Telefónica

Data Flow Aggregation for Smarter Network Security

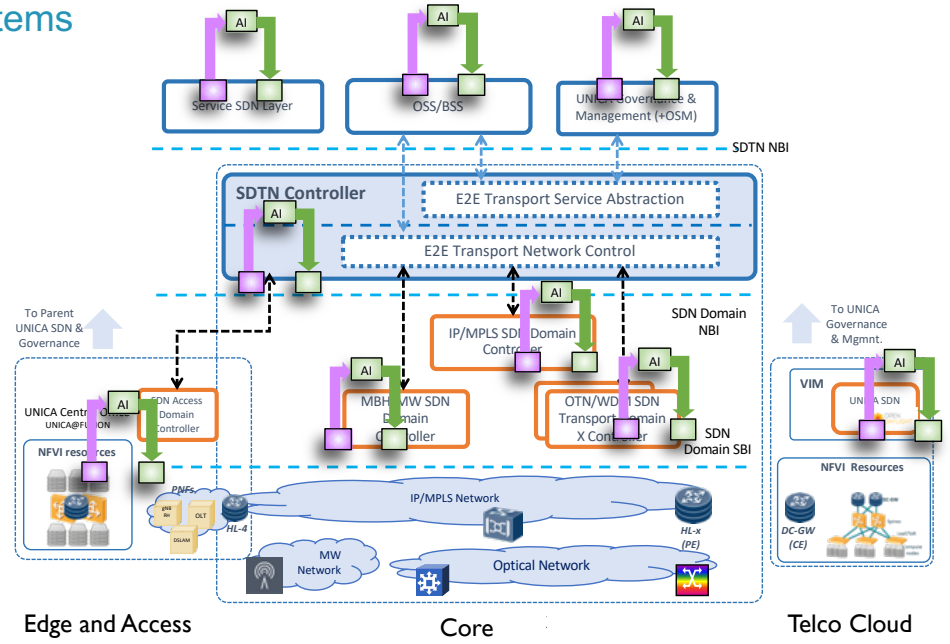
Diego R. López
Telefónica

The Autonomous Security Aspiration

- Security is one of the key areas for applying Autonomous Network technologies
- Based on a definition of policies and goals
  - Security intents
  - Policy enforcement and propagation
- Detect and deter miscreants
  - Dynamic identification
  - Adaptable response
- Based on closed loop technologies
  - Around for a long time
  - With AI to derive further insights and improve policy mapping

