Telefónica

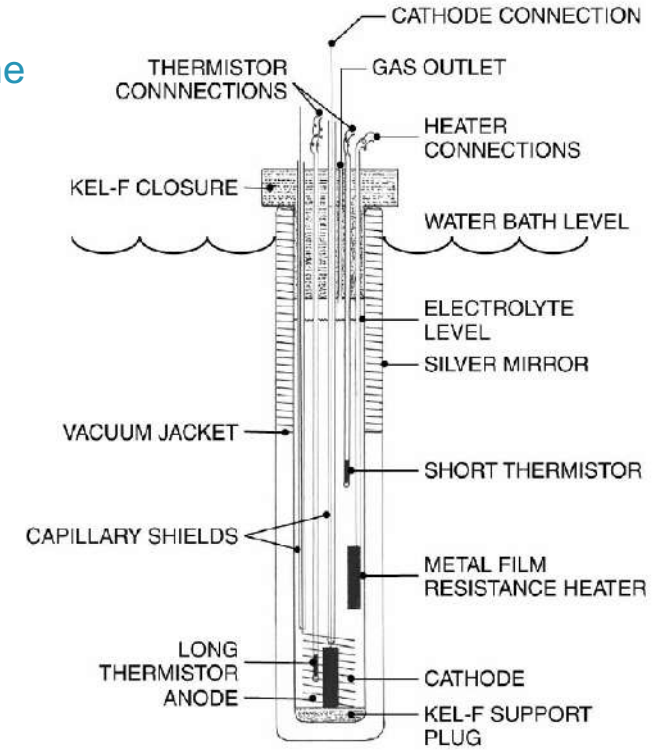
Serious Science, Serious Engineering

The Way of Software-Based Network Experimentation

Diego R. López
Telefónica

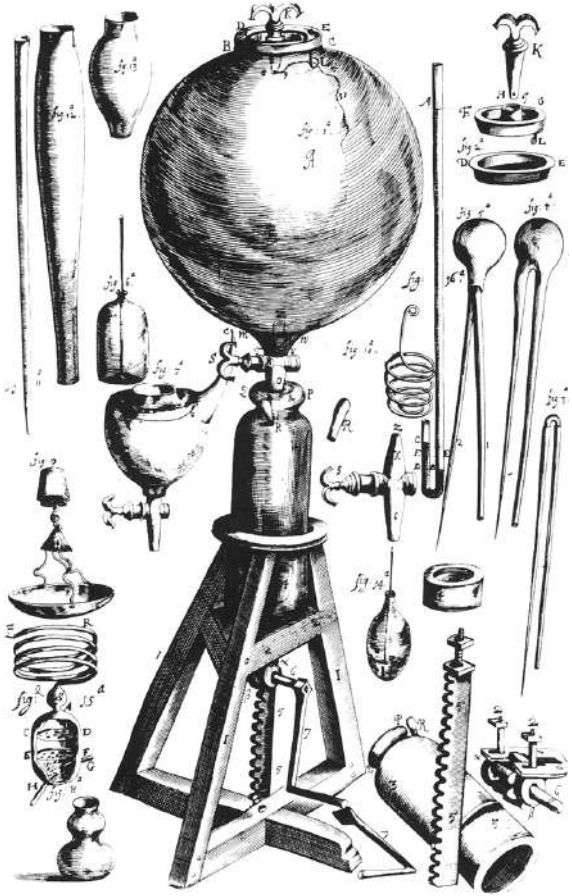
Making Serious Science (and Engineering)

- Independent verification and *reproducibility* are essential to the scientific method
  - *“Non-reproducible single occurrences are of no significance to science” (K. Popper)*
- Complicated in many cases because different reasons
  - Ethical
  - Nature of the research field
- Recent computing and network results
  - Complexity
  - Disparate conditions



