Self-driving Networks

Raouf Boutaba

David R. Cheriton School of Computer Science
University of Waterloo

KHU, Seoul, South Korea, May 16, 2019

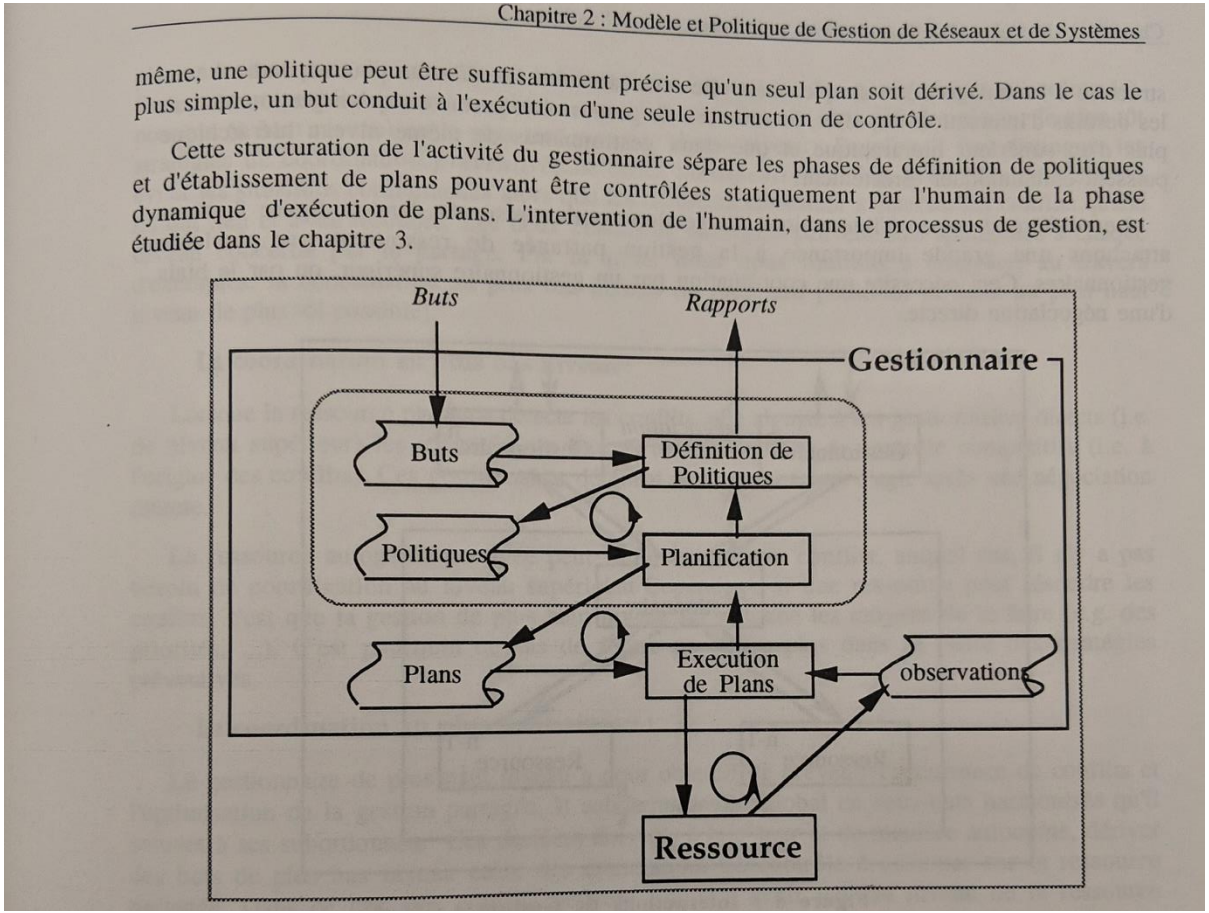
Outline

- Self-driving Networks
- A Walk Down the Memory Lane
- Recent Incarnations
- Can it Happen this Time ?
- Future Research Directions
- Take-away

Self-driving Networks

- What?
  - Networks capable to autonomously monitor their status, analyze problems, make decisions, and execute corrective actions with minimal to no human intervention.
- Why?
  - Traditionally network management has been human-centric
    - Costly, error-prone, and slow to adapt to changes
    - Cannot cope with the increasing complexity due to
      - Large number and diversity of network devices
      - Future application requirements, e.g., high-capacity, ultra-low latency, very high reliability, and massive connectivity.

