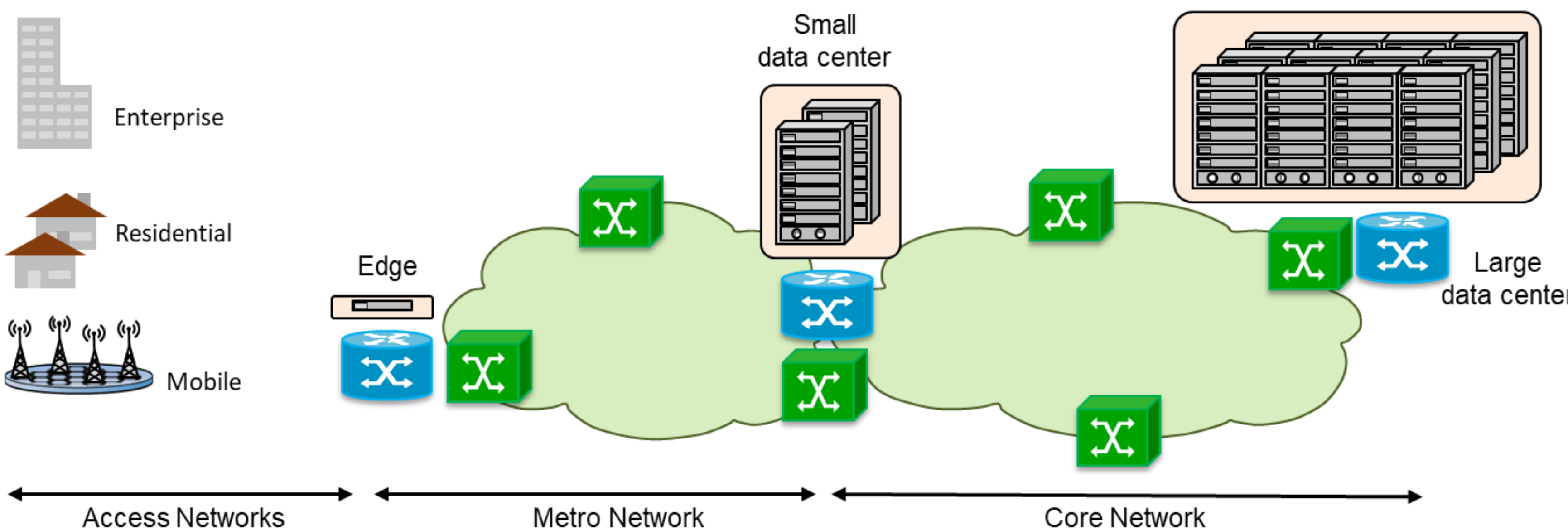
Luis Velasco
lvelasco@ac.upc.edu

Forum Numerica, Sophia-Antipolis, 2020

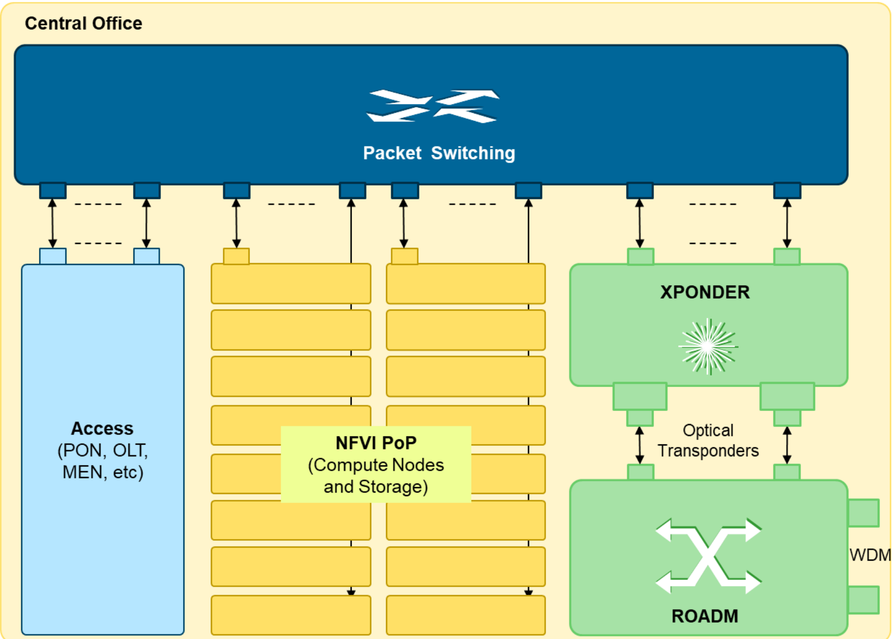
Motivation

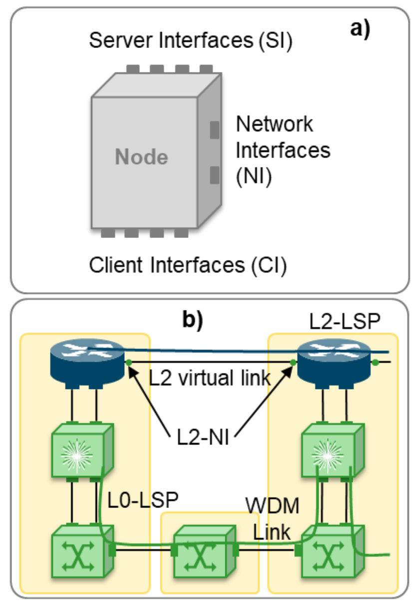
- Operators' network management continuously measure network health by collecting data from the deployed network devices.
  - Data used mainly for performance **reporting,** diagnosing problems **after failures**, and to **predict future traffic growth** for planning.
- Network management is typically **reactive** and requires significant human effort and skills to operate effectively.
- As optical networks evolve to fulfil highly flexible connectivity and dynamicity requirements, they must also provide **reliable connectivity** and **increased** network resource **efficiency**.
- Future optical networks must support **fully automated management**, providing:
  - **dynamic** resource re-optimization to rapidly adapt network resources based on **predicted** conditions and events
  - identify service **degradation** conditions that will impact connectivity and **highlight** critical devices and links for further inspection
  - **Activate recovery** if a failure is predicted or detected and facilitate **resource optimization after restoration** events.

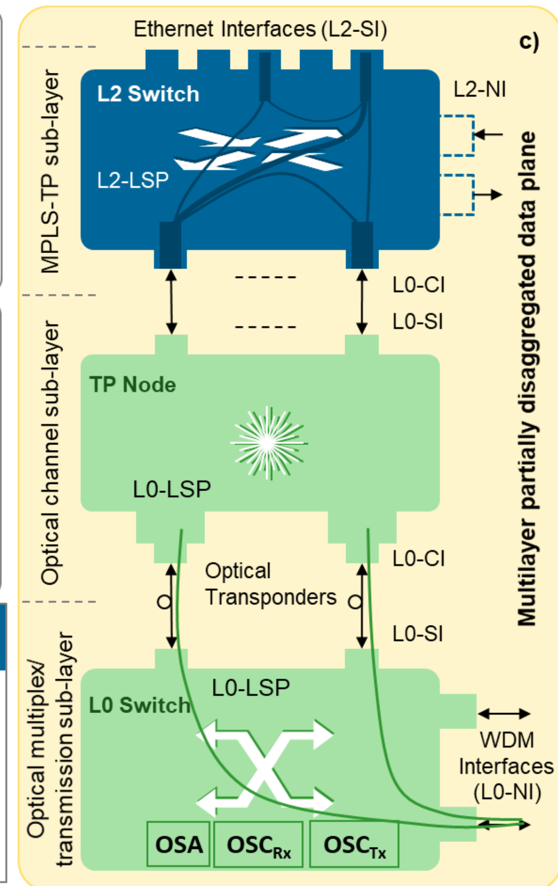
Heterogeneous optical network architecture



CO Architecture and OPs







Monitorable components and data measured d)

1. L2-SI: Ethernet aggregated traffic
2. L2-NI: L2 aggregated traffic
3. L0-CI: Average optical power
4. L0-SI: Average optical power
5. L0-NI: Average optical power
6. L2-LSP: L2-LSP traffic
7. L0-LSP: L0-LSP BER and opt. power

Monitoring and Data Analytics for Optical Networking

L. Velasco, A. Chiadò Piat, O. González, A. Lord, A. Napoli, P. Layec, D. Rafique, A. D'Errico, D. King, M. Ruiz, F. Cugini, and R. Casellas, "Monitoring and Data Analytics for Optical Networking: Benefits, Architectures, and Use Cases," IEEE Network Magazine, vol. 33, pp. 100-108, 2019.

Use Case 1: Network planning and provisioning with reduced margins

Description	Expected Benefits	Modeling and Parameters involved
Application of **just enough margin** in the network design and in lightpaths provisioning.	**CAPEX saving** opportunity by avoiding or postponing unnecessary investments at a given time.	Attenuation, dispersion and other **fiber parameters**, the **noise figure** of amplifiers, **WSS passband**, the **sensitivity of TPs**, etc. Those parameters can be used together with an **analytical model** to estimate the QoT of lightpaths accurately. ML-based methods to **predict** the **probability** that the **QoT** of a candidate lightpath will not **exceed** a defined **threshold**.

Use Case 2: Dynamic Network adaptation

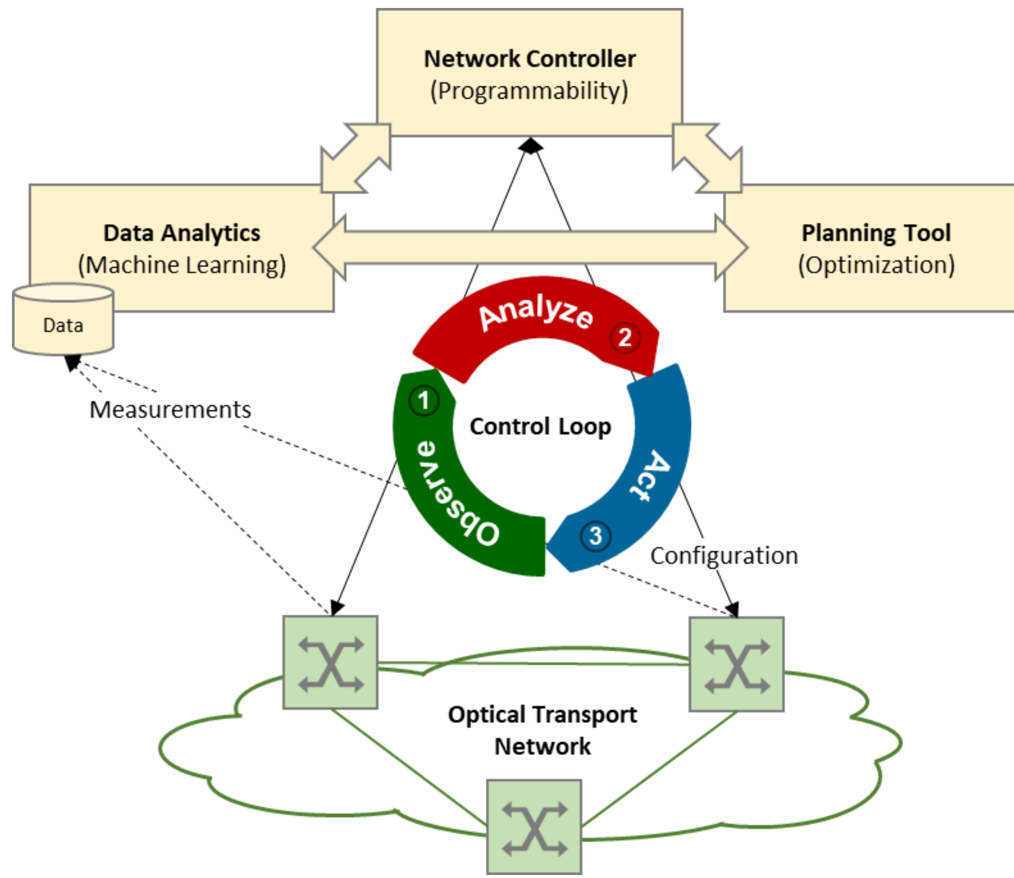
Description	Expected Benefits	Modeling and Parameters involved
Leveraging on **configurable TPs** the allocation of just enough data rate for any connection at any time to cope with **traffic dynamics** at **minutes** or **hours** scale.	Better exploitation of network resources and potential savings by **reducing** the typical **overprovisioning** of static allocation.	Use of **models** to evaluate the expected **QoT** of a lightpath at any **new TP configuration**. Use of models for **traffic analysis** to evaluate traffic trends and periodicity.

Use Case 3: Lightpath degradation and failure localization

Description	Expected Benefits	Modeling and Parameters involved
QoT reduces over time due to network and device **degradation** (e.g., fiber cuts and repairs), **ageing**, or **load increasing**.	Degradation **anticipation** allows appropriately **tune systems' parameters** before alarm triggering. **Localizing** the element responsible for a failure **facilitates** network **maintenance** by planning a human intervention.	**Predictive analysis** based on QoT evolution. **Localization** based on the **per-system analysis**. Algorithms that find the potential cause of the failure.

MDA enables OAA control loop implementation

- The analysis of the collected data can **discover knowledge** and use it to proactively **self-configure and self-tune** the network in a cost-effective (near) real-time manner by adapting resources to future conditions.
- OAA control loops can be enabled, where **outcomes of data analysis** can be used for event notifications together with **recommended actions** to the SDN controller.
- ML models can be estimated from monitoring data to **feed planning tools** to compute optimal solutions for the expected future conditions.

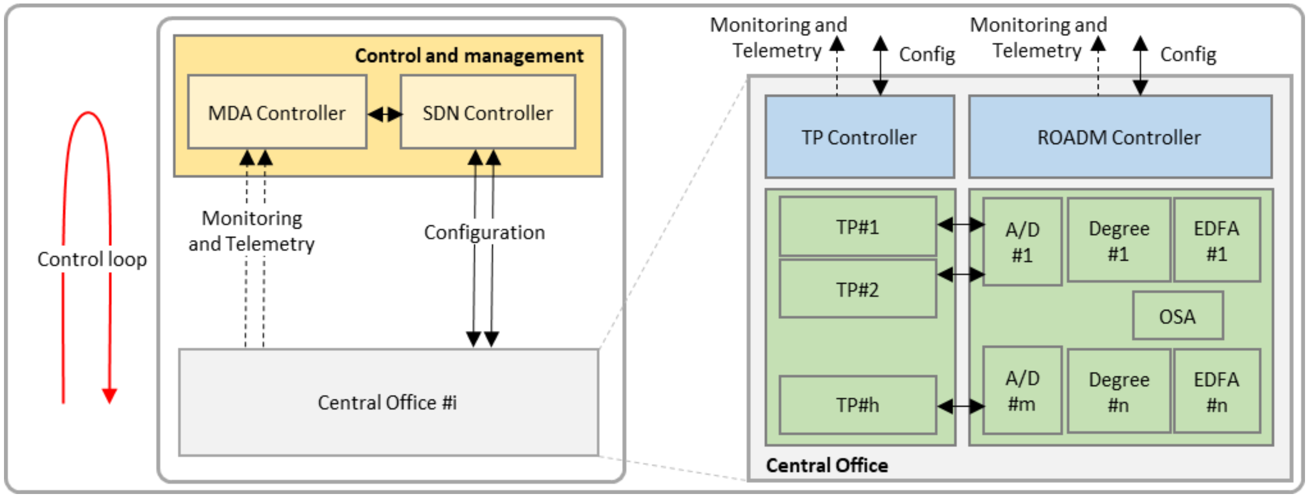


Monitoring and Telemetry protocols

Prot.	Description	Data types	Scalability and performance
IPFIX	• It was developed in IETF for typical **IP networks applications**. • It works in **push mode** and supports a many-to-many relationship between OPs and collectors.	• The structure of the protocol messages is **based on templates** that enable to export any type of data.	• Scalability is considered in the design of the protocol, with a requirement of **hundreds** of different **exporting processes** to be supported.
gRPC	• gRPC uses Protocol Buffers for **serializing structured data**. • It supports **data streaming** based on **request/response**.	• **Specific data structures can be defined**; a compiler generate source code representing the data and methods to serialize them.	• It is **carried over HTTP/2** and leverages on effective binary framing and header compression that improve data transfer efficiency.
Thrift	• It is an open source software library and set of code-generation tools. • It is **stream-oriented**.	• It allows the **definition of datatypes** and generates all the necessary code.	• Data transfer efficiency is **comparable** to that of gRPC.

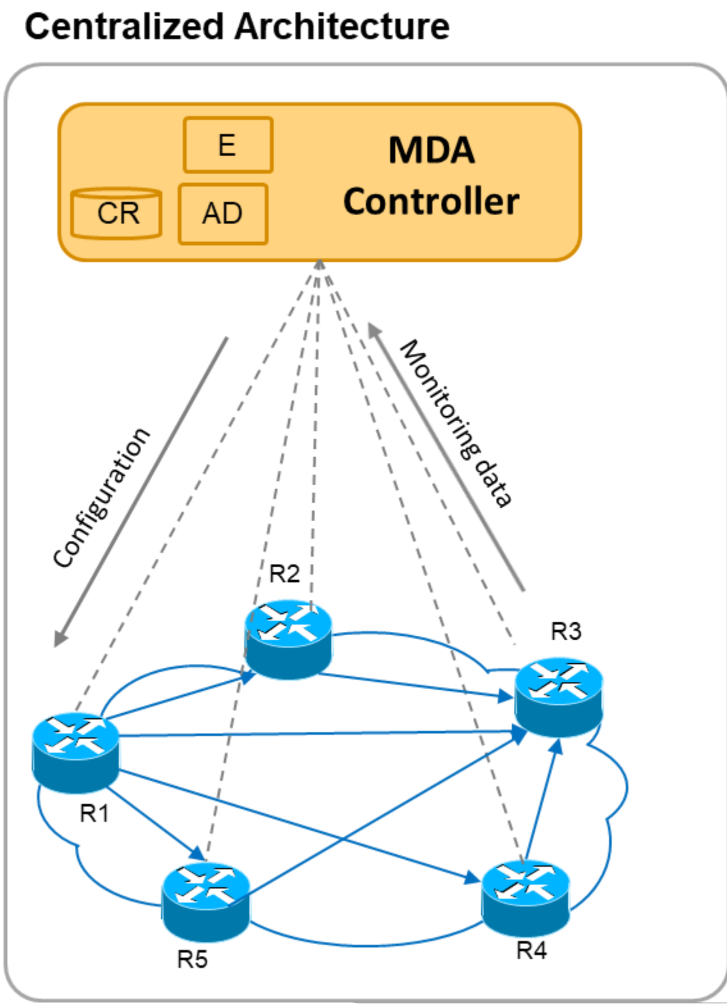
- ❖ With these protocols, the collection period is **not limited to 15 minutes** (the shorter the collection period, the shorter the event that can be detected, as well as the shorter the time to detect degradations)
- ❖ Reducing the period increases the amount of data to be collected, stored, and analyzed.
- ❖ An approach to reduce the amount of data is to rely on **monitoring** (collection period of minutes) and **activate telemetry on demand**.

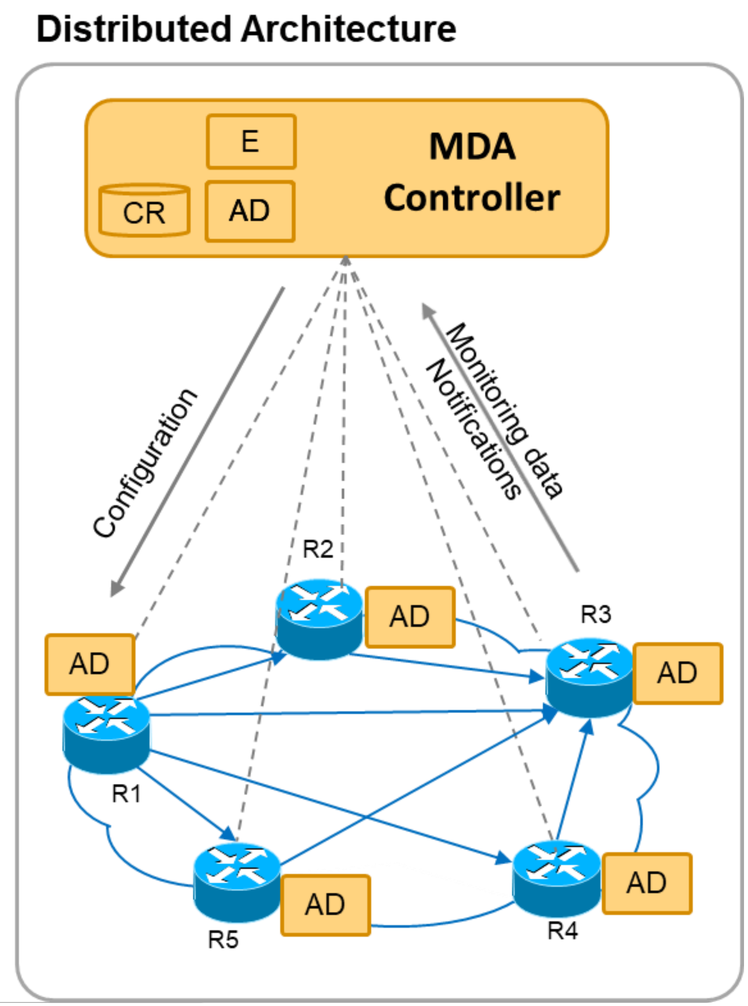
Centralized MDA Architecture



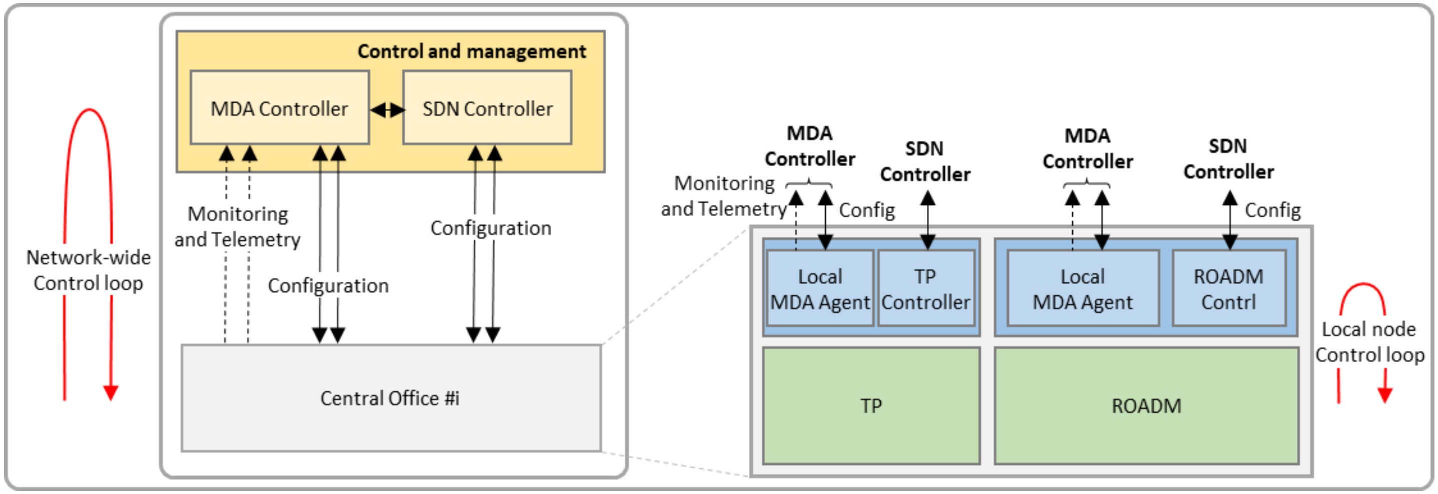
Features	Strengths	Weaknesses
• Includes a **centralized** MDA system with a data repository for monitoring/telemetry data where data analytics can be applied. • Monitoring and telemetry **activation and deactivation** can be managed by an external system, e.g., the NMS.	• Data analytics results can be used for network **self-adaptation** to changing conditions. • **Interfaces** with the SDN controller (and NFVO) can be **easily standardized**.	• Different monitoring / telemetry protocols need to be **available at the MDA controller**. • The **amount of data** to be collated increases exponentially to keep low reaction times against degradations. • **Configuration tuning** through the **SDN controller only**.