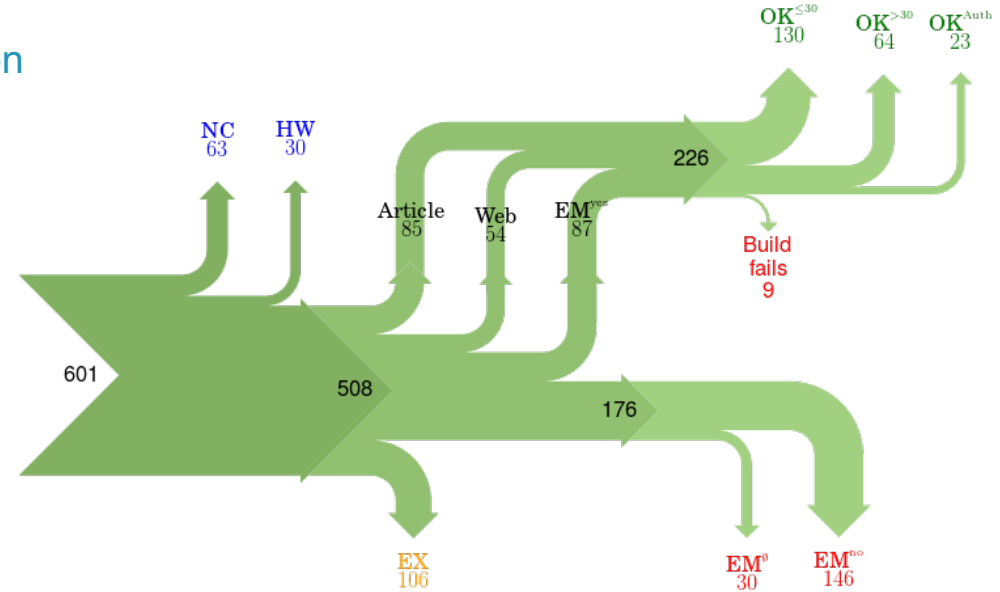
The Essential Goals

- Corroboration
  - Avoid (un)intentional idiosyncratic results
  - Repeatable results
- Transparency
  - Avoid (un)intentional biases
  - Environment and measurements
  - Repeatable methods
- Robustness
  - Avoid (un)intentional best-of-breed results
  - Repeatable causes



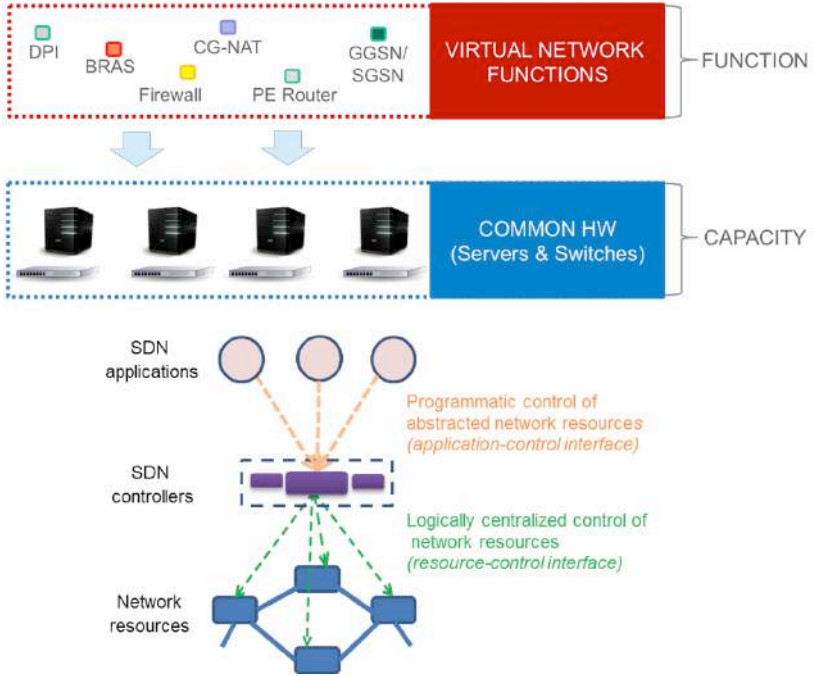
The ICT Front

- Reproducible result reports are not common
  - Complex environments
  - Incomplete descriptions
  - Too-focused measurements
  - Limited variations
- This is not just an academic issue
  - Technology evaluation
  - Strategic planning
  - Multi-domain environments
  - While shortening lifecycles



http://repeatability.cs.arizona.edu/

How Software Networks Come to Help



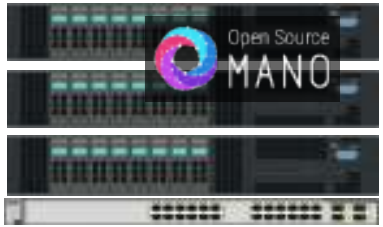
- More regular infrastructure
  - Easier to match environmental parameters
  - Homogeneous measurement points
  - On-demand measures
- Model-based approaches
  - Consistent descriptions
  - Scalable verification
  - For environments and experiments
- And, well, software based
  - Virtualization as a transportation guarantee
  - Open source

The OSM Experience

LOCAL DEVELOPMENT &TESTING

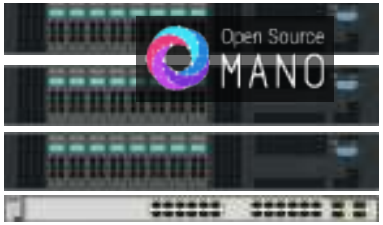
- Open development environment
- Functional tests
- Low cost
- Integration from the beginning

TEST POOL FOR DEVELOPERS



- Real servers and switches
- Performance tests (EPA can be enforced)
- Cost-effective shared infrastructure
- Move the value to VNF services

