DE CIX

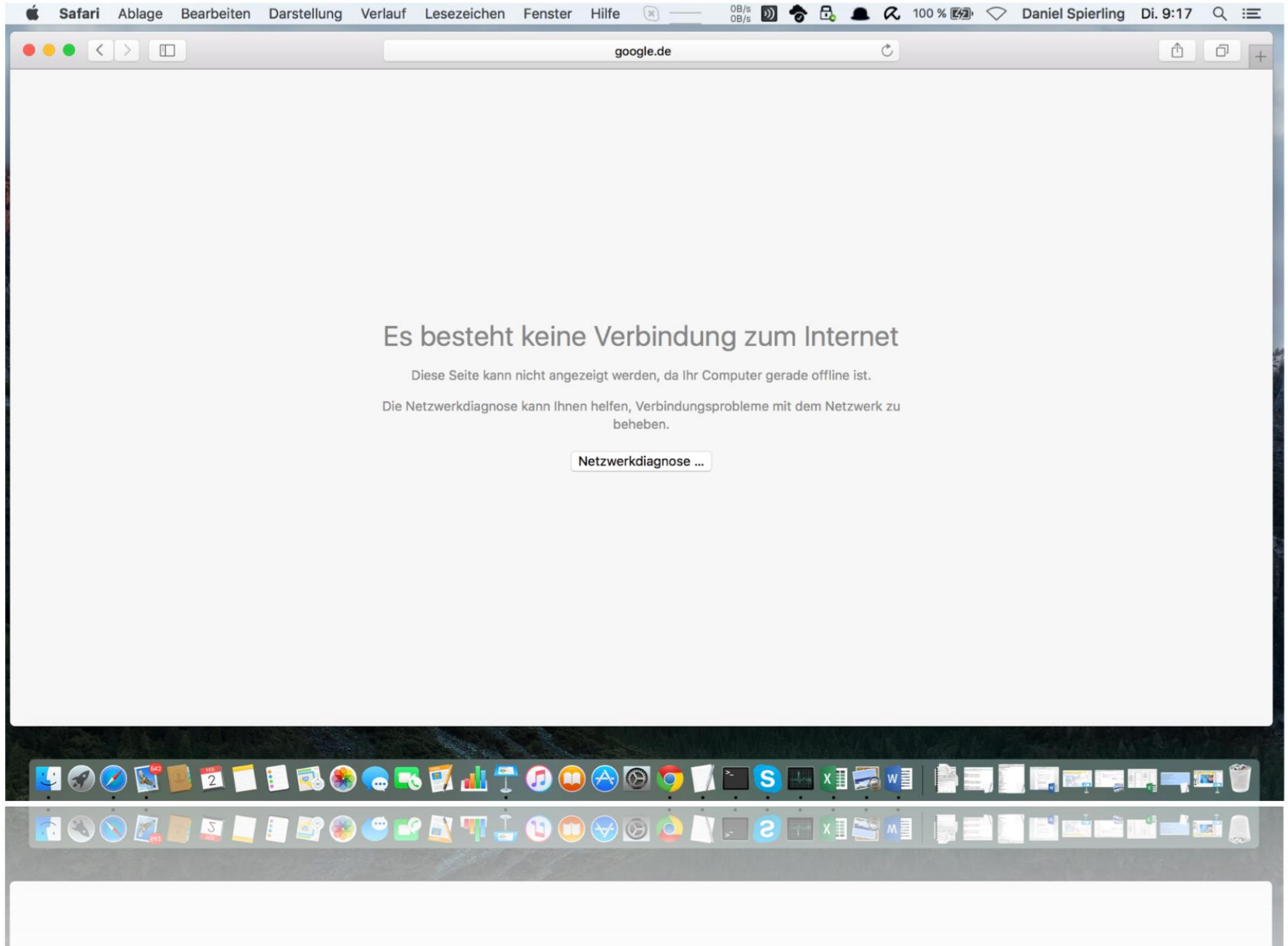
Safe and Secure Operation of an IXP

Christoph Dietzel

Head of Research & Development

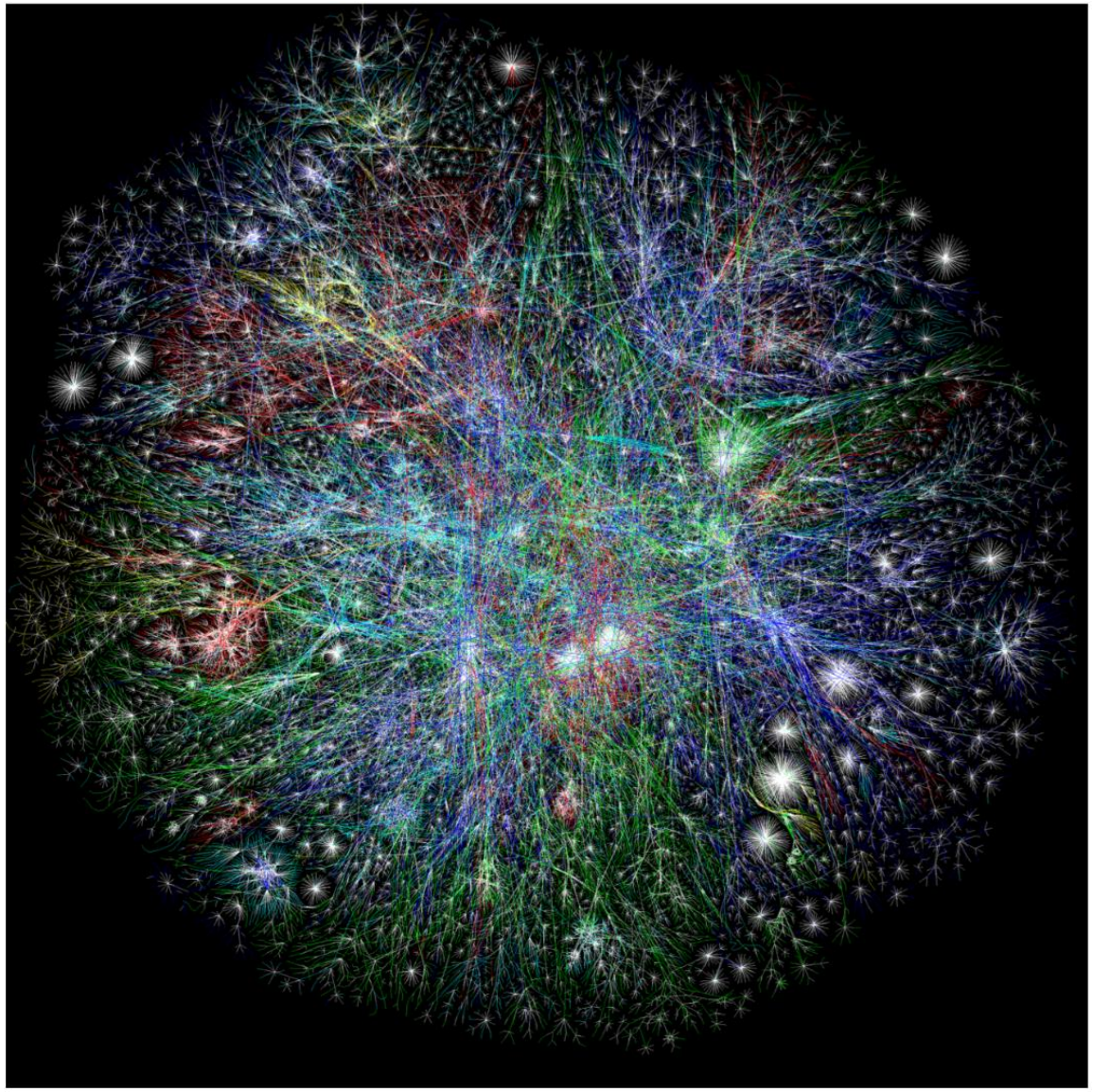
Where networks meet

www.de-cix.net



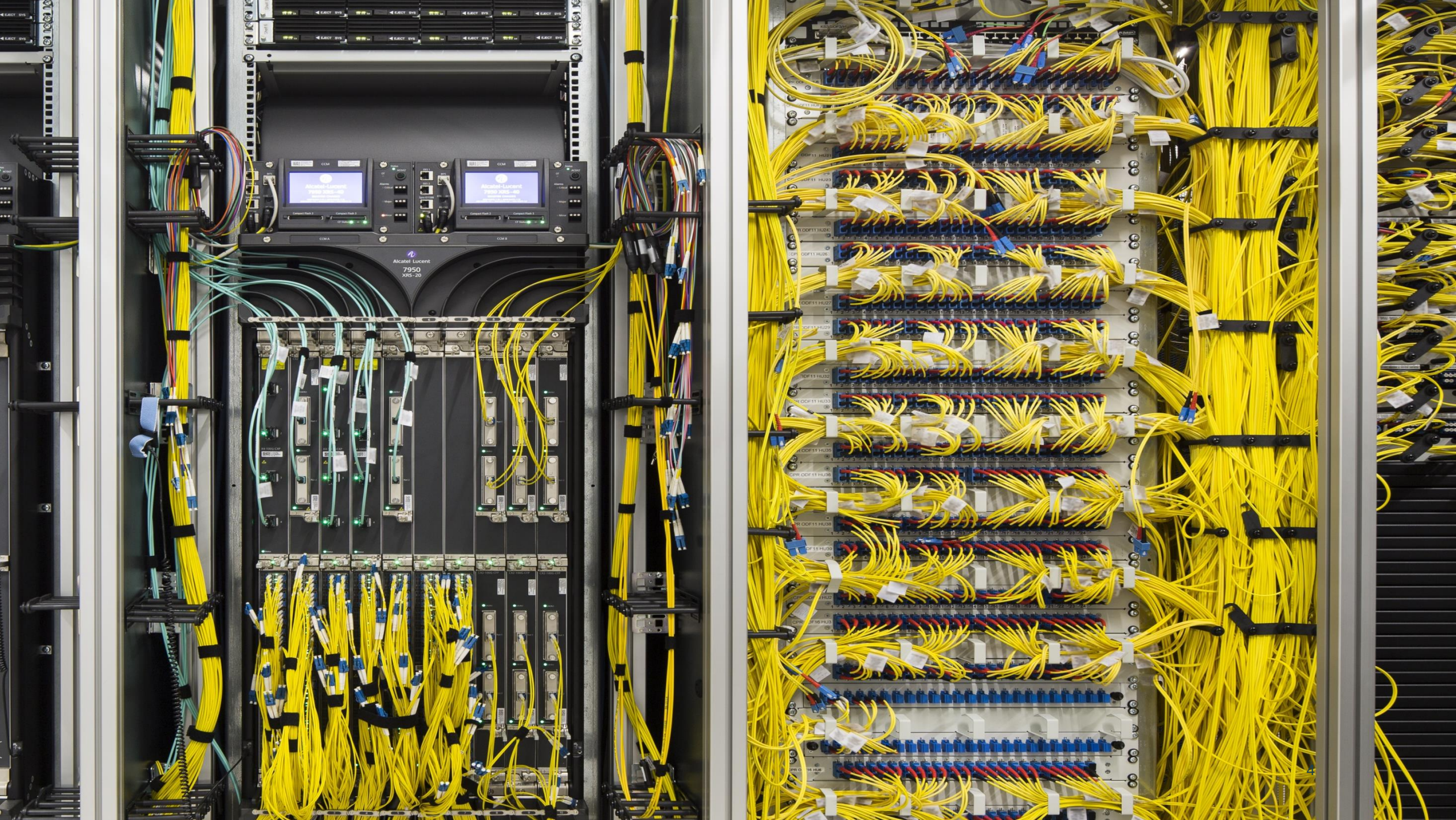