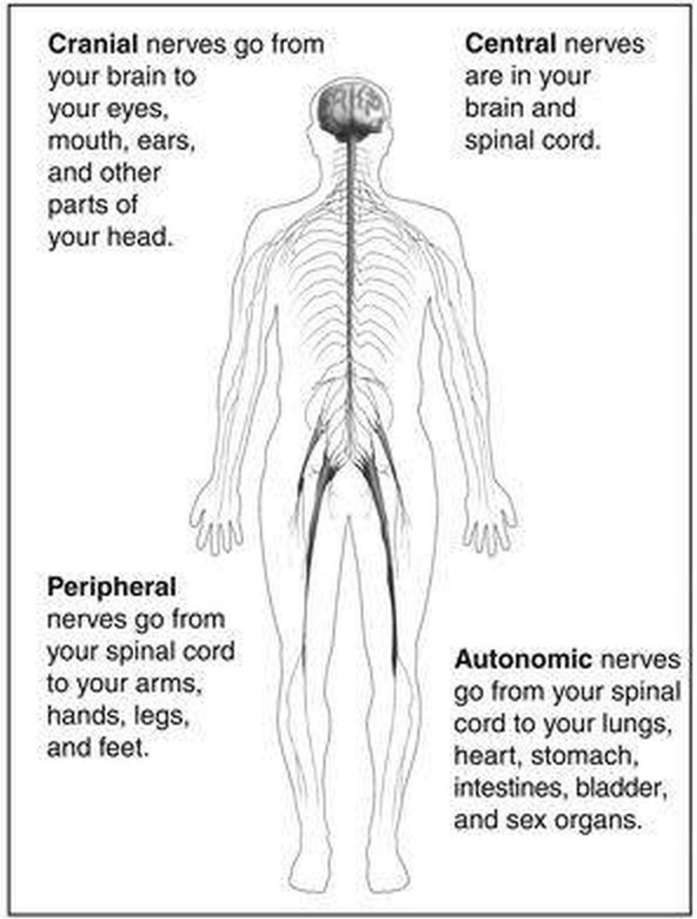
The Architectural Mapping

- Networks are critical and naturally distributed systems
  - A distributed AI for managing them
- The nature of distribution
  - Aggregation of knowledge
  - Accumulation of decisions
  - Cooperative vs independent vs selfish
  - Fixed vs mobile vs roaming
- Protocols
  - Specific knowledge and policy exchanges
  - Reuse stream mechanisms
  - Apply good-ole BGP and others of its kin
- Topologies
  - The mapping on the network topology
  - Depth and breadth
  - Nervous system approaches



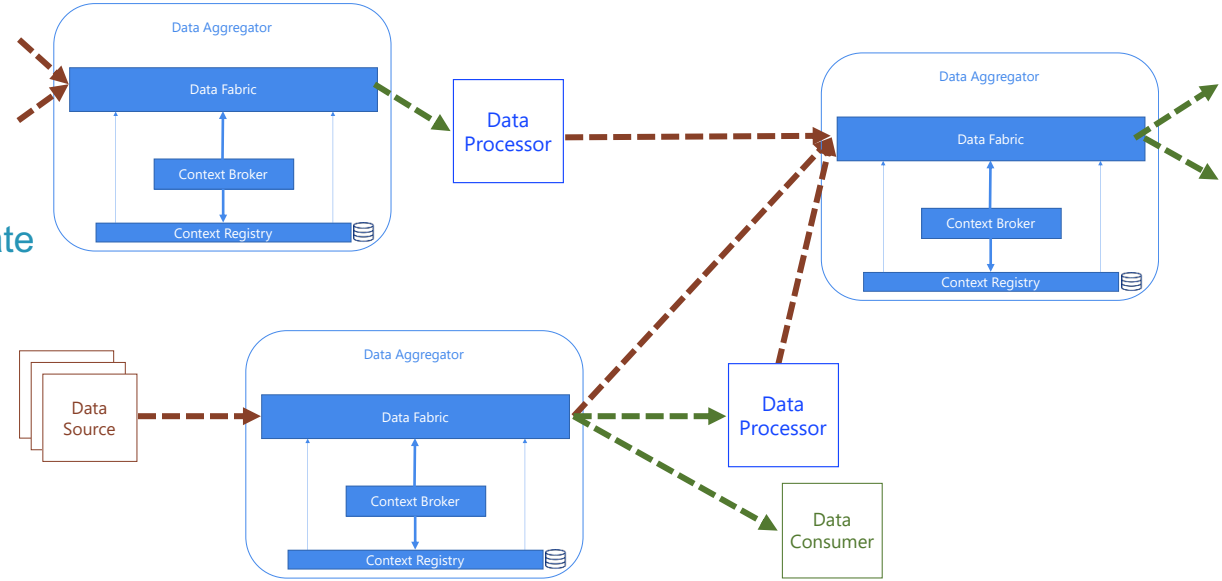
The Nervous System Paradigm

- Combine distributed architectures and holistic approaches
- Local loops
  - Detailed analysis
  - Fast response
  - Dynamic deployment
- Central loop(s)
  - Cumulated analysis
  - Integral view
  - Explainabiliity
  - Local loop orchestration
- All using a common impulse for all kind of interactions
  - Central elements receive and process aggregate information
  - A common data infrastructure for forwarding and aggregation