SERVICE PROVIDER

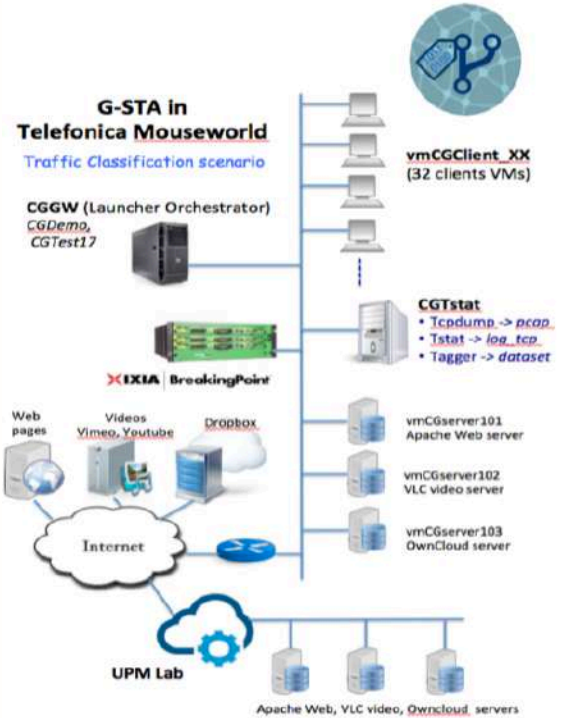


- Production/pre-production environment
- Real network scenarios
- Final service configuration
- Fast deployment
- Low final integration cost

- Development and testing
- Experiment description and sharing

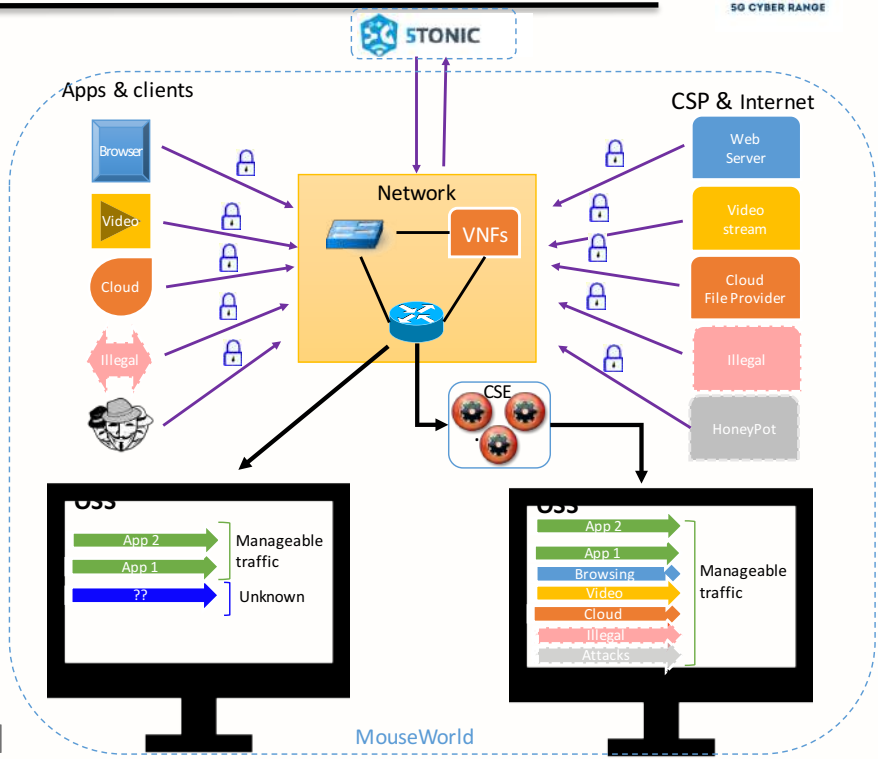
Models at Work: The *Mouseworld*

- A lab environment able to incorporate network applications, functions and topologies
  - Virtualized network functions
  - Physical network functions or their emulation counterparts
  - Traffic traces (at any plane) to be injected.
- Experiments defined by models
  - Descriptors to be combined
  - Boundary conditions
  - Pre- and post-processing of data
- Reproducibility enablers
  - Simple reconfiguration
  - Differential analysis of alternate scenarios



Applying the Mouseworld

- Originally conceived as an environment for security experimentation
  - Anomaly detection
  - Threat characterization
- Generation of sound, bespoke datasets
  - For any kind of network conditions
  - As result of general experiments
- Acknowledge privacy implications
  - Most of data is generated from synthetic traffic
  - Pre-processing allows full anonymization whenever required
  - Other techniques (such as GAN) being considered