Where networks meet

www.de-cix.net

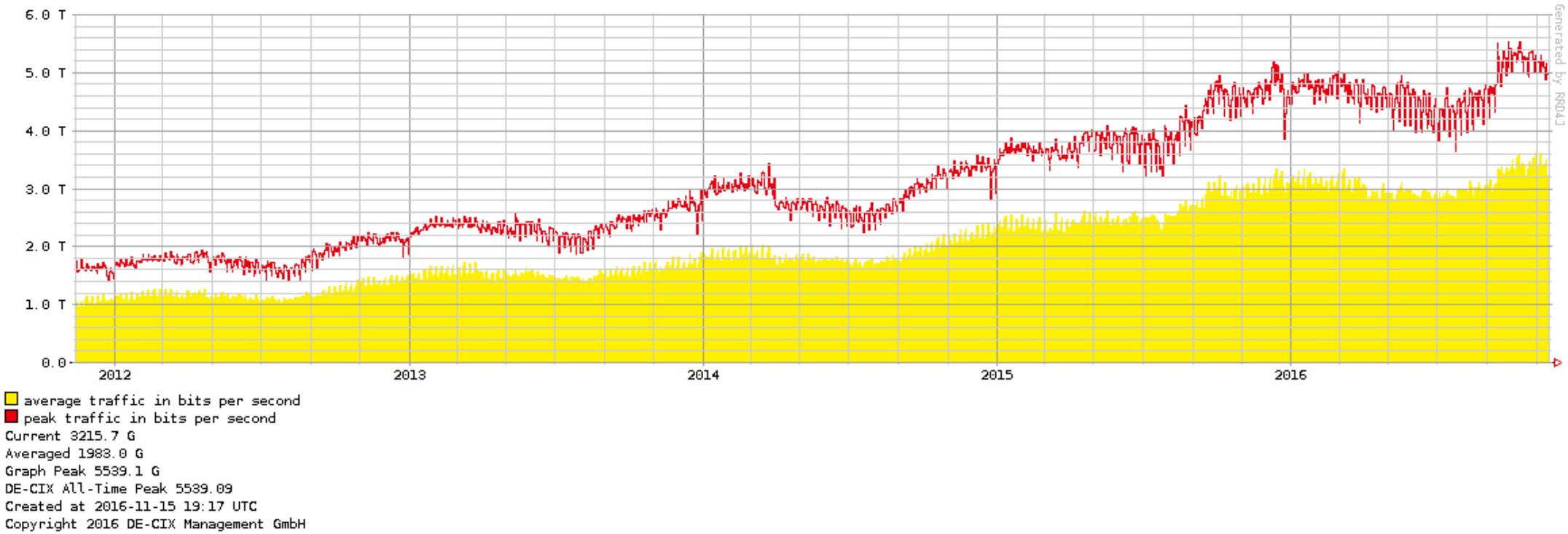


Where networks meet

www.de-cix.net


Alcatel Lucent
7950
XRS-20

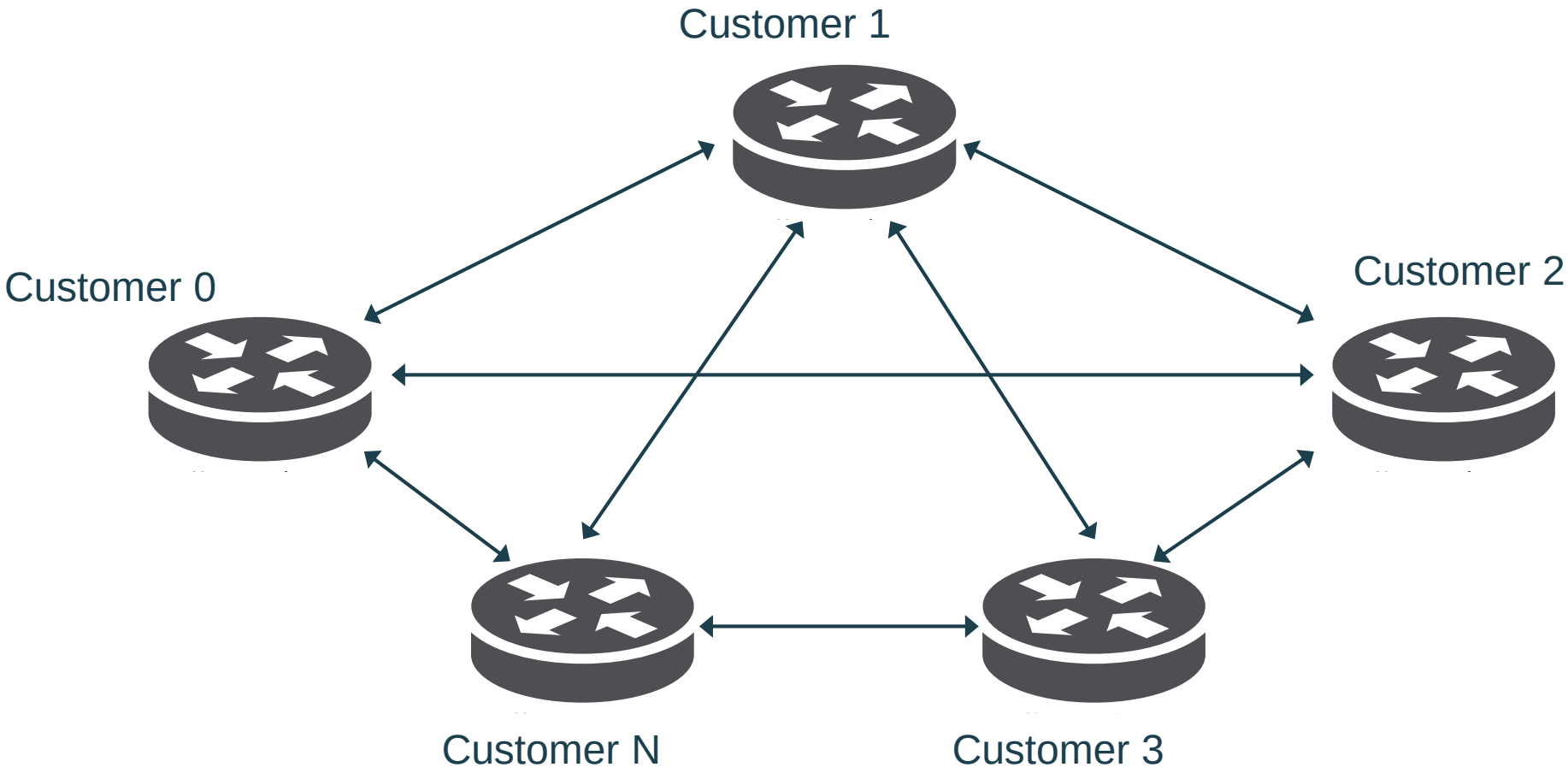
DE-CIX Frankfurt last 5 years traffic - statistic



