InterNetworking – made physical

Nikhef

ICTRecht visit
2018-11-30

David Groep
davidg@nikhef.nl

On Needles and Haystacks

Zachte botsingen	10^8
$W^{\pm} \rightarrow e^{\pm} \nu$	15
$Z^0 \rightarrow e^+ e^-$	1
Top-anti-top quarks	1
$bb \rightarrow \mu + X$	10^3
QCD jets, $p_T > 150$ *GeV*	10^2

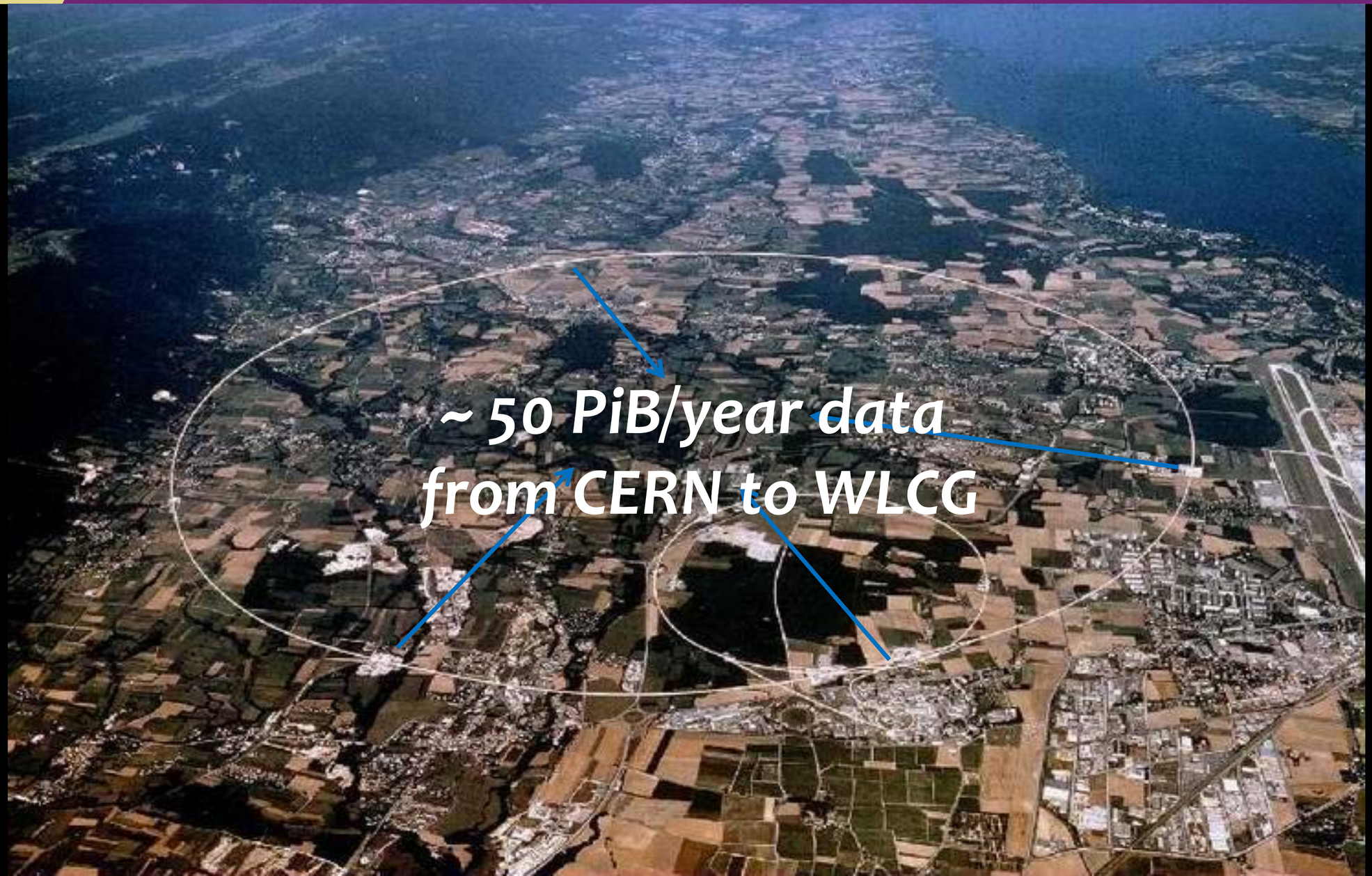
Higgs deeltje: ~ 1 per dag

event rates from LHC Run 1

Selectie interessante gebeurtenissen uit 'achtergrond' van 1 op 10^{13} gebeurtenissen

