Italian Networking Workshop | Network Softwarization
Cavalese (TN), 30 January 2020

A Model-based Approach Towards Real-time Analytics in NFV Infrastructures

Raffaele Bolla, Roberto Bruschi, Franco Davoli, Jane Frances Pajo

IEEE TGCN
10.1109/TGCN.2019.2961192

ACKNOWLEDGEMENT
MATILDA

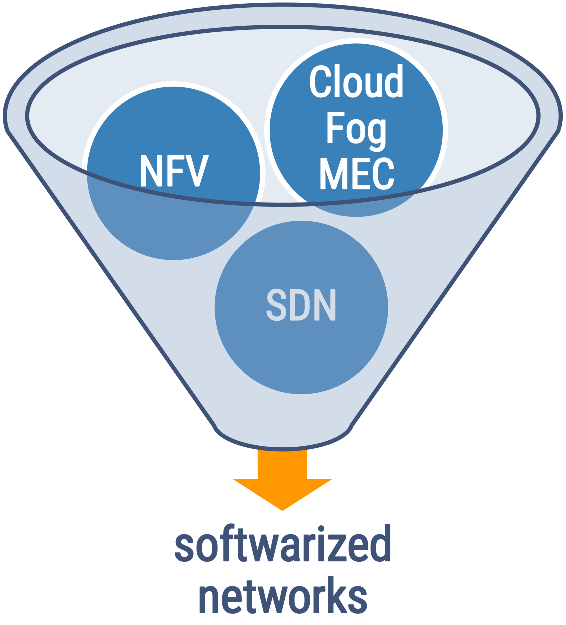
Outline

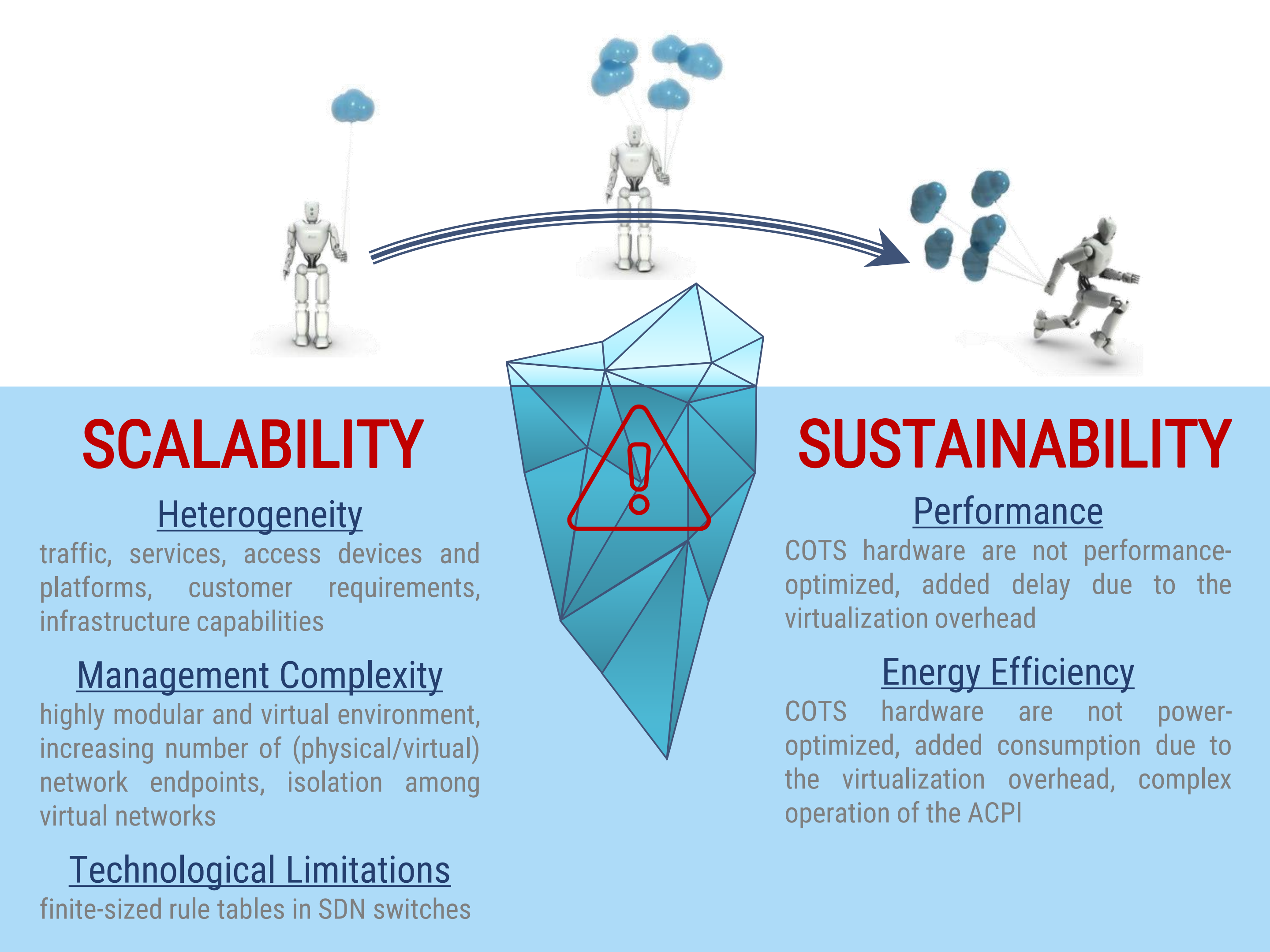
Network Softwarization

A telecom architecture **paradigm shift**

"…from boxes to functions, and from protocols to APIs…" [1]

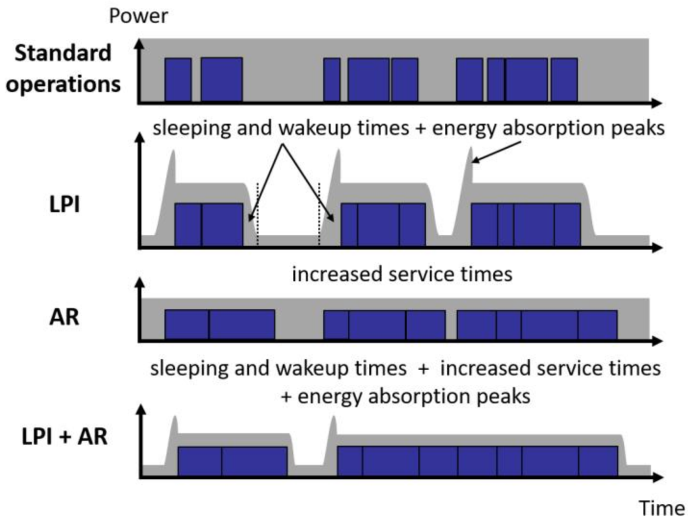
- **Cloud-Fog-MEC interplay** results in the *pooling of computing, storage and networking resources* in a virtualized (geo-distributed) infrastructure
- **NFV** consolidates *various types of network appliances as VNFs* on COTS hardware (in the Cloud-Fog-MEC domain)
- **SDN** decouples the control and data planes to enable a *highly programmable network behavior*




SCALABILITY
Heterogeneity
traffic, services, access devices and platforms, customer requirements, infrastructure capabilities
Management Complexity
highly modular and virtual environment, increasing number of (physical/virtual) network endpoints, isolation among virtual networks
Technological Limitations
finite-sized rule tables in SDN switches
SUSTAINABILITY
Performance
COTS hardware are not performance-optimized, added delay due to the virtualization overhead
Energy Efficiency
COTS hardware are not power-optimized, added consumption due to the virtualization overhead, complex operation of the ACPI

Power Management in COTS Hardware

The massive introduction of **COTS hardware** in softwarized environments would tend to **increase power requirements** with respect to specialized hardware solutions, in the absence of specific control actions.



