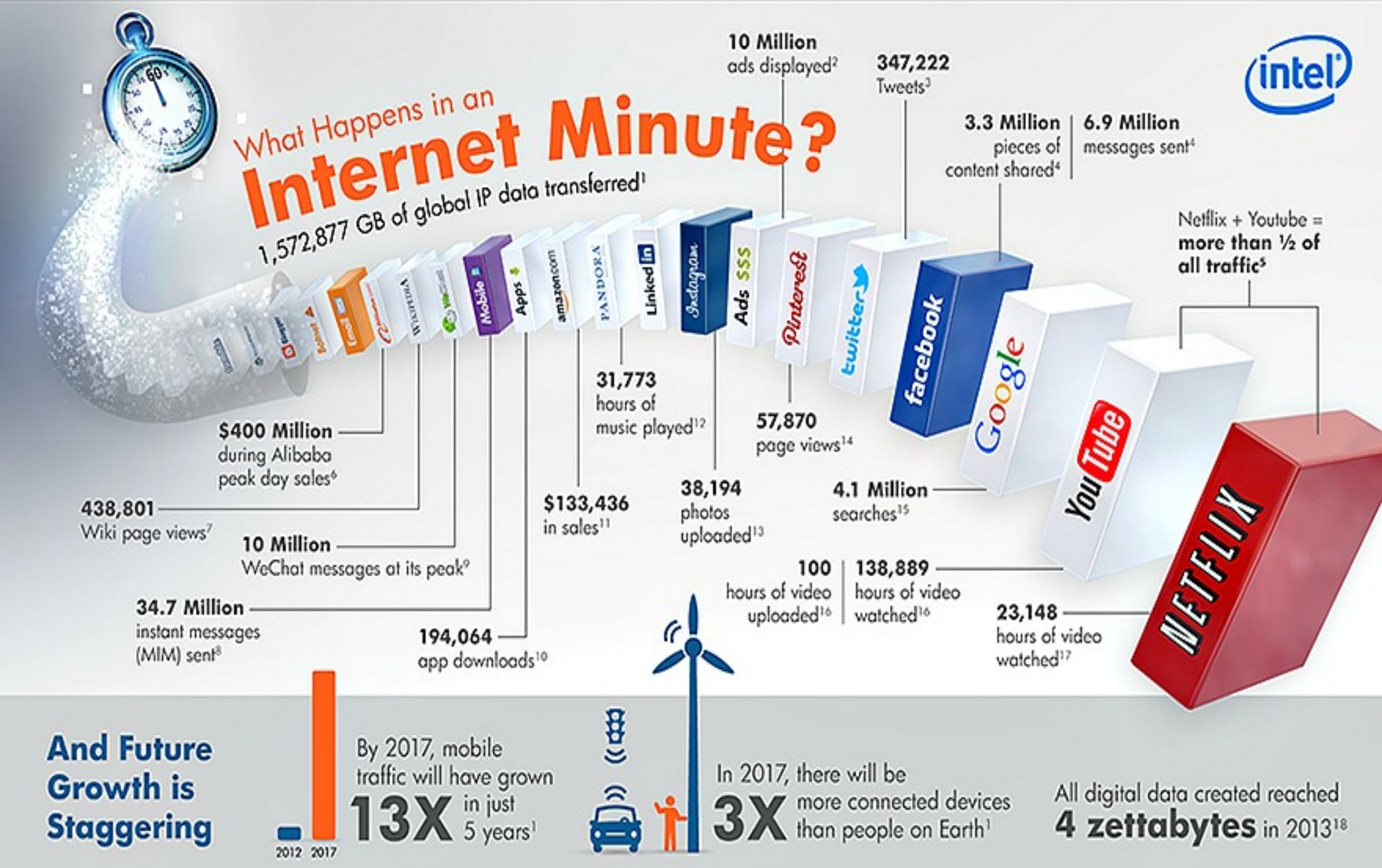
Smart and Secure Cyber Infrastructure.

Cees de Laat
System & Network Engineering
University of Amsterdam

COMMIT/



1,572,877 GByte/minute =
(8*1,572,877*10^9/60 bit/s)/(10*10^12 bit/s per fiber) =
21 fibers with each about 100 * 100 Gb/s channels

Trucks of Tapes

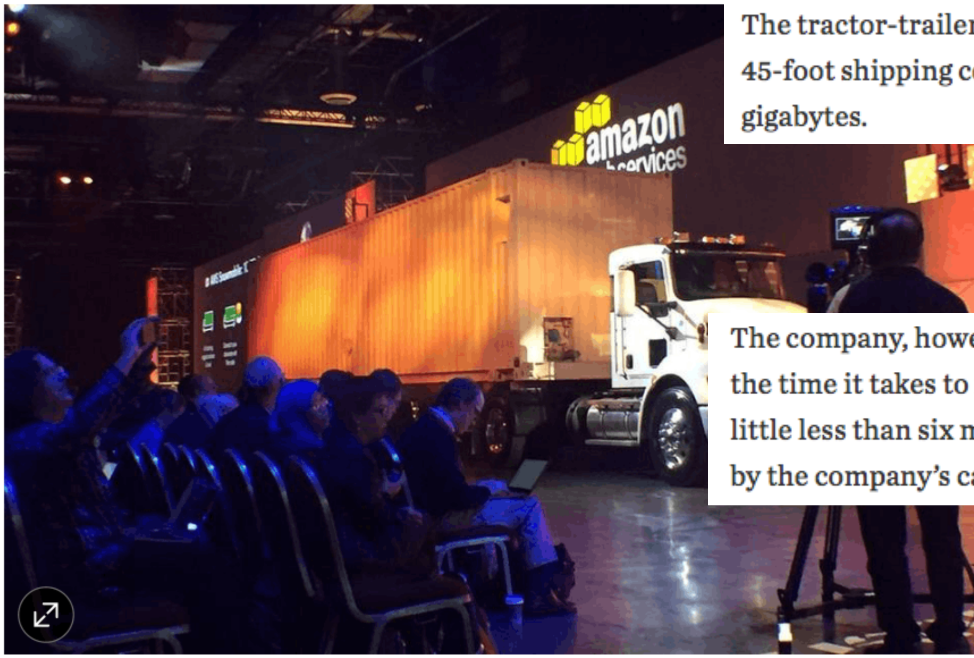
THE WALL STREET JOURNAL.

TECH

Amazon Uses Trucks to Drive Data Faster

Cloud-computing unit, Amazon Web Services, unveils new offerings at annual conference in Las Vegas

The tractor-trailer hauls a massive storage device, dubbed Snowmobile, in the form of a 45-foot shipping container that holds 100 petabytes of data. A petabyte is about 1 million gigabytes.



Amazon unveiled the 'Snowmobile' service on Wednesday in Las Vegas. *PHOTO: AMAZON WEB SERVICES*

The company, however, isn't promising lightning speed. Ten Snowmobiles would reduce the time it takes to move an exabyte from on-premises storage to Amazon's cloud to a little less than six months, from about 26 years using a high-speed internet connection, by the company's calculations.