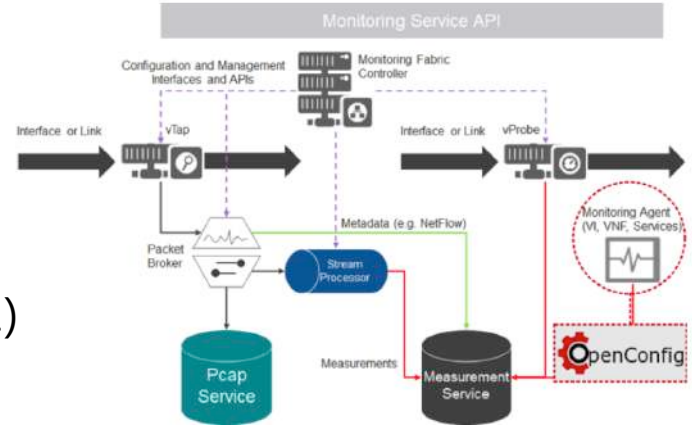


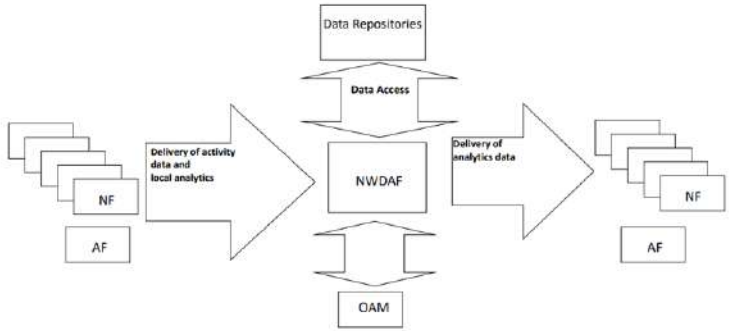
Expanding the Mouseworld

- AI training and validation
  - Labelled datasets for arbitrary conditions
  - Automating ML application
- Cyber-ranges
  - One-step deployment
  - Simpler replays, including variations
  - External source integration (traces, adversarial or support AI…)
- A core component for network digital twins
  - Though not totally there yet
  - Real-time data ingestion
  - Integration with online data analytics
  - Direct control actions

Data

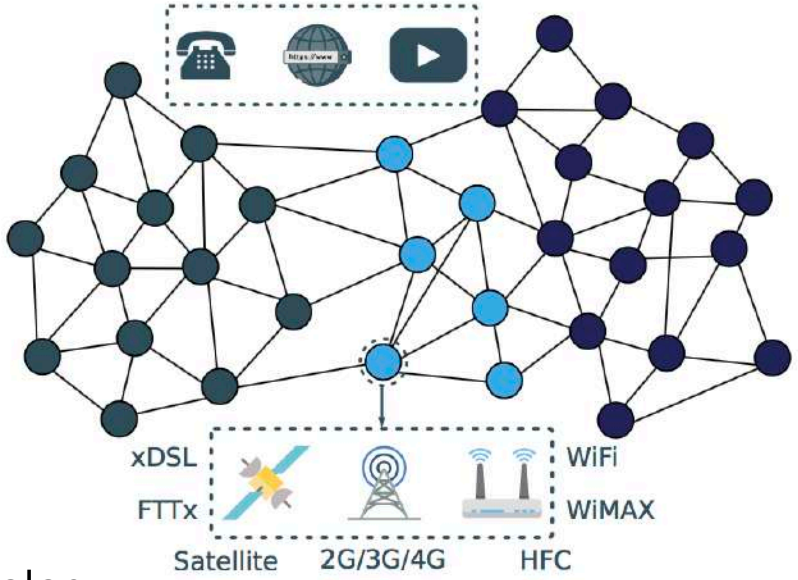
- Data is the main result of any experiment
  - Usable: Adaptation (formats, scales…)
  - Sufficient: Topology (sources, aggregators…)
  - Safe: Provenance (origin, timestamps…)
  - Steady: Continuity (pace, availability…)
- An enhanced data fabric seems the logical approach
  - Supporting resource, orchestration and function sources
  - Combining current network monitoring tools and recent telemetry developments