ACPI Specification[2]

- **Low Power Idle (LPI)**
  via the **power states (C_x)** – comprising the active state C_0 and sleeping states $\{C_1, \dots, C_X\}$
- **Adaptive Rate (AR)**
  via the **performance states (P_y)** – corresponding to different processing performances $\{P_0, \dots, P_Y\}$ at C_0

Goal

To enable **scalable** and **sustainable** softwarized networks through **real-time** analytics that exposes the **power** and **performance** KPIs in COTS hardware according to VNF workloads.

State-of-the-art [3-8]

- system modeling and analytics
- power-performance tradeoff optimization

machine learning, queueing theory

ACPI vs burstiness: packet-level analysis

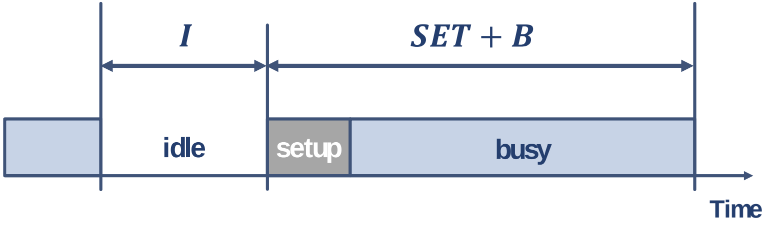
Contributions in the paper

- Modeling cores in the NFVIs as $M^X/G/1/SET$ queues
- Exposing model parameters from available PMCs
- VNF workload profiling and estimation of network KPIs

$M^X/G/1/SET$ Queueing Model

A generalization of the well-known $M^X/G/1$ model for **batch arrivals** that also covers the cases in which an additional **setup period** is necessary before service can be resumed…

- incoming traffic as a **Batch Markov Arrival Process (BMAP)**: **exponentially-distributed batch inter-arrival times**
- **generally-distributed setup times** due to wakeups and reconfigurations
- single server with **generally-distributed service times**
- system works as a **renewal process** of idle (I) and delay busy $(SET + B)$ periods



Power Modeling

Renewal Theory

$$\Phi = \frac{I_{(1)}}{R_{(1)}}\Phi_i + \frac{\tau_{(1)}}{R_{(1)}}\Phi_w + \frac{B_{(1)}}{R_{(1)}}\Phi_a$$

average renewal cycle $\boldsymbol{R}_{(1)} = \boldsymbol{I}_{(1)} + \boldsymbol{\tau}_{(1)} + \boldsymbol{B}_{(1)}$

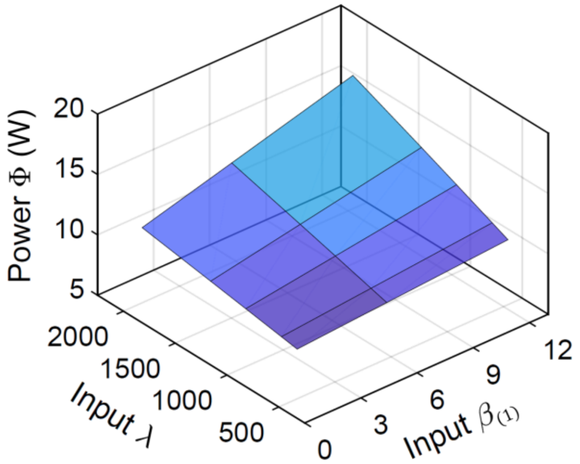
$$\Phi = \frac{1-\rho}{1+\lambda\tau_p}\Phi_i + \frac{\lambda\tau_p(1-\rho)}{1+\lambda\tau_p}\Phi_w + \rho\Phi_a$$