By JAY GREENE By LAURA STEVENS

4 COMMENTS

Updated Nov. 30, 2016 7:19 p.m. ET

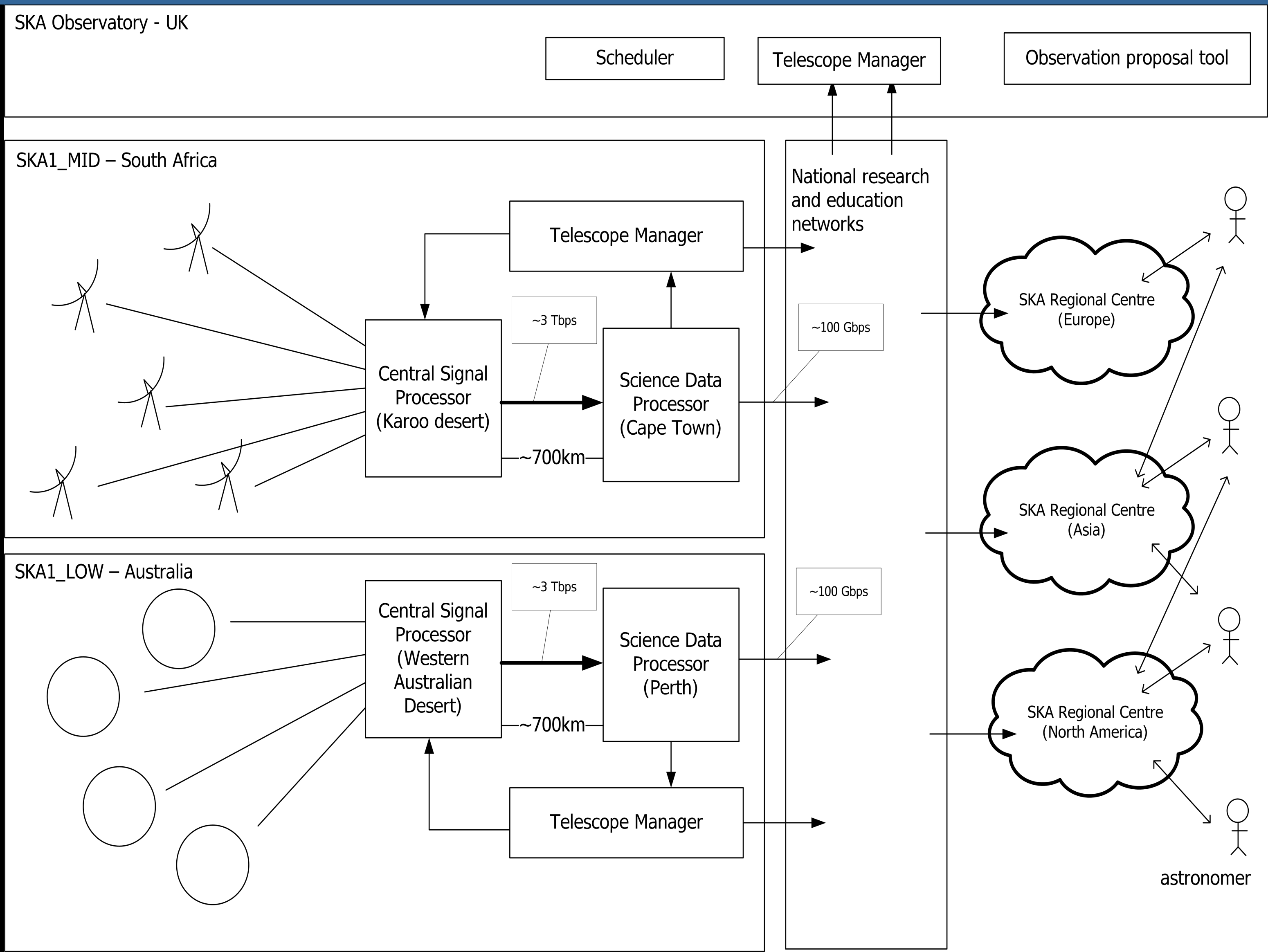
LAS VEGAS—In Amazon Web Services, Amazon.com Inc. has built one of the most powerful computing networks in the world, on pace to post more than $12 billion in revenue this year.

But the retail giant on Wednesday proposed a surprising way to move data from large corporate customers' data centers to its public cloud-computing operation: by truck.

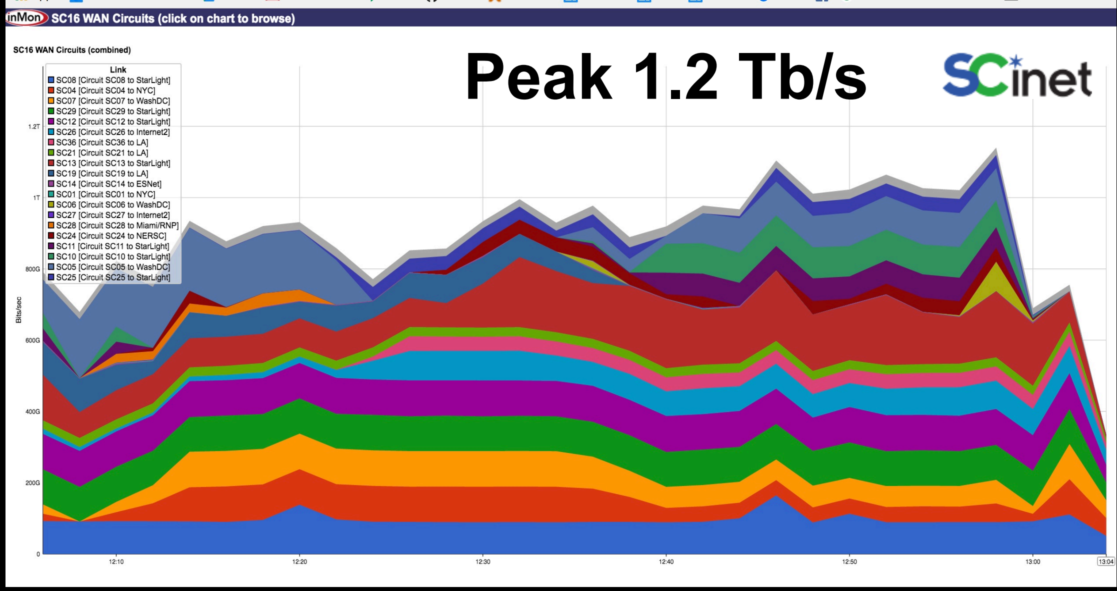
Networks can move massive amounts of data only so fast. Trucks, it turns out, can move it faster.

1 fiber does about 16 Tbit/s
= 2 Tbyte/s
⇒ 500000 s/ExaByte
⇒ One week/ExaByte
Or stick Joe and Harvey in a RV for 2 months.