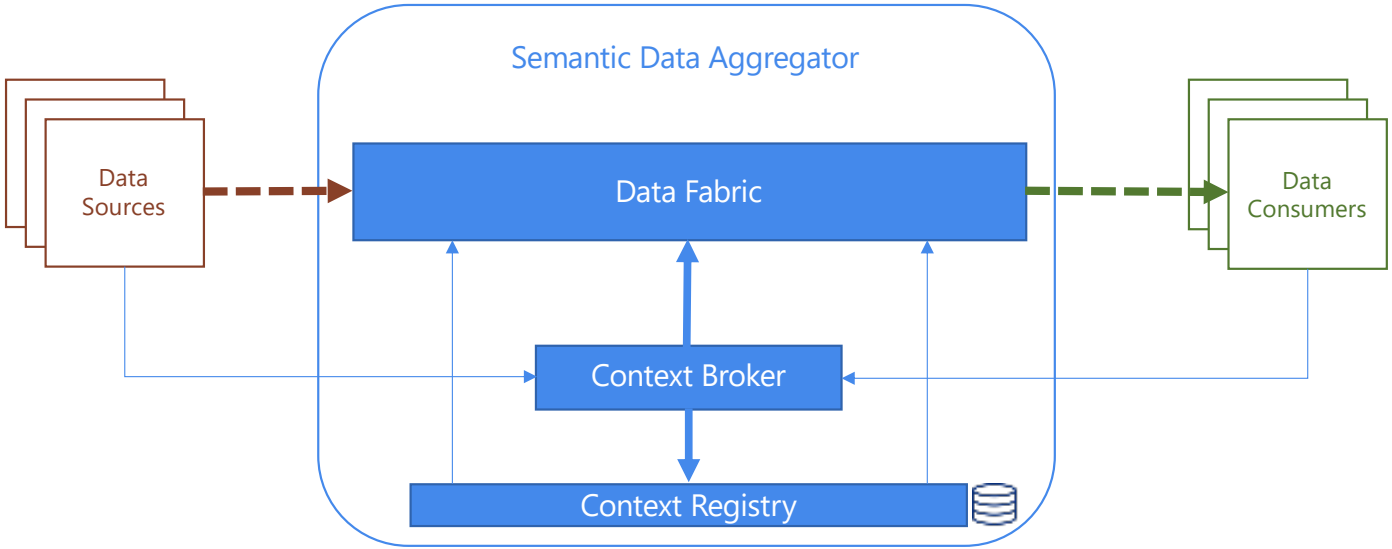


The Data Aggregation Scenario

- Support the integration of different data flows
  - Open
  - Automated
  - Secure
  - Scalable
- Deal with heterogeneity at all levels
  - Data producers and consumers
  - Data models
  - Deployment styles
  - Supporting infrastructures
- Not just data
  - Metadata becomes essential
  - What seems to claim for a data stream ontology



The Context Game

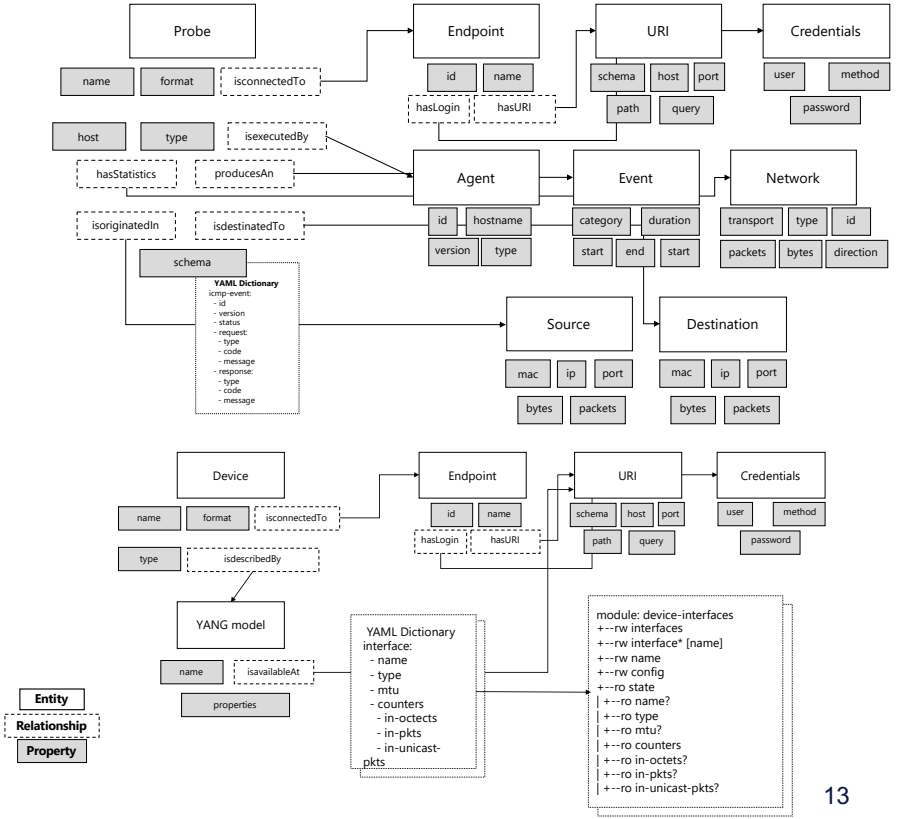


- Combine data flows
- Model-based
  - Sources and consumers
  - Integrated with descriptors

- A semantic metadata framework for telemetry data
  - Founded on the current results in data model space