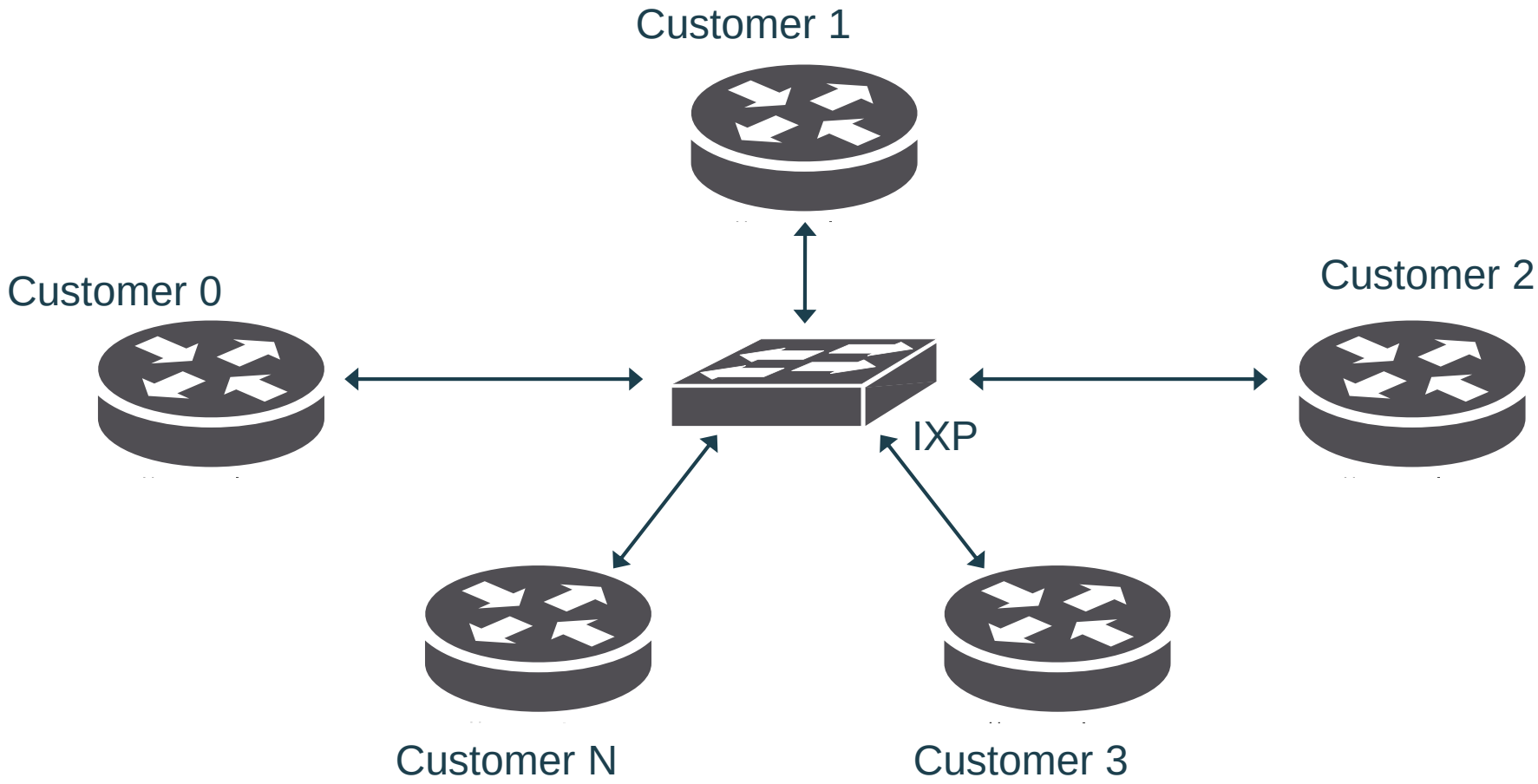
Where networks meet

www.de-cix.net

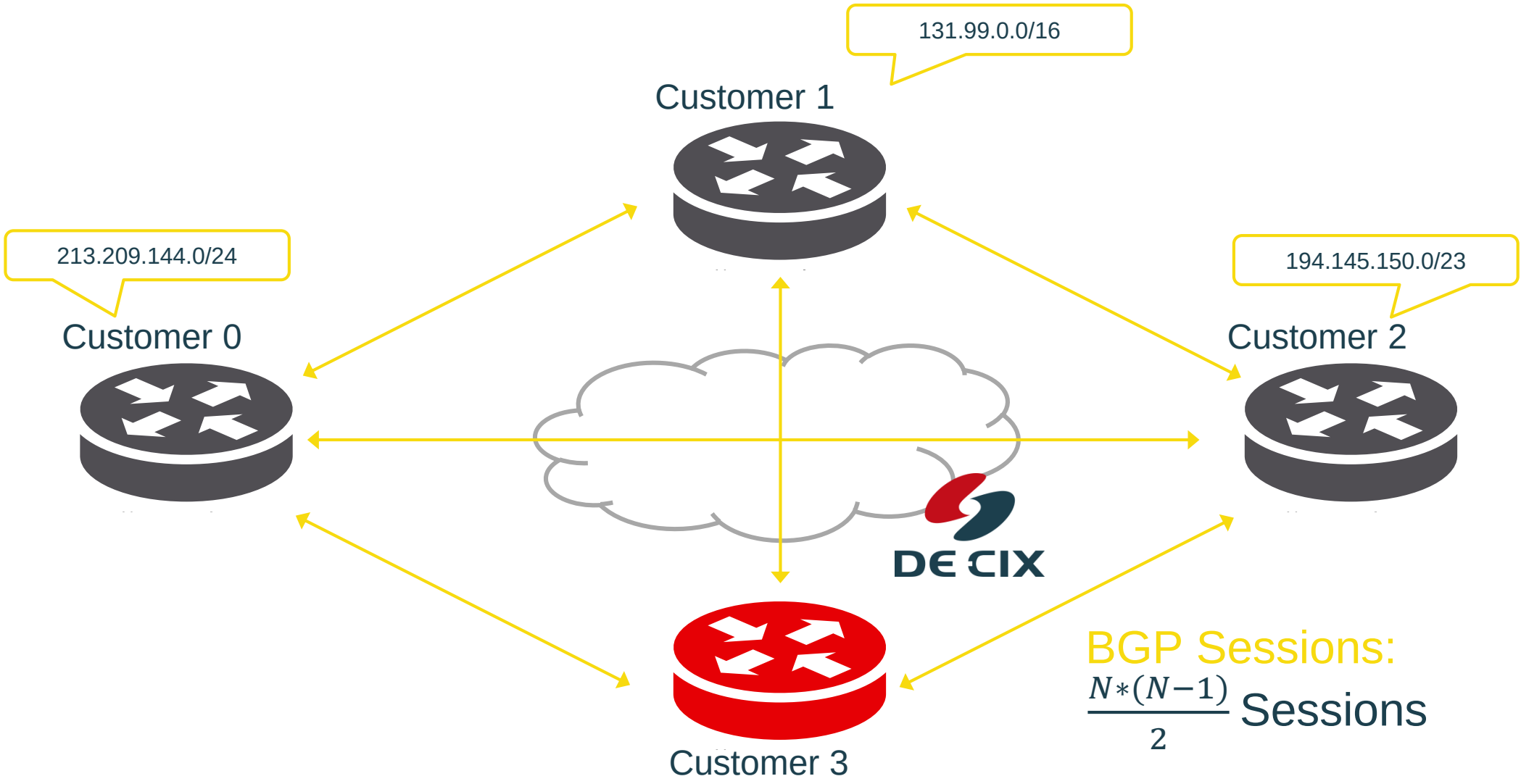
Generic IXP Topology – Layer 3 View



Generic IXP Topology – Layer 2 View



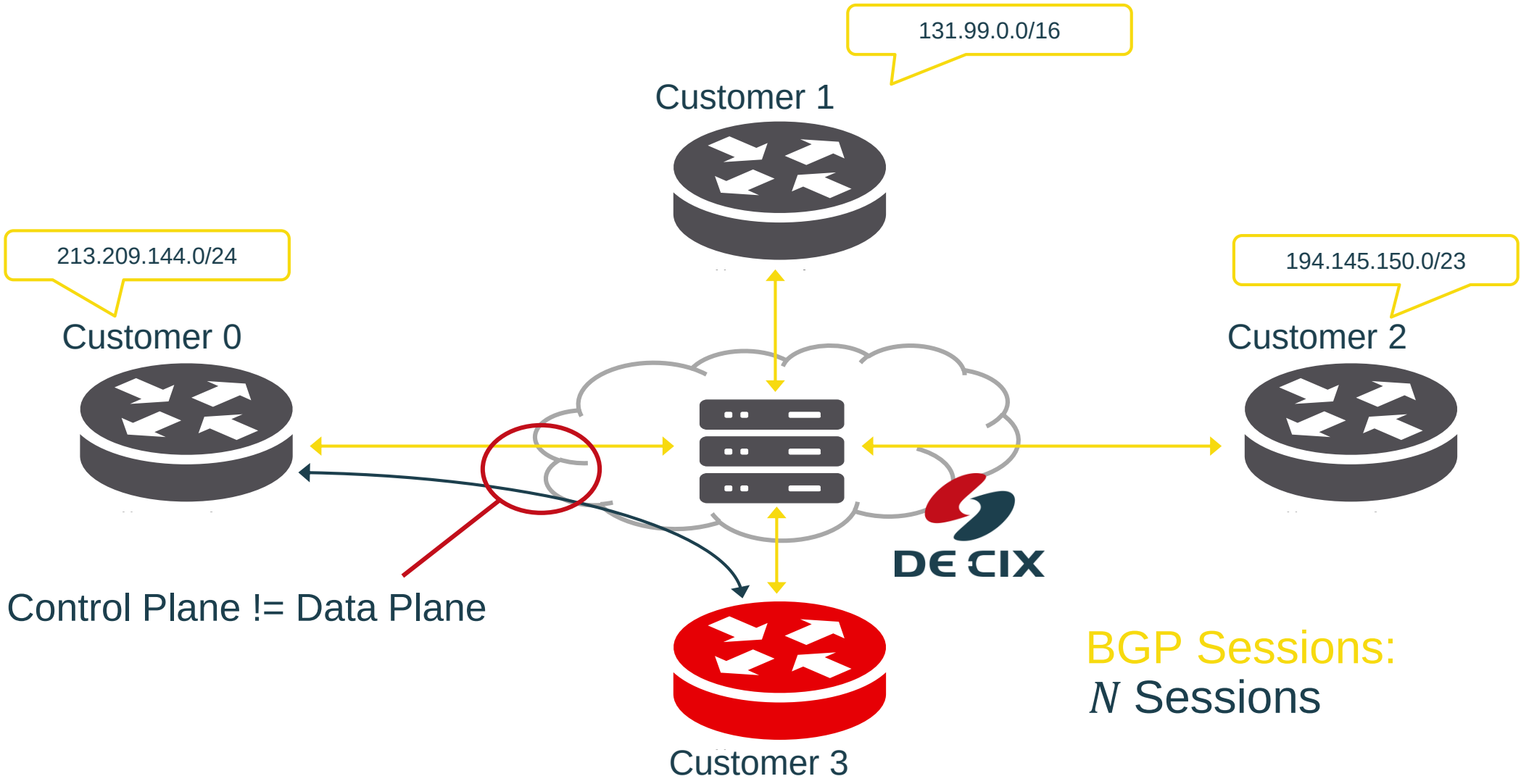
Route Server I



BGP Sessions:

$\frac{N*(N-1)}{2}$ Sessions

Route Server II



BGP Sessions:
N Sessions

Where networks meet

www.de-cix.net

Safe and Secure Route Server Software

Where networks meet