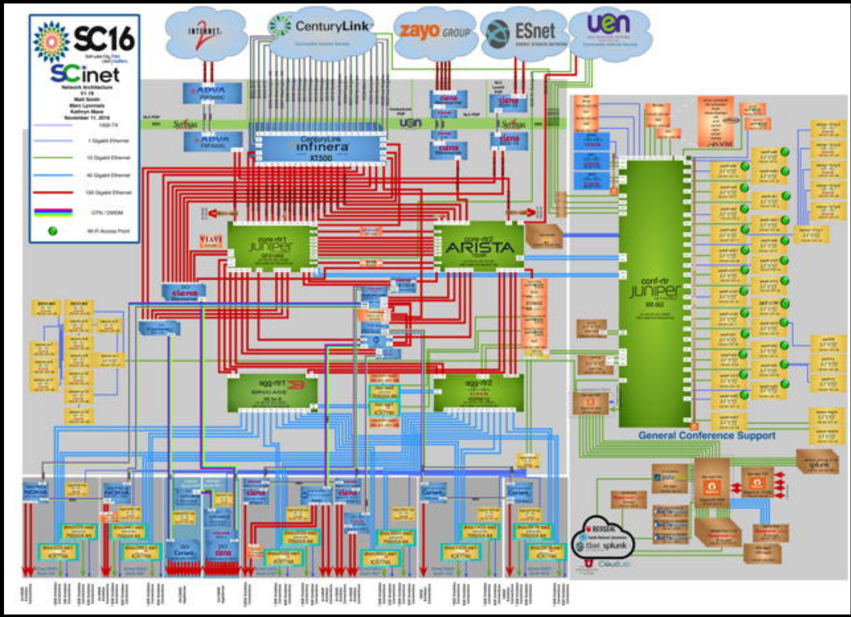
SKA: Depending on analysis load & physics mode they want to investigate to use SDN in real time to direct bursts of data to different compute resources and do load balancing.



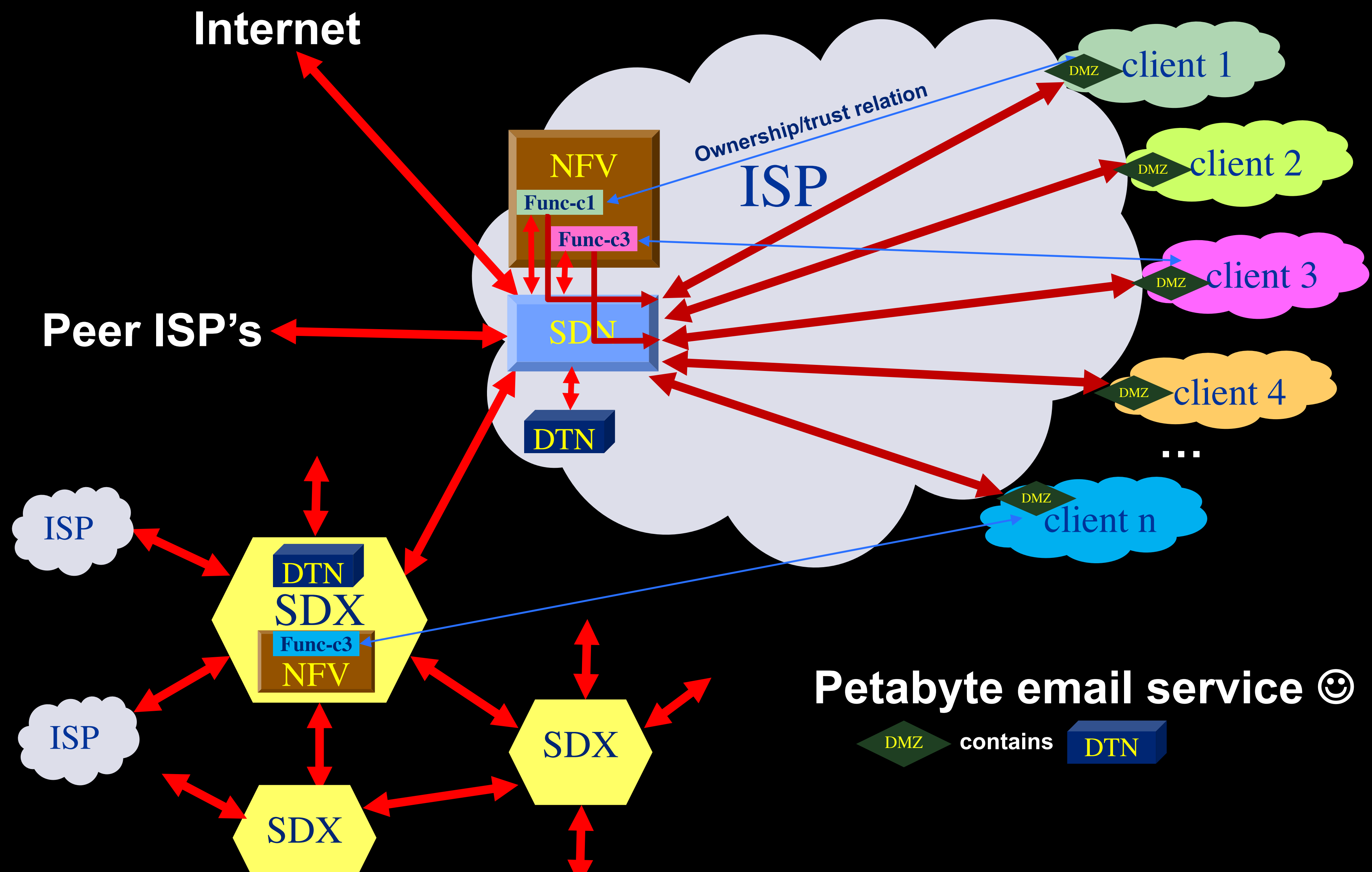
Learned from Scinet & INDIS



- 2013 - 2016
  - SDN
  - Security
  - Traffic management, policing, control
  - Hybrid – optical ring - approach to reach Tb/s
- 2017 - 2020
  - NFV
  - SDX
  - DTN @ core ➔ petabyte email network
  - Data abstractions (e.g. NDN)



Networks of ScienceDMZ's & SDX's

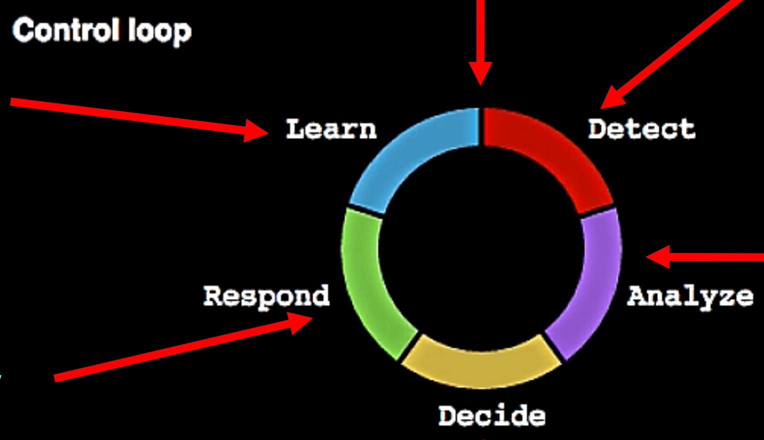


Status SARNET Operational Level

Laboratory: ExoGeni & PRP
Fieldlab with KLM & CIENA
OSA-Optical Forum Conference paper [1]

CoreFlow
Berkeley Internship 2016
SC16 INDIS workshop paper [2]