A Walk Down the Memory Lane

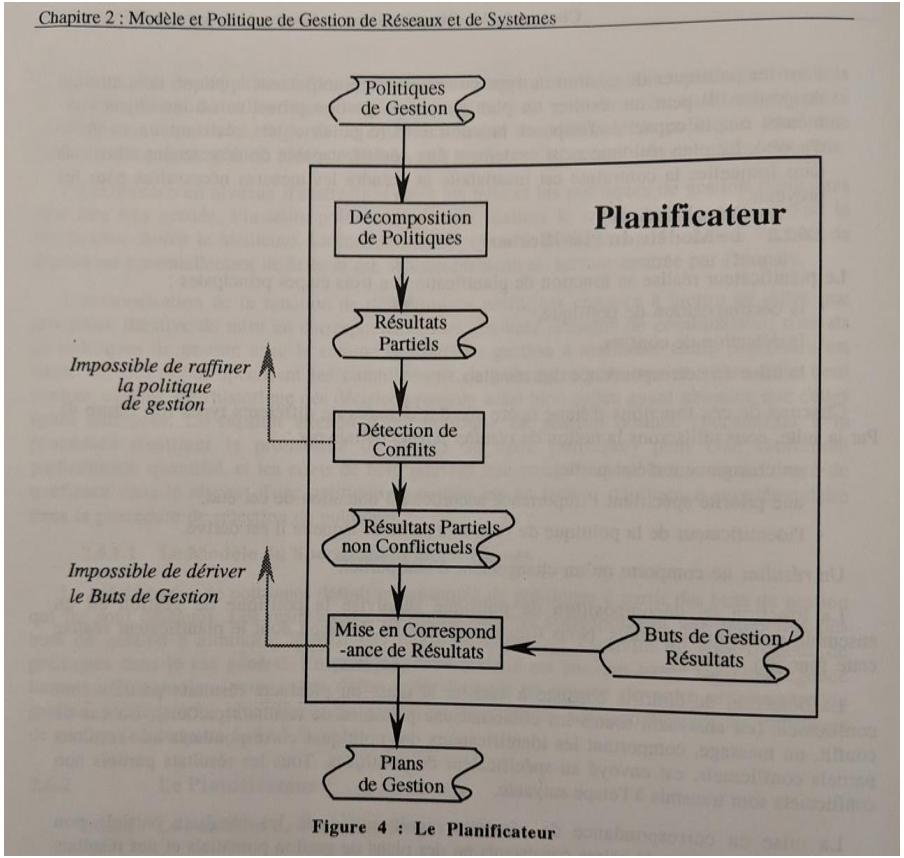


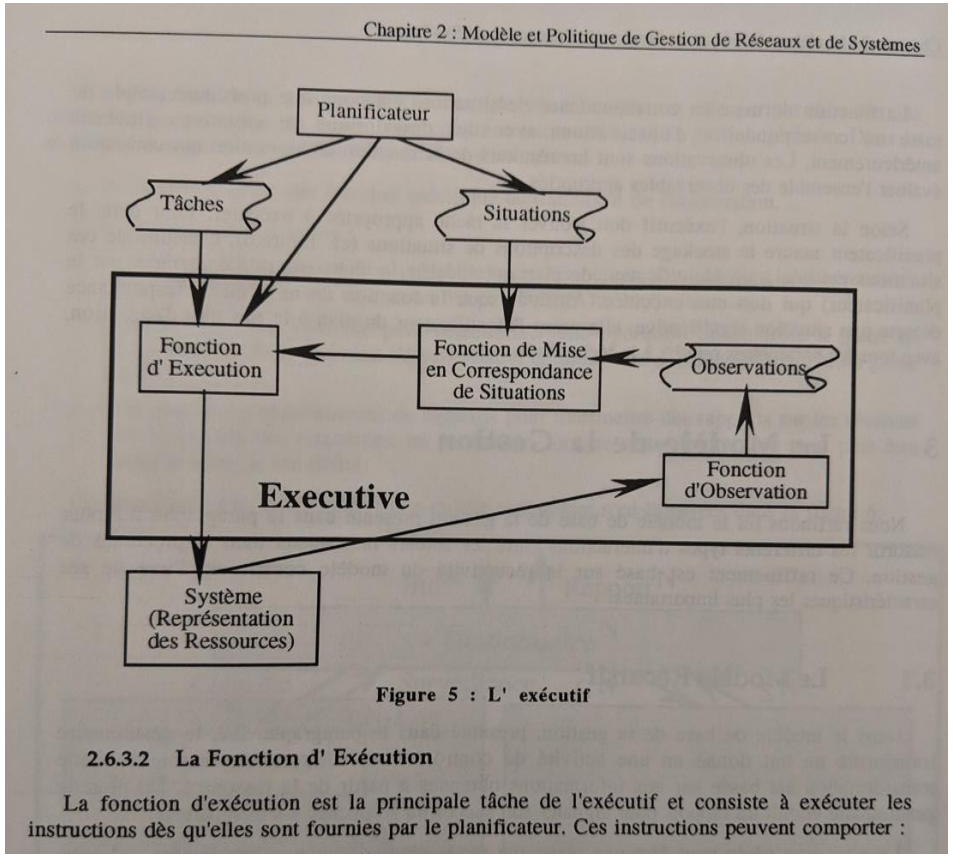
même, une politique peut être suffisamment précise qu'un seul plan soit dérivé. Dans le cas le plus simple, un but conduit à l'exécution d'une seule instruction de contrôle.

Cette structuration de l'activité du gestionnaire sépare les phases de définition de politiques et d'établissement de plans pouvant être contrôlées statiquement par l'humain de la phase dynamique d'exécution de plans. L'intervention de l'humain, dans le processus de gestion, est étudiée dans le chapitre 3.