Bringing Data Analytics to the Network Nodes



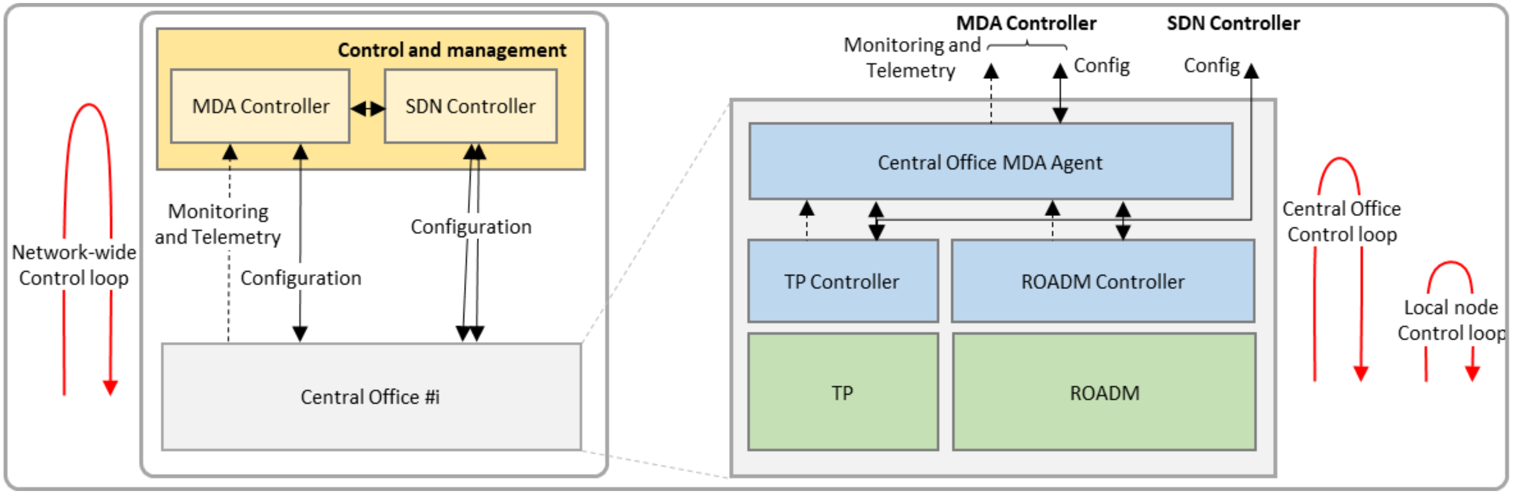


Distributed MDA Architecture



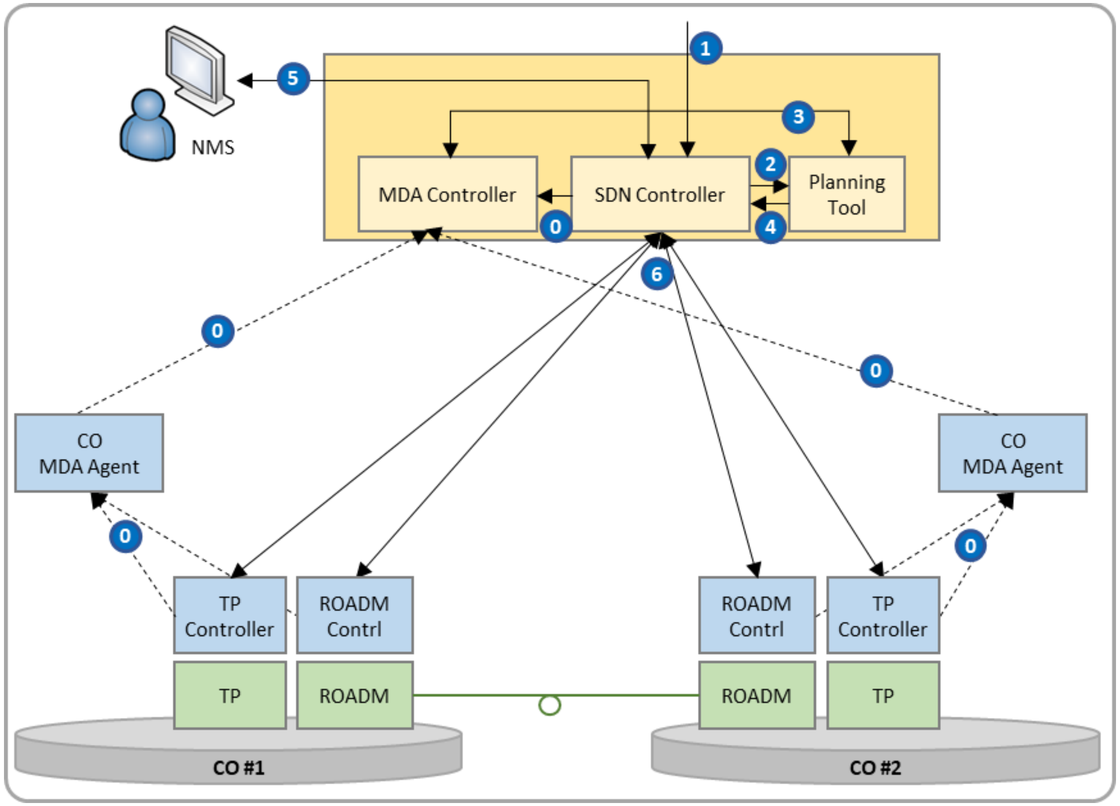
Features	Strengths	Weaknesses
• Allows **data analytics** to be **applied** within the MDA agents, **close to the network nodes**. **Control loops can be implemented locally at the node level**. • Monitoring and telemetry activation / deactivation is **managed by the MDA controller**.	• **Supports configuration tuning**. • It **reduces data to be conveyed** to the MDA controller since patter recognition can be done in the MDA agents. • MDA agents expose **one single monitoring and telemetry interface** to the MDA controller.	• A **configuration interface** needs to be defined between the MDA controller and the agents. • More **complex MDA controller** as more features are added, like monitoring/telemetry control, and configuration tuning.

Hierarchical MDA Architecture



Features	Strengths	Weaknesses
• It includes a **per-CO MDA agent** that connects to all the nodes in the CO.	• **Control loops** can be implemented at the **node**, as well as at the **CO level** involving more than one node. • **Appropriate for node disaggregation scenarios**, where monitoring devices can be installed in one node, but configuration tuning needs to be done in a different node. • It **reduces** the total **number of agents** and the **number of interfaces** toward the MDA controller.	• Requires **more complex MDA agents** to consider complex relations among nodes.

Use Case 1: Lightpath provisioning with reduced margin



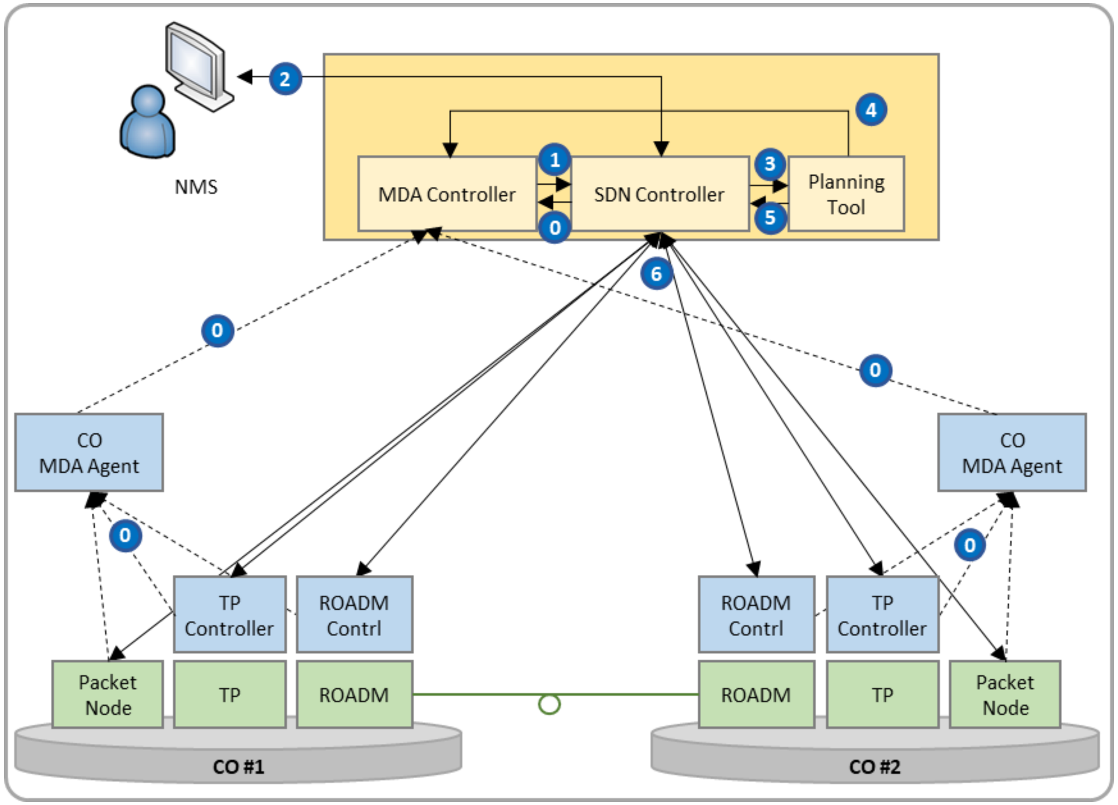
- It might happen that the lightpath can be established **only if other lightpaths are changed**.
- The PT returns the optimal configuration found for the requested lightpath, together with a (*reactive*) **recommended action** for the SDN controller to reconfigure a subset of already established lightpaths.

❖ **OSNR estimation** at optimal lunch power and in nonlinear regime **requires data** from monitoring that are already available in the MDA controller.

❖ When a lightpath set-up request arrives at the SDN controller, the latter relies on the **PT for the computation of the RMSA** and other parameters that contribute **to minimize the system margin** while guaranteeing its QoT.

❖ In order to compute an optimal solution, **the PT needs to access data from the MDA controller**; once a solution has been found, it is sent back to the SDN controller.

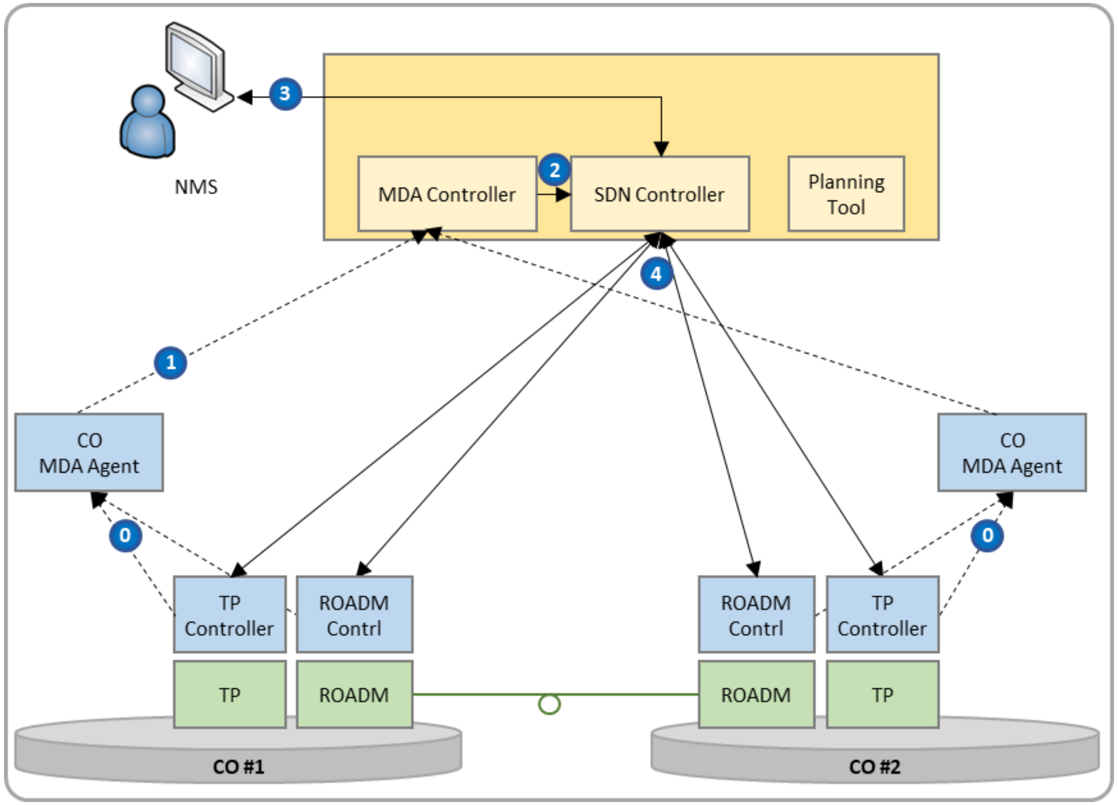
Use Case 2: Dynamic network adaptation



- The MDA controller issues ***proactive* recommendations** to the SDN controller as a result of analyzing data and aiming at **anticipating relevant events**.
- ML algorithms running in the MDA controller can use the **measured packet traffic volume** to determine a traffic model for the traffic between every origin and destination CO.
- When the **expected traffic** for the near future is close to the allocated capacity, the MDA controller notifies the SDN controller to reconfigure the capacity.

- The SDN controller might inform the operator in the NMS and then, request the PT to compute the optimal capacity configuration for the detected event. The PT needs data from the MDA controller, e.g., the **expected traffic matrix**.
- When the optimal solution is computed, the PT responds to the SDN controller, which re-configures the network.

Use Case 3: Lightpath degradation detection and modulation format adaptation



If both TPs need to be **simultaneously re-configured**, the MDA agent can send a notification to the controller that evaluates the capabilities of the TPs. The degradation detection together with a recommendation is sent to the SDN controller that implements it in the devices, (the operator in the NMS can be checked).

- BER measurements are collected by the MDA agents from the TPs.
- A ML algorithm in the MDA agents can detect BER trends to anticipate QoT degradation.
- In the case of QoT degradation detection, a **decision** can be **locally made** without the intervention of the MDA controller.
  - Modern TPs identify the MF of the received signal by means of DSP.
  - A change in the MF can be initiated in one of the TPs and the end TP will automatically realize of such change and carry out the same in the opposite direction

Data Analytics at the Optical Layer

A. P. Vela, M. Ruiz, F. Fresi, N. Sambo, F. Cugini, G. Meloni, L. Potí, L. Velasco, and P. Castoldi, "BER Degradation Detection and Failure Identification in Elastic Optical Networks," IEEE/OSA Journal of Lightwave Technology (JLT), vol. 35, pp. 4595-4604, 2017.

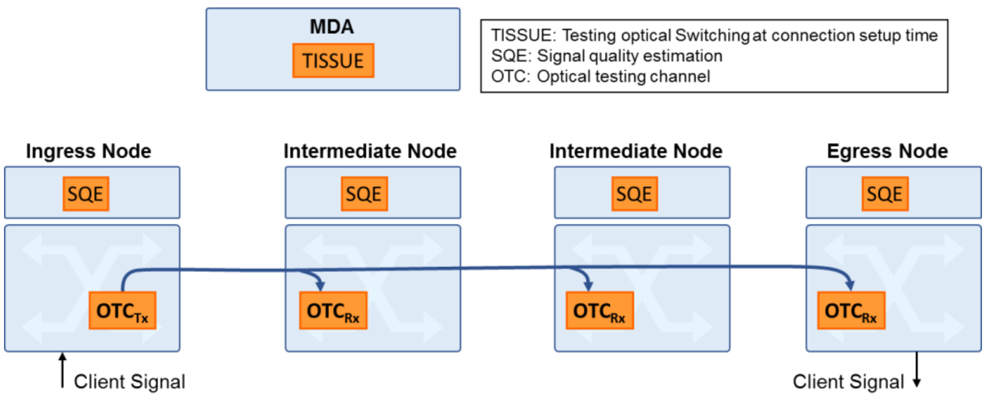
A. P. Vela, B. Shariati, M. Ruiz, F. Cugini, A. Castro, H. Lu, R. Proietti, J. Comellas, P. Castoldi, S. J. B. Yoo, and L. Velasco, "Soft Failure Localization during Commissioning Testing and Lightpath Operation [Invited]," IEEE/OSA Journal of Optical Communications and Networking (JOCN), vol. 10, pp. A27-A36, 2018.

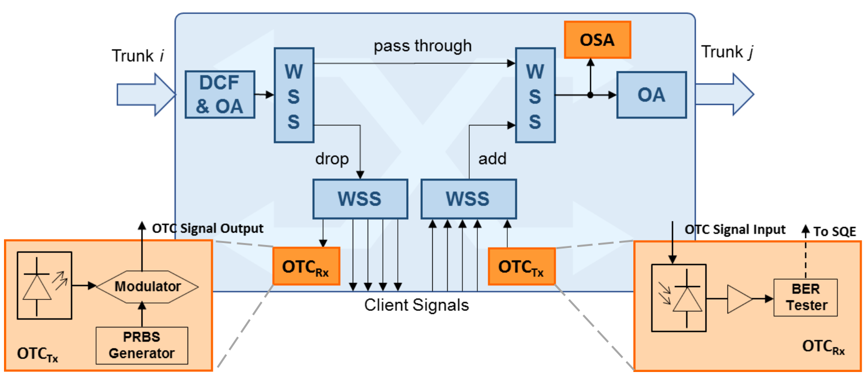
B. Shariati, M. Ruiz, J. Comellas, and L. Velasco, "Learning from the Optical Spectrum: Failure Detection and Identification [Invited]," IEEE/OSA Journal of Lightwave Technology (JLT), vol. 37, pp. 433-440, 2019.

Commissioning testing

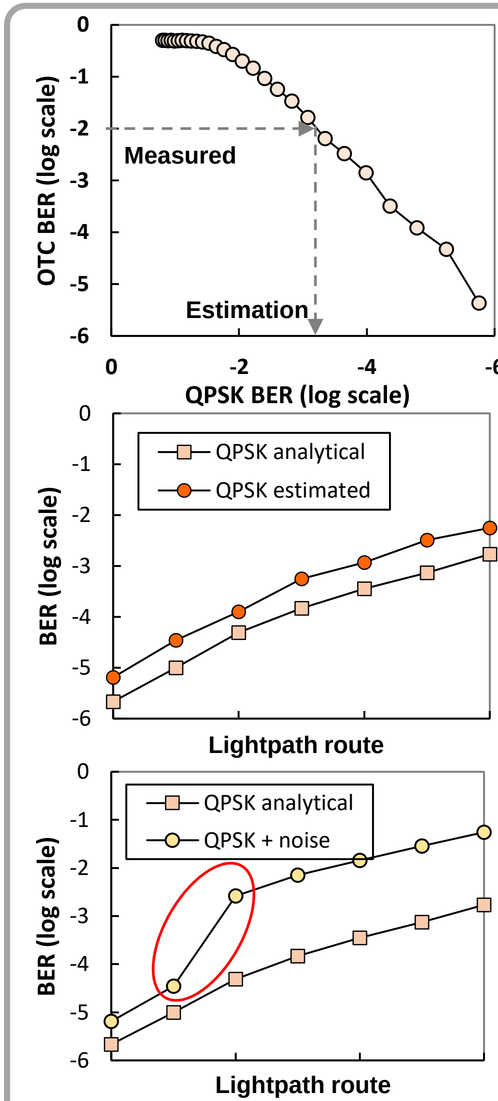
Active monitoring

Lightpath is created but client signal is not yet connected. Estimate BER that the client signal would experience.





Low-cost optical testing channel (**OTC**) modules (UCDavis)



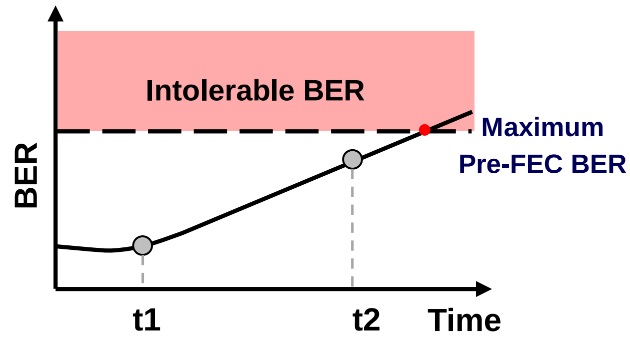
OTC BER Estimation (SQE)

Analytical vs Estimation (along the route) (TISSUE)

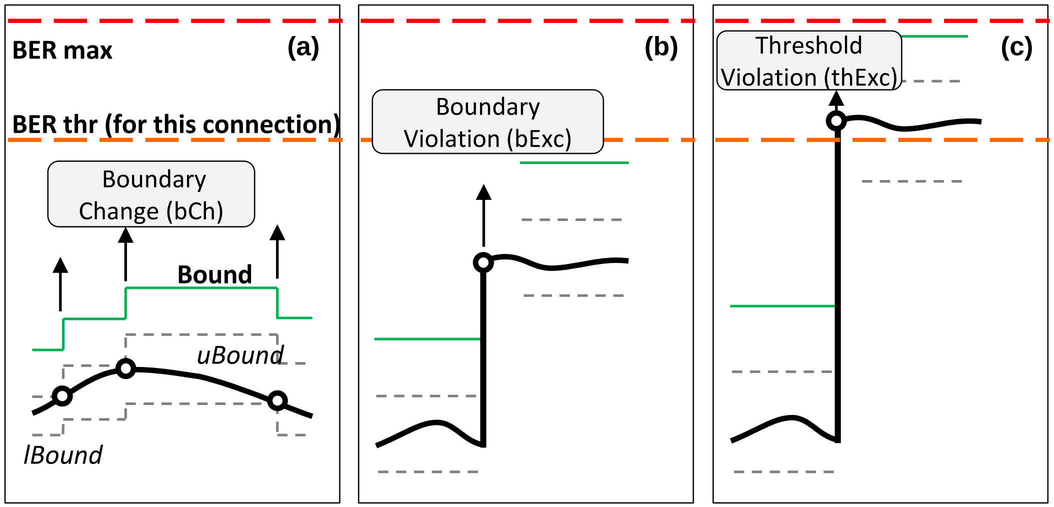
Localization!