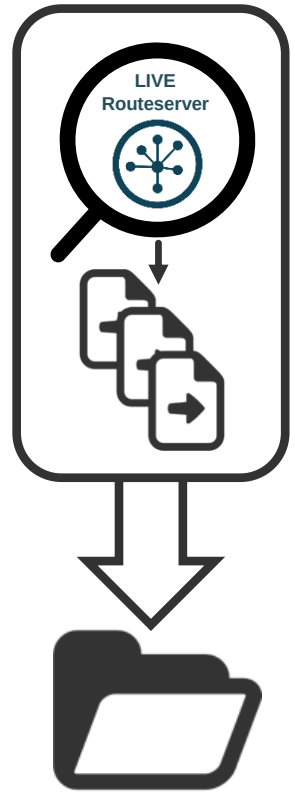
www.de-cix.net

Realistic Peer Generation & Simulation

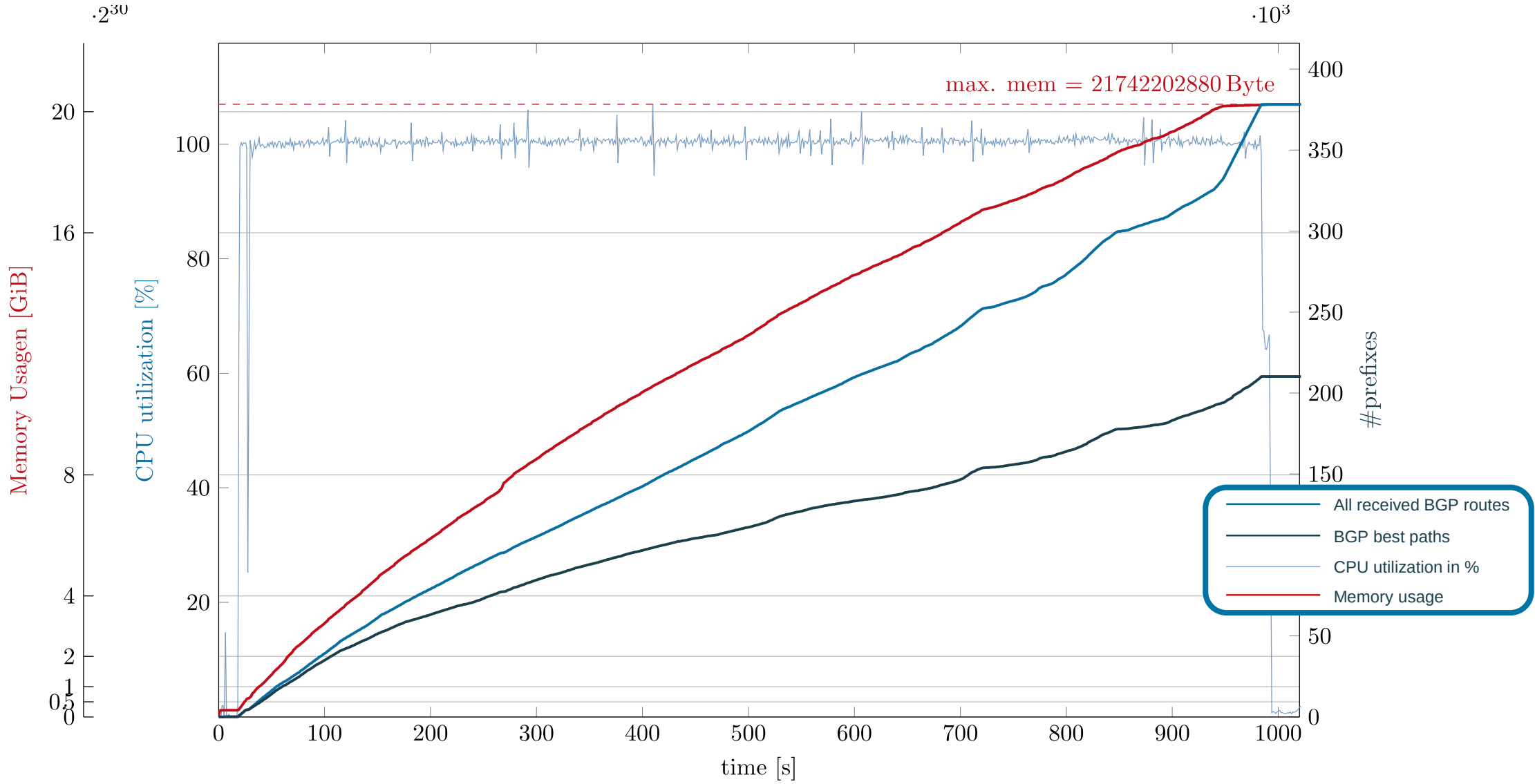
- Emulation of peers with ExaBGP (https://github.com/Exa-Networks/exabgp)
- One ExaBGP process per peer

- Real world peer snapshots from DE-CIX route servers
- Auto generated ExaBGP configs incl.:
  - ➔ Session Hold timers
  - ➔ Announced prefixes
    - ➔ AS-Path, BGP next hop, local pref, (extended) BGP communities, ...