Context Information Management

- Mechanisms to deal with context information
  - From many different sources
  - For many different consumers
- Shared through interoperable metadata publication platforms
- Based on an information model describing entities and relationships
  - For each class of data endpoint



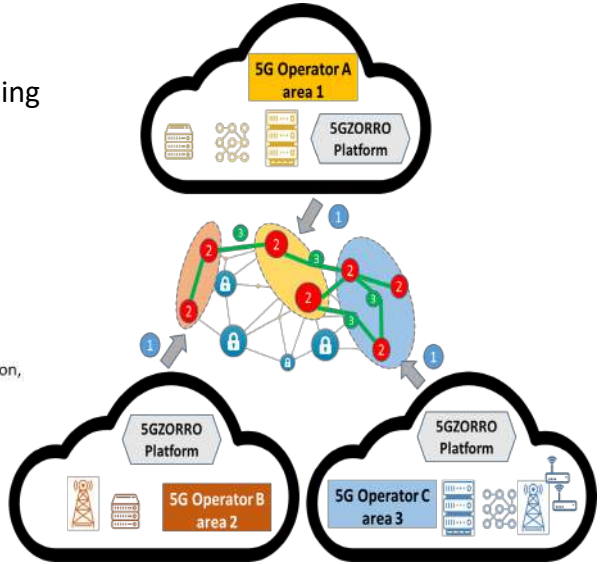
Trusted Data (and Metadata)

- *A common platform available at each participating domain*
  - Operational data lakes for service discovery, brokering and SLA monitoring
  - DLTs and Smart Contracts for auditability, licensing and disintermediated trading
  - Trusted Execution Environment to support trust without privacy loss
- *Applicable to support experimental repeatability*
  - Infrastructure selection
  - Applicable tool
  - Data verification
- *And to V&V (Validation and Verification)*
  - Testing is a kind of experiment

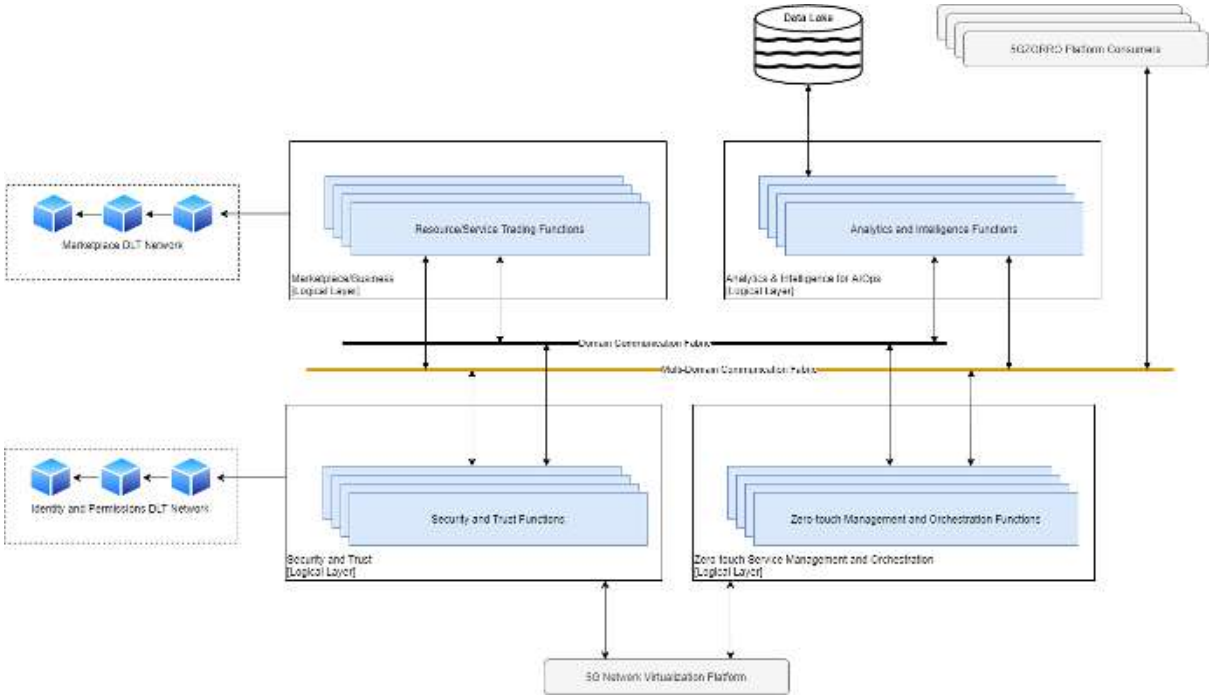
1 Zero Touch Resource Discovery using DLT/BC