R. Boutaba. PhD thesis. 1994
Université Pierre et Marie Curie
(Now Sorbonne Université).

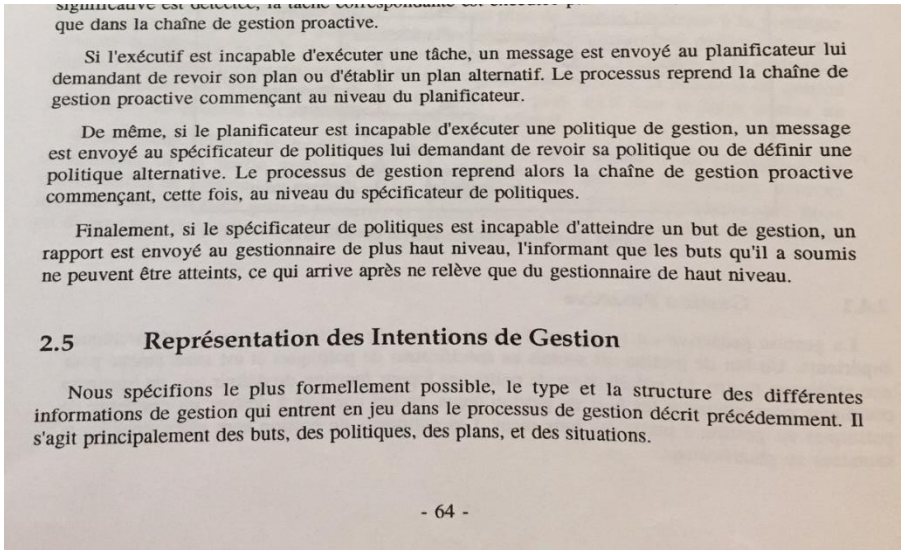
A Walk Down the Memory Lane





R. Boutaba. PhD thesis. 1994, Université Pierre et Marie Curie (Now Sorbonne Université).

A Walk Down the Memory Lane



que dans la chaîne de gestion proactive.

Si l'exécutif est incapable d'exécuter une tâche, un message est envoyé au planificateur lui demandant de revoir son plan ou d'établir un plan alternatif. Le processus reprend la chaîne de gestion proactive commençant au niveau du planificateur.

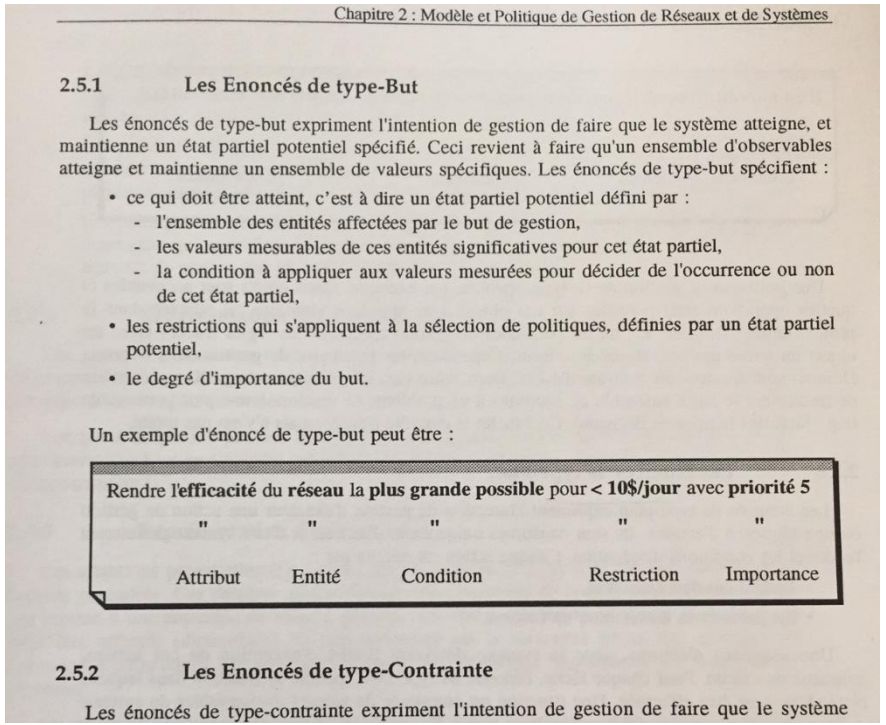
De même, si le planificateur est incapable d'exécuter une politique de gestion, un message est envoyé au spécificateur de politiques lui demandant de revoir sa politique ou de définir une politique alternative. Le processus de gestion reprend alors la chaîne de gestion proactive commençant, cette fois, au niveau du spécificateur de politiques.

Finalement, si le spécificateur de politiques est incapable d'atteindre un but de gestion, un rapport est envoyé au gestionnaire de plus haut niveau, l'informant que les buts qu'il a soumis ne peuvent être atteints, ce qui arrive après ne relève que du gestionnaire de haut niveau.

2.5 Représentation des Intentions de Gestion

Nous spécifions le plus formellement possible, le type et la structure des différentes informations de gestion qui entrent en jeu dans le processus de gestion décrit précédemment. Il s'agit principalement des buts, des politiques, des plans, et des situations.

- 64 -



Chapitre 2 : Modèle et Politique de Gestion de Réseaux et de Systèmes

2.5.1 Les Enoncés de type-But

Les énoncés de type-but expriment l'intention de gestion de faire que le système atteigne, et maintienne un état partiel potentiel spécifié. Ceci revient à faire qu'un ensemble d'observables atteigne et maintienne un ensemble de valeurs spécifiques. Les énoncés de type-but spécifient :

- ce qui doit être atteint, c'est à dire un état partiel potentiel défini par :
  - l'ensemble des entités affectées par le but de gestion,
  - les valeurs mesurables de ces entités significatives pour cet état partiel,
  - la condition à appliquer aux valeurs mesurées pour décider de l'occurrence ou non de cet état partiel,
- les restrictions qui s'appliquent à la sélection de politiques, définies par un état partiel potentiel,
- le degré d'importance du but.