- Export from per-customer RIBs
  - ➔ Includes all filtered prefixes as well
  - ➔ ~ 720.000 routes in ExaBGP configs

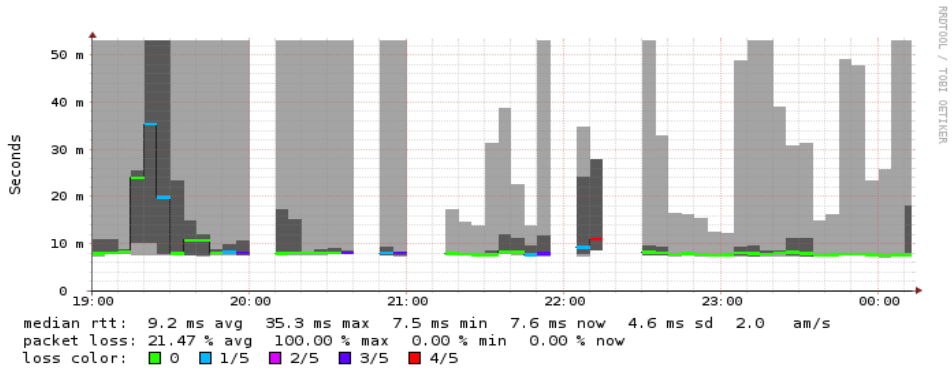


What do we measure?



Simulation of L2 problems

- Emulate packet loss and delay with an existing tool (https://github.com/tylertreat/comcast)
  - Makes use of iptables and tc (on Linux)
- Simulate L2 problems and emerging peer flaps
  - High Loss leading to missed keepalives
  - Will result in peer flaps
- Example: simulate entire switch / linecard failures
  - generate 100% packet loss for a given time
  - No flaps, but high number of sessions go down
  - RS needs to calculate new best paths / send withdraws

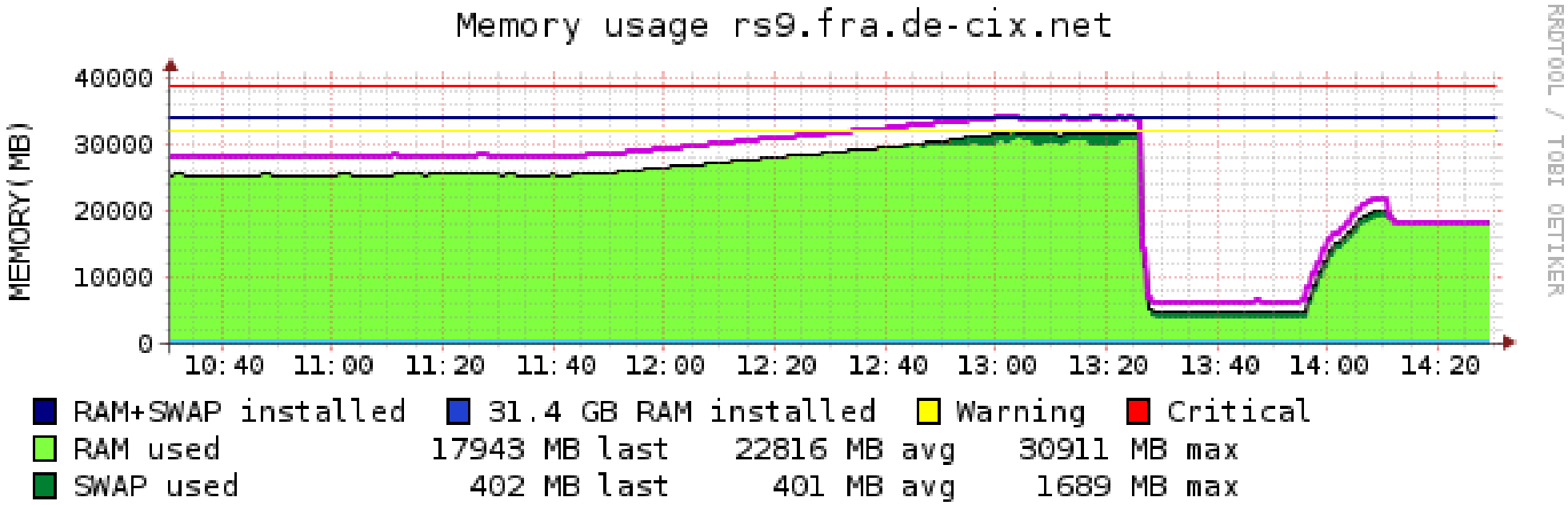


BIRD Memory Leak / Cisco Bug

Example I

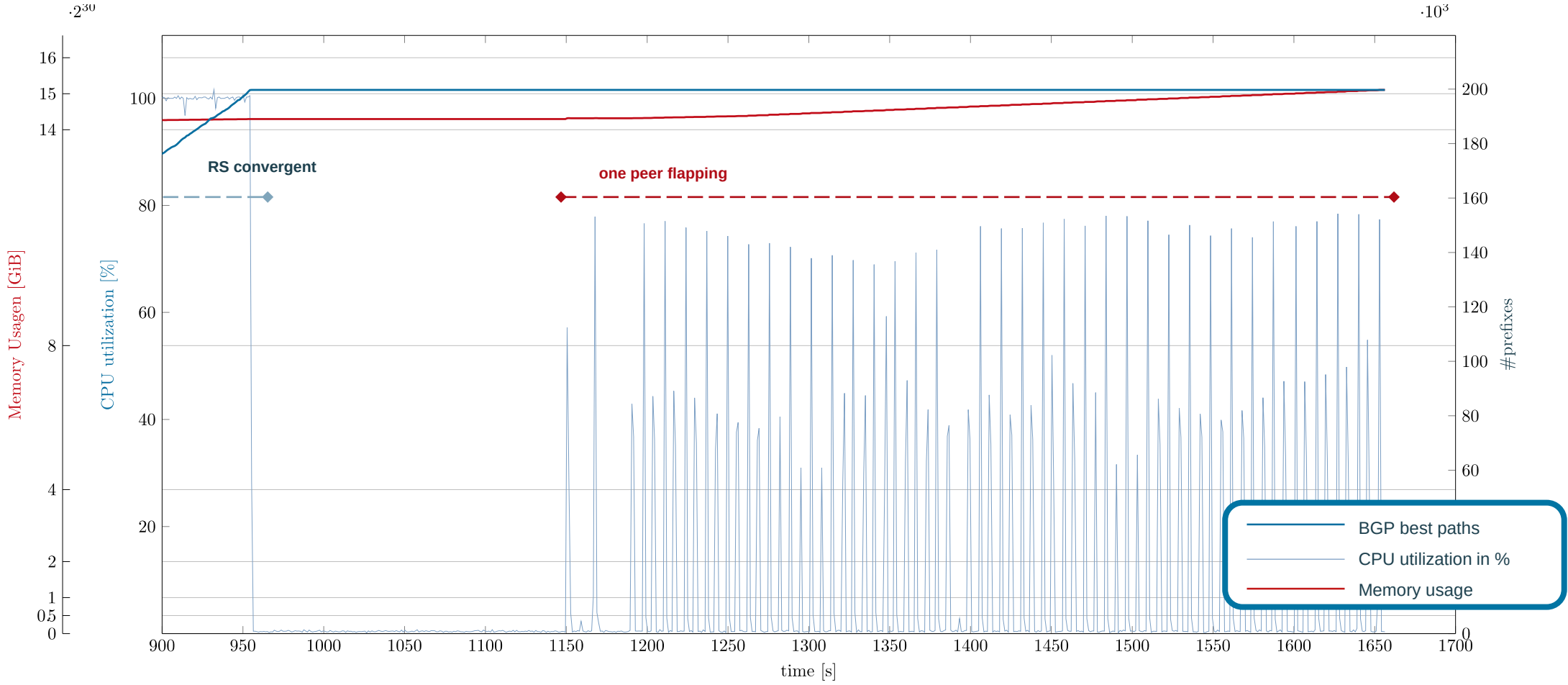
Detection and investigation of a memory leak in BIRD

- Customer facing Cisco bug CSCus56036, graceful restart (4s)
- Memory leak, BIRD process killed by OoM killer
- Comunicate with developers, bug fixed in BIRD v1.6.3
- Reproduce scenario and test effectiveness of fix