ToDO ‘18



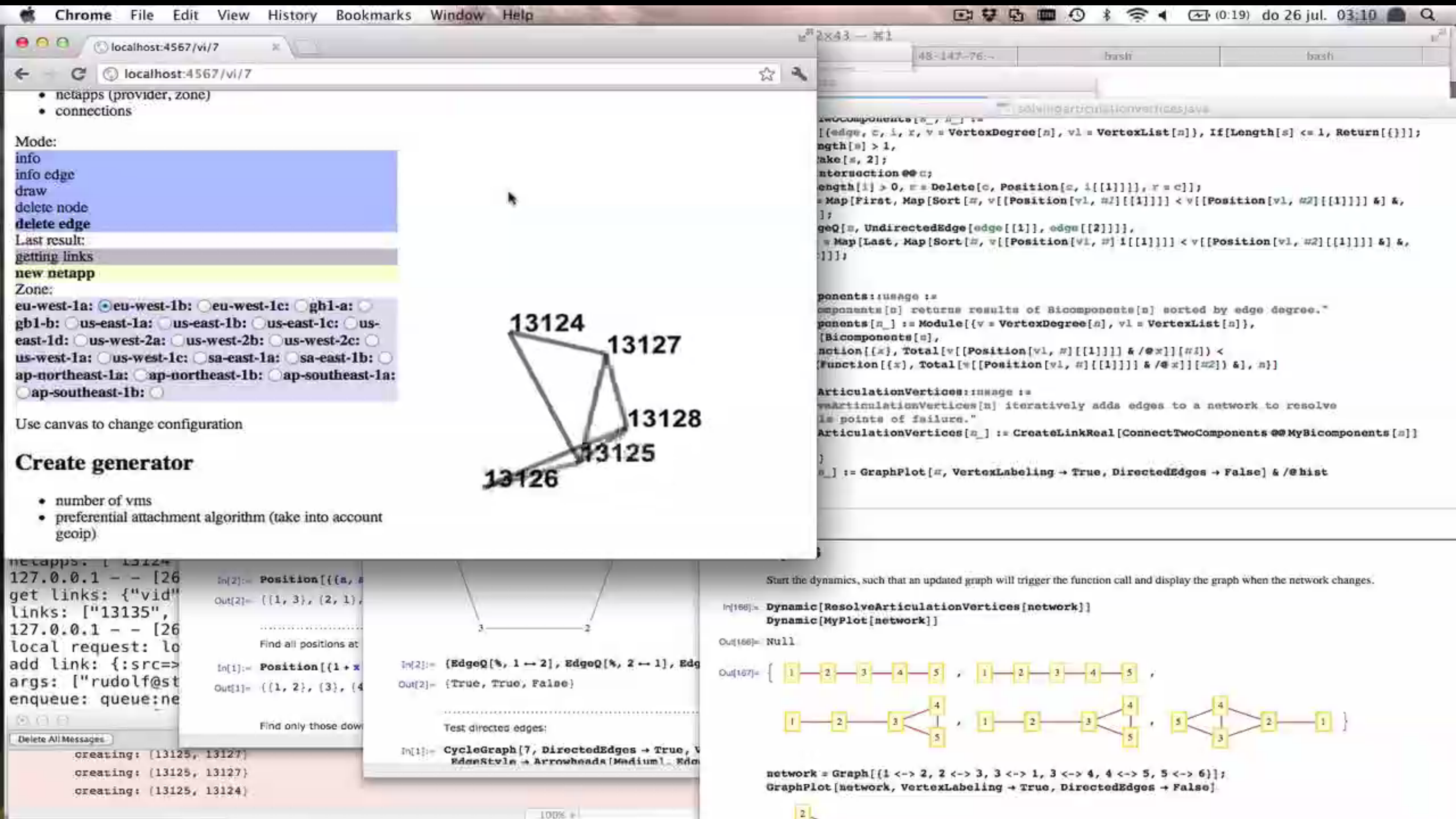
SC16 demo plus poster
Salt Lake City (UT)
IEEE Sec-Virtnet 2016 paper [3]

SC15 demo plus poster
Austin (TX)

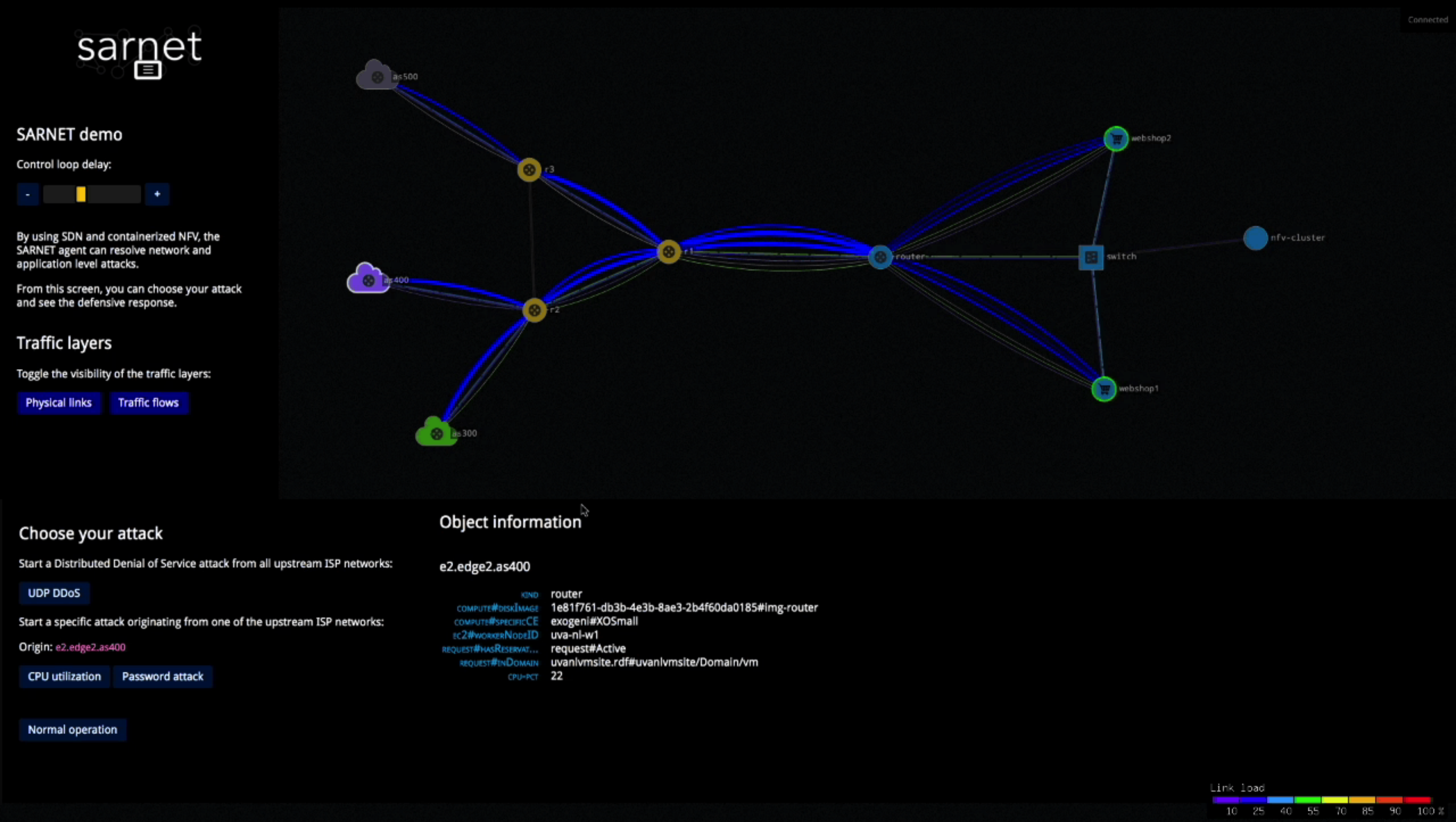
SC16 demo plus poster
Salt Lake City (UT)