2 Intelligent 3rd party resource selection, request and access/usage

3 Trust establishment among multi-parties



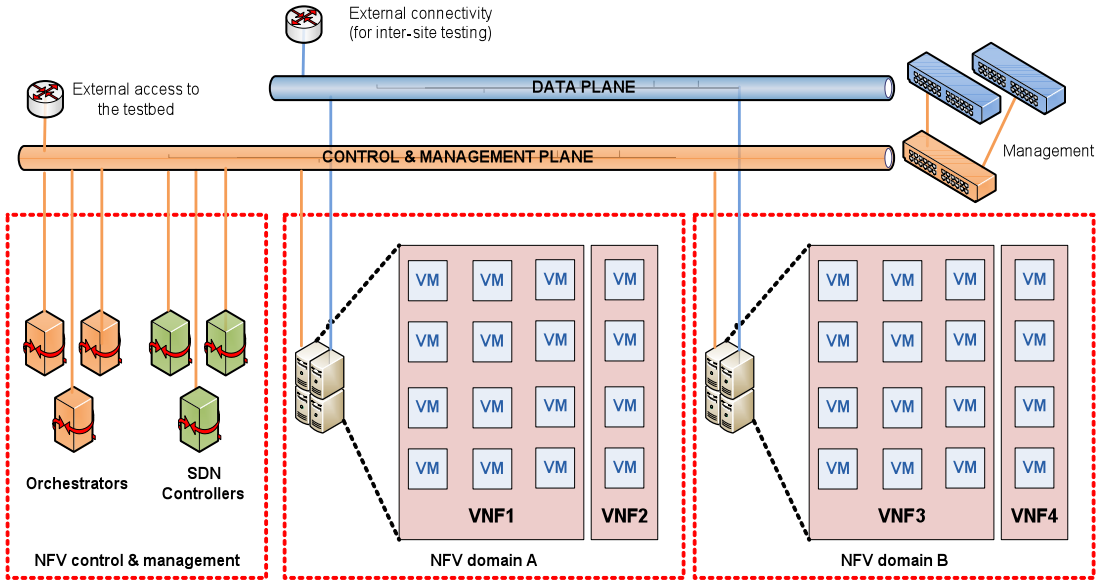
A Fabric for Multi-Domain Experiments



- *Different sources for infrastructure, components and tools*
  - Including licensing mechanisms
- *Shared data repositories*
  - Collaborative exploitation
  - Independent verification
  - Privacy preservation
- *Loosely-coupled, service-based architecture*
  - Intra- and inter-domain
  - Zero-touch operation

5TONIC

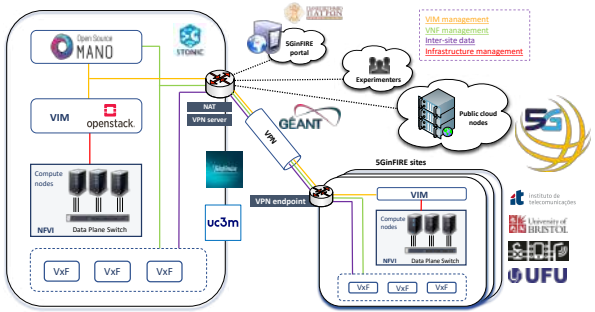
- An open research and innovation ecosystem
  - Focused on (beyond) 5G products and services
  - Created in 2015, based on the direct collaboration between Telefónica and IMDEA Networks
- Intended to become a hub for knowledge sharing and collaboration in the area of next-generation network technologies

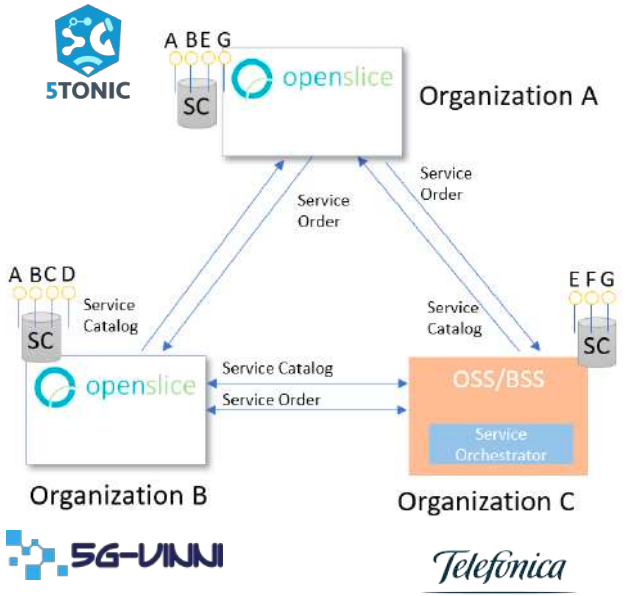


- 5TONIC infrastructure follows Software Network principles
  - The only exception is dedicated physical devices for access technologies: radio, fiber…