Graceful Restart bug

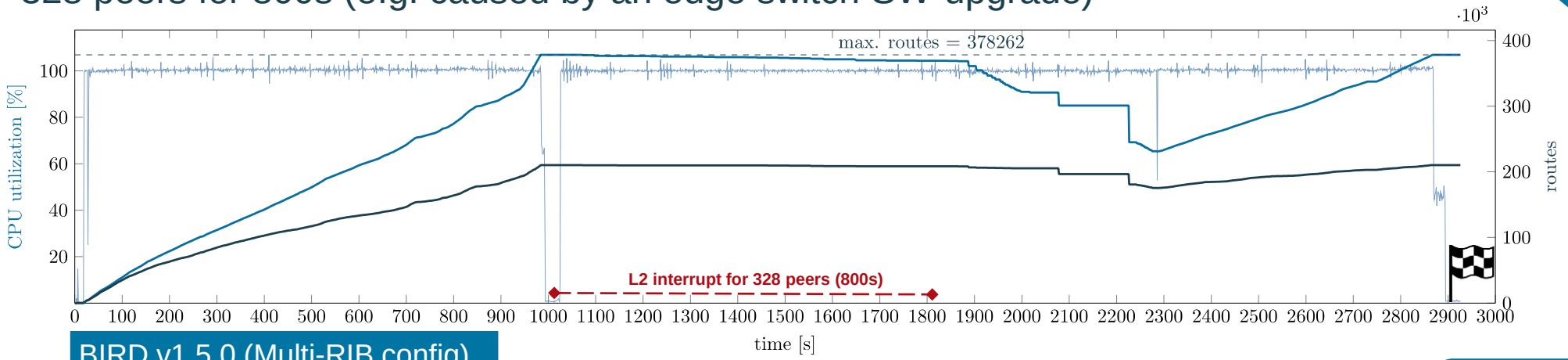
Example I



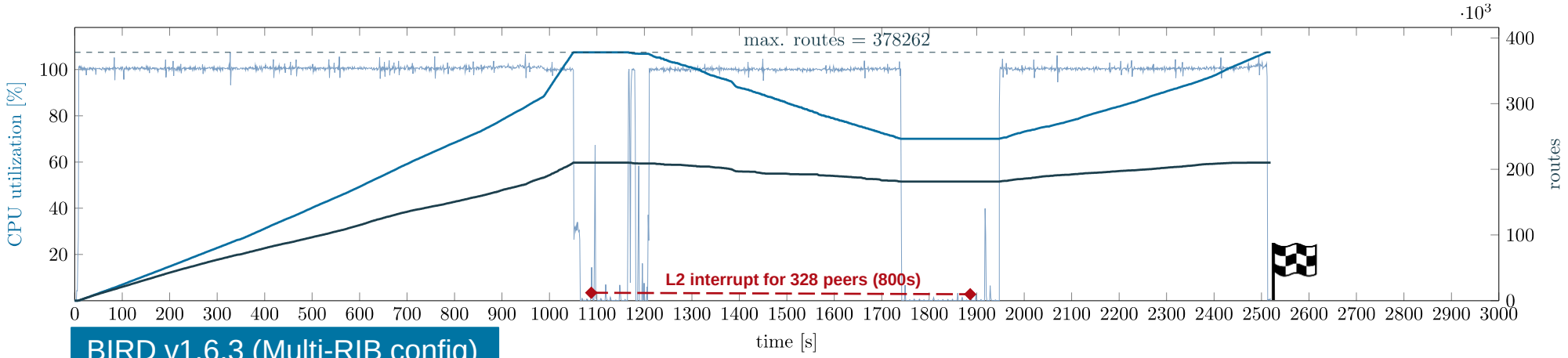
Simulation of a realistic L2 disruption

Example II

328 peers for 800s (e.g. caused by an edge switch SW-upgrade)



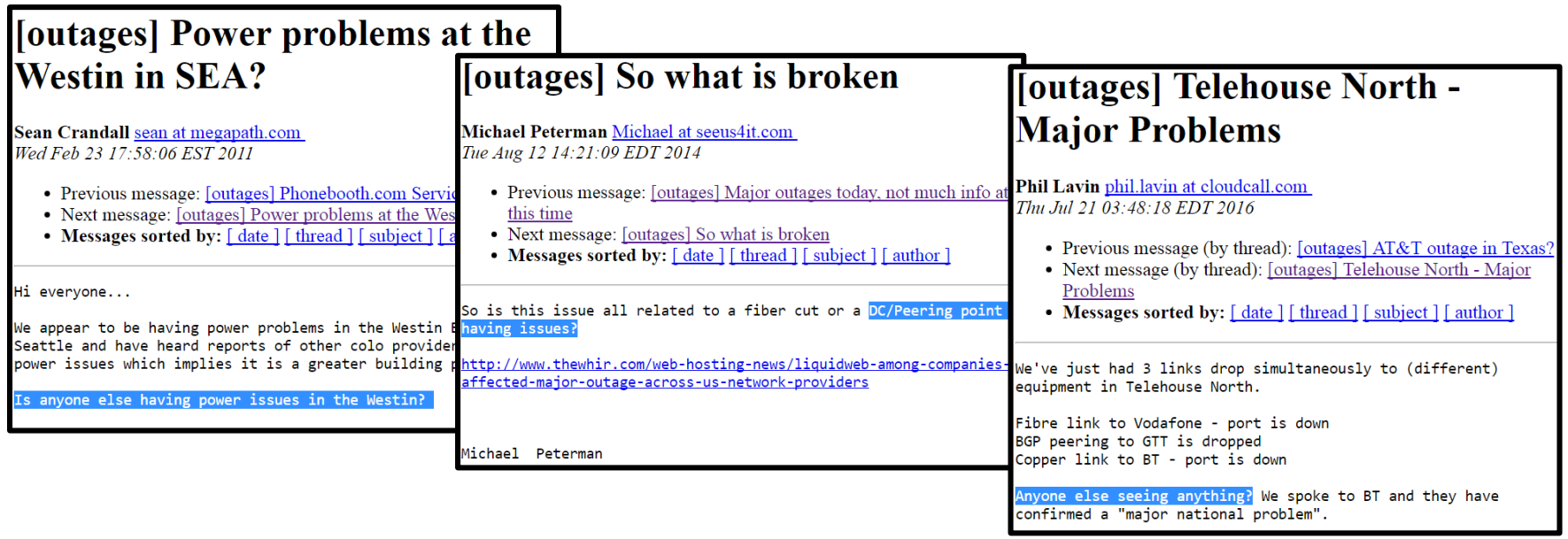
BIRD v1.5.0 (Multi-RIB config)



BIRD v1.6.3 (Multi-RIB config)

All received BGP routes
BGP best paths
CPU utilization in %

State of Internet Robustness



[outages] Power problems at the Westin in SEA?

Sean Crandall sean at megapath.com
Wed Feb 23 17:58:06 EST 2011

- Previous message: [outages] Phonebooth.com Servi
- Next message: [outages] Power problems at the Wes
- **Messages sorted by:** [date] [thread] [subject] [a

Hi everyone...

We appear to be having power problems in the Westin B
Seattle and have heard reports of other colo provider
power issues which implies it is a greater building p

Is anyone else having power issues in the Westin?

[outages] So what is broken

Michael Peterman Michael at seeus4it.com
Tue Aug 12 14:21:09 EDT 2014

- Previous message: [outages] Major outages today, not much info at this time
- Next message: [outages] So what is broken
- **Messages sorted by:** [date] [thread] [subject] [author]

So is this issue all related to a fiber cut or a DC/Peering point having issues?

http://www.thewhir.com/web-hosting-news/liquidweb-among-companies-affected-major-outage-across-us-network-providers

Michael Peterman

[outages] Telehouse North - Major Problems

Phil Lavin phil.lavin at cloudcall.com
Thu Jul 21 03:48:18 EDT 2016

- Previous message (by thread): [outages] AT&T outage in Texas?
- Next message (by thread): [outages] Telehouse North - Major Problems
- **Messages sorted by:** [date] [thread] [subject] [author]

We've just had 3 links drop simultaneously to (different) equipment in Telehouse North.