1. Paper: R. Koning, A. Deljoo, S. Trajanovski, B. de Graaff, P. Grosso, L. Gommans, T. van Engers, F. Fransen, R. Meijer, R. Wilson, and C. de Laat, "Enabling E-Science Applications with Dynamic Optical Networks: Secure Autonomous Response Networks ", OSA Optical Fiber Communication Conference and Exposition, 19-23 March 2017, Los Angeles, California.
2. Paper: Ralph Koning, Nick Buraglio, Cees de Laat, Paola Grosso, "CoreFlow: Enriching Bro security events using network traffic monitoring data", SC16 Salt Lake City, INDIS workshop, Nov 13, 2016.
3. Paper: Ralph Koning, Ben de Graaff, Cees de Laat, Robert Meijer, Paola Grosso, "Analysis of Software Defined Networking defences against Distributed Denial of Service attacks", The IEEE International Workshop on Security in Virtualized Networks (Sec-VirtNet 2016) at the 2nd IEEE International Conference on Network Softwarization (NetSoft 2016), Seoul Korea, June 10, 2016.

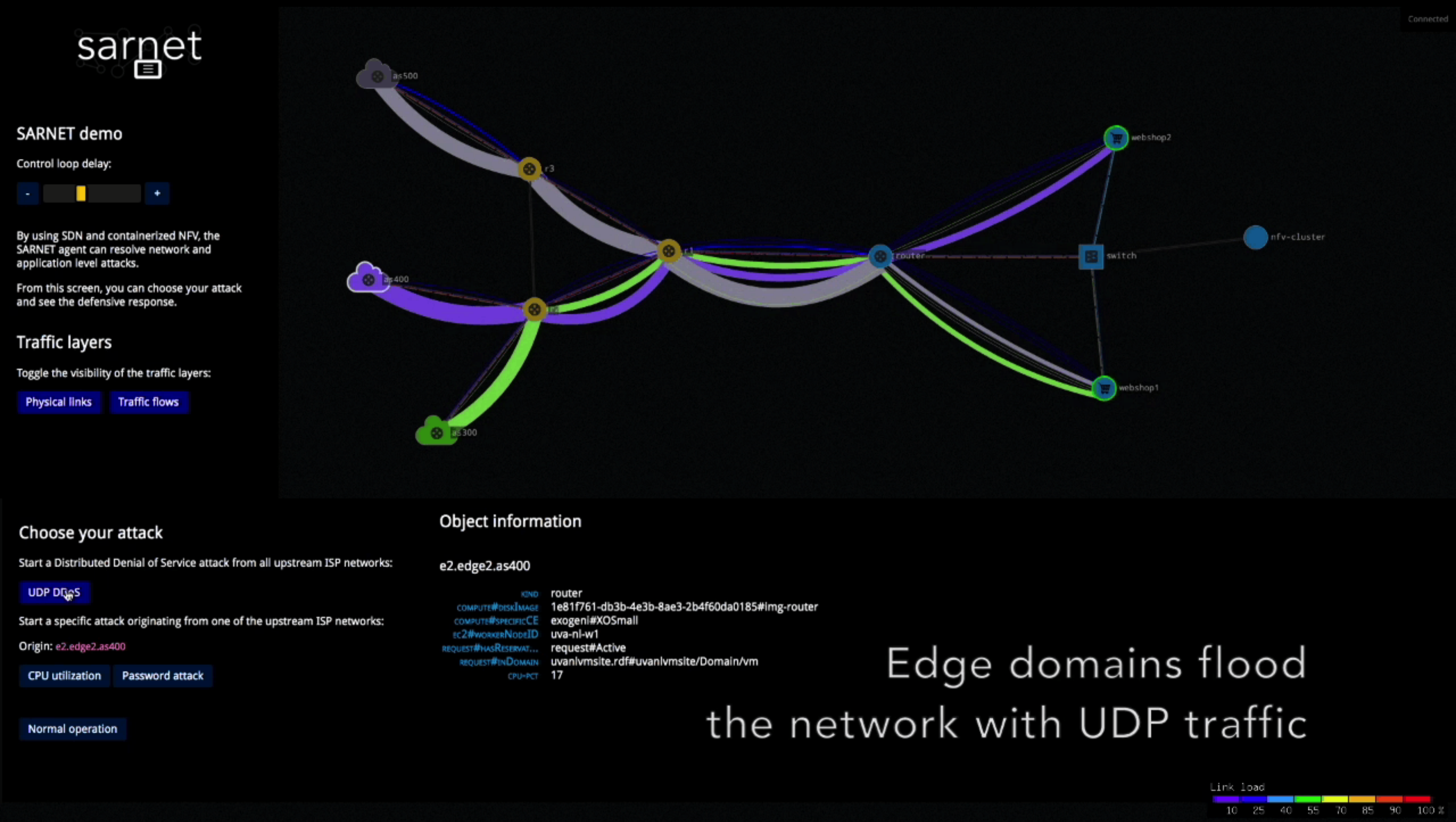
Basic operating system loop



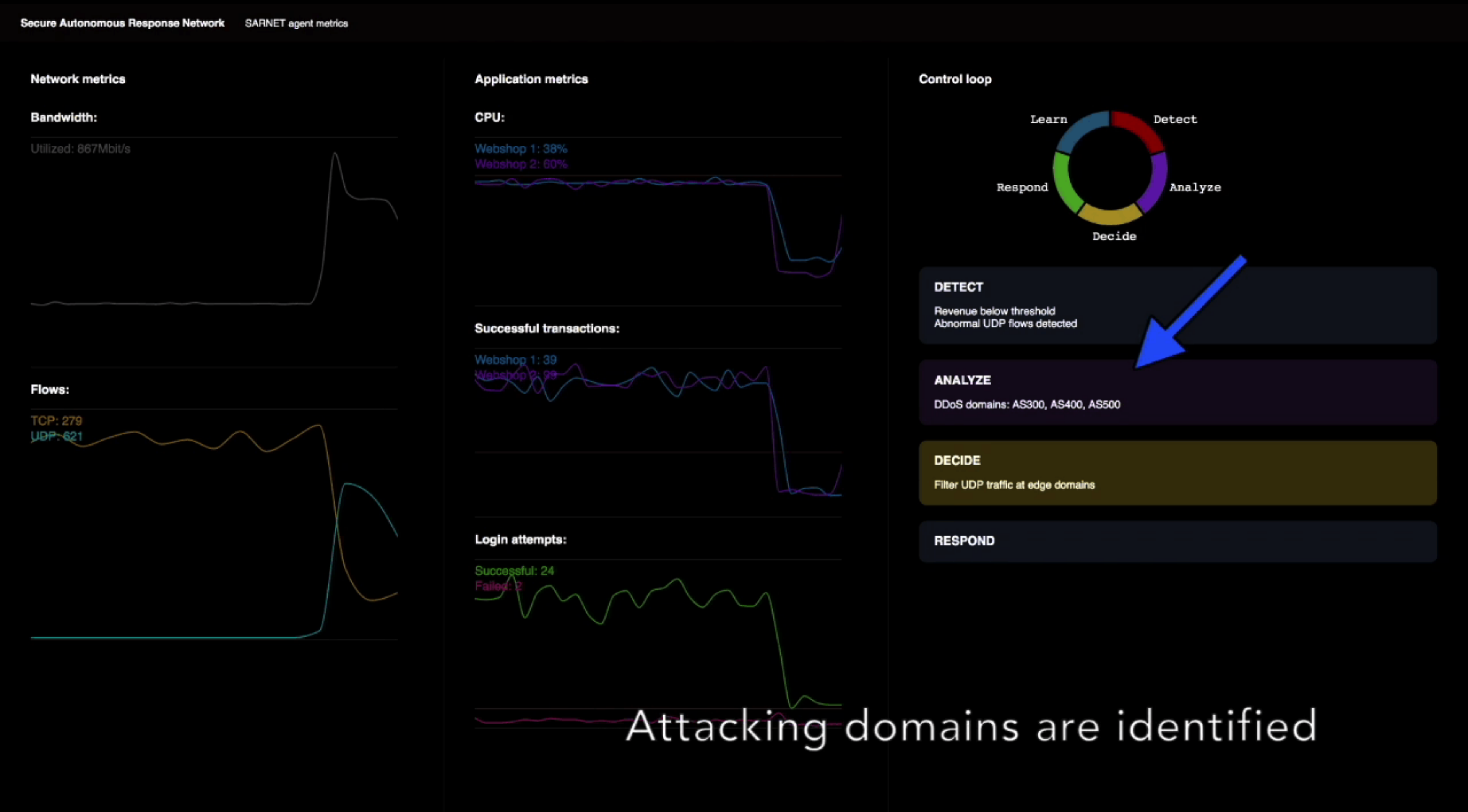
SC16 DEMO SARNET Operational Level



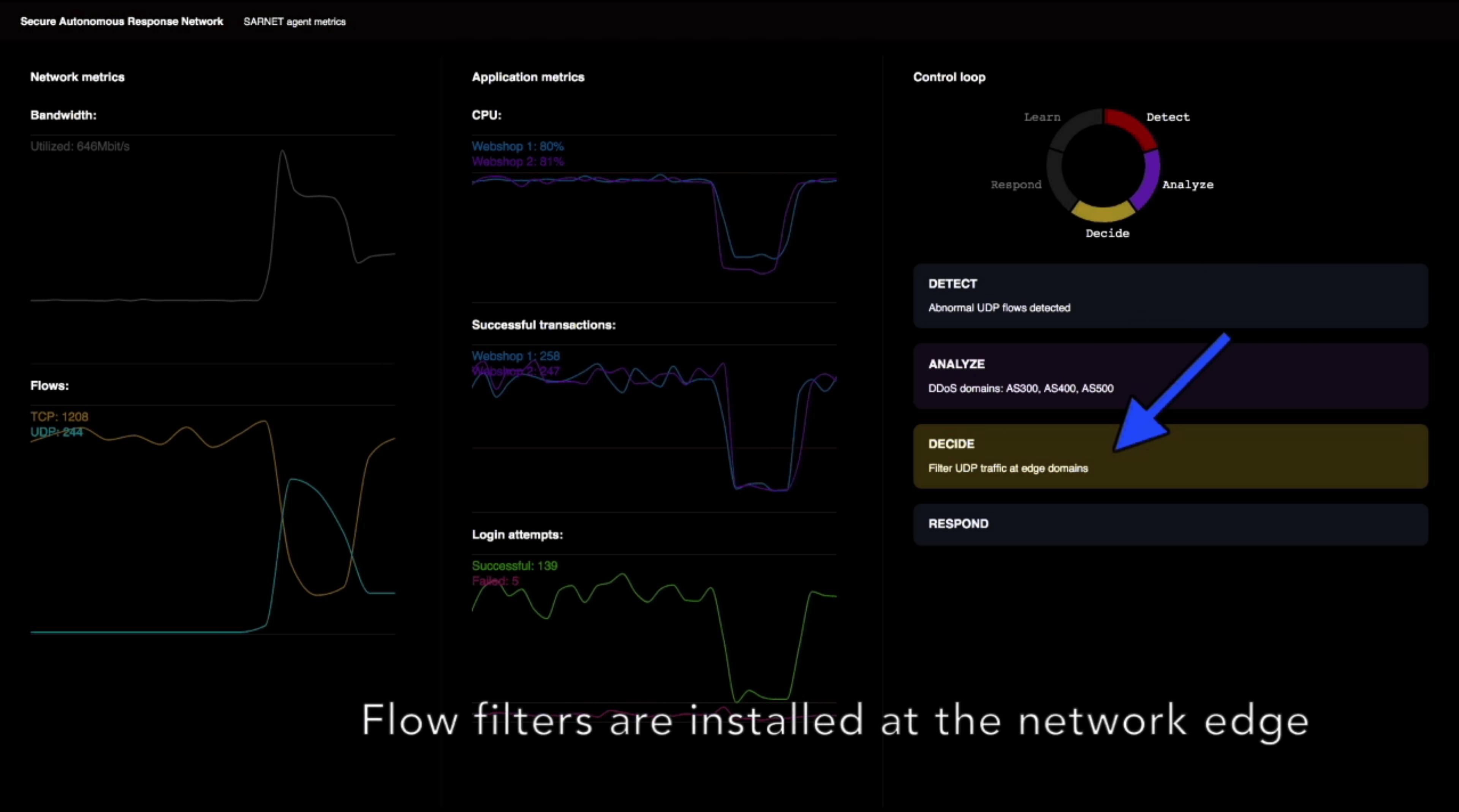
SC16 DEMO SARNET Operational Level