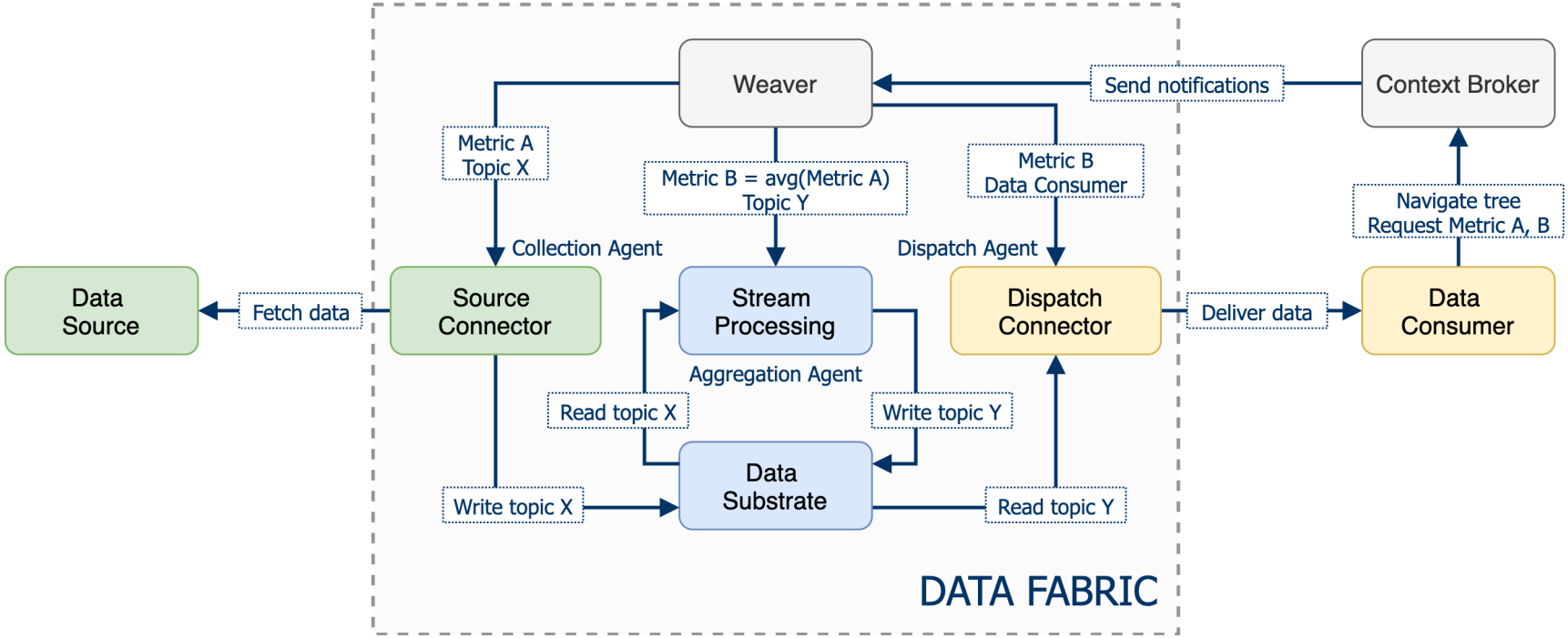


Building the Nervous Data Infrastructure

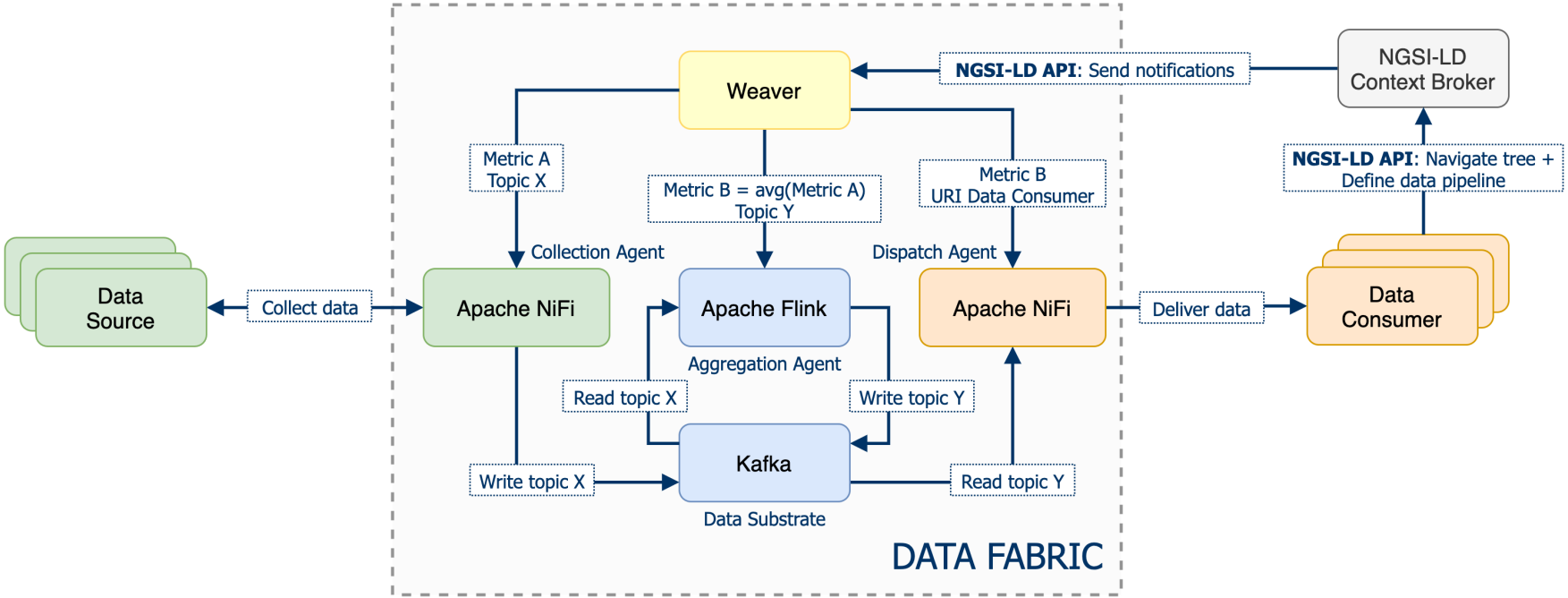
- Rely on aggregation nodes
  - Sources feed data
  - Consumers receive them
  - Aggregators map and integrate
- Based on metadata
  - Dynamic composition
  - Transport protocol agnostic
  - Telemetry data models
  - Knowledge ontologies
- Compositional patterns
  - Any element can play any role
  - AI / ML supported anywhere