Going Multisite

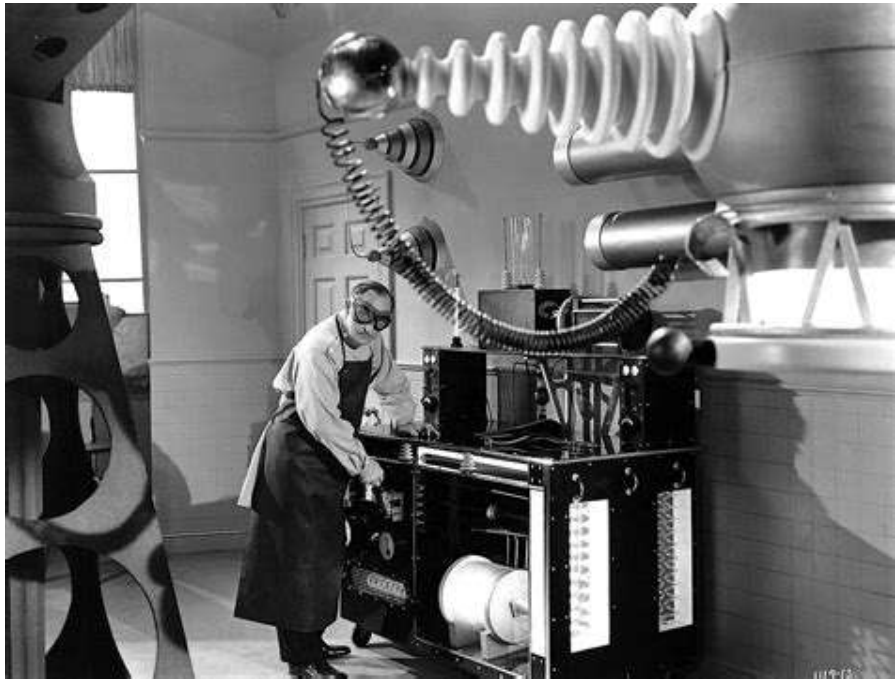
- Several projects focused on cooperative experimentation
  - Centralized: Single orchestrator and common infrastructure abstractions
  - Multi-domain: Federation schemas
- Learning interesting lessons
  - Security and trust issues, especially in multi-domain scenarios
  - Procedures for new participants
  - SDN in the WAN
  - Experimenter access
  - The SDN on NFV issues
- Model-based
  - Driven by experiment and service descriptors
  - Constructs for monitoring, measurement, telemetry...
  - Service continuity and model-based testing





Empowering Experimentation

- The Software Network opens several paths to improve network experimentation
  - And make it *real science* again
  - Avoid networking cold fusion
- Not only academic goals
  - Real engineering requires real science
- Towards model-driven experimentation

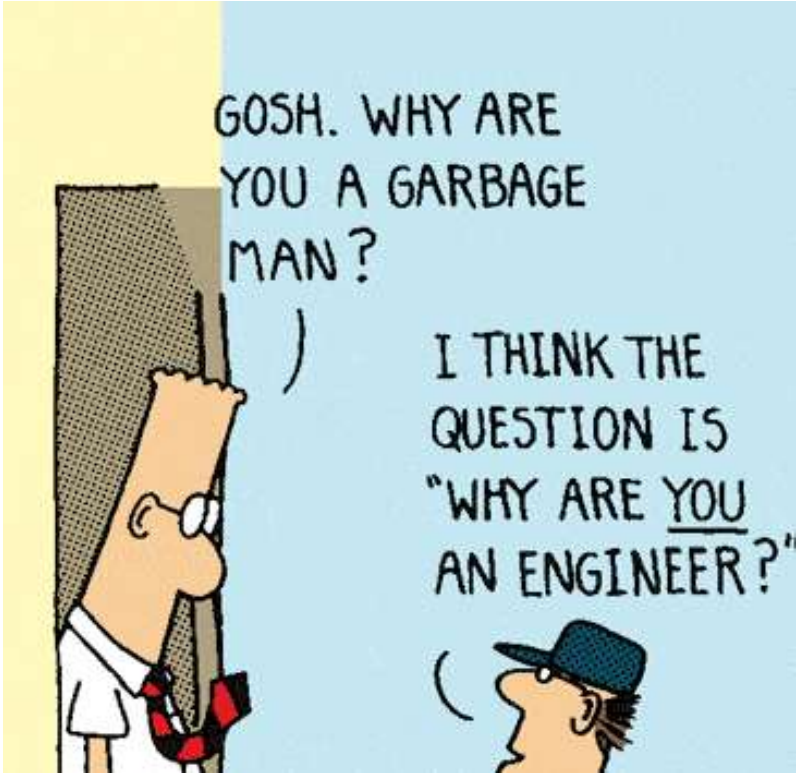


Better Answers to Difficult Questions

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

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.



5G PPP PUBLIC-PRIVATE PARTNERSHIP

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 871533

#RECONECTA

Telefónica