Un exemple d'énoncé de type-but peut être :

Rendre l'**efficacité** du **réseau** la **plus grande possible** pour **< 10$/jour** avec **priorité 5**

" " " " "

Attribut Entité Condition Restriction Importance

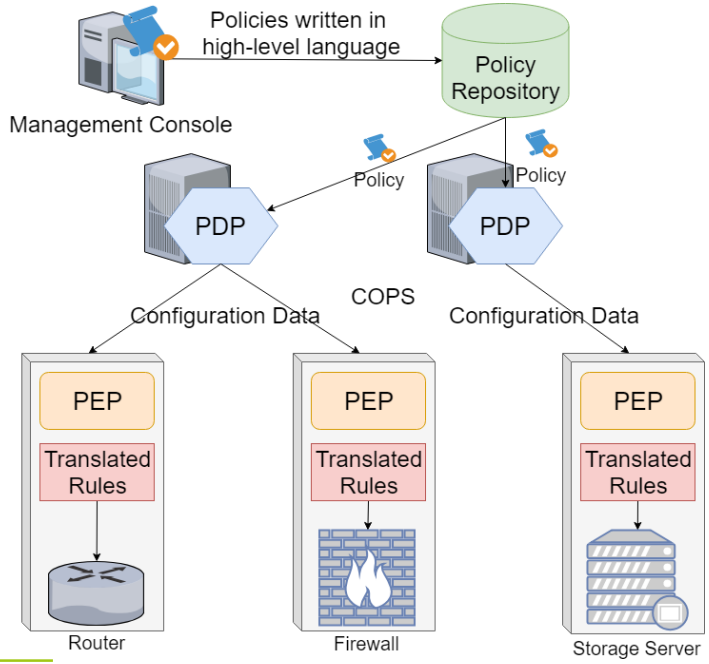
2.5.2 Les Enoncés de type-Contrainte

Les énoncés de type-contrainte expriment l'intention de gestion de faire que le système

R. Boutaba. PhD thesis. 1994, Université Pierre et Marie Curie (Now Sorbonne Université).

A Walk Down the Memory Lane

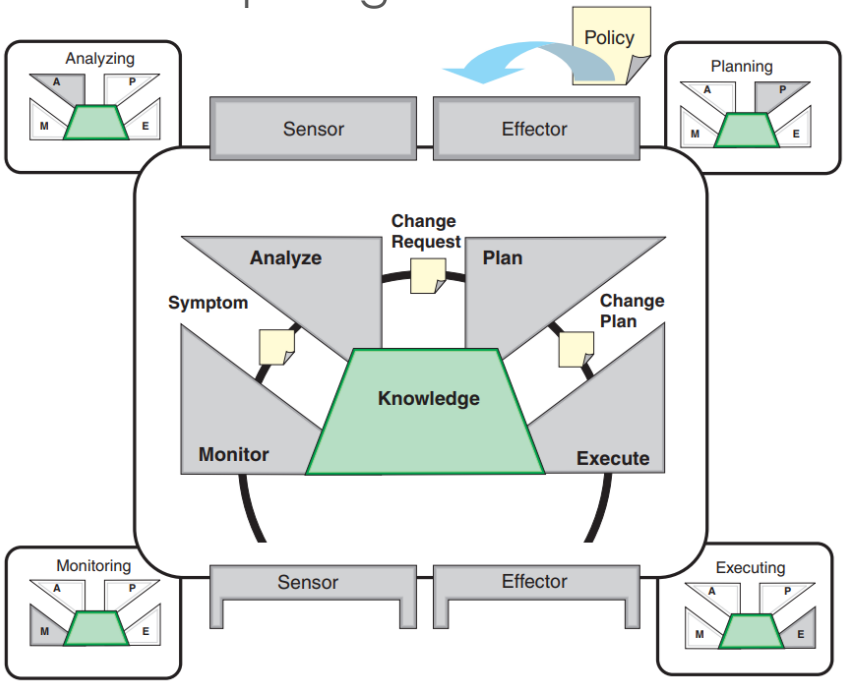
- Policy-based Management (PBM)



The COPS (common open policy service) protocol. IETF RFC 2748, 2000.

A Walk Down the Memory Lane

- IBM Autonomic Computing



Jeffrey O. Kephart, David M. Chess. The vision of autonomic computing, IEEE Computer, 36(1): 41-50. 2003

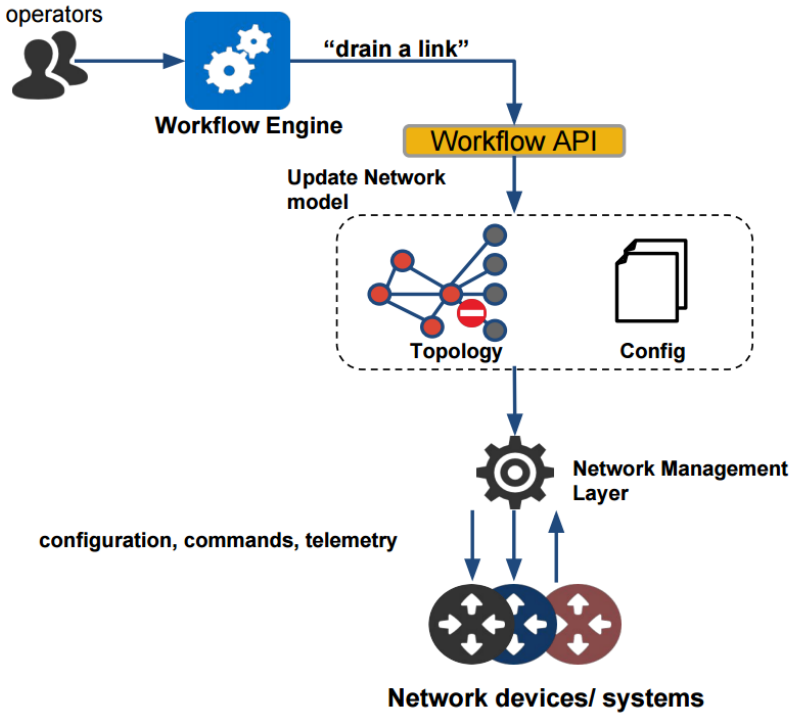
Why Practical Deployments Remained Unrealized?