BANDO: BER Anomaly Detection

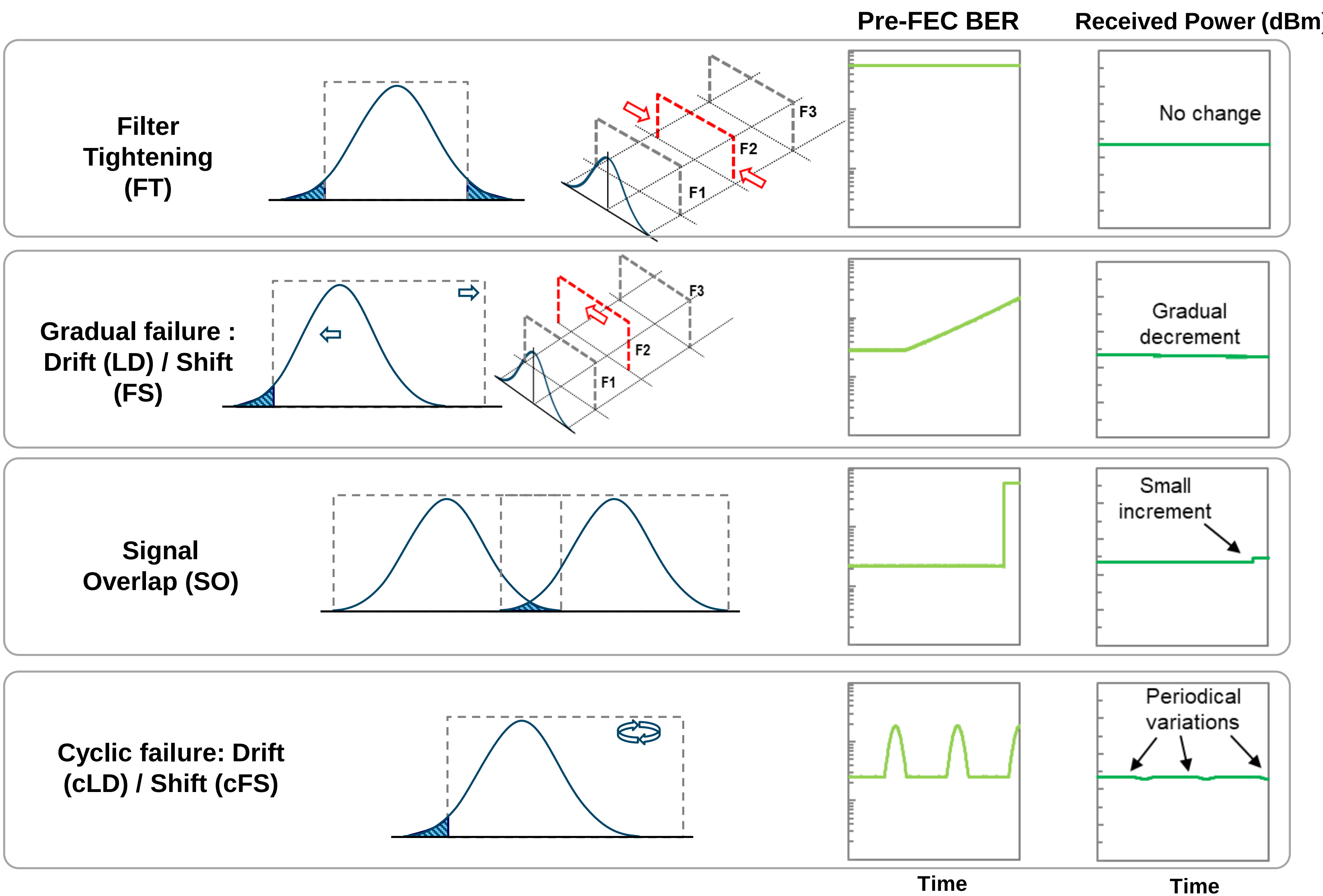
- Receives metered **pre-FEC BER** and **power** for every connection at a given rate.
- Analyzes pre-FEC BER to detect **gradual changes before reaching intolerable BER**.
- Detect **soft-failures** before the connection is disrupted.



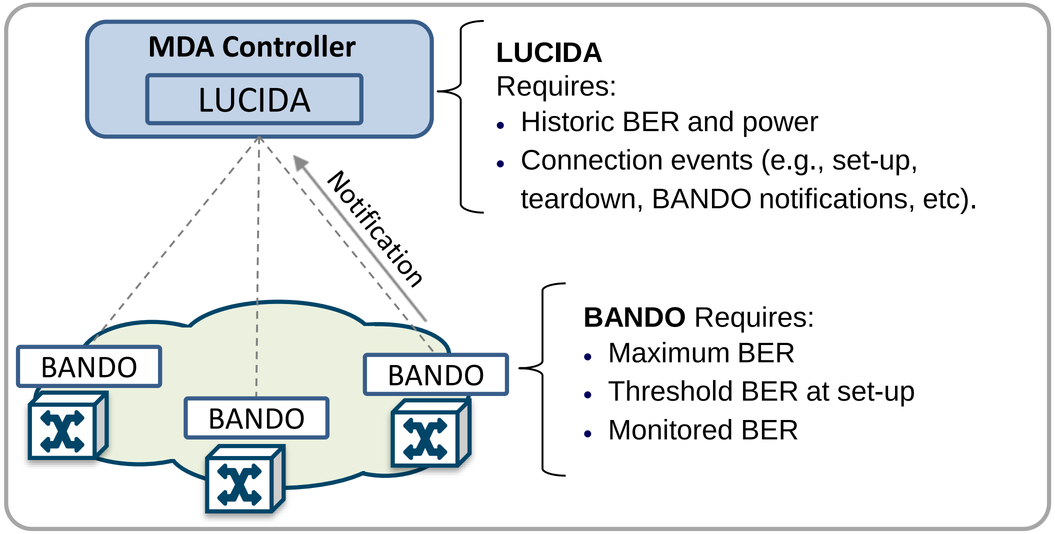
Three cases of BER evolution with time



Failure identification



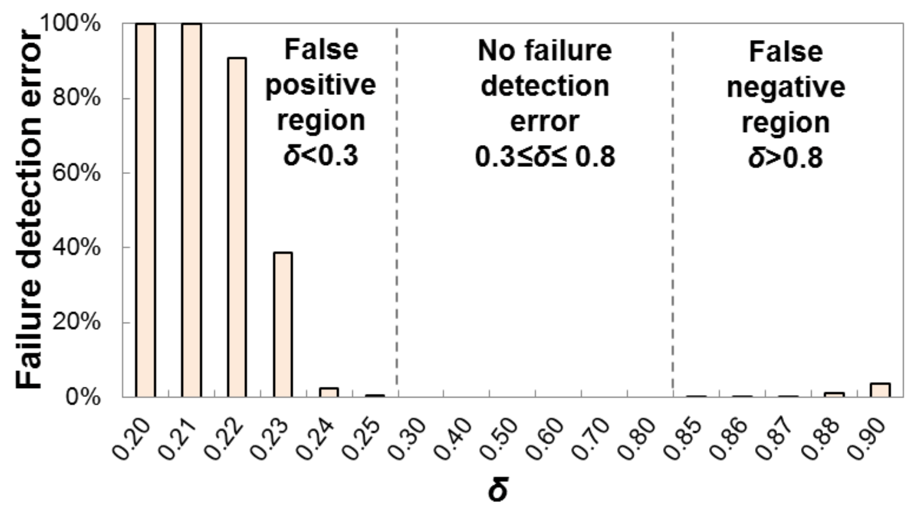
LUCIDA: Failure Identification



- Identifies the **most probable** cause of failure.
- Predicts whether **BER threshold will be violated** in the near future.
- Detects **pre-FEC BER variations** on individual optical connections.

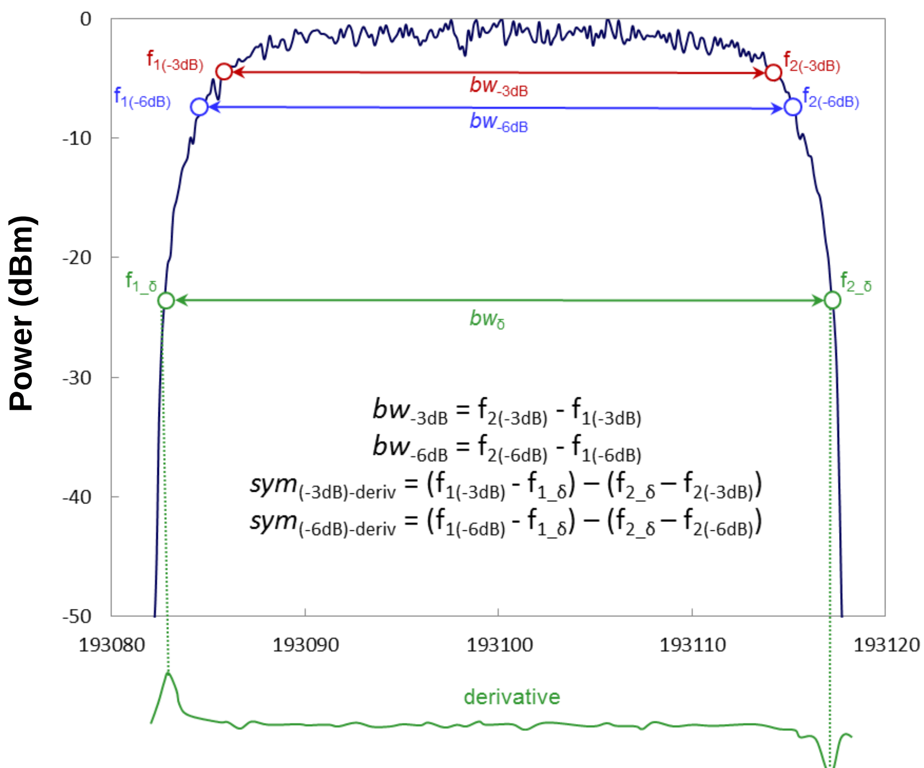
```
INPUT: notif, α, β_qh, δ
OUTPUT: <class, prob, timeToMaxBER>
connId ← notif.conn_id
lastBer ← notif.getLast().ber
threshold ← notif.threshold
storeNotif(connId, notif)
if (lastBer / threshold) < δ then return <'no',-,->
M.D={<t,ber,P_Rx>} ← getData(connId)
M.N={<t,type,data>} ← getNotif(connId)
P_H ← computeFeatureProbs(M, α)
P_Q ← computeFailureProbs(P_H, β_qh)
failureClass ← arg max_(q ∈ Q)(P_Q)
tmax ← -
if P_H(BERTrend) > 0 then
  tmax ← computeTMax (M)
return <failureClass, P_Q(failureClass), tmax>
```

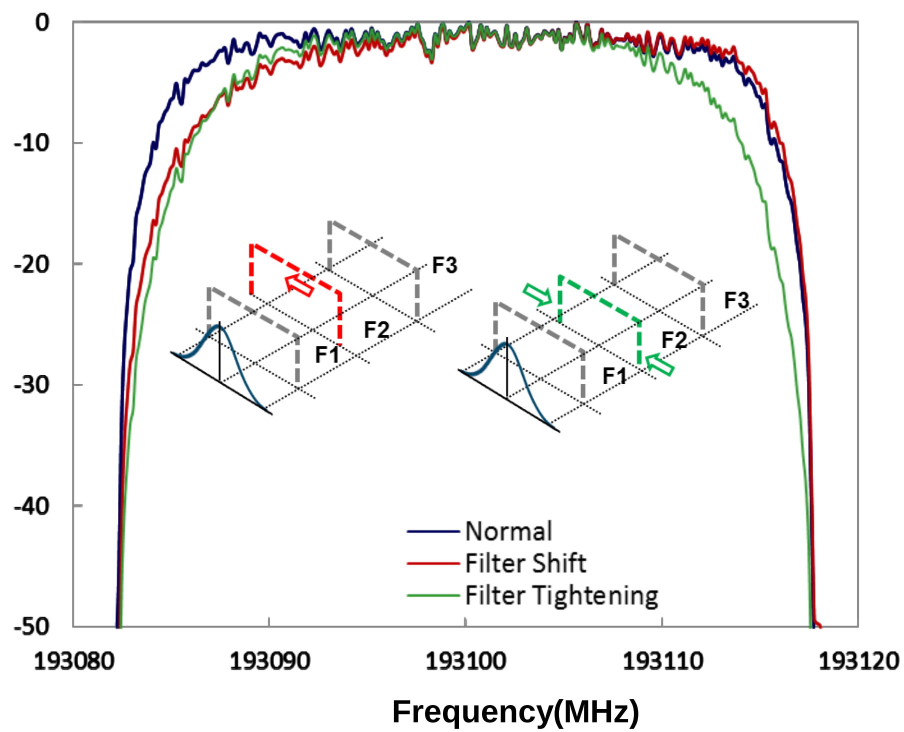
Selection of *δ*



In-operation Failure Localization / Identification (OSA)

- The signal becomes asymmetrical in the case of Filter Shift (*).
- Its edges get noticeably rounded in the case of Filter Tightening (*).

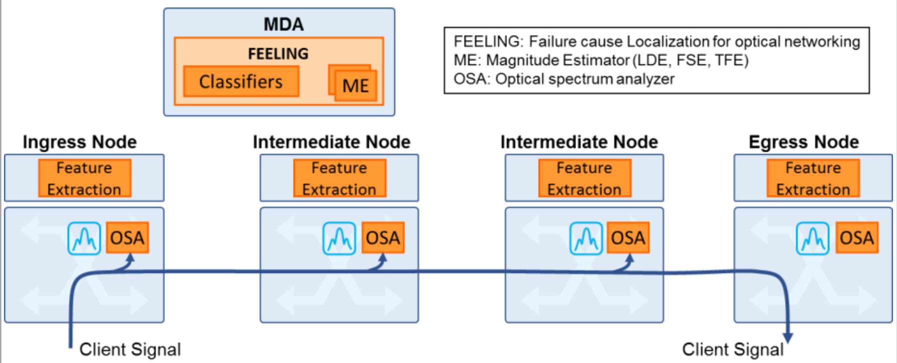




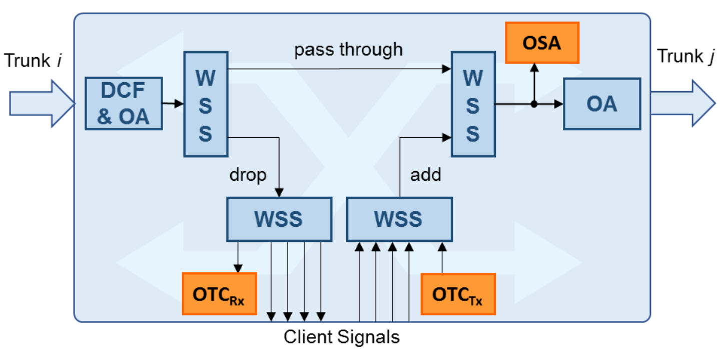
*** From B. Shariati's PhD Thesis**

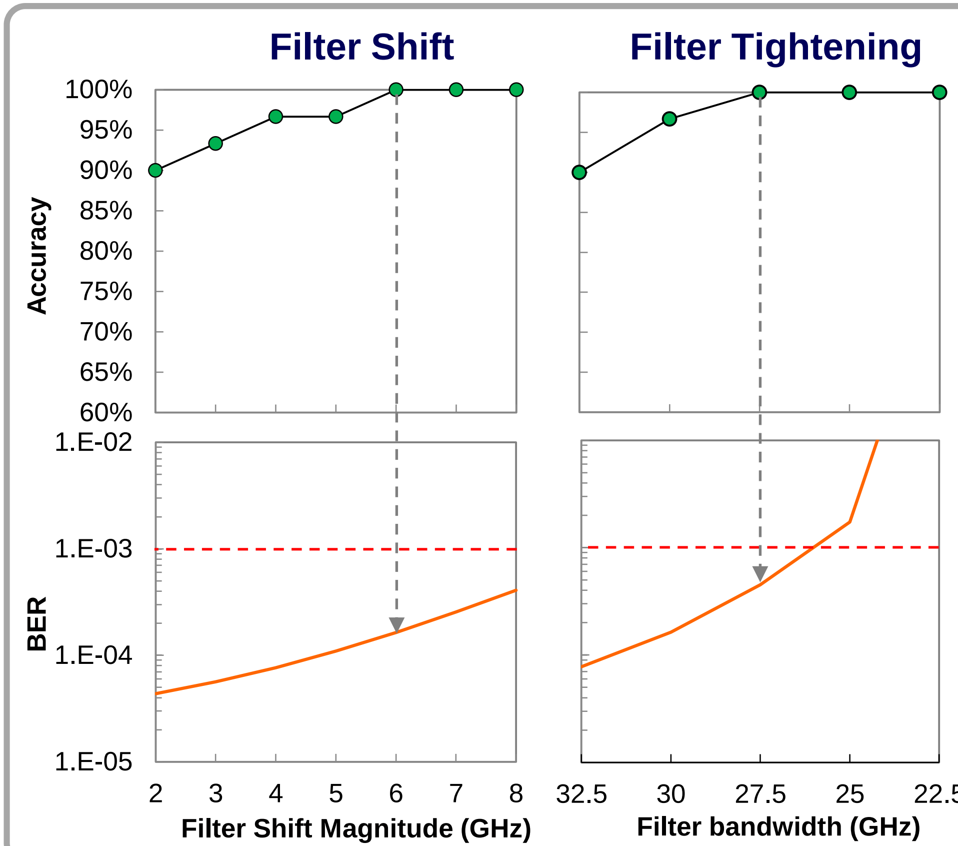
In-operation failure localization

Passive Monitoring



Relies on cost-effective **OSAs** installed in the outgoing links in the optical nodes

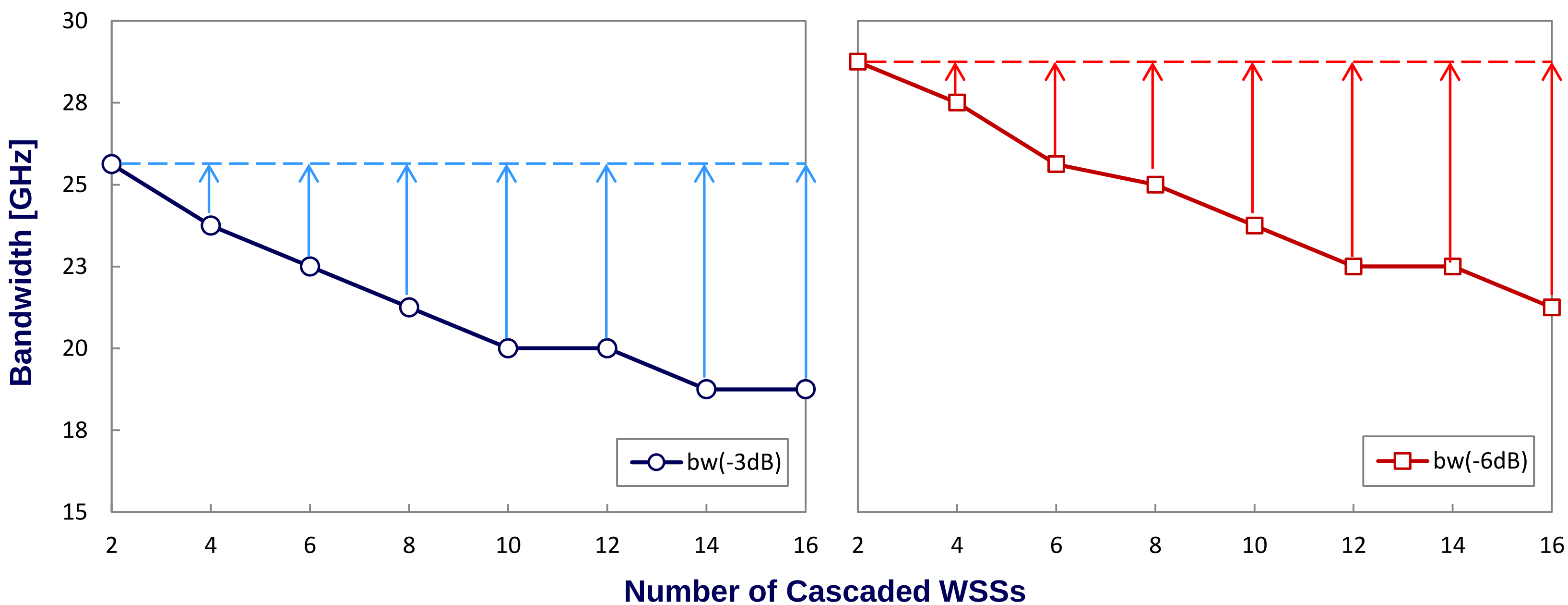




Correction Mask Calculation

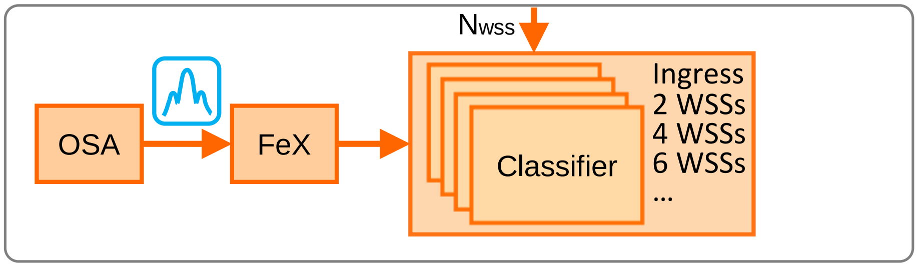
- **Filter cascading** strongly affect some of the features that the classifier uses to make the predictions.