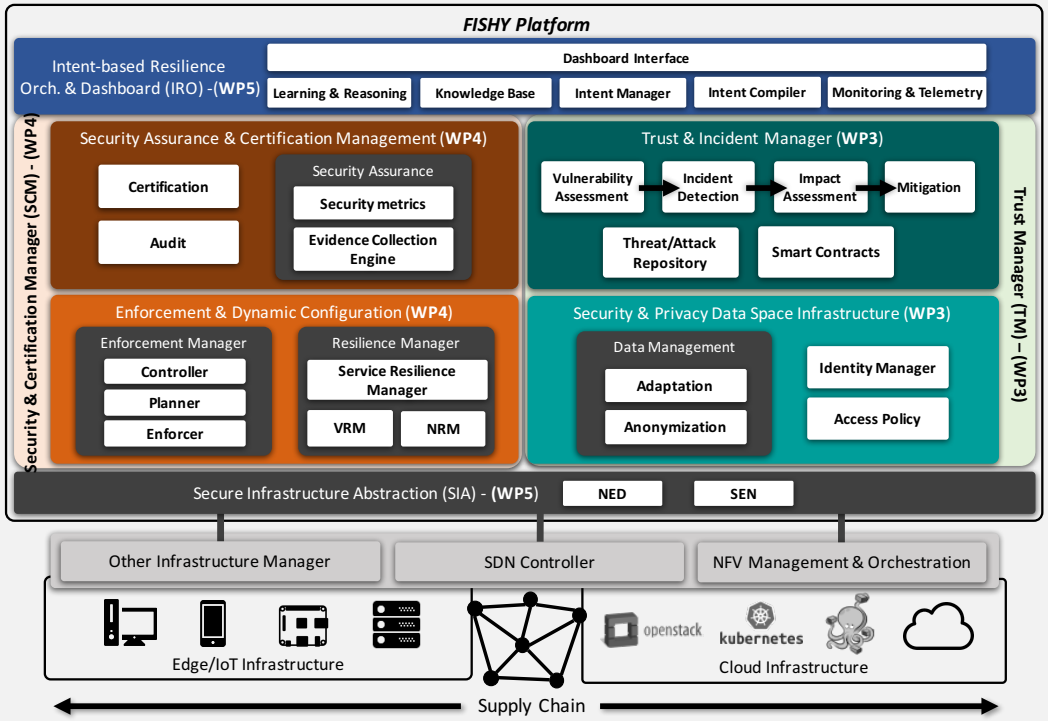
The Core Element: 5Growth Data Aggregator



Implementing the 5Growth Data Aggregator



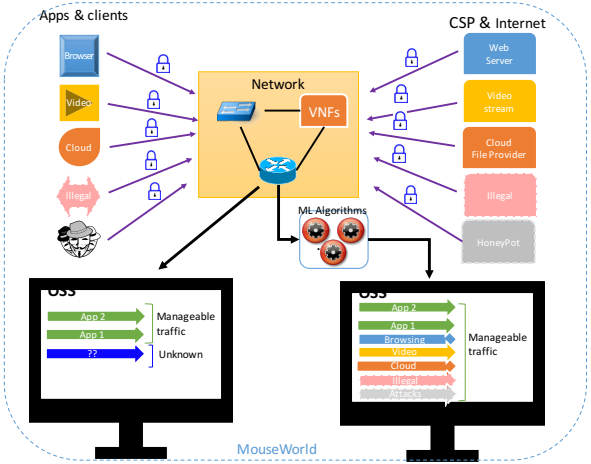
ICT Supply Chain Security: The FISHY Architecture



- Data consumers and processors
  - Intent-based Orchestrator and Dashboard
  - Assurance and Dynamic Configuration
  - Trust and Incident Management
- Two layers of data aggregators
  - Security Data Space Infrastructure
  - Secure infrastructure Abstraction (SIA)