- Reliance on proprietary hardware with little to no programmability
- Lack of global visibility restricting network-wide optimizations
- Inability to extract knowledge from network monitoring data at scale

Re(cent)incarnations

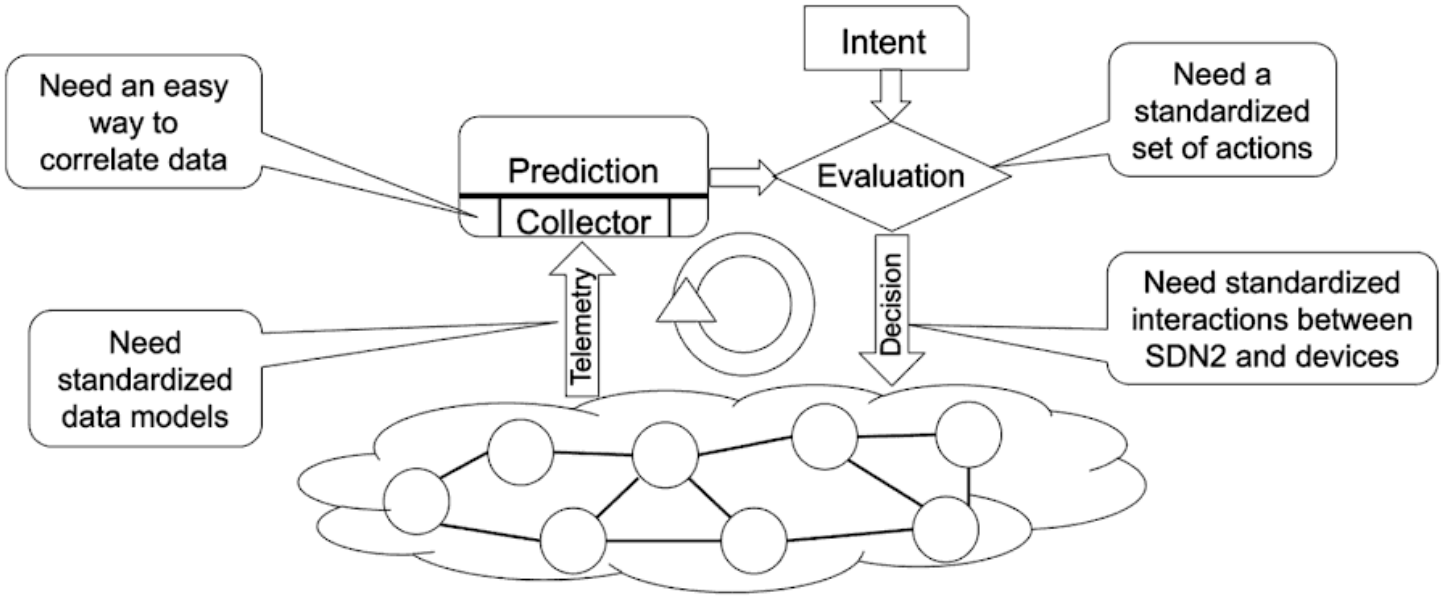
- Google Zero Touch Network
- Juniper Self-Driving Networks
- Open Network Automation Platform (ONAP)
- ETSI Zero-touch Network and Service Management (ZSM) Industry Standards Group
- Knowledge-defined Networking
- …and more

Google's Zero Touch Network



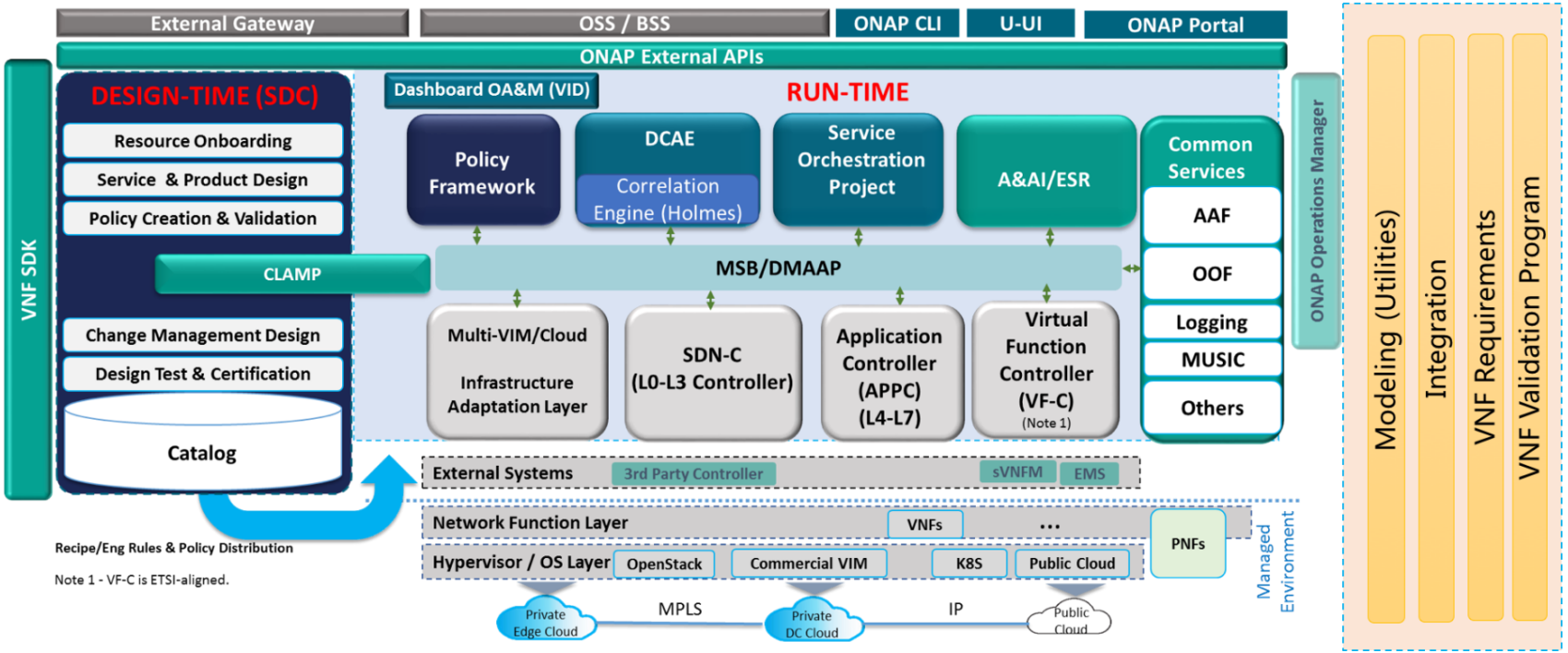
Bikash Koley, The Zero Touch Network, CNSM 2016 Keynote