- The feature values of a signal, which passes **N filters**, can be corrected by adding/subtracting the differences between the features of a properly configured signal at that node w.r.t. its feature values in the ingress node.

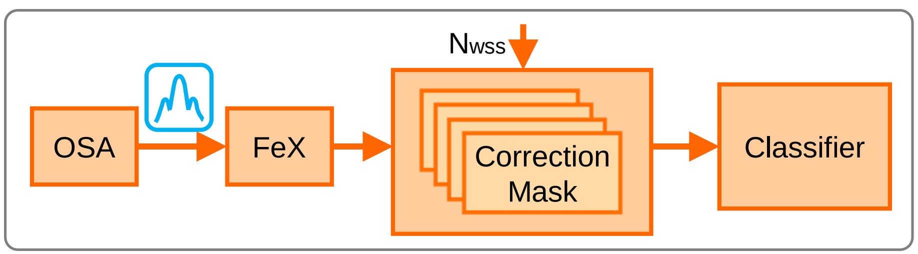


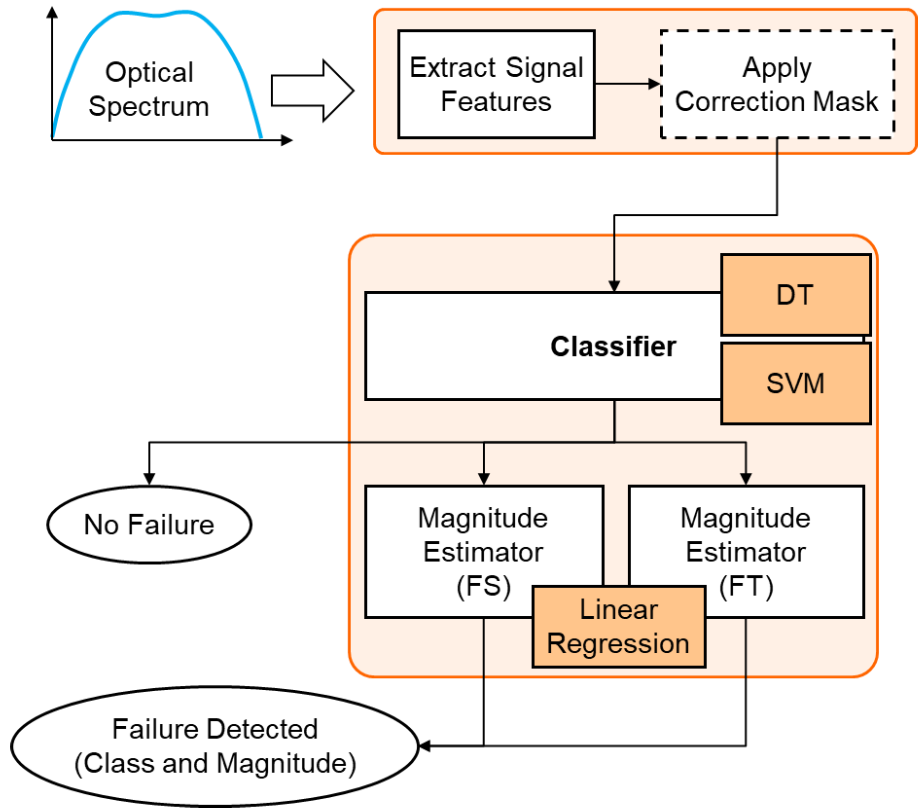
Analyzing the optical spectrum

Multi-Classifier Approach



Single-Classifier Approach

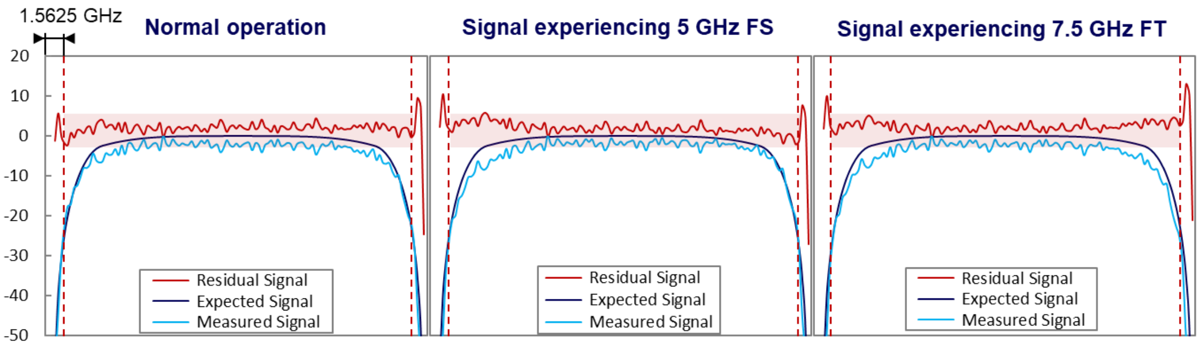


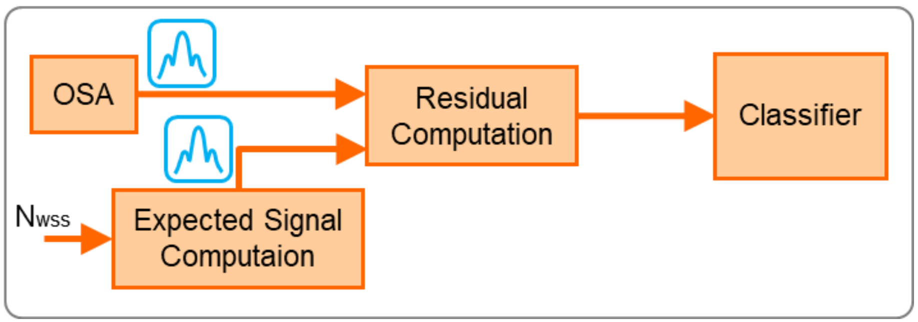


Summary

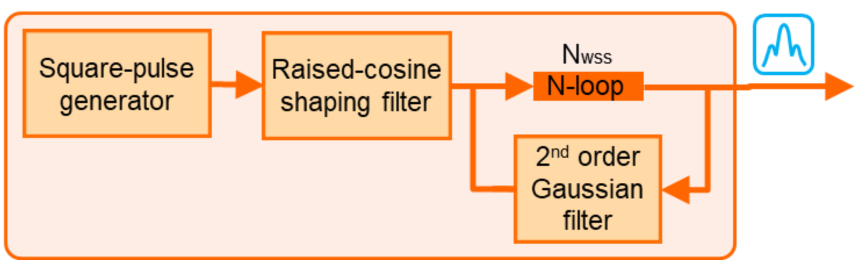
	pre-processing	training phase	# classifiers	availability at the node level	Classif. method	accuracy
Multi Classifier	does not require	observations of **every level** of filter cascading	# of nodes to support	all classifiers	DT	medium
					SVM	**high**
Single Classifier	pre-processing of the features	observations of just a **single level** of filter cascading	1	one classifier + all correction masks	DT	medium
					SVM	**high**

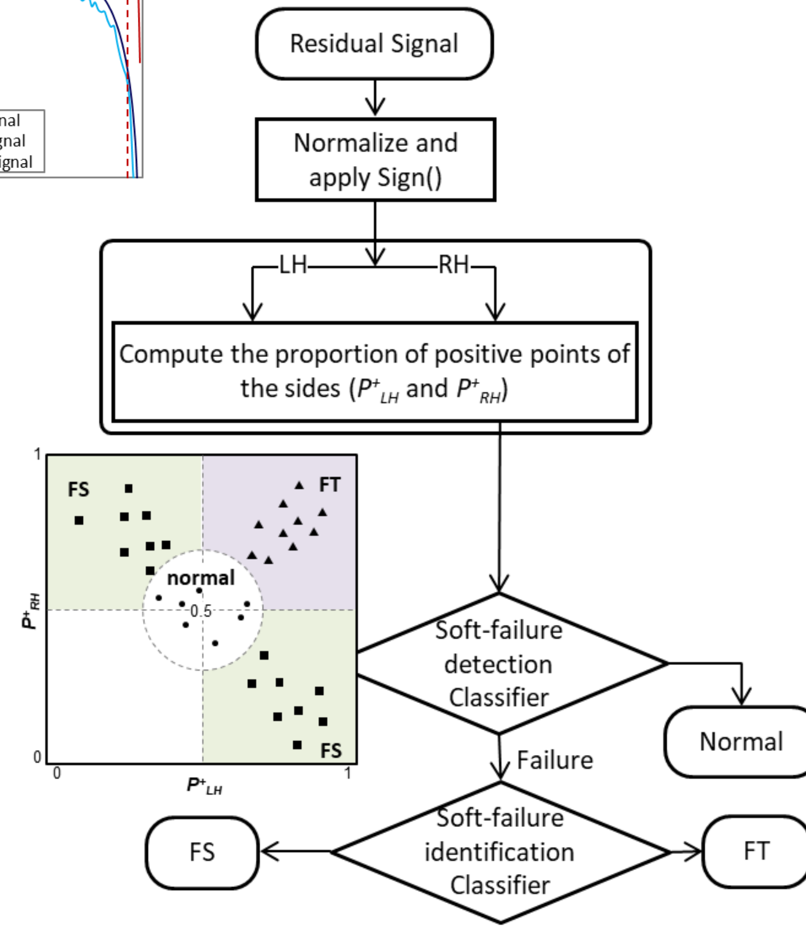
Residual-Based Approach





Expected Signal Calculator (ESC)

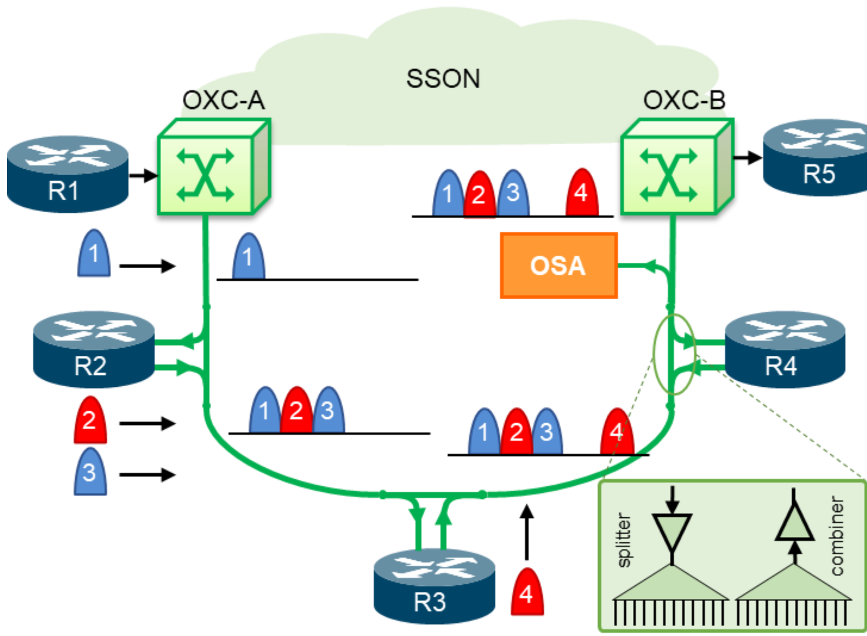




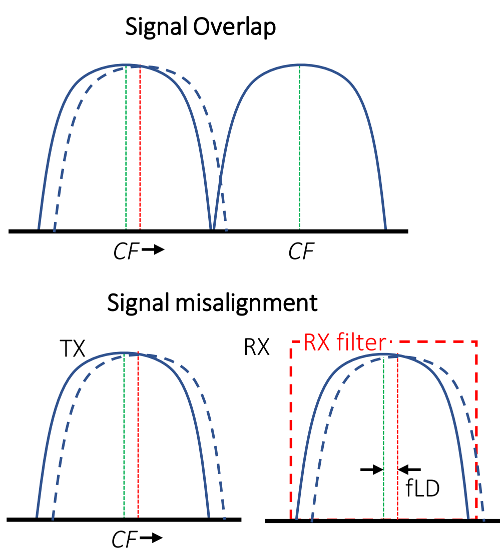
Summary of Approaches

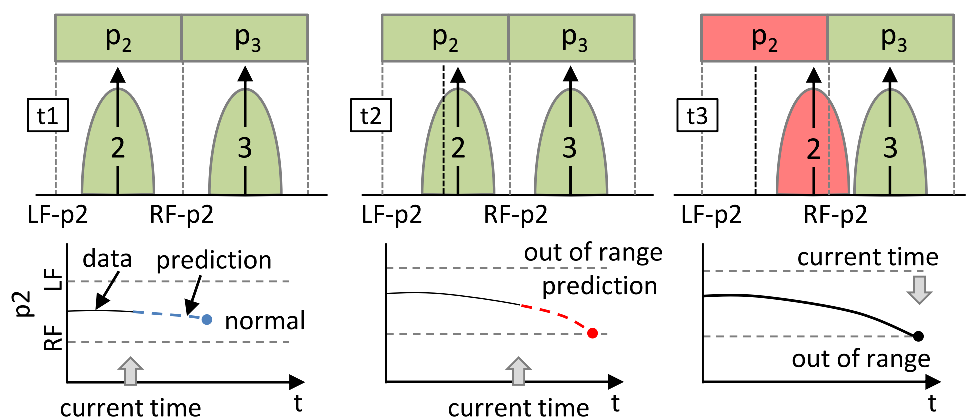
	pre-processing	classification method	training phase	number of classifiers	availability at the node level	accuracy	robustness w.r.t # nodes
Multi Classifier	**does not require**	SVM	requires observations of every level of filter cascading	# of nodes to support	all classifiers	good	good
Single Classifier	features	SVM	requires observations of just a single level of filter cascading	1	one classifier + all correction masks	good	good
Residual Based	optical spectrum	hypothesis testing	**does not require**	1	expected and residual signal computations modules	**very good**	**very good**

Filterless Optical Metro Networks

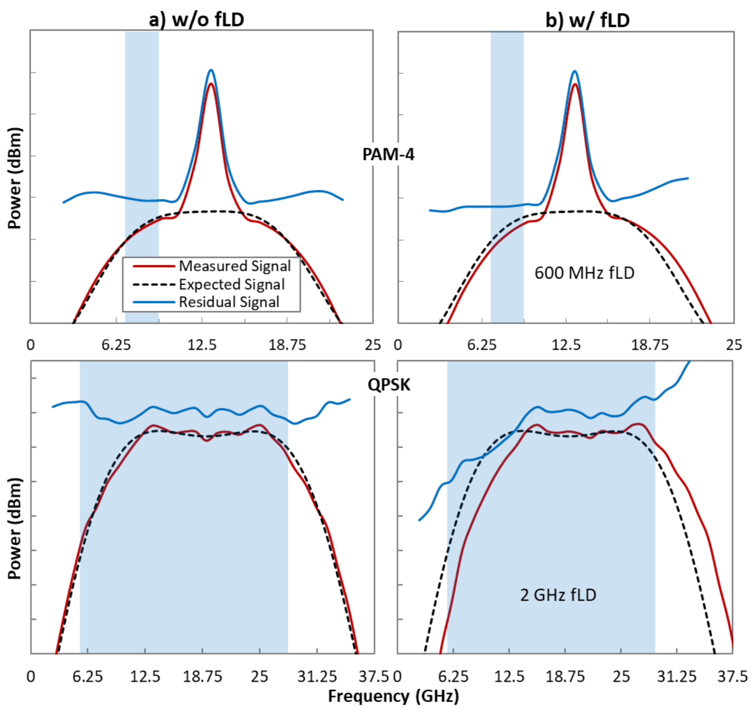


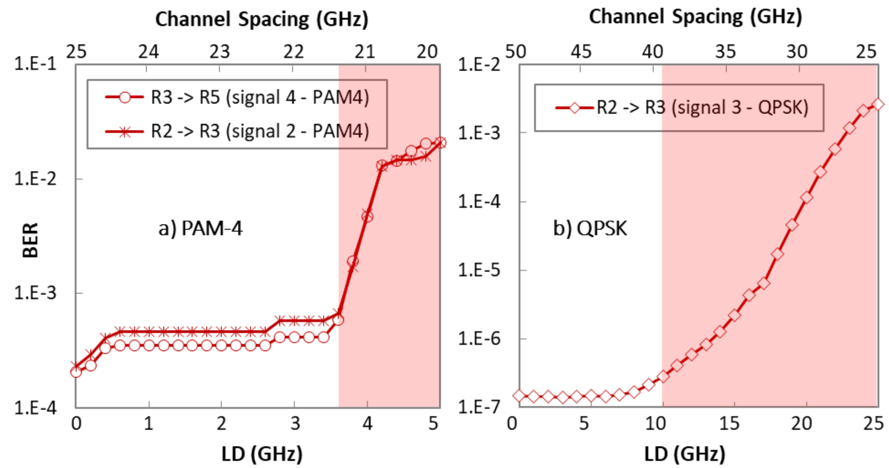
- **Direct-detection** optical transmission systems can be used in metro FONs, which reduces costs, as compared to coherent-detection systems.
- The **residual-based** approach is proposed for optical signal tracking to detect small frequency **laser drift** problems and enable safely **reducing channel spacing**.

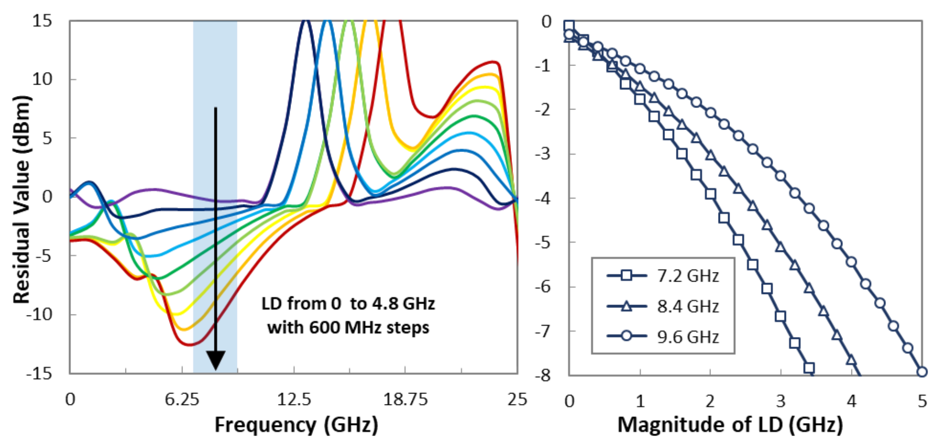




Some results







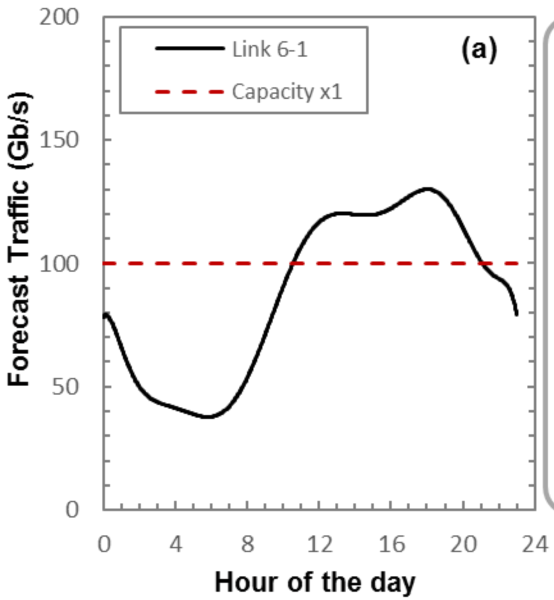
Data analytics at the Packet layer

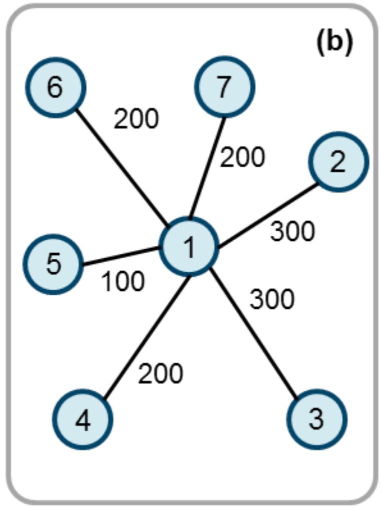
F. Morales, Ll. Gifre, F. Paolucci, M. Ruiz, F. Cugini, P. Castoldi, and L. Velasco, "Dynamic Core VNT Adaptability based on Predictive Metro-Flow Traffic Models," IEEE/OSA Journal of Optical Communications and Networking (JOCN), vol. 9, pp. 1202-1211, 2017.

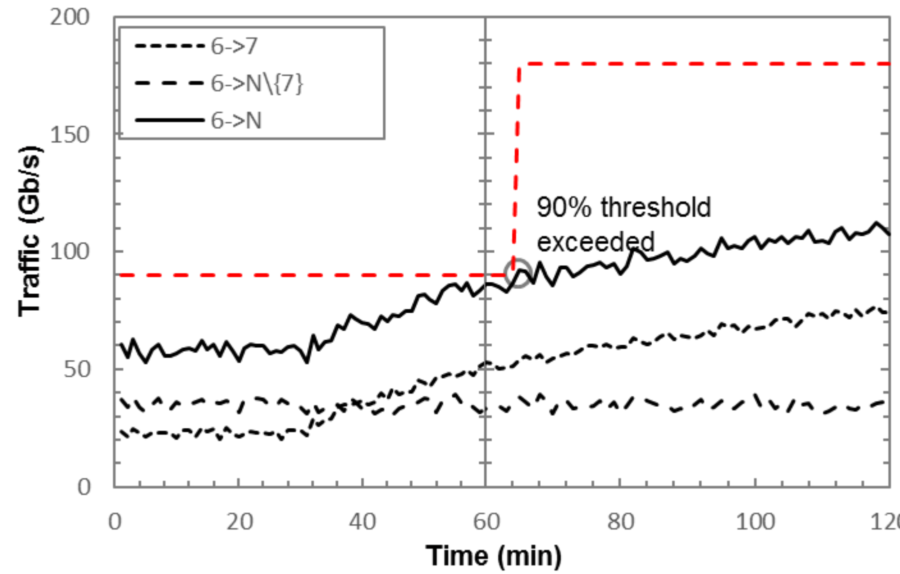
F. Morales, M. Ruiz, Ll. Gifre, L. M. Contreras, V. López, and L. Velasco, "Virtual Network Topology Adaptability based on Data Analytics for Traffic Prediction," (Invited) IEEE/OSA Journal of Optical Communications and Networking (JOCN), vol. 9, pp. A35-A45, 2017.