SC16 DEMO SARNET Operational Level

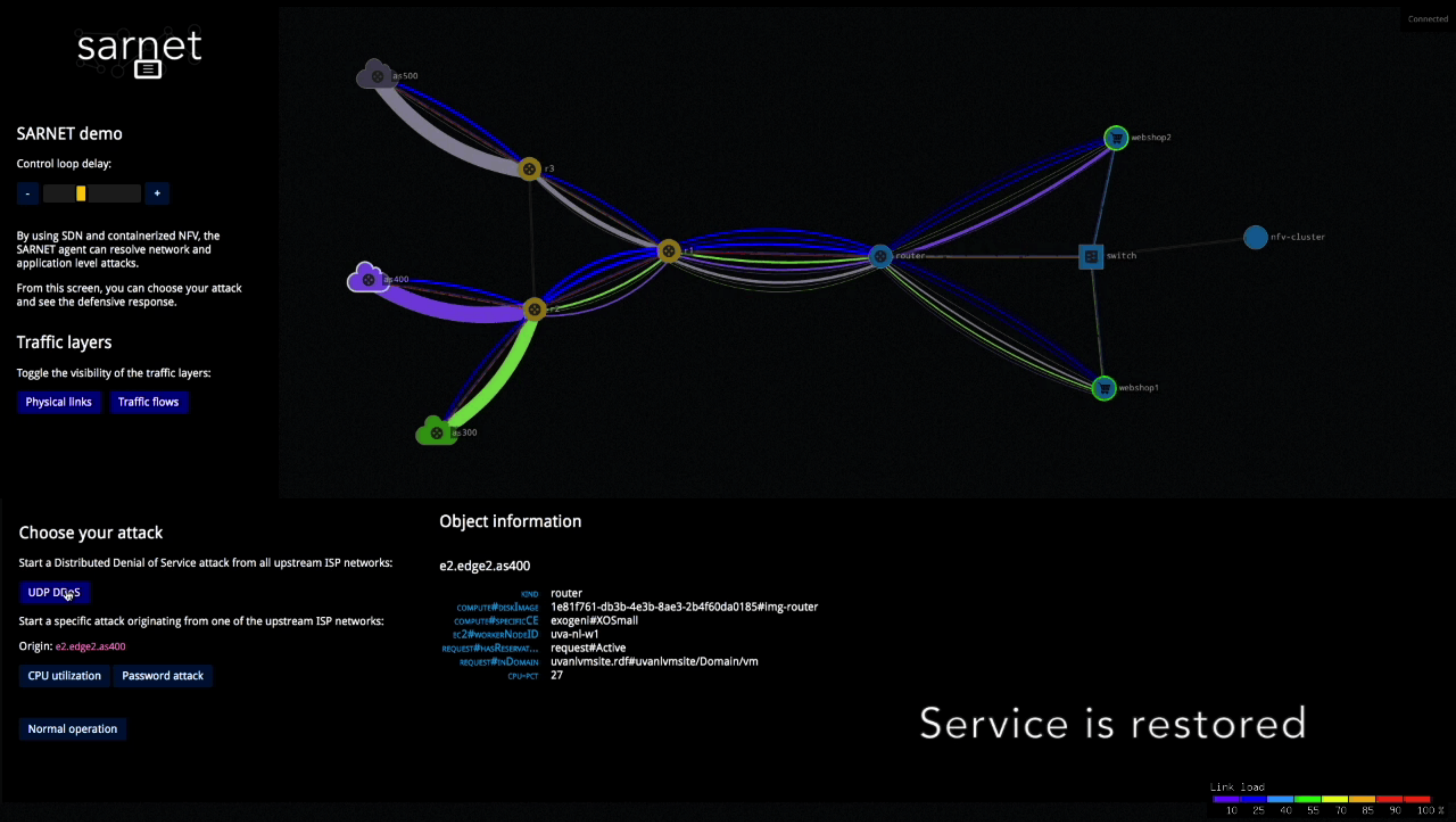


SC16 DEMO SARNET Operational Level



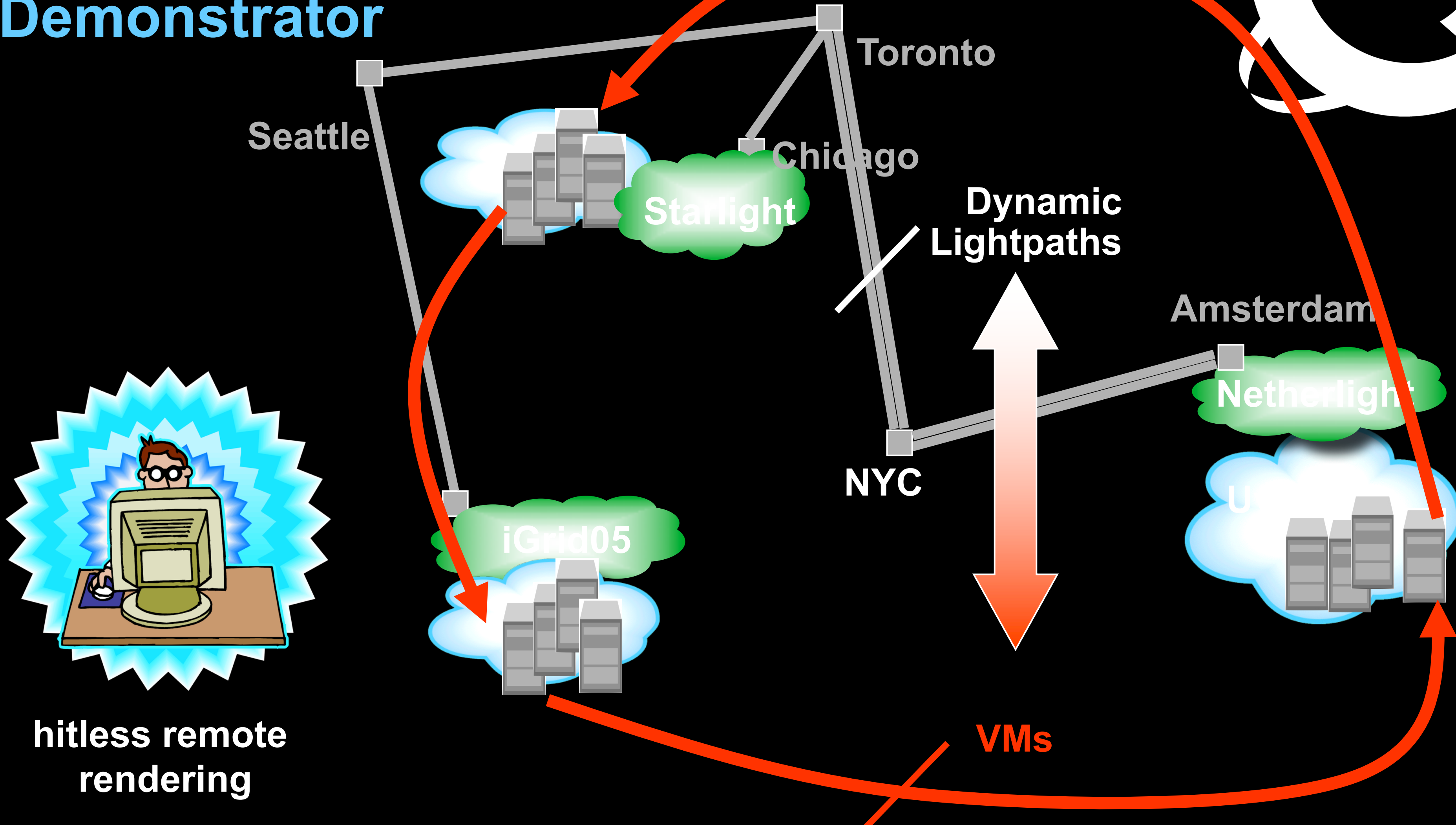
Flow filters are installed at the network edge

SC16 DEMO SARNET Operational Level



The VM Turntable Demonstrator

iGrid2005
SC2005



The VMs that are live-migrated run an iterative search-refine-search workflow against data stored in different databases at the various locations. A user in San Diego gets hitless rendering of search progress as VMs spin around

Experiment outcomes
Note, this was in 2005!