Juniper's Self-Driving Network



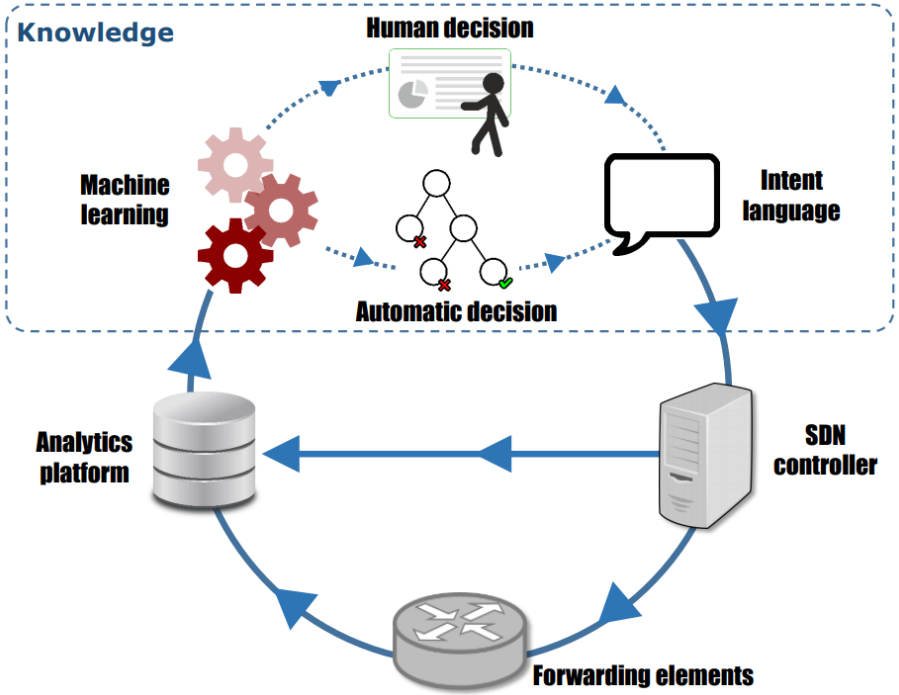
Kompella, Kireeti. "Self-Driving Networks." Emerging Automation Techniques for the Future Internet. IGI Global, 2019. 21-44.

Open Network Automation Platform



https://www.onap.org/platform-2

Knowledge-Defined Networking



A. Mestres, *et al*. Knowledge-Defined Networking. SIGCOMM CCR, 2017, 47:3, 2-10.

Can it Happen this Time?

- The stars are now aligned due to recent technological developments:
  - Network Softwarization
    - Enables flexible monitoring and control of networks
    - Facilitates network-wide optimizations and the deployment of network services on the fly.
  - Machine Learning
    - Enables knowledge extraction from monitoring data and automated decision making
  - Large-scale data processing
    - Enables network data analytics for large and complex networks with many users and applications

Network Softwarization

- Emerging networking paradigm where software controls the treatment of network flows and adds value to these flows by software processing.
- Two expressions of network softwarization:
  - Software-Defined Networking
    - Decouples the network's control and data planes for better programmability
  - Network Function Virtualization
    - Moves packet processing from purpose-built middleboxes to software appliances running on commodity hardware

Machine Learning

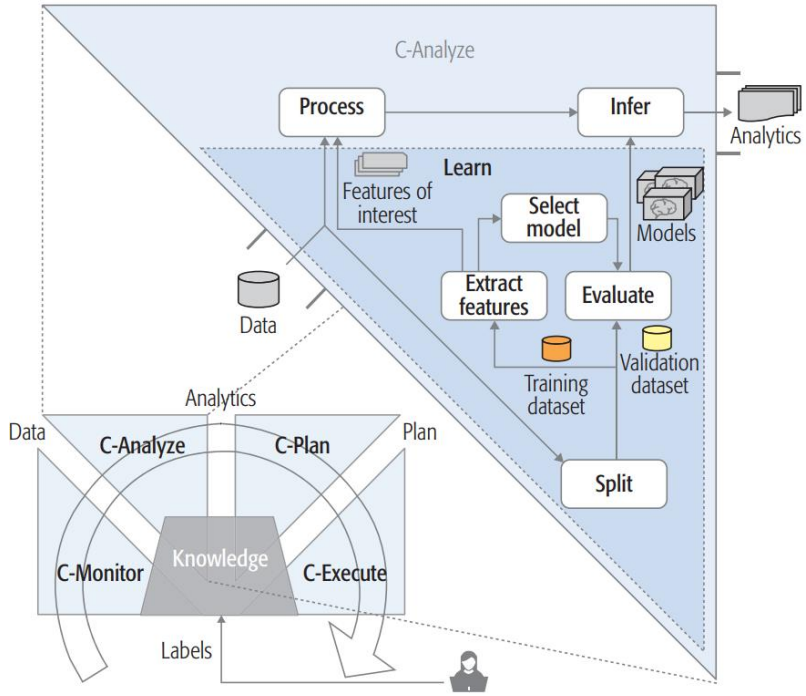
- Knowledge extraction
  - Recent success of Deep Learning
    - image & speech recognition, natural language processing
  - Proliferation of Machine Learning tools
    - TensorFlow, Torch, Keras
  - Availability of large volumes of data (aka Big Data)
- Automated decision making
  - (Deep) Reinforcement Learning has been successful in automating decision making processes
    - cluster resource management, web service configuration, recommendation systems, and robotics