τ_p setup time due to core wakeup

Φ_i idle power consumption

Φ_w wakeup power consumption

Φ_a active power consumption



Latency Modeling

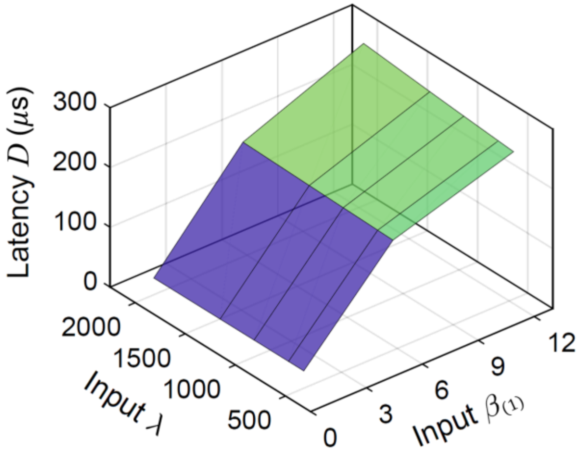
Little's Law

$$W = \frac{L}{OL}, \text{ where } L = \left.\frac{dP(z)}{dz}\right|_{z=1} - \rho$$

PGF of the # of packets in the system $P(z)$

$$D = \underbrace{\frac{\rho\beta_{(1)}+\beta_{(2)}}{2\beta_{(1)}\mu(1-\rho)} + \frac{2\tau_l+\lambda\tau_l^2}{2\beta_{(1)}(1+\lambda\tau_l)}}_{W} + \frac{1}{\mu}$$

τ_l setup time due to interrupt coalescing, core wakeups (τ_p) and reconfiguration



Estimation with PMCs

- $\hat{\rho} = \frac{\widetilde{OL}}{\mu}$
- $\hat{\lambda} = \frac{\alpha_1}{\widetilde{T}_{C_x}} + \alpha_0$
- $\hat{\beta}_{(1)} = \frac{\widetilde{OL}}{\hat{\lambda}}$
- $\hat{\beta}_{(2)} = \max\left(0, \left(\frac{\mu^2}{1+\hat{\lambda}\tau_l}\right)\left(\hat{B}_{(2)}(1-\hat{\rho})^3 - \frac{\hat{\beta}_{(1)}(1+\hat{\lambda}\tau_l)}{\mu^2} - \frac{\hat{\rho}(1-\hat{\rho})\hat{\beta}_{(1)}(2\tau_l+\hat{\lambda}\tau_l^2)}{\mu}\right)\right)$

$\tilde{\rho} = \frac{\widetilde{T}_{P_y}}{\widetilde{T}_{P_y} + \widetilde{T}_{C_x} + \widetilde{T}_{POLL}}$ → has busy period overhead due to other operations

$\tilde{\lambda} = \frac{1}{\widetilde{T}_{C_x} + \widetilde{T}_{POLL}}$ → $\widetilde{T}_{POLL}$ has high variance

VnStat

$\widetilde{OL}$

Idlestat

$\widetilde{T}_{P_y}$

$\widetilde{T}_{C_x}$

$\widetilde{T}_{POLL}$

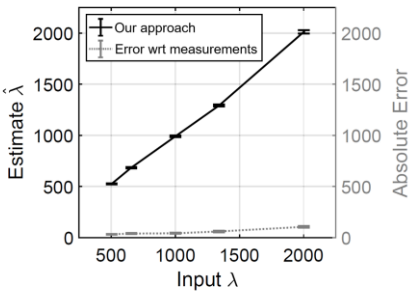
Supposing that $B^n_{(1)}, n = 1, \dots, \eta$, are iid, where $B^n_{(1)} \approx \widetilde{T}_{P_y} - \Delta T$

$$\hat{B}_{(1)} = \frac{1}{\eta}\sum_{n=1}^{\eta} B^n_{(1)} \qquad \widehat{var}(B_{(1)}) = \frac{1}{\eta-1}\sum_{n=1}^{\eta}\left(B^n_{(1)} - \hat{B}_{(1)}\right)^2$$