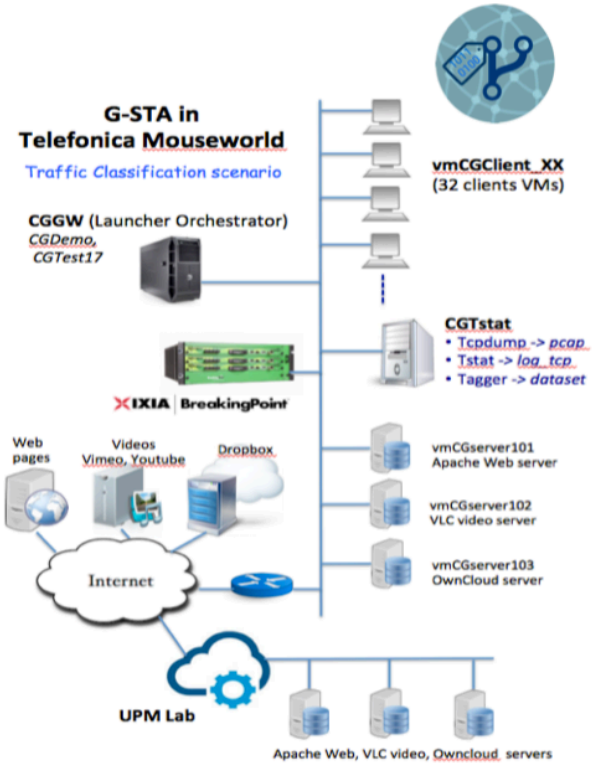
Trustworthy Datasets

- A serious lack of usable datasets
  - For training or validation
  - Data as an asset
  - Privacy concerns
  - None or limited tagging
- Generation of synthetic datasets
  - Traffic samples generated in a controlled way
  - Configurable mixes of synthetic and real traffic
- And metadata management
  - Different scenarios, from high loads to security threats
  - Training and validation loops
- Rely on the Data Infrastructure
  - Repeatability and reproducibility
  - Controlled variations



The *Mouseworld* – Synthetic Traffic and Beyond

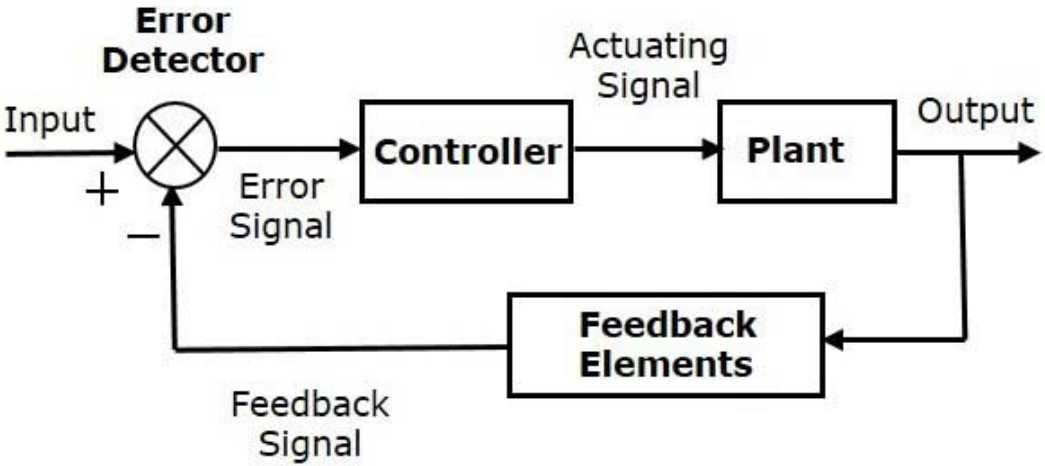
INSPIRE-5Gplus



- Traffic at all network segments
- Clients, servers, middleboxes and network functions of many natures
  - Plus raw traffic captures and other external sources
- Traffic analysis to produce (labelled) datasets
  - Flow aggregation and composition
- Train and validate
  - ML solutions, supervised and unsupervised
  - Data-driven modules (AI, Analytics…)
- Repeatable and controlled conditions and variants
  - SDN/NFV
  - Data infrastructure orchestration

Making the Autonomous Closed-Loop Feasible

- Signal flows in network architectures
  - Matching controllers to plants
- Beware the network differential facts
  - Topology (and geometry!) awareness
  - The conservation principle
  - Openness
  - Integrity and auditability
  - Isolation
- A few steps ahead
  - Metadata distribution
  - Aggregator orchestration
  - The application to the other kind of signals (an *Action Infrastructure*?)



Acknowledgment:
The research conducted by INSPIRE-5Gplus receives funding from the European Commission H2020 programme under Grant Agreement Nº 871808. The European Commission has no responsibility for the content of this presentation.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 856709.

This project has received funding from the European Union's H2020 research and innovation programme under the grant agreement No. 952644