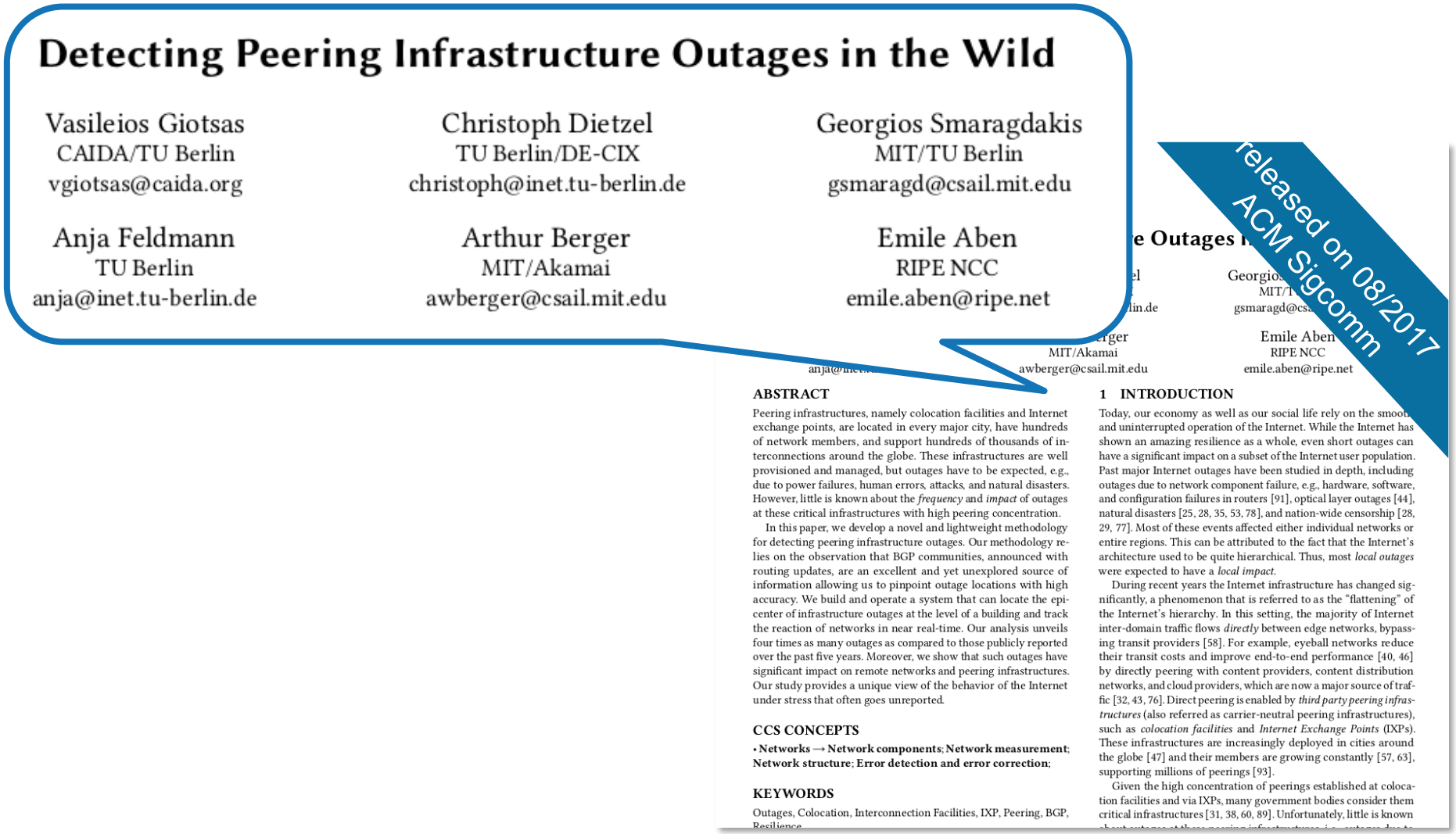
Fibre link to Vodafone - port is down
BGP peering to GTT is dropped
Copper link to BT - port is down

Anyone else seeing anything? We spoke to BT and they have confirmed a "major national problem".

Research Goals

- Outage detection
- Outage localization
- Outage tracking

Detecting Peering Infrastructure Outages in the Wild



Detecting Peering Infrastructure Outages in the Wild

Vasileios Giotsas
CAIDA/TU Berlin
vgiotsas@caida.org

Christoph Dietzel
TU Berlin/DE-CIX
christoph@inet.tu-berlin.de

Georgios Smaragdakis
MIT/TU Berlin
gsmaragd@csail.mit.edu

Anja Feldmann
TU Berlin
anja@inet.tu-berlin.de

Arthur Berger
MIT/Akamai
awberger@csail.mit.edu

Emile Aben
RIPE NCC
emile.aben@ripe.net

Released on 08/2017
ACM Sigcomm

ABSTRACT

Peering infrastructures, namely colocation facilities and Internet exchange points, are located in every major city, have hundreds of network members, and support hundreds of thousands of interconnections around the globe. These infrastructures are well provisioned and managed, but outages have to be expected, e.g., due to power failures, human errors, attacks, and natural disasters. However, little is known about the *frequency* and *impact* of outages at these critical infrastructures with high peering concentration.

In this paper, we develop a novel and lightweight methodology for detecting peering infrastructure outages. Our methodology relies on the observation that BGP communities, announced with routing updates, are an excellent and yet unexplored source of information allowing us to pinpoint outage locations with high accuracy. We build and operate a system that can locate the epicenter of infrastructure outages at the level of a building and track the reaction of networks in near real-time. Our analysis unveils four times as many outages as compared to those publicly reported over the past five years. Moreover, we show that such outages have significant impact on remote networks and peering infrastructures. Our study provides a unique view of the behavior of the Internet under stress that often goes unreported.

CCS CONCEPTS

• **Networks** → **Network components**; **Network measurement**; **Network structure**; **Error detection and error correction**;

KEYWORDS

Outages, Colocation, Interconnection Facilities, IXP, Peering, BGP, Resilience

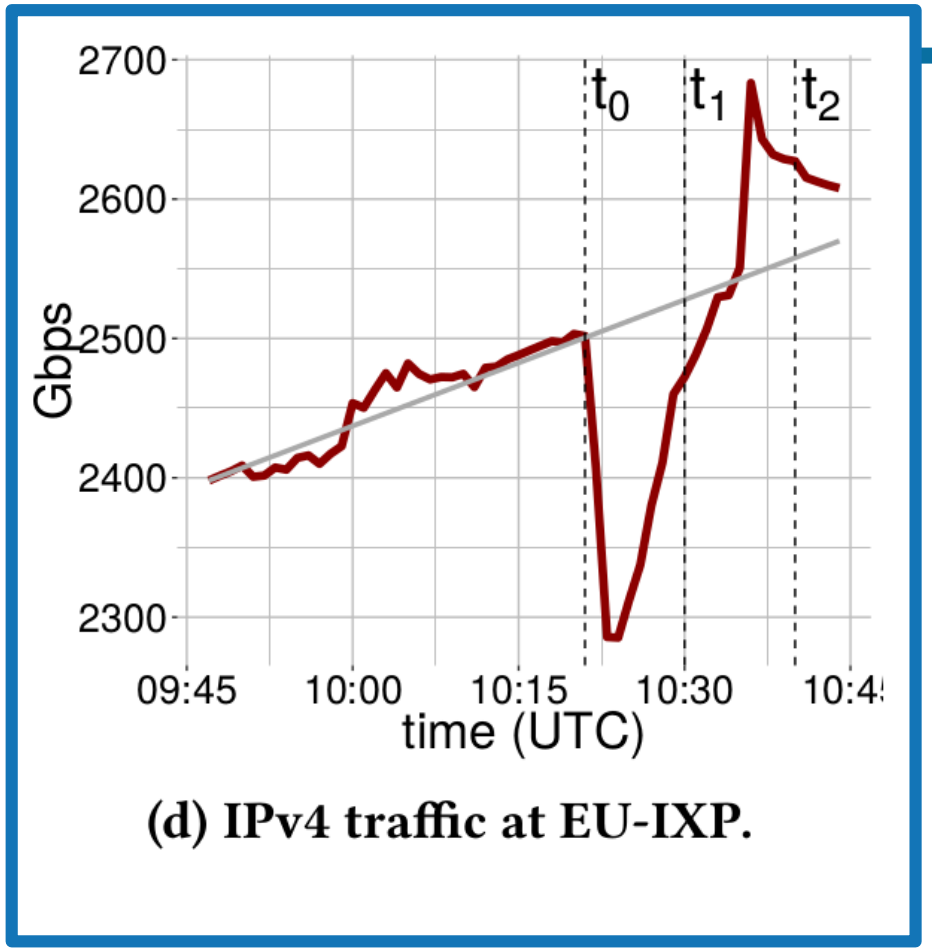
1 INTRODUCTION

Today, our economy as well as our social life rely on the smooth and uninterrupted operation of the Internet. While the Internet has shown an amazing resilience as a whole, even short outages can have a significant impact on a subset of the Internet user population. Past major Internet outages have been studied in depth, including outages due to network component failure, e.g., hardware, software, and configuration failures in routers [91], optical layer outages [44], natural disasters [25, 28, 35, 53, 78], and nation-wide censorship [28, 29, 77]. Most of these events affected either individual networks or entire regions. This can be attributed to the fact that the Internet's architecture used to be quite hierarchical. Thus, most *local outages* were expected to have a *local impact*.