Large-scale Data Processing

- Availability of infrastructure and platforms for data ingestion, storage, and analysis at large-scale
  - Cheap computing and storage
  - Massive parallelization using GPUs
  - Software platforms
    - Spark, Storm, Kafka
- Data processing as a cloud service
  - Amazon EMR, Amazon Kinesis, Azure Stream Analytics

Blue-print of a Cognitive MAPE

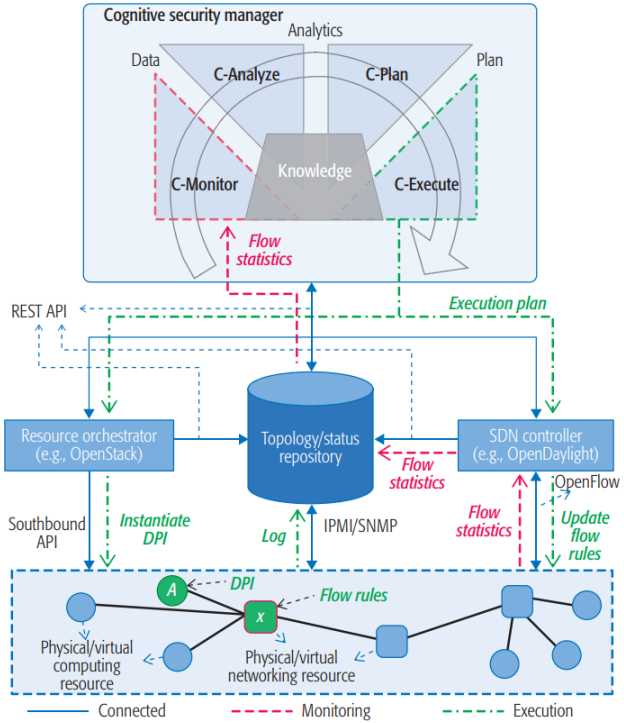
- C-MAPE: Cognitive control loop blue-print for automated network management
- Incorporates ML at each stage
  - *C-Monitor*: performs intelligent probing
  - *C-Analyze*: detects and predicts changes in networks
  - *C-Plan*: automated planning engine to react to changes
  - *C-Execute*: Optimal scheduling for plan execution



S. Ayoubi, N. Limam, M.A. Salahuddin, N. Shahriar, R. Boutaba. Machine Learning for Cognitive Network Management. IEEE Communications Magazine. Vol. 56(1), pp. 158-165, Jan 2018.

Blue-print of a Cognitive MAPE

- Use case: Cognitive Security Manager - Security anomaly detection and mitigation
  - Collects and analyses network statistics to detect security anomalies using ML
  - Uses reinforcement learning to generate a mitigation plan
  - Executes the plan leveraging network softwarization



S. Ayoubi, N. Limam, M.A. Salahuddin, N. Shahriar, R. Boutaba., *et al*. Machine Learning for Cognitive Network Management. IEEE Communications Magazine. Vol. 56(1), pp. 158-165, Jan 2018.

Future Research Directions

- Programmable network monitoring leveraging ML and network softwarization
- Predictive machine learning for automated management decision making
- On-demand orchestration of network services

ML-aided Network Monitoring

Background: Accuracy-overhead trade-off in network monitoring – constructing an accurate network view incurs high overhead

	High Overhead	Low Overhead
Spatial coverage (switch, flow)	Accurate network view (monitor all flows)	Approximate network view (sample flows)
Temporal coverage (query freq.)	Captures even short-lived events (microbursts)	Misses short-lived events

Research Questions:

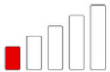
- *When* and *what* to query?
- Can we *predict* some measurements without querying the network?

Research Direction: Devise a predictive ML model for SDN controllers to: (i) decide when and what to query; and (ii) capture patterns in monitoring data & predict future measurements to reduce monitoring overhead

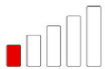
Monitoring Probe Distribution for Increased Network Traffic Visibility

Background: Programmable switches and general purpose servers can measure complex statistics beyond simple counters in the data plane (*e.g.*, flow size distribution, heavy hitters, ect.)

Visibility into network traffic



Resources (CPU, flow table)



Visibility into network traffic

Resources (CPU, memory)

Research Question:
Given a set of measurement queries, where should we perform traffic monitoring?

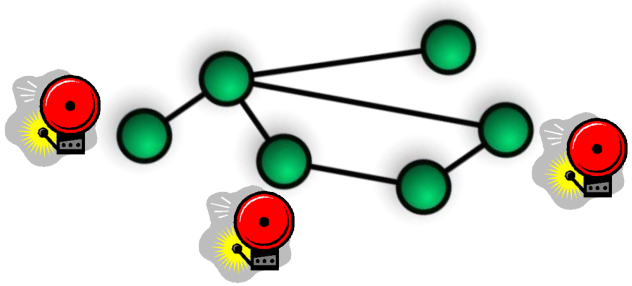
Research Direction: Optimally distribute monitoring tasks on end-hosts and programmable switches for maximizing network visibility under resource constraints.