A. P. Vela, M. Ruiz, L. Velasco, "Distributing Data Analytics for Efficient Multiple Traffic Anomalies Detection," Elsevier Computer Communications, vol. 107, pp. 1-12, 2017.

VNT Design and Reconfiguration Options

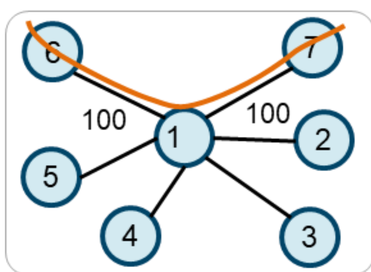


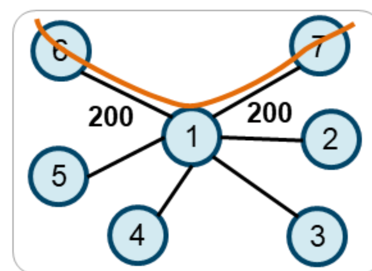




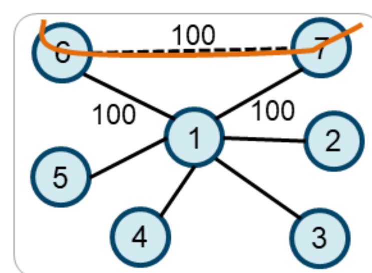
1 **Static VNT design**

Current VNT



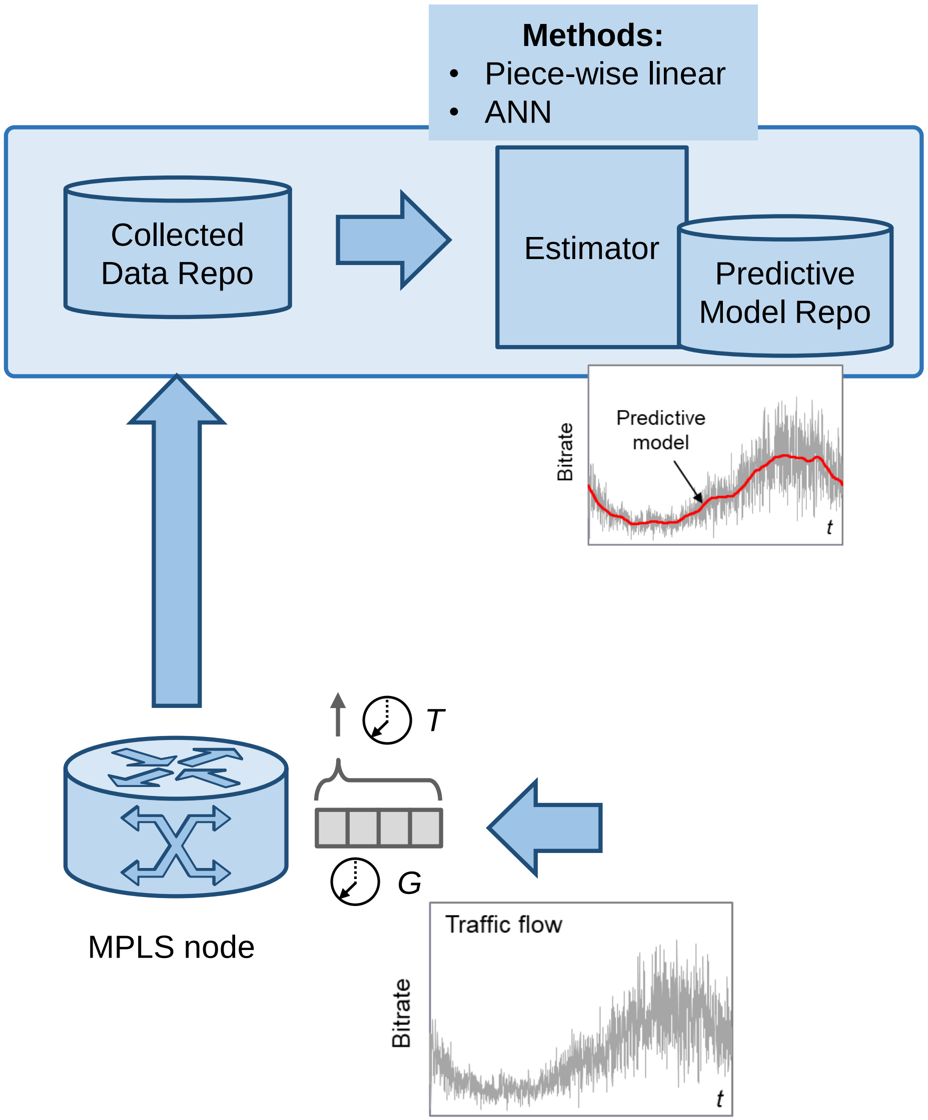


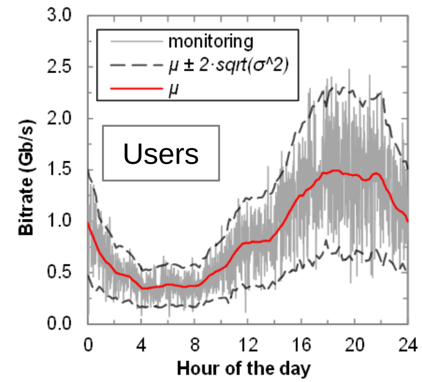
2 **Threshold-based reconfiguration**

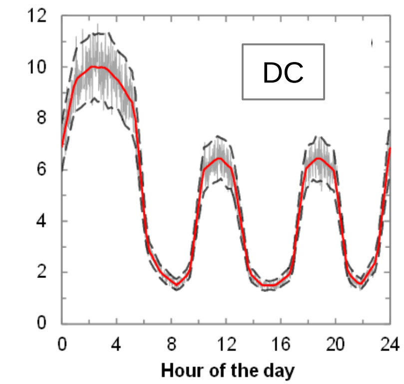


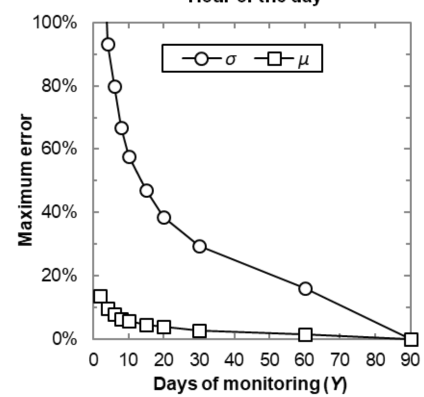
3 **Reconfiguration based on OD traffic prediction**

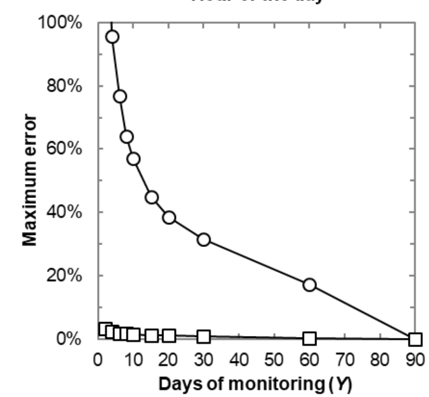
Modeling OD traffic pairs

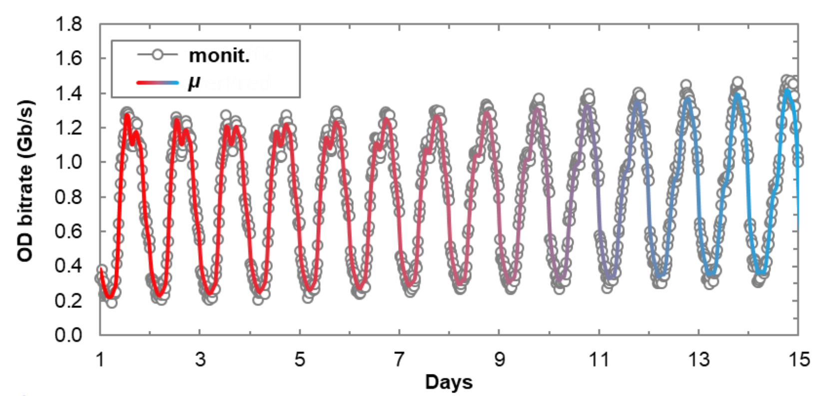




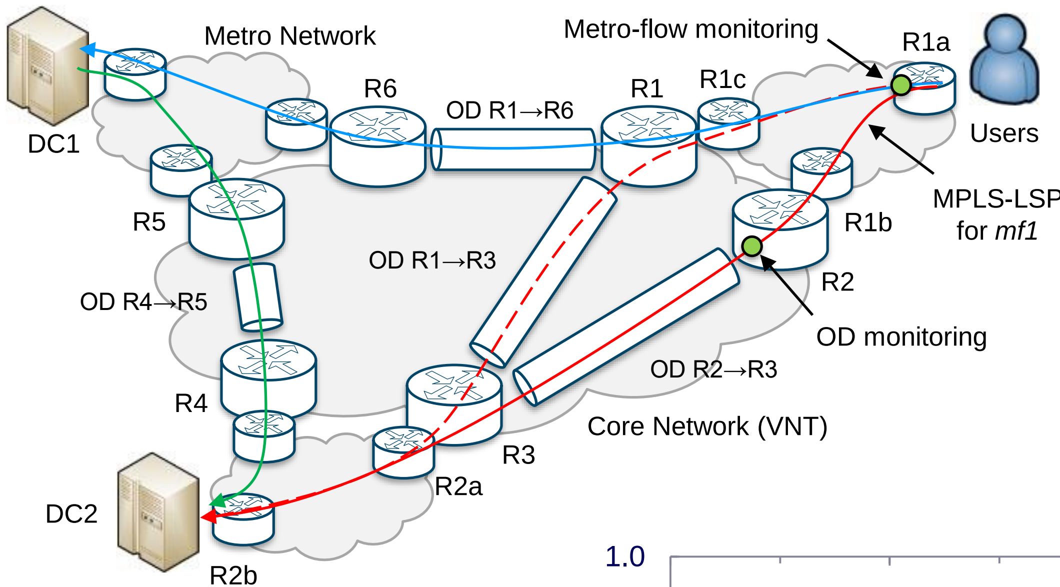




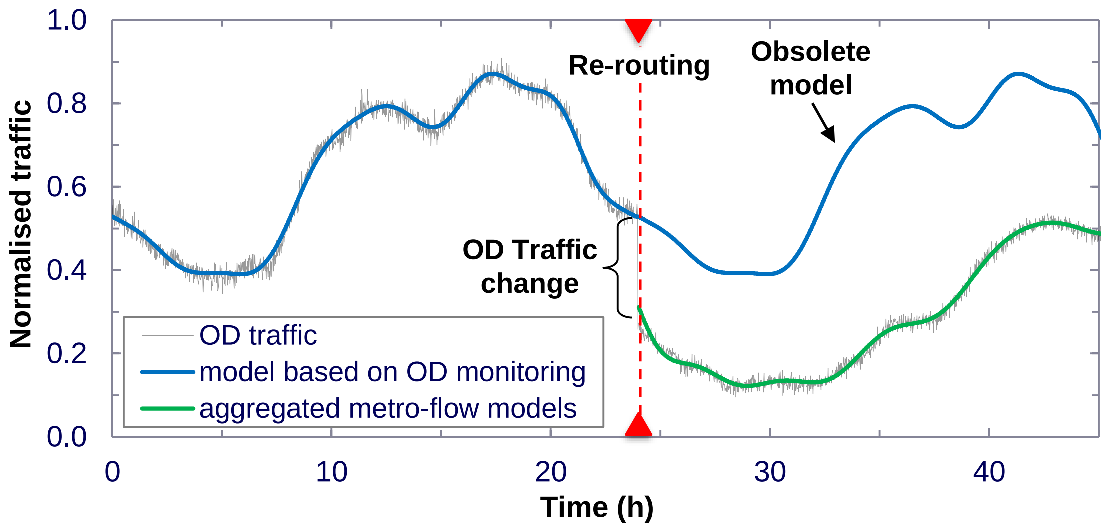




Flow traffic prediction under changing traffic



- OD traffic **patterns** in the core can **change** because of **metro-flow rerouting**.
- **Aggregating** metro-flow traffic **models** can produce quality, **new OD models**.



Metro-flow model aggregation

	Piece-wise linear model	ANN model
Pros	• Addition results in a new piece-wise linear function.	• Adapts to evolutionary traffic. • No assumptions about underlying traffic distribution.
Cons	• Does not adapt to evolutionary traffic. • Assumes random variables for the traffic.	• The addition of ANNs is not defined.

- We target at predicting the **maximum bitrate** of the **aggregation** of several (e.g. hundreds) metro flows.
- **Adding maximums** largely **overestimates** the maximum of the aggregate.
- We can use μ(t) and σ(t) models of each metro-flow to estimate the maximum of the aggregate.

$$\max_{od}(t) \approx \mu_{od}(t) + k\sqrt{\sigma^2_{od}(t)}$$

$$\mu_{od}(t) = \sum_{f \in F(od)} \mu_f(t)$$

$$\sigma^2_{od}(t) = \sum_{f \in F(od)} \sigma^2_f(t)$$

$$\mu_{od}(t) += \sum_{f \in F_{IN}(od)} \mu_f(t) - \sum_{f \in F_{OUT}(od)} \mu_f(t)$$

$$\sigma^2_{od}(t) += \sum_{f \in F_{IN}(od)} \sigma^2_f(t) - \sum_{F_{OUT}(od)} \sigma^2_f(t)$$

Traffic Anomaly detection

Traffic-based method

Uses predictive models to detect sequences of consecutive atypical traffic values

Normalize value of y(t) w.r.t. the average model (mean and standard deviation)

$$\hat{y}(t) = \frac{y(t) - \mu(t)}{\sigma(t)}$$

Score-based method

Probabilistic classifier with 2 labels (Norm/Anom)

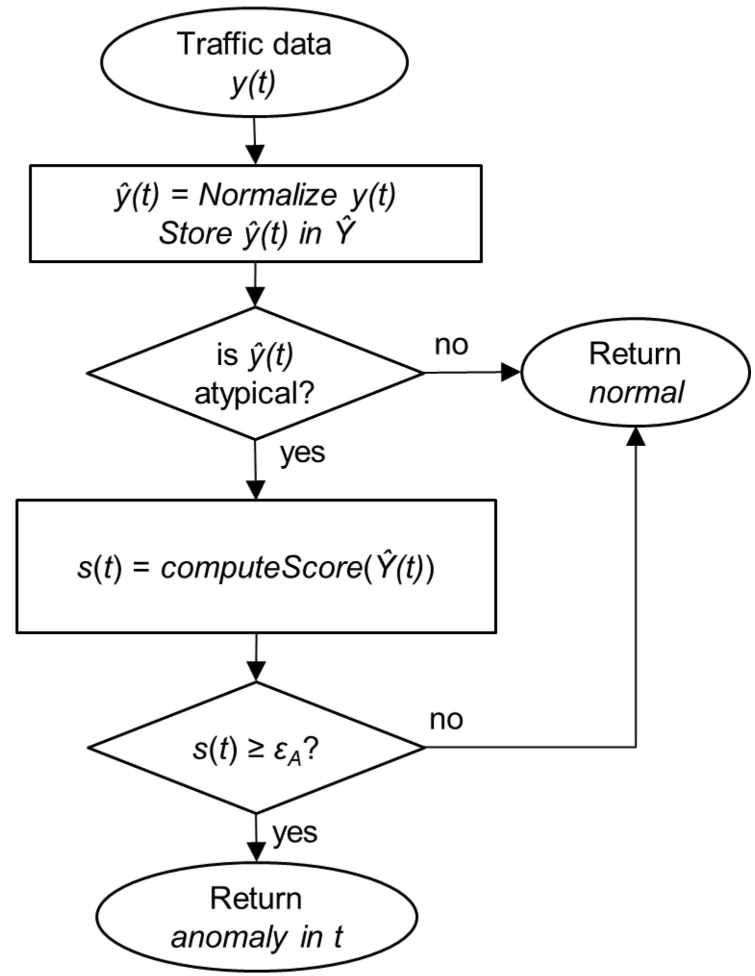
We define probability as an indicator of **how likely** is to consider ŷ as a **normal traffic value**

$$p(i) = 1 - P(z \leq |\hat{y}(t-i)|)$$

Score function s(t) to compute how likely is that a data series $\hat{Y}(t)$ does **not belong** to the **normal class.**

$$s(t) = \frac{1}{\sqrt[m]{\prod_{i=0..m-1} \left[1 - P\left(z \leq \left|\hat{y}(t-i)\right|\right)\right]}}$$

Score-based algorithm



MDA Architecture and Use Cases

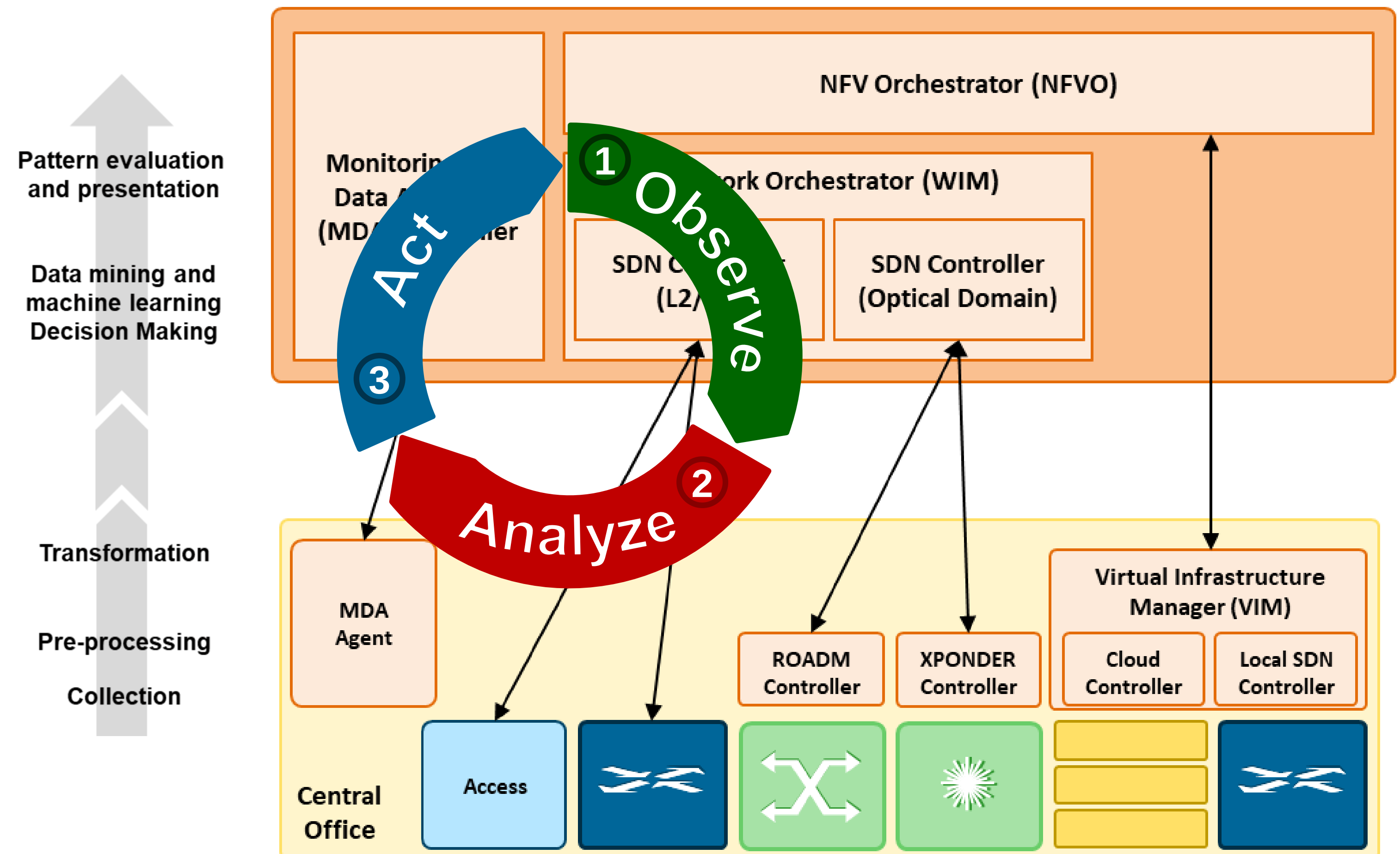
L. Velasco, A. Sgambelluri, R. Casellas, Ll. Gifre, J.-L. Izquierdo-Zaragoza, F. Fresi, F. Paolucci, R. Martínez, and E. Riccardi, “Building Autonomic Optical Whitebox-based Networks,” IEEE/OSA Journal of Lightwave Technology (JLT), vol. 36, pp. 3097-3104, 2018.

Ll. Gifre, J.-L. Izquierdo-Zaragoza, M. Ruiz, and L. Velasco, “Autonomic Disaggregated Multilayer Networking,” IEEE/OSA Journal of Optical Communications and Networking (JOCN), vol. 10, pp. 482-492, 2018.