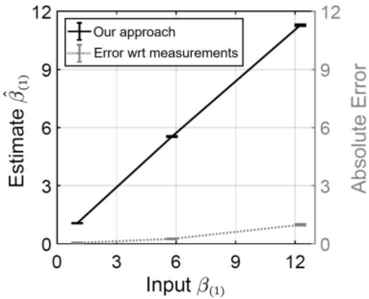
$$\hat{B}_{(2)} = \Delta t\, \widehat{var}(B_{(1)}) + \left(\hat{B}_{(1)}\right)^2$$

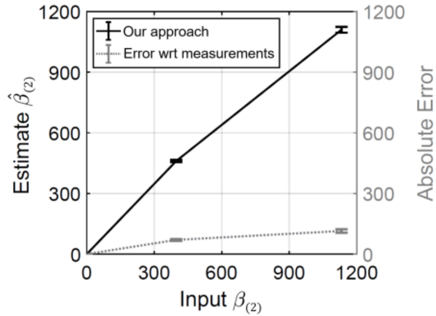
Results from Emulated BMAP Traffic

Batch arrival rate

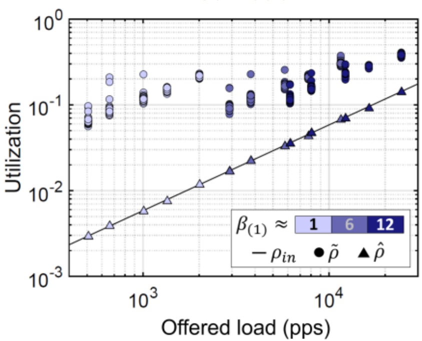


Factorial moments of the batch size

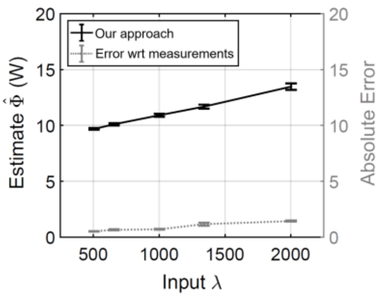




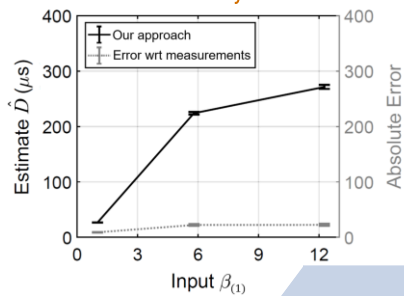
Utilization



Power

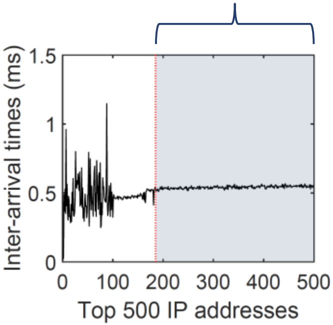


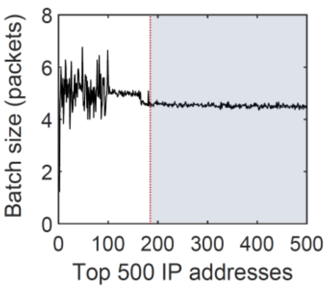
Latency



Results from Facebook Traces[9]

5 IP addresses randomly chosen





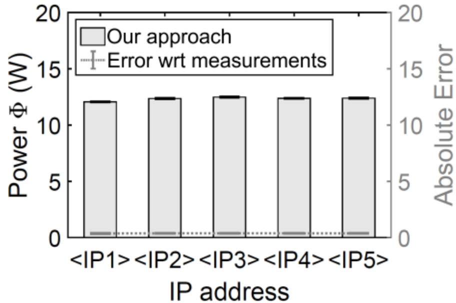
Fitting batch inter-arrival times to exponential distributions:

<IP1>:	<IP2>:	<IP3>:	<IP4>:	<IP5>:
R=95.28%	R=95.49%	R=95.67%	R=95.44%	R=95.64%
R^2=90.79%	R^2=91.19%	R^2=91.52%	R^2=91.08%	R^2=91.46%

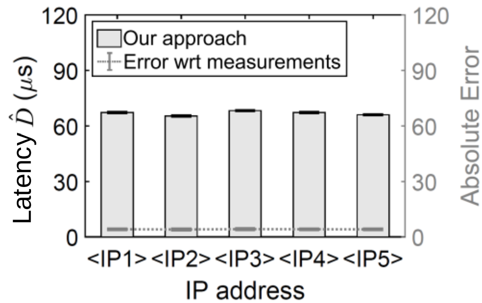
Fitting batch sizes to generalized Pareto distributions:

<IP1>:	<IP2>:	<IP3>:	<IP4>:	<IP5>:
R=99.85%	R=99.85%	R=99.91%	R=99.86%	R=99.9%
R^2=99.7%	R^2=99.7%	R^2=99.81%	R^2=99.72%	R^2=99.8%

Power



Latency



Conclusions

- **Network softwarization**: Cloud-Fog-MEC interplay, NFV, SDN
- **Scalability and sustainability issues**: massive use of COTS hardware, highly heterogeneous, modular and virtual environment; increasing # of network endpoints
- **Infrastructure modeling**: cores as $M^X/G/1/SET$ queues
- **Real-time analytics in NFVIs**: VNF workload profiling and estimation of network KPIs based on PMCs
- **Validation based on emulated traffic and Facebook's dataset**: high scalability, good estimation accuracies
- **Power tool**: augmenting VIM capabilities, development of next-generation resource/service provisioning solutions

Thank you!

Any questions?

References

[1] C. Tao et al. Vision on Software Networks and 5G. 5GPPP SN WG White Paper. Jan. 2017. URL: https://5g-ppp.eu/wp-content/uploads/2014/02/5GPPP_SoftNets_WG_whitepaper_v20.pdf.

[2] Advanced Configuration and Power Interface Specification. URL: http://www.acpi.info/DOWNLOADS/ACPIspec40a.pdf

[3] C. Möbius, W. Dargie, and A. Schill. "Power Consumption Estimation Models for Processors, Virtual Machines, and Servers". In: IEEE Trans. Parallel Distrib. Syst. 25.6 (June 2014), pp. 1600–1614.

[4] M. Dayarathna, Y. Wen, and R. Fan. "Data Center Energy Consumption Modeling: A Survey". In: IEEE Commun. Surveys Tuts. 18.1 (2016), pp. 732–794.

[5] M. Dimitrov et al. "Coordinated Optimization: Dynamic Energy Allocation in Enterprise Workload". In: Intel® Technol. J. 16.2 (2012), pp. 32–51.

[6] R. Bolla et al. "Green Networking with Packet Processing Engines: Modeling and Optimization". In: IEEE/ACM Trans. Netw. 22.1 (Feb. 2014), pp. 110–123.

[7] R. Bolla et al. "Corrections to: "Green Networking With Packet Processing Engines: Modeling and Optimization"". In: IEEE/ACM Trans. Netw. (2018). DOI: 10.1109/TNET.2017.2761892.

[8] L. Duan, D. Zhan, and J. Hohnerlein. "Optimizing Cloud Data Center Energy Efficiency via Dynamic Prediction of CPU Idle Intervals". In: Proc. 8th IEEE Int. Conf. Cloud Comput. (CLOUD). New York, NY, June 2015, pp. 985–988.

[9] Facebook, "FBFlow Dataset." URL: https://www.facebook.com/network-analytics