Future Research Directions

- Programmable network monitoring leveraging ML and network softwarization
- Predictive machine learning for automated management decision making
- On-demand orchestration of network services

Root Cause Analysis of Network Anomalies

Background: Network anomaly is a departure from a network's desired behavior, *e.g.*, packet drops, link failures, DDoS attack.



Why these alarms went off?
(Misconfiguration? DDoS Attack?)

Challenges:

- Anomaly detection is half the battle
- Existing ML-based solutions cannot scale to the high-dimensional network state space for root cause localization*

Research Direction: Devise predictive models to uncover hidden correlations between a large number of high-dimensional network states to identify root causes of network anomalies.

* R. Boutaba, *et al*. A comprehensive survey on machine learning for networking: evolution, applications and research opportunities. Journal of Internet Services and Applications, 9(16), Jun 2018.

Automated Generation of Mitigation Workflows

Background: Once an anomaly and its root causes have been identified, the next step is to automatically decide a mitigation workflow

Policy-driven approach:
If-condition-then-action policies:

If-"packet drops on link (1,2) due to routing blackhole"-then-"reroute along path avoiding (1,2)"

Problem:
Modern networks are far too complex to generate workflows for all possible network conditions

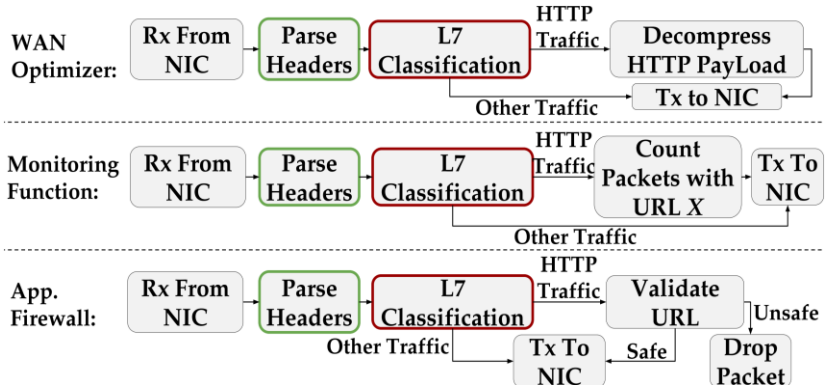
Research Direction: Leverage RL to automatically derive mitigation workflows based on past experience and current observations to bring the network to a *"good working state"* while scaling to the high-dimensional *"state-action"* space of modern networks.

Future Research Directions

- Programmable network monitoring leveraging ML and network softwarization
- Predictive machine learning for automated management decision making
- On-demand orchestration of network services

Re-architecting VNFs

Background: Current practice in NFV is to replace hardware middleboxes with monolithic software VNFs



*Functional decomposition of NFs**

Problems with monolithic VNFs:

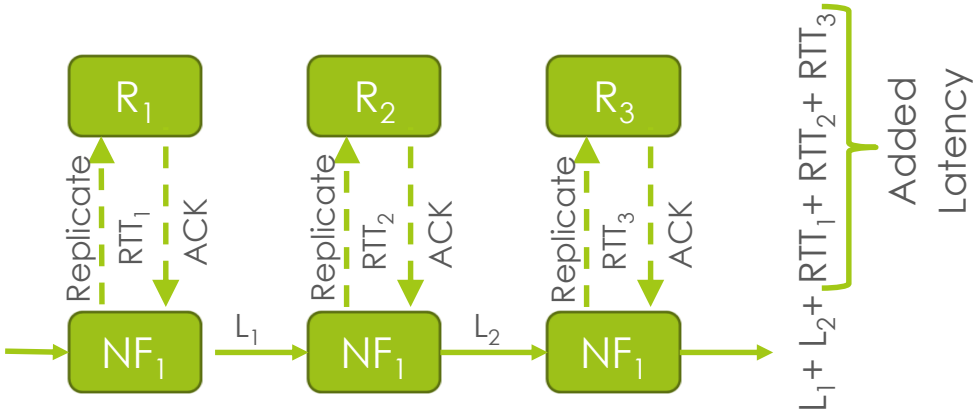
- Redundant development of common tasks
- Coarse-grained resource allocation & scaling
- Wasted CPU resources when VNFs are chained

Research Direction: Re-think VNF architectures to allow for more modular service composition, and finer-grained resource allocation and scaling

* S.R. Chowdhury, M.A. Salahuddin, N. Limam, R. Boutaba. Re-architecting NFV Ecosystem with Microservices: State-of-the-art and Research Challenges. IEEE Network, 2019.

Fault-tolerant Service Function Chaining

Background: NFV has significantly higher reliability requirement (five nines) than traditional cloud applications (four nines)



State-of-the-art: Make individual VNFs fault-tolerant through state replication and VM snapshots

Problem:
Increased latency due to per-VNF independent state replication & VM check-points

Research Direction: Design chain-wide fault-tolerance protocols for fast failure-recovery without adding significant delay during normal operations

Take-away

- Realizing the long-term vision of autonomous networks is even more critical today
  - increasing complexity of contemporary networks
  - stringent performance and reliability requirements of emerging applications
- Despite many attempts over the years, technological barriers prevented the realization of autonomous networks
- Stars are now aligned to achieve "self-driving" networks

Self-driving Networks

Can you make it happen this time ?