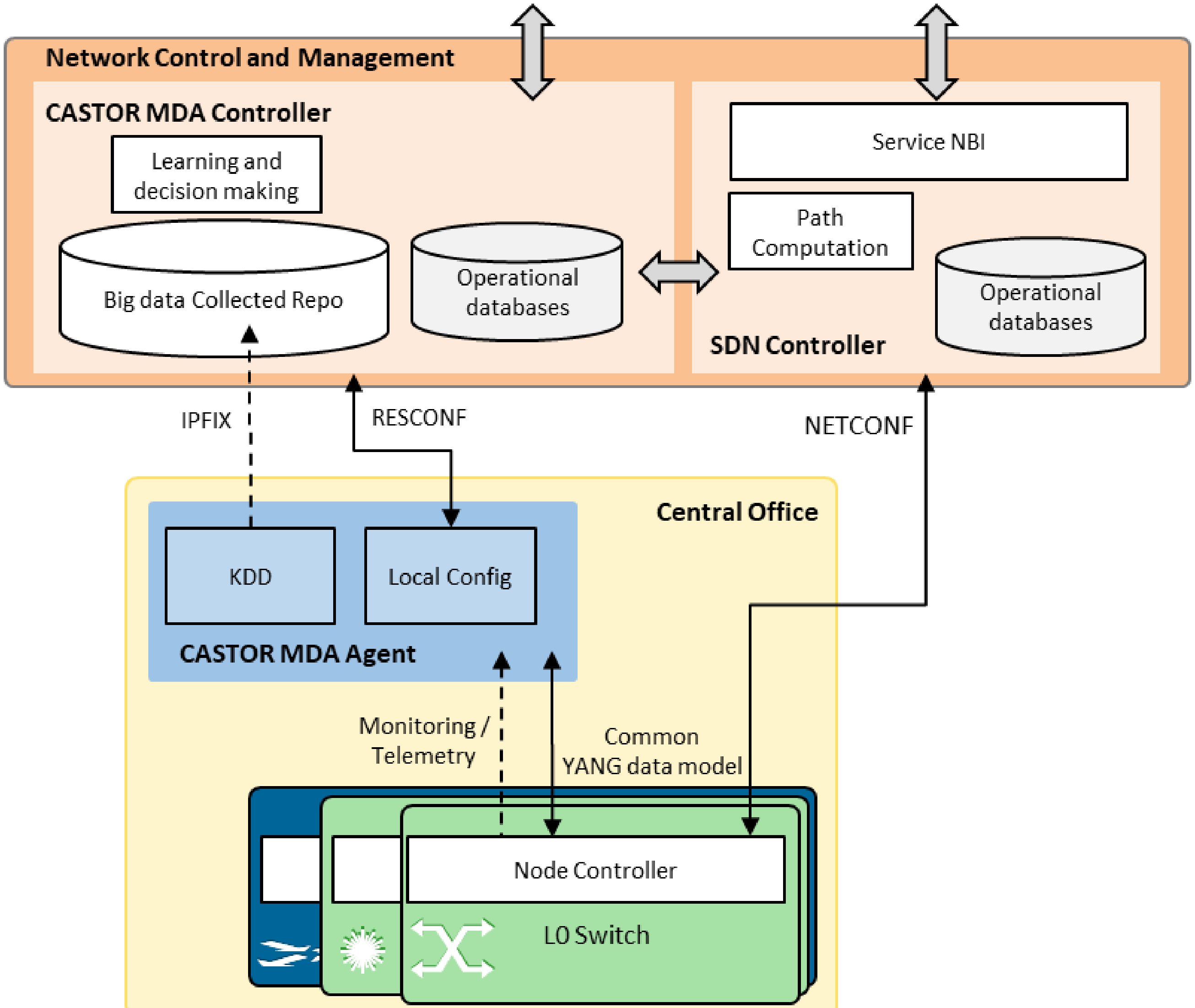
L. Velasco, Ll. Gifre, J.-L. Izquierdo-Zaragoza, F. Paolucci, A. P. Vela, A. Sgambelluri, M. Ruiz, and F. Cugini, “An Architecture to Support Autonomic Slice Networking [Invited],” IEEE/OSA Journal of Lightwave Technology (JLT), vol. 36, pp. 135-141, 2018.

L. Velasco, R. Casellas, S. Llana, Ll. Gifre, R. Martinez, R. Vilalta, R. Muñoz, and M. Ruiz, "A Control and Management Architecture Supporting Autonomic NFV Services," Springer Photonic Network Communications, vol. 37, pp. 24-37, 2019.

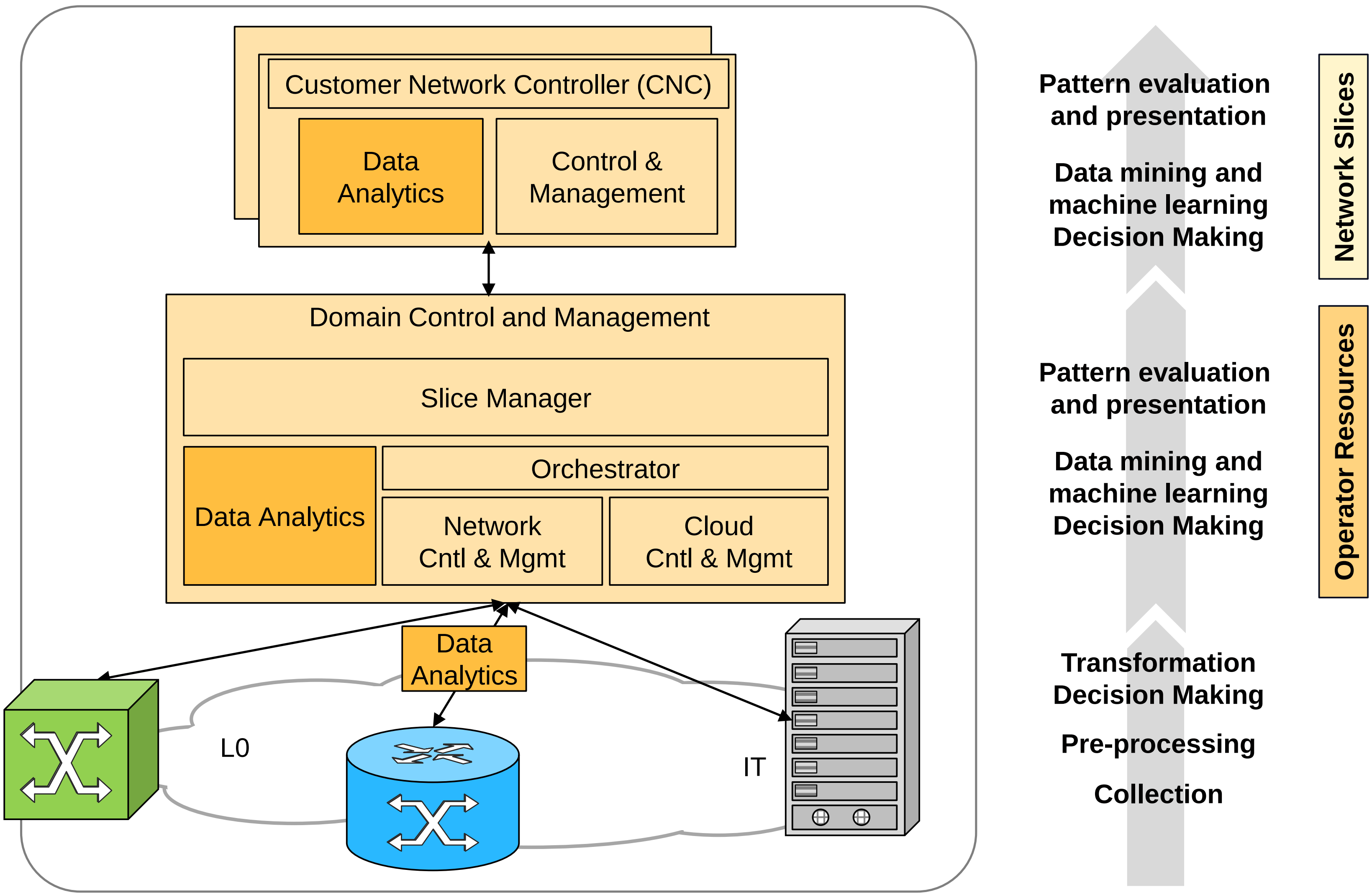
Conceptual Control, Orchestration and Management Architecture