We have demonstrated seamless, live migration of VMs over WAN

For this, we have realized a network service that

- Exhibits predictable behavior; tracks endpoints
- Flex bandwidth upon request by credited applications
- Doesn't require peak provisioning of network resources

Pipelining bounds the downtime in spite of high RTTs

- San Diego – Amsterdam, 1GE, RTT = 200 msec, downtime <= 1 sec
- Back to back, 1GE, RTT = 0.2-0.5 msec, downtime = ~0.2 sec*
- **Clark et al. NSDI 05 paper. Different workloads*

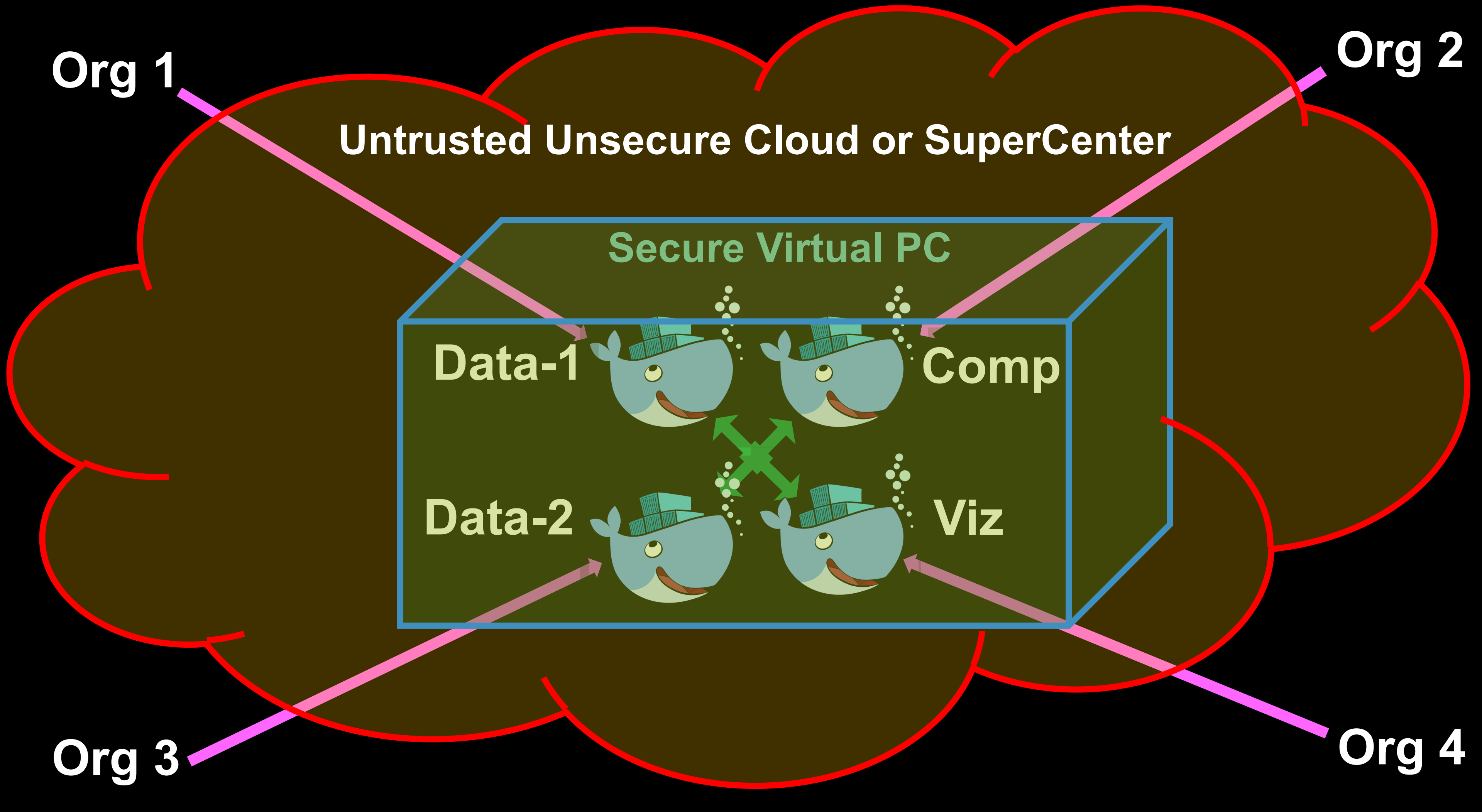
VM + Lightpaths across MAN/WAN are deemed a powerful and general alternative to RPC, GRAM approaches

We believe it's a representative instance of active cpu+data+net orchestration

F. Travostino, P. Daspit, L. Gommans, C. Jog, C.T.A.M. de Laat, J. Mambretti, I. Monga, B. van Oudenaarde, S. Raghunath and P.Y. Wang, "Seamless Live Migration of Virtual Machines over the MAN/WAN", Future Generation Computer Systems, Volume 22, Issue 8, October 2006, Pages 901-907.

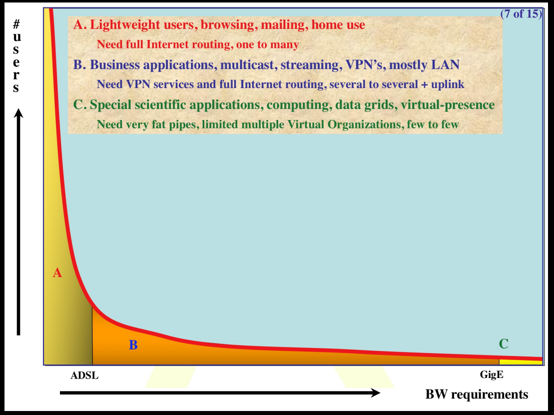
Secure Policy Enforced Data Processing

- Bringing data and processing software from competing organisations together for common goal
- Docker with encryption, policy engine, certs/keys, blockchain and secure networking
- Data Docker (virtual encryped hard drive)
- Compute Docker (protected application, signed algorithms)
- Visualization Docker (to visualize output)



Areas of research

- Each domain its own AI on networks.
  - Multiple AI's fighting on my behalf?
- A-B-C slide
  - Where makes what AI sense?



- Many layers of complexity and abstraction.
  - Can AI help to understand and debug?
  - Can it explicitly understand? Reveal a model?
- Probabilities are badly understood in AI
  - How to deal with false positives?
  - Ethical issues?
  - Trust issues?
  - Intention issues?

Critical notes

- We created complexity
- Huge number of actors (devices)
- Millions of lines of codes
- We have shrinking trust in the Internet
- Let’s throw in another hunderd-thousend lines of code! Good luck…
- Complexity encapsulation
- Do we have enough information for RL - ML?
- Do we understand what the Machine needs to learn?

Acknowledgements

WP 20.3 SeSI: Security of (Virtual) e-Science Infrastructure

Cees de Laat
Matthijs Koot, Leon Gommans, Guido van 't Noordeinde

overhead
production

☺

WP 20.11 SARNET: Security Autonomous Response with programmable NETworks

Cees de Laat
Leon Gommans, Rodney Wilson, Rob Meijer
Tom van Engers, Marc Lyonnais, Paola Grosso, Frank Fransen, Ameneh Deljoo, Ralph Koning, Ben de Graaff, Gleb Polevoy

overhead

production

☹

Why?

Because we can!