During recent years the Internet infrastructure has changed significantly, a phenomenon that is referred to as the "flattening" of the Internet's hierarchy. In this setting, the majority of Internet inter-domain traffic flows *directly* between edge networks, bypassing transit providers [58]. For example, eyeball networks reduce their transit costs and improve end-to-end performance [40, 46] by directly peering with content providers, content distribution networks, and cloud providers, which are now a major source of traffic [32, 43, 76]. Direct peering is enabled by *third party peering infrastructures* (also referred as carrier-neutral peering infrastructures), such as *colocation facilities* and *Internet Exchange Points* (IXPs). These infrastructures are increasingly deployed in cities around the globe [47] and their members are growing constantly [57, 63], supporting millions of peerings [93].

Given the high concentration of peerings established at colocation facilities and via IXPs, many government bodies consider them critical infrastructures [31, 38, 60, 89]. Unfortunately, little is known

Detecting Peering Infrastructure Outages in the Wild



(d) IPv4 traffic at EU-IXP.

Algorithm 1: Overview of our outage detection and investigation algorithm *explain T_{fail} in the text*

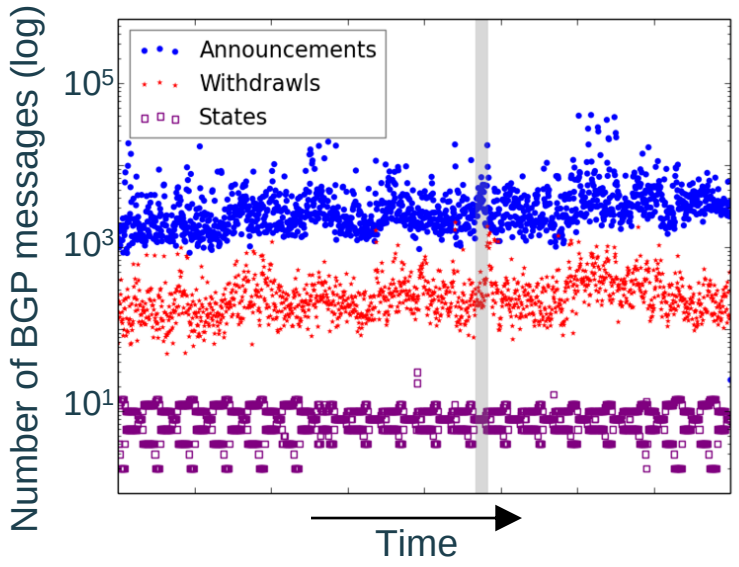
Input: (BGP paths, BGP Community Dictionary, Colocation Map, Targeted Active Measurements)
Output: Location, Time and Duration of a PoP-level Outage

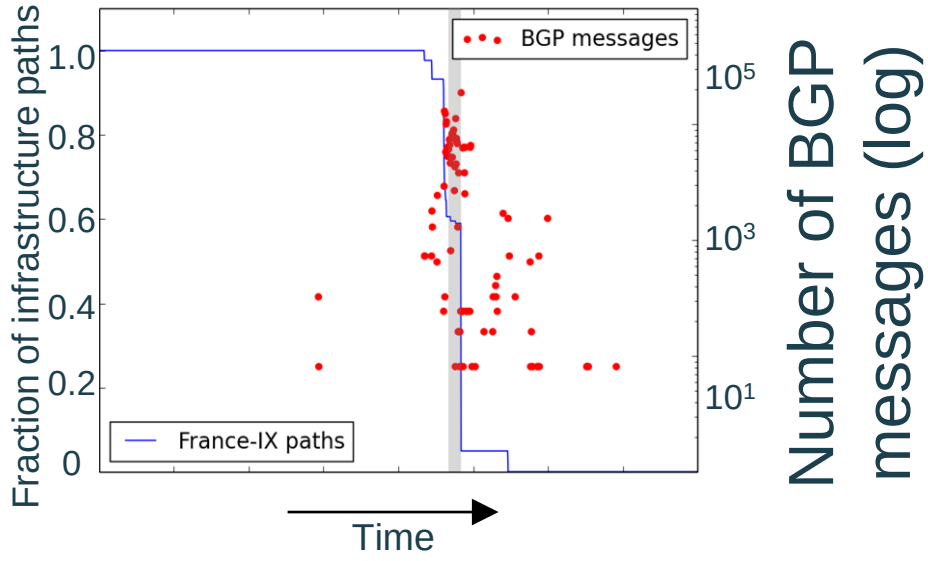
$Paths_{mapped}$ ⟵ Map BGP paths to traversed PoPs based on the attached Communities meta-data;
$Paths^{stable}_{mapped}$ ⟵ Filter-out transient paths;
for *BGP updates in new measurement interval* **do**
  $Paths^{diverted}_{mapped}$ ⟵ calculate how many paths diverted from the PoP in the stable baseline;
  if $\frac{Paths^{diverted}_{mapped}}{Paths^{stable}_{mapped}} > T_{fail}$ **then**

    Signal investigation
    $Signal_{type}$ ⟵ Infer the type of outage signal based on the number of affected ASes and AS links;
    if $Signal_{type}$ *is PoP* **then**

      POP^{BGP}_{type} ⟵ Determine the type of PoP based on the colocation map;
      POP^{trace}_{type} ⟵ Confirm the affected PoP through traceroute queries;
      if $POP^{BGP}_{type} \equiv POP^{trace}_{type}$ **then**
        while $Outage_{state}$ *is True* **do**
          $duration$ ⟵ record the duration of the outage
        return $Outage(time, POP, duration)$

De-noising of BGP Routing Activity

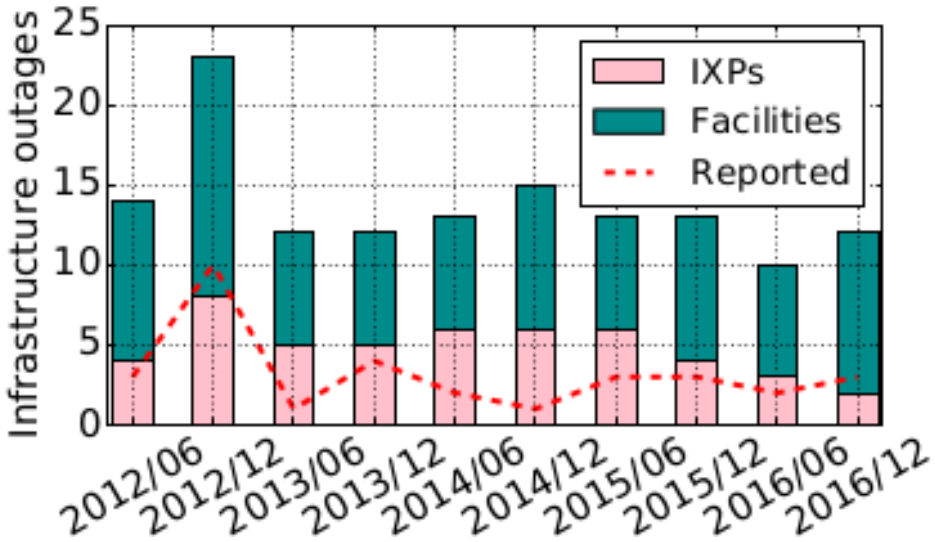


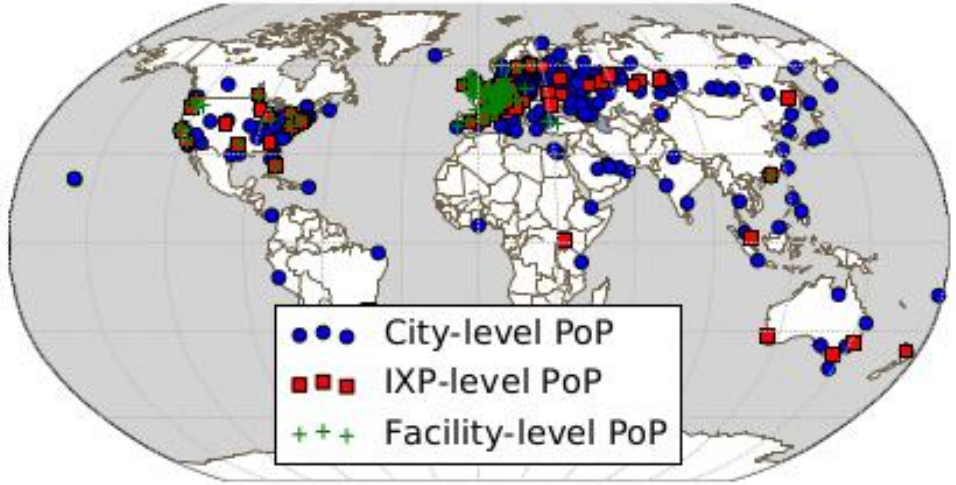


The aggregated activity of BGP messages (updates, withdrawals, states) provides no outage indication.

The BGP activity filtered using communities provides **strong outage signal**.

Detecting Peering Infrastructure Outages in the Wild





Conclusion

- Crucial to understand important software (and it’s limitation)
- Simulation/emulation is key to understand distruptions
- Help to improve the entire Internet eco-system