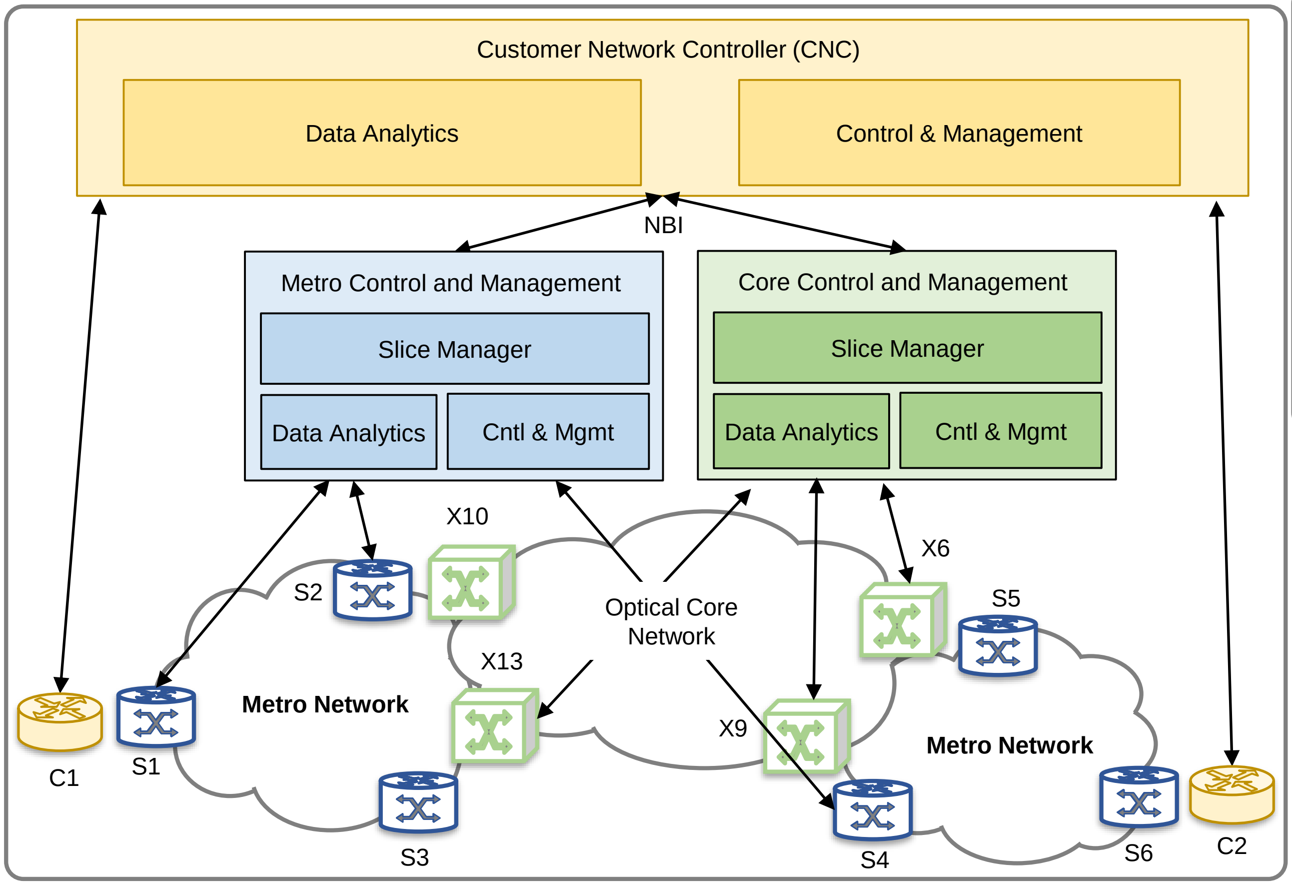
Architecture and Interfaces

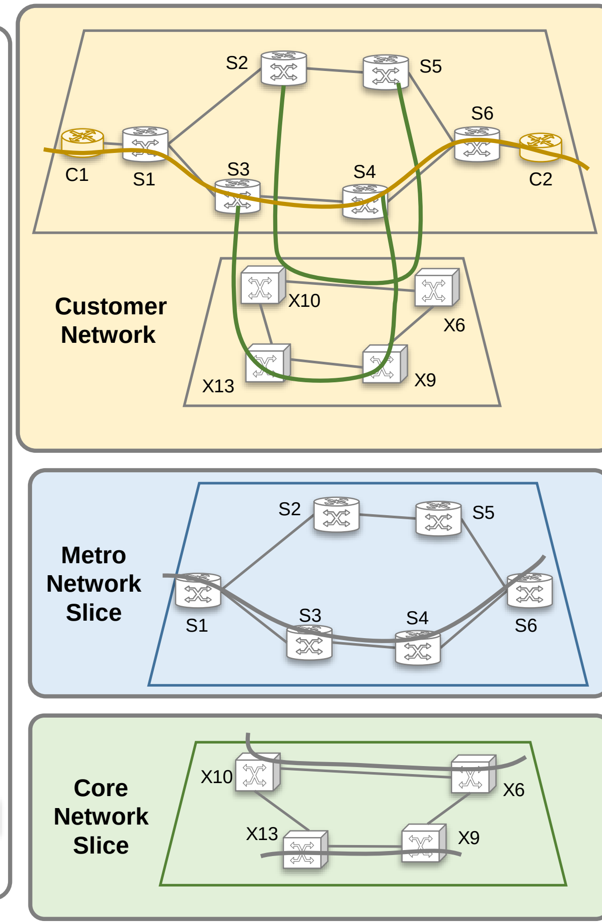


Network Slicing Support

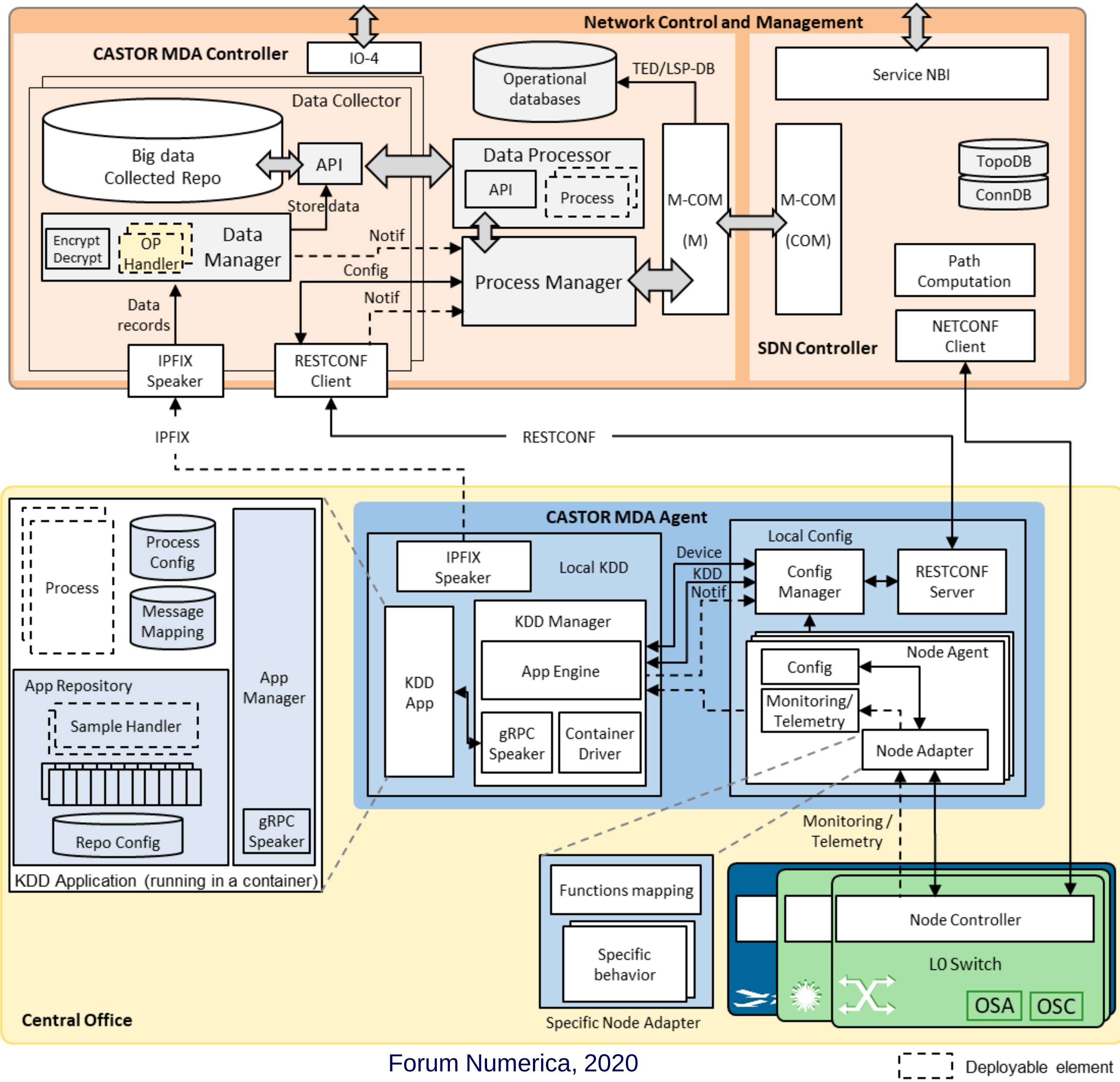


Example of Network Slicing

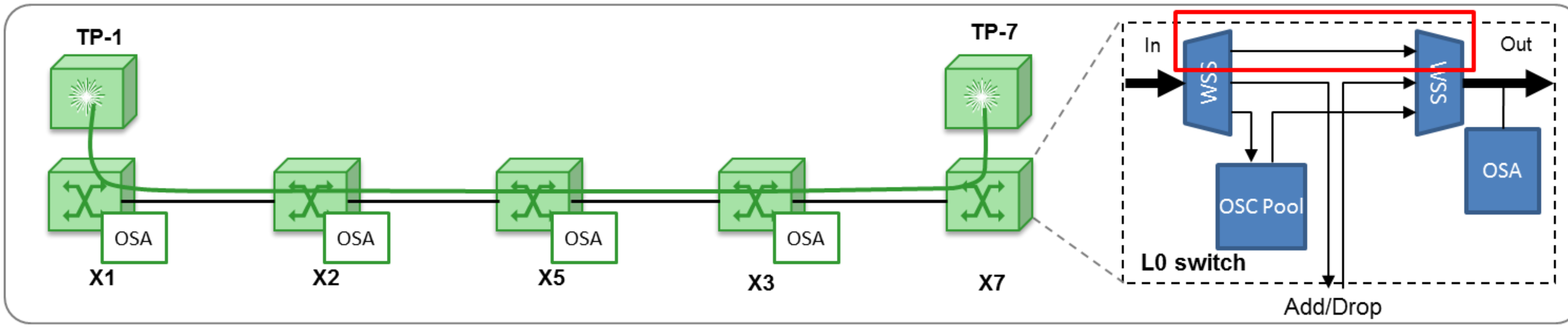


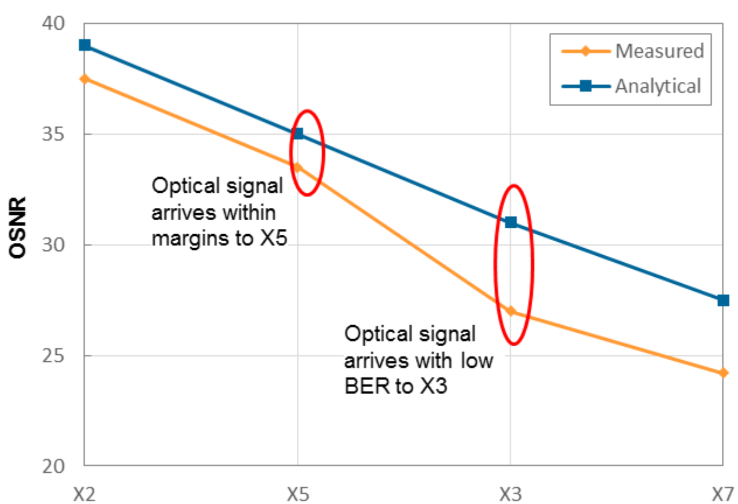


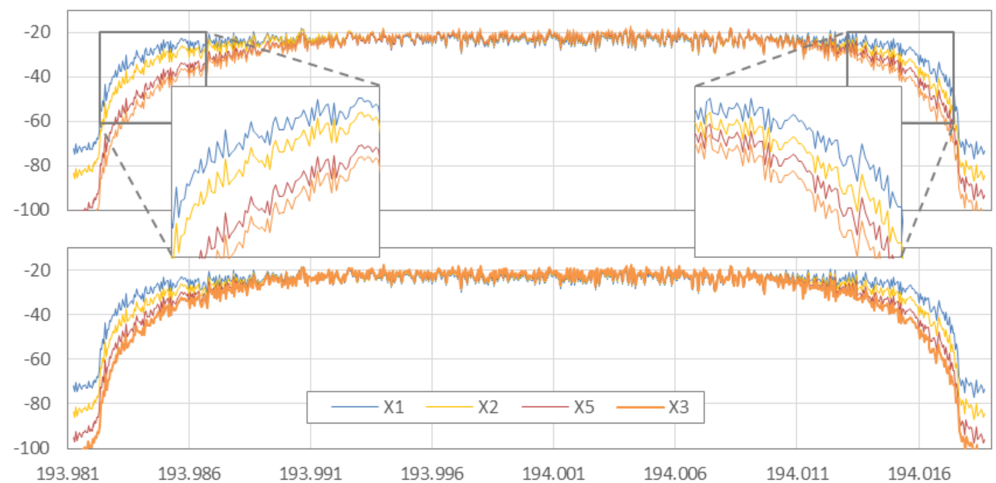
Architecture



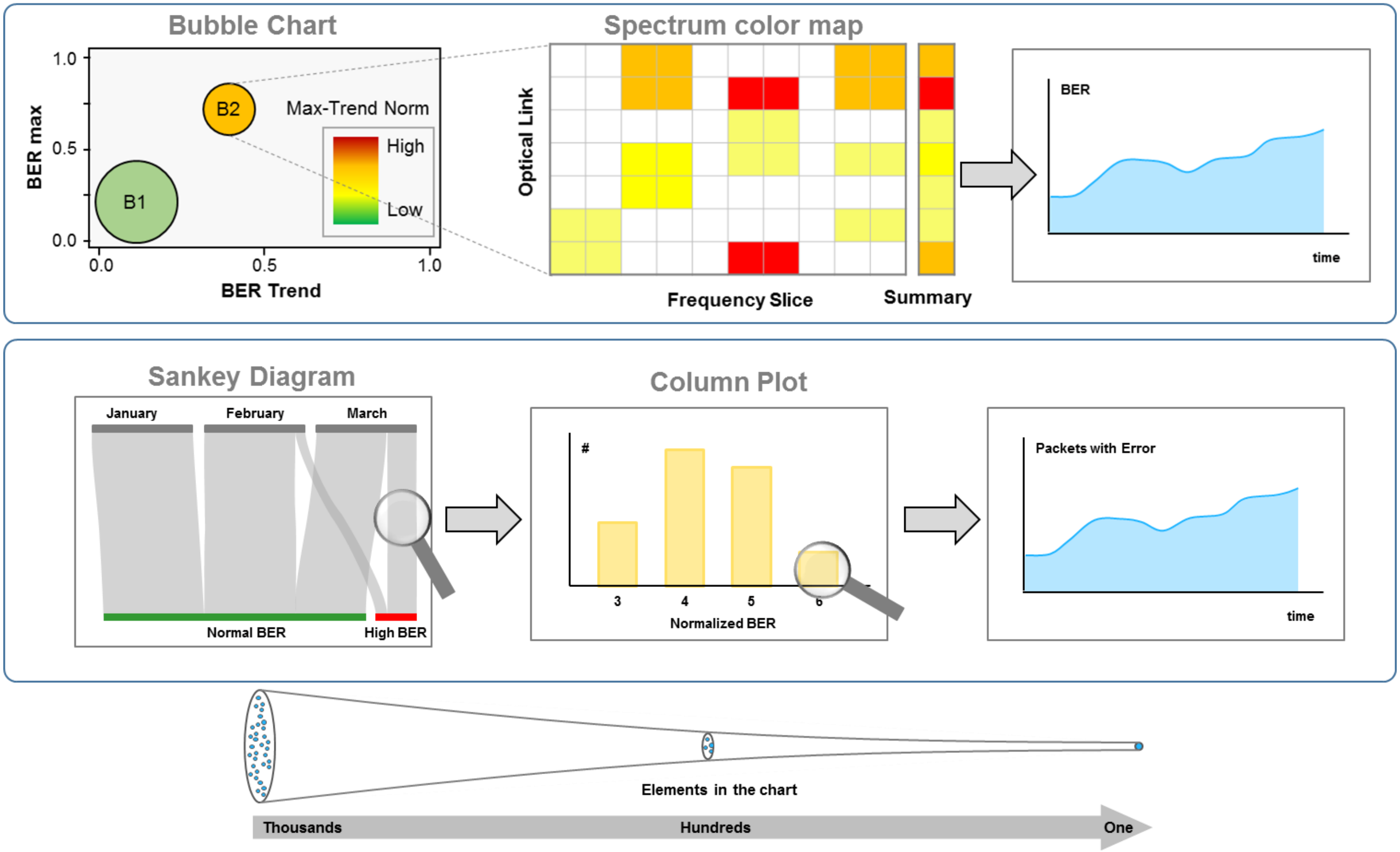
Failure Localization



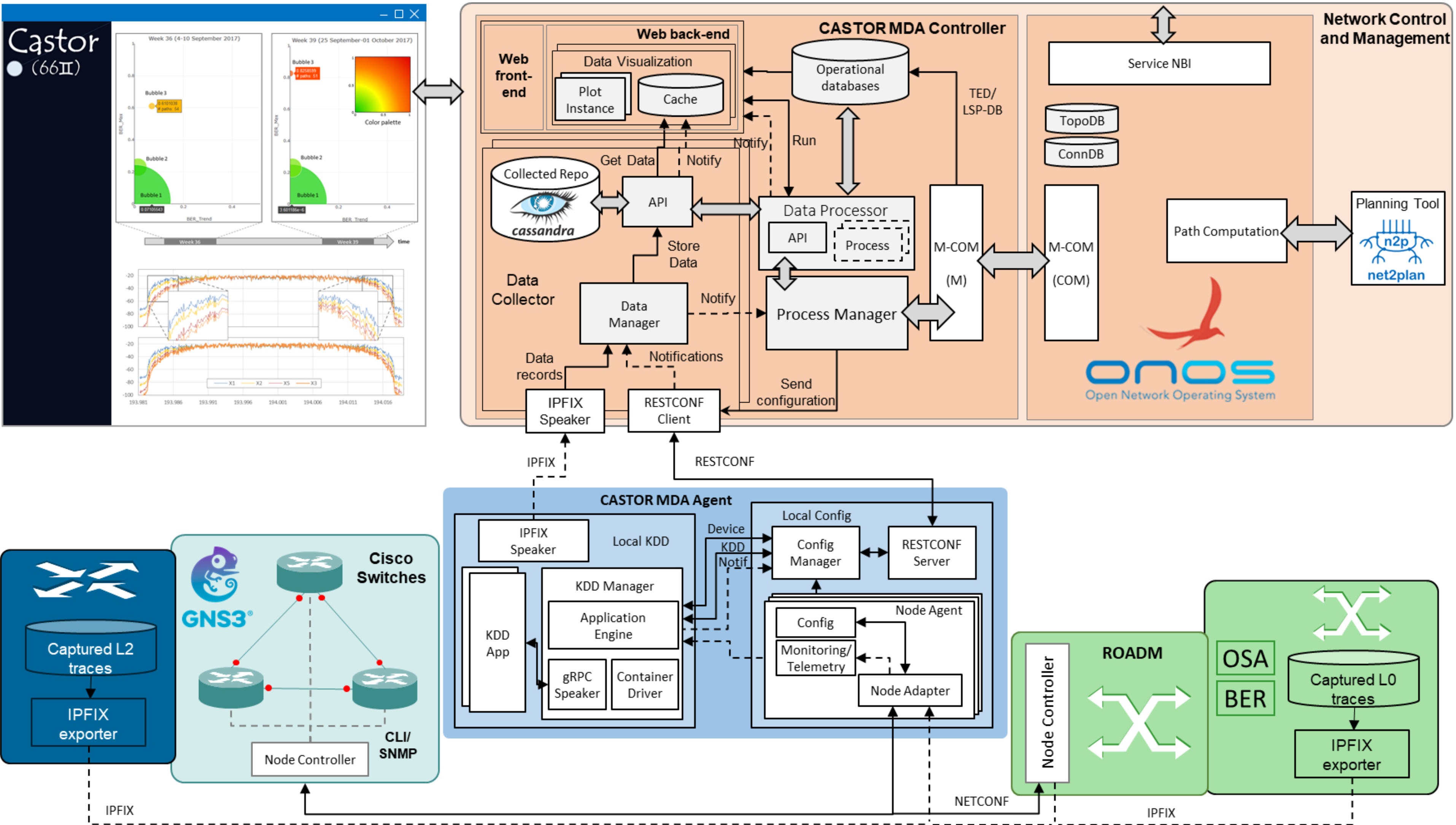




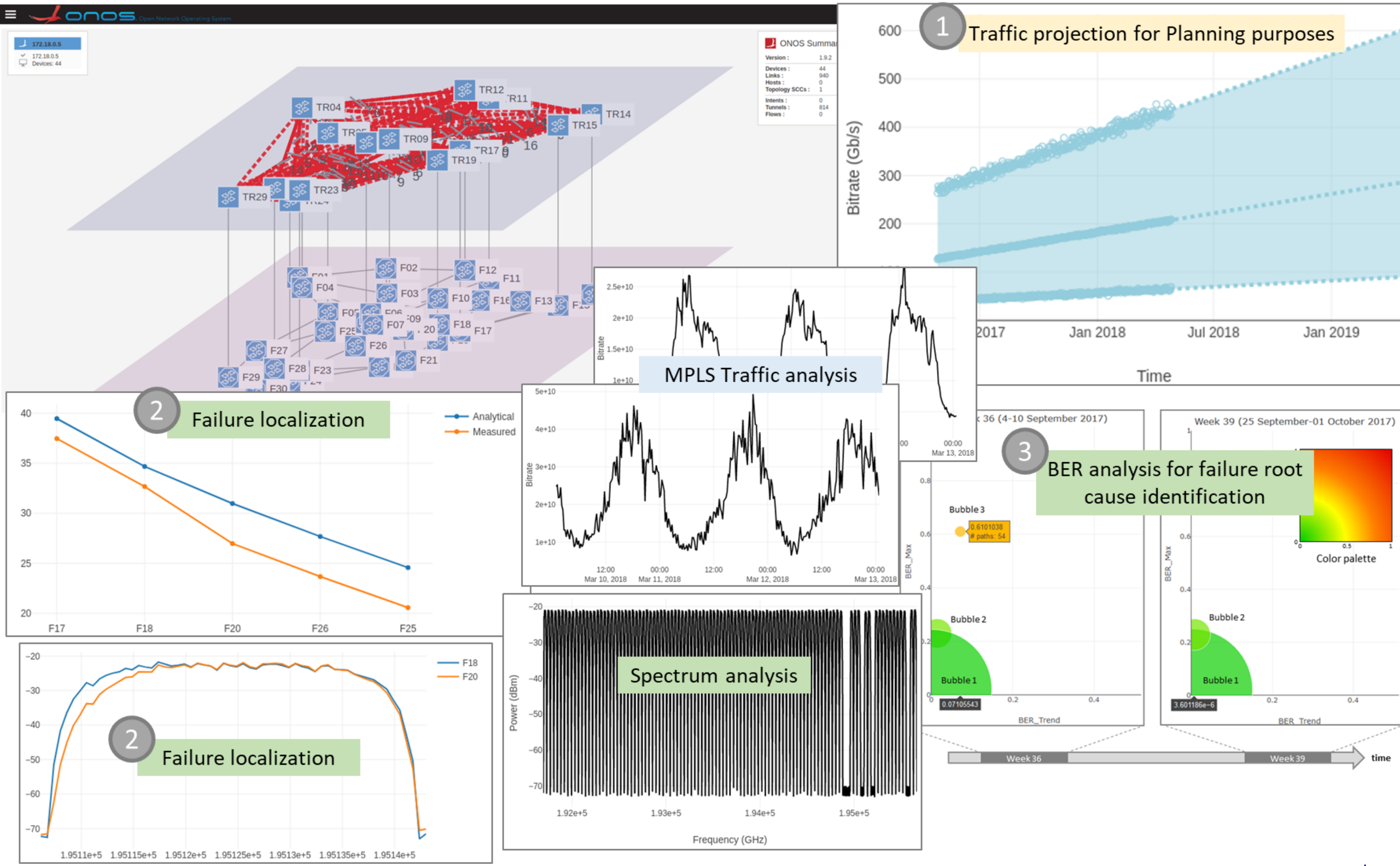
Task-oriented Visualization



OFC2018 Demo Set-up

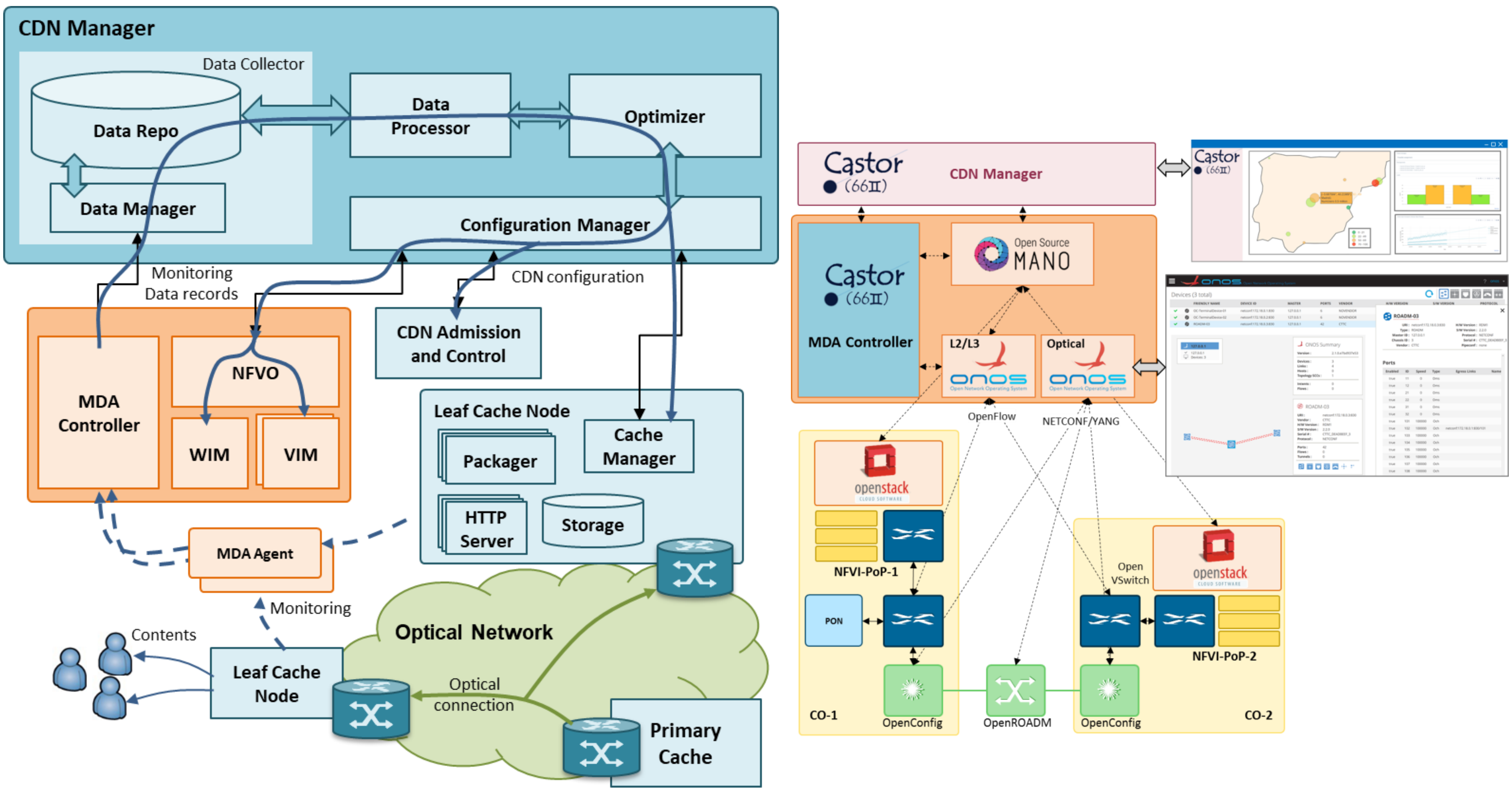


Examples



ONOS
1 Traffic projection for Planning purposes
Bitrate (Gb/s)
Time
2017
Jan 2018
Jul 2018
Jan 2019
MPLS Traffic analysis
Bitrate
Mar 10, 2018
Mar 11, 2018
Mar 12, 2018
Mar 13, 2018
2 Failure localization
Analytical
Measured
F17
F18
F20
F26
F25
2 Failure localization
F18
F20
Spectrum analysis
Power (dBm)
Frequency (GHz)
3 BER analysis for failure root cause identification
Week 36 (4-10 September 2017)
Week 39 (25 September-01 October 2017)
Bubble 3
Bubble 2
Bubble 1
Color palette
BER_Max
BER_Trend
Week 36
Week 39
time

NFV Support and Demo OFC 2019



ITU-T FG-ML5G-ARC5G

- Multiple sources of data to take advantage of correlations in data (ML-unify-001).
- Multiple technologies and network layers to achieve end-to-end user experience (ML-unify-002)
- Multi-level and distributed instantiation of the ML pipeline (ML-unify-003)
- ML-pipeline and network service will be as decoupled as possible (ML-unify-004)
- Flexible implementation in terms of splitting of logical nodes in the ML-pipeline (ML-unify-005), placement (ML-unify-008) (ML-unify-012), including third-parties (ML-unify-013), meeting the defined constraints (ML-unify-015), and chaining (ML-unify-020)
- Transferring data and trained ML models through Interface 8 (ML-unify-006, ML-unify-007)
- Heterogeneous Interfaces based on existing or extended interfaces (ML-unify-009)
- ML-ML for specifying the use case and translating such specifications into intents (ML-unify-010, ML-unify-011) including time constraints (ML-unify-014)
- Dynamic ML model selection based on data from the source (ML-unify-016)
- Defining sandboxes for model training and host simulators in the sandbox (ML-unify-017)
- Enabling control loops (ML-unify-018)
- Using MLFOs for monitoring and managing ML-pipelines (ML-unify-019)
- Plug-in/out data sources to a running ML-pipeline (ML-unify-021)
- Sharing data between nodes in the ML-pipeline (ML-unify-022)

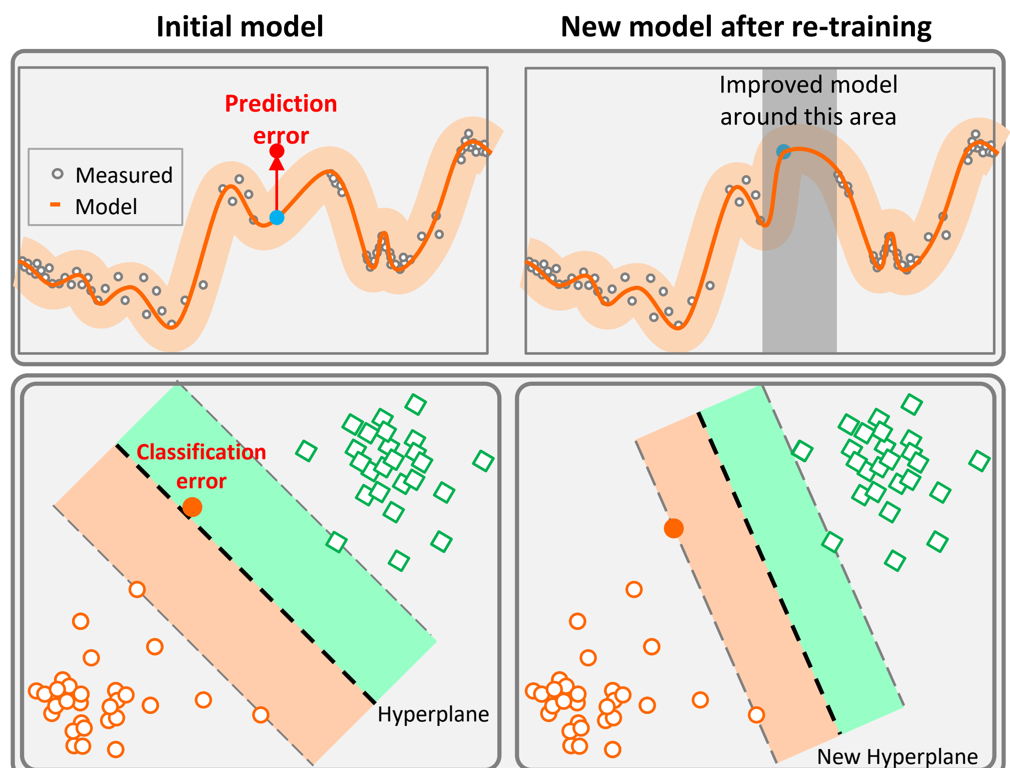
Collaborative Self-Learning

M. Ruiz, F. Tabatabaeimehr, and L. Velasco, "Knowledge Management in Optical Networks: Architecture, Methods and Use Cases [Invited]," IEEE/OSA Journal of Optical Communications and Networking, vol. 12, pp. A70-A81, 2020.

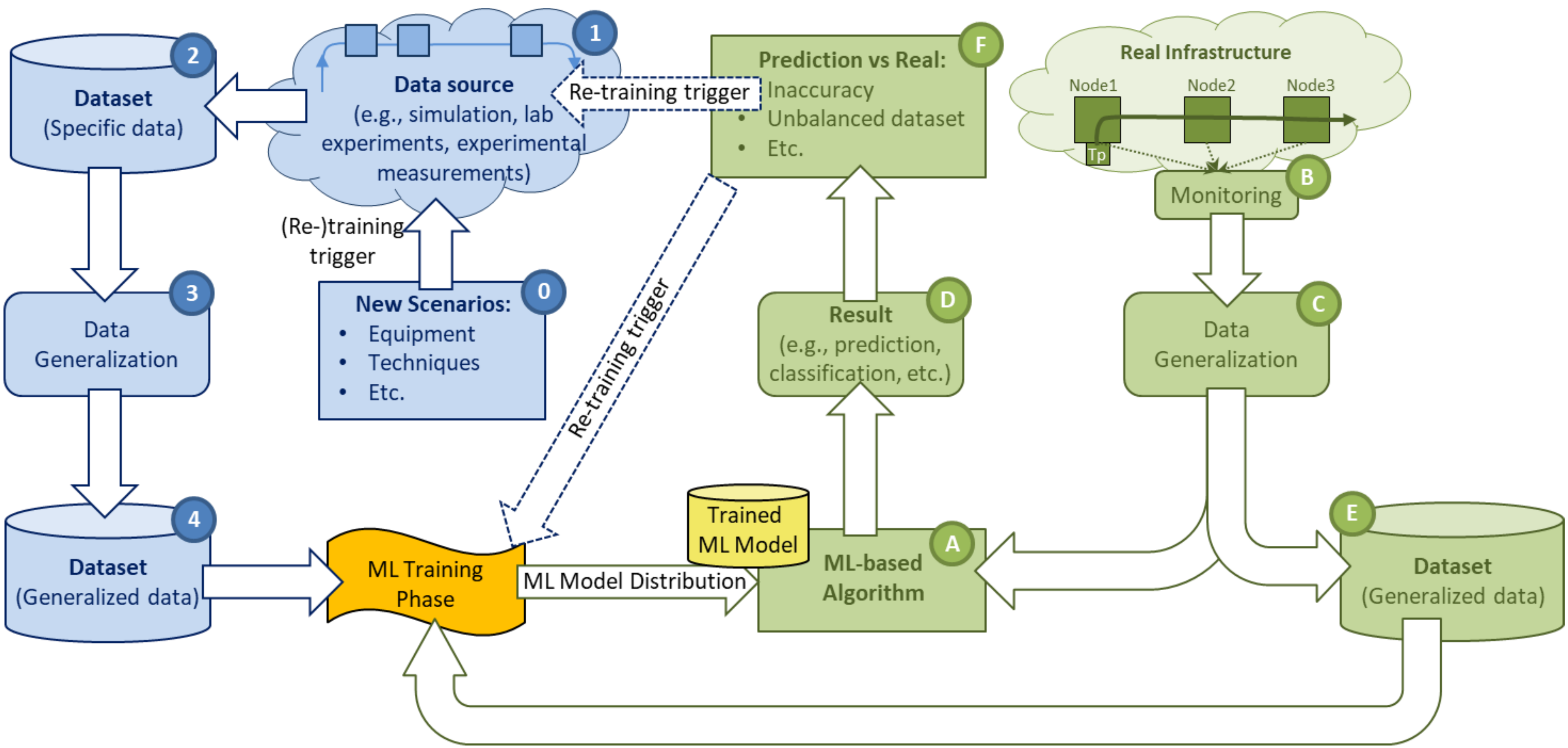
L. Velasco, B. Shariati, F. Boitier, P. Layec, and M. Ruiz, "A Learning Life-Cycle to Speed-up Autonomic Optical Transmission and Networking Adoption," in IEEE/OSA Journal of Optical Communications and Networking, vol. 11, pp. 226-237, 2019.

Motivation

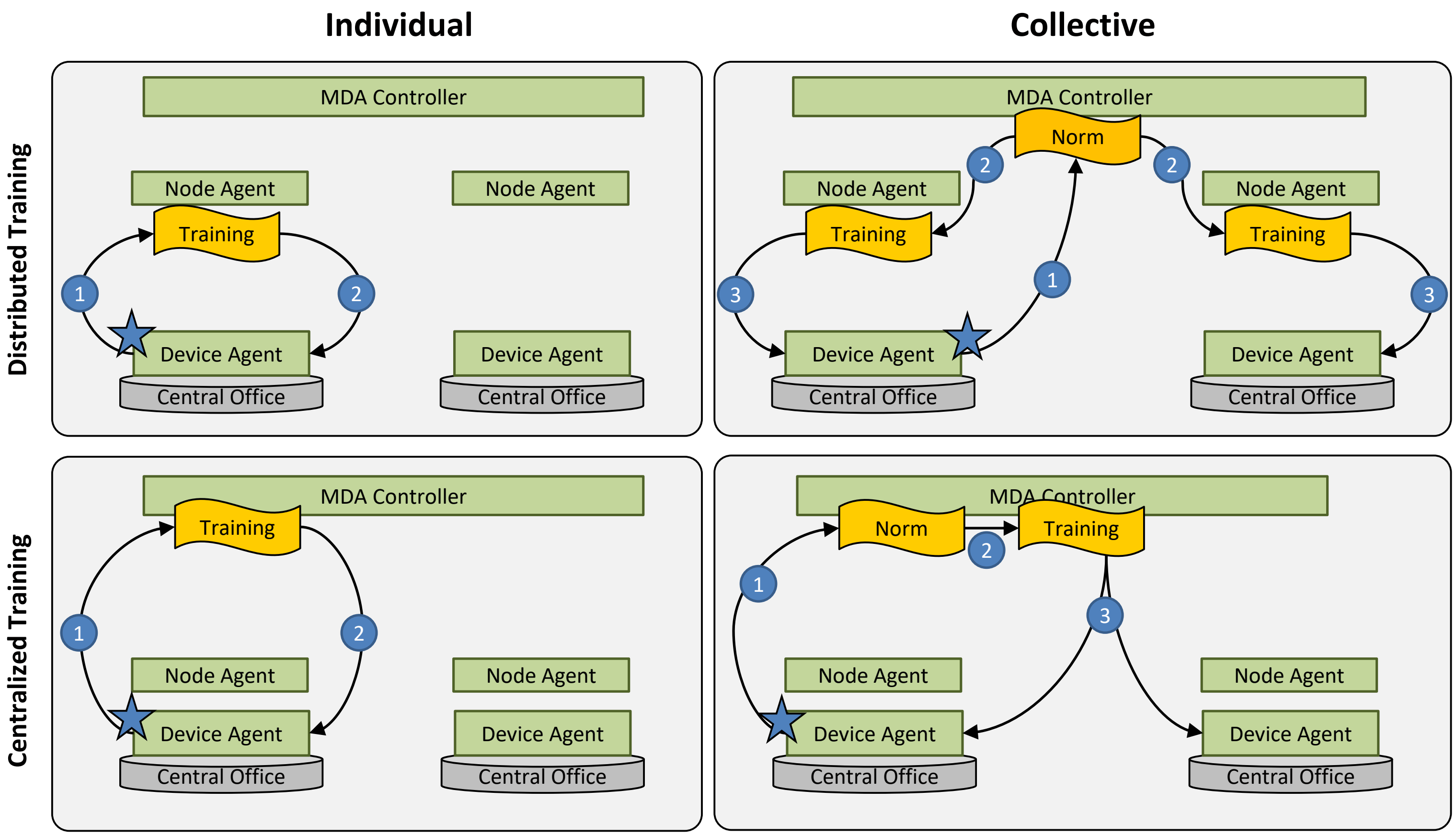
- **Autonomic operation** of optical transmission and networking requires from **ML-based algorithms**.
  - **ML models** have to be **trained** with datasets covering the **whole features space** to produce accurate ML models.
  - The **availability** of enough **data** is **rarely ensured**
  - Training datasets **cover just partially the features space**, hence reducing ML models accuracy.
- Datasets can be initially populated for ML training. Once models are generated, **ML re-training** can **improve** their **precision**.



ML-based algorithm life-cycle



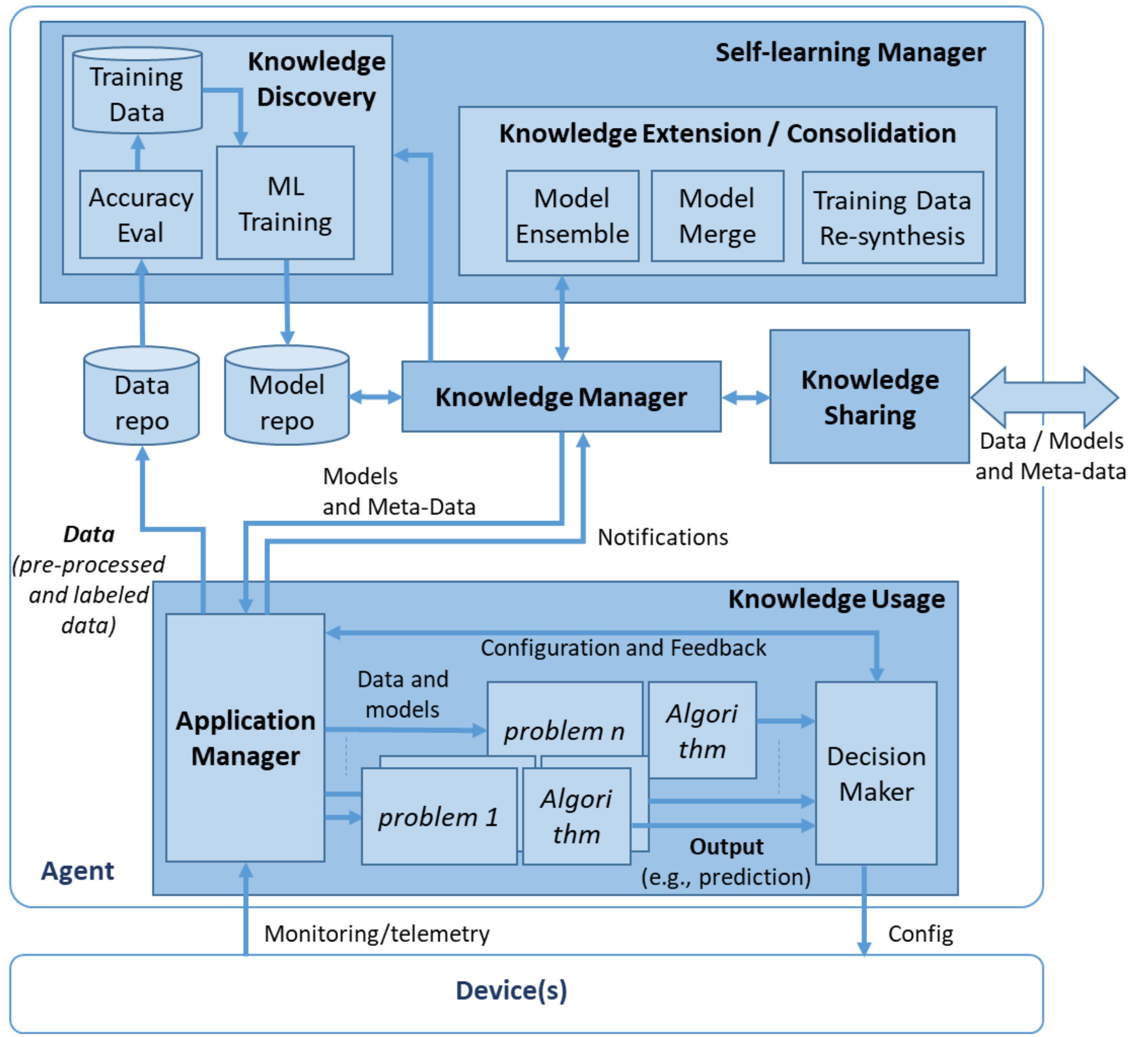
Self-learning



- **Collaborative** self-learning outperforms **individual** strategies…
- …. at the cost of **increasing** both **data** to be **exchanged**.

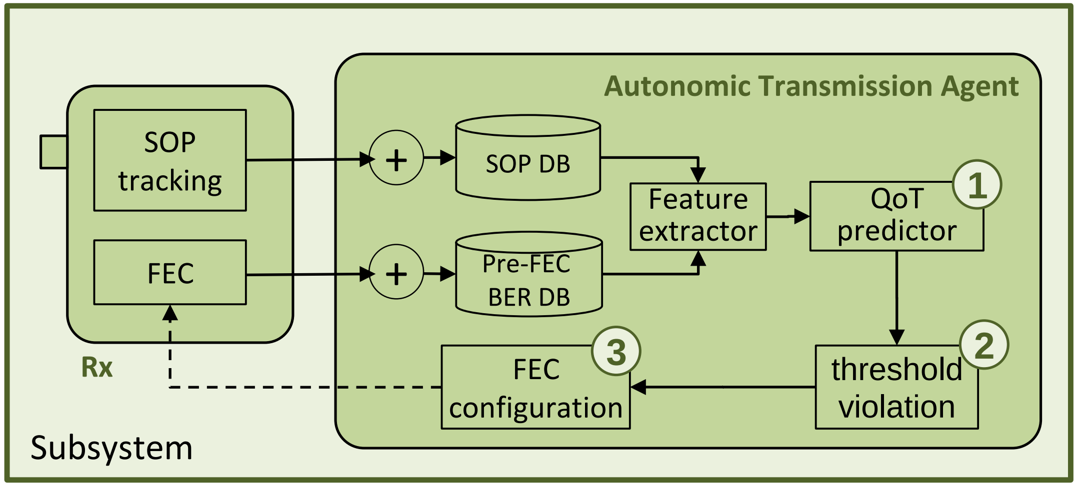
Collective Self-Learning based on Sharing Models

- Collective self-learning based on ML model sharing and combination can **reduce the amount of data** being shared among agents.
- The architecture allows a wide range of ML model combination alternatives
  - **ML training** can be executed either in the **node agent** or in the **controller**
  - **Once trained**, ML models can be **deployed** to the **device agents**.



The Autonomic Transmission use case

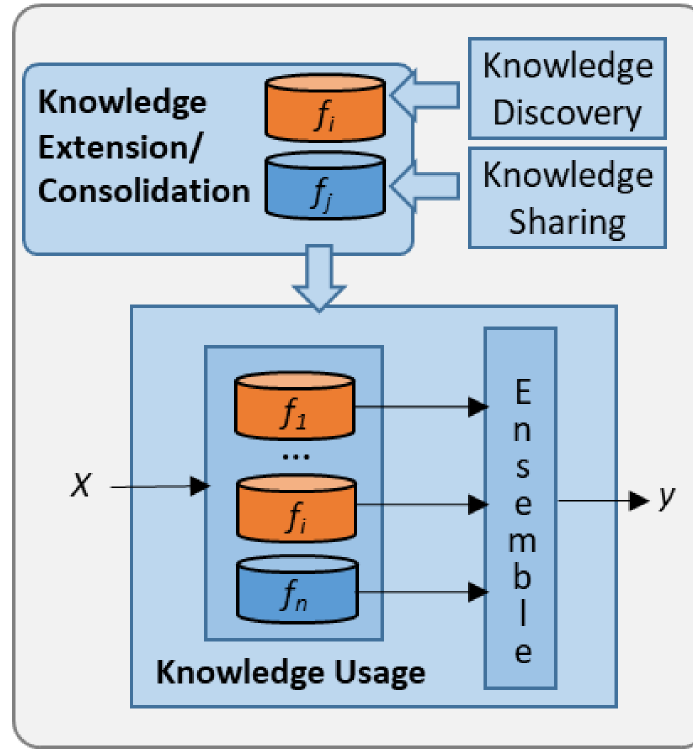
- We assume scenarios where **low-resolution ADC** are used.
  - The **evolution** of the **SOP** and the **pre-FEC BER** (last *w* samples) are used for **dynamic receiver configuration** and anticipating **BER degradation**.
- Three different ML-based problems need to be continuously solved:
  1) a **regression** model that estimates future pre-FEC BER
  2) a **probabilistic estimator** of the chance of **violating** a given BER threshold
  3) a **classifier** to determine e.g., the number of iterations of soft-decision FEC to run.



- **Combination** methods will be applied to each of the problems.

Model Ensemble

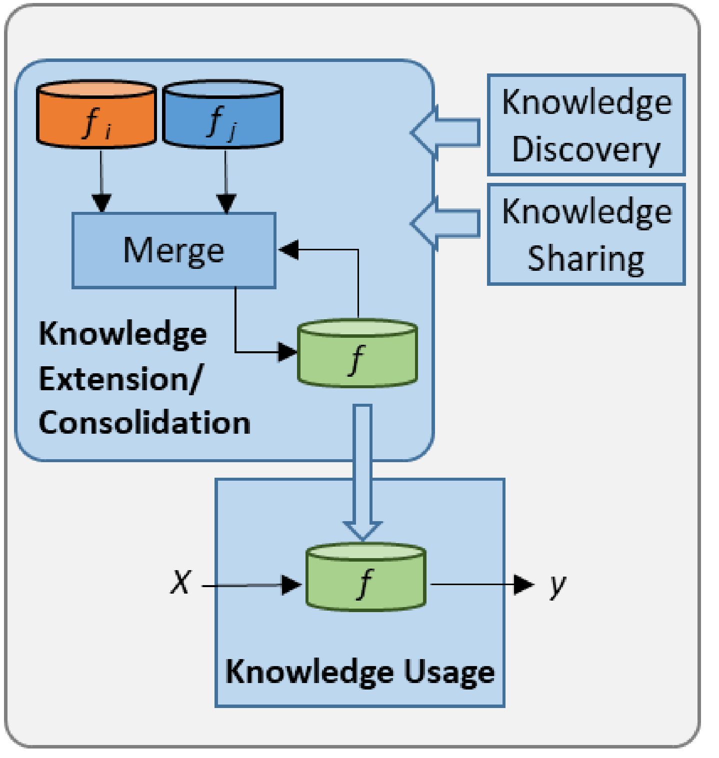
- It combines the prediction of several individual ML models and returns one single output.
  - Meta-data can include weights and/or the features range observed during model training.
- It can be applied to **any ML technique** or a **combination** of them.
- It requires **low computational** effort to apply collective self-learning, and **small additional storage** for individual ML model's persistency.



- We apply model ensemble for the local **receiver configuration classifier**, implemented as an SVM.
  - the individual ML models can be seen as **weak classifiers** that are combined into a **strong** (accurate) classifier.

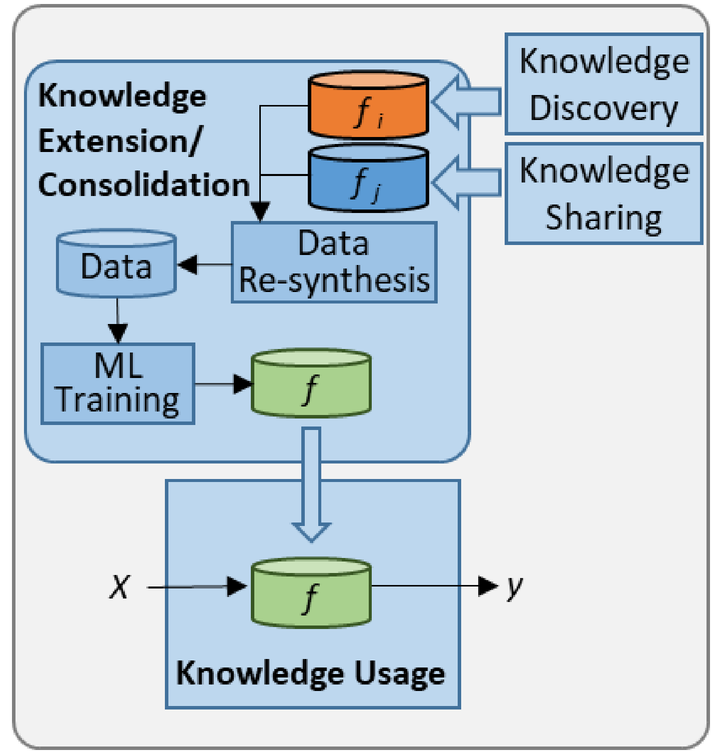
Model Merge

- It merges individual ML models to obtain **one single enhanced model**.
- It can provide **benefits** when model **parameters** can be **partially updated** without affecting the robustness and accuracy of the non-updated part.
- We use model merging in the probabilistic **BER threshold violation** estimator that it is based on a **DT**.
  - Model merge can enhance the ML model, e.g., **by updating probabilities of leaf nodes** and/or **extending a new sub-tree** (branching) from a leaf node.



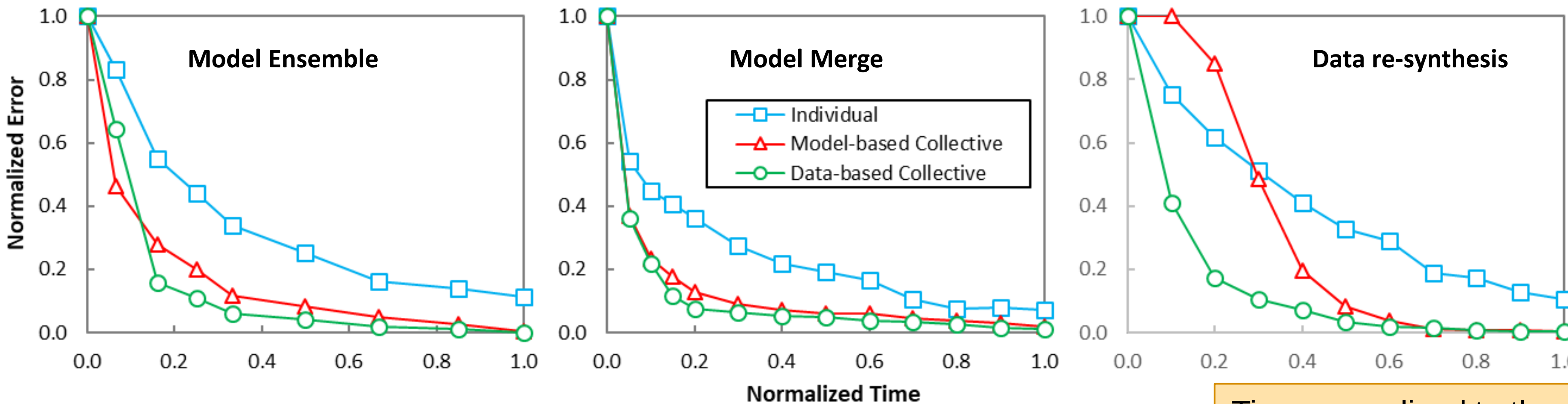
Retraining with data re-synthesis

- It consists in **generating the response** from the shared **individual ML models** in the given features range to obtain a **synthetic training dataset** from which a new ML model is trained.
- The local data re-synthesis from ML models **avoids exchanging** large amounts of monitored **data** among nodes and/or to the controller.
- We use re-training with data re-synthesis for the ANN-based pre-FEC BER estimator.



Prediction Error vs Time

Prediction error normalized to the error of the initial trained models.

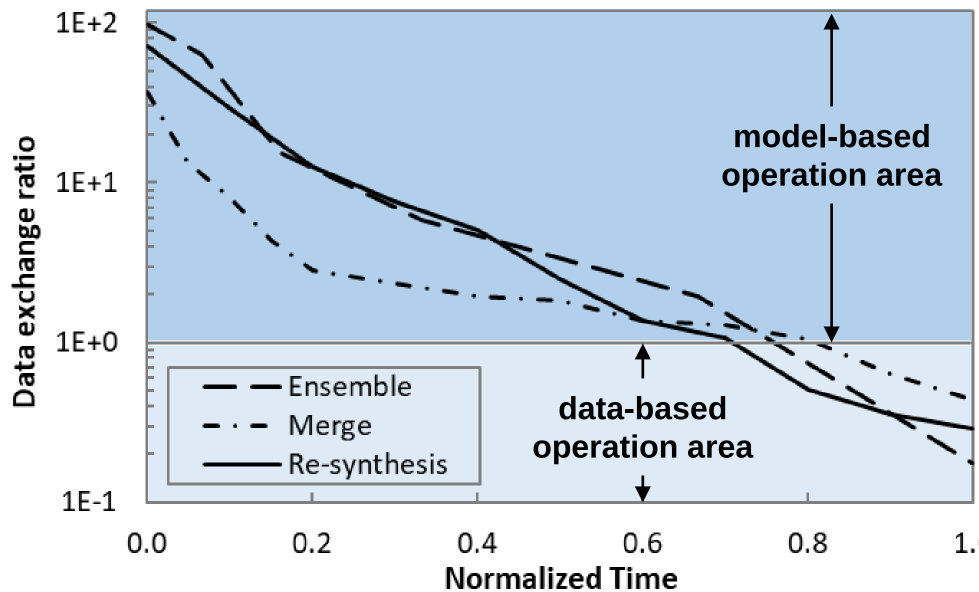


Time normalized to the time when the most accurate approach reached error <0.5%.

- ❖ **Collective self-learning** clearly **over-performs the individual approach** and gets **noticeably close** to that of the **data-based collective** one.
- ❖ **Exception**: during the early stage of operation in the case of data re-synthesis.
  - For data re-synthesis, **ML models** should **start being sharing** once they provide **moderated accuracy**.

Data exchanged comparison

Data exchange ratio:
amount of **data exchanged** under the **data-based approach** (data for training) **over** that under the **model-based** one (model's parameters and meta-data).



- The model-based approach requires **far less data exchange** to converge to accurate models (error <1%).
  - When error < 1%, exchanging just data is the best approach.

- Absolute estimated size of data exchange of a single device during the very first stage of operation.

	Ensemble	Merge	Re-synthesis
Data-based	220 KB	200 KB	1,510 KB
Model-based	2.3 KB	5.5 KB	21 KB

  - In a deployment with **hundreds of de**vices exchanging data and under **dynamic conditions**, the model-based collective self-learning approach would reduce the requirements of the control plane DCN.

Conclusions

- B5G requires radical changes to networks
- **Resilient networking for supporting improved and elastic reliability**
  - **Smart degradation detection**, predicting and resisting negative effects before they happen and/or rapidly recovering if negative effects cannot be avoided
  - **Elastic network technologies**, e.g., support for changing network topologies at runtime
- **AI/ML-powered adaptative network operations**
  - **Intent-based networking**, monitoring and reacting in real time to changing network conditions
  - **Closed-loop AI/ML mechanisms** to make automatically actionable decisions
  - **Proactive** (rather than reactive) resource allocation decisions
  - **Decision modules** as software control elements realizing an **adaptive control** over the network resources

UNIVERSITAT POLITÈCNICA DE CATALUNYA

Optical Communications Group

Thank you!

Optical Network Automation

Luis Velasco
lvelasco@ac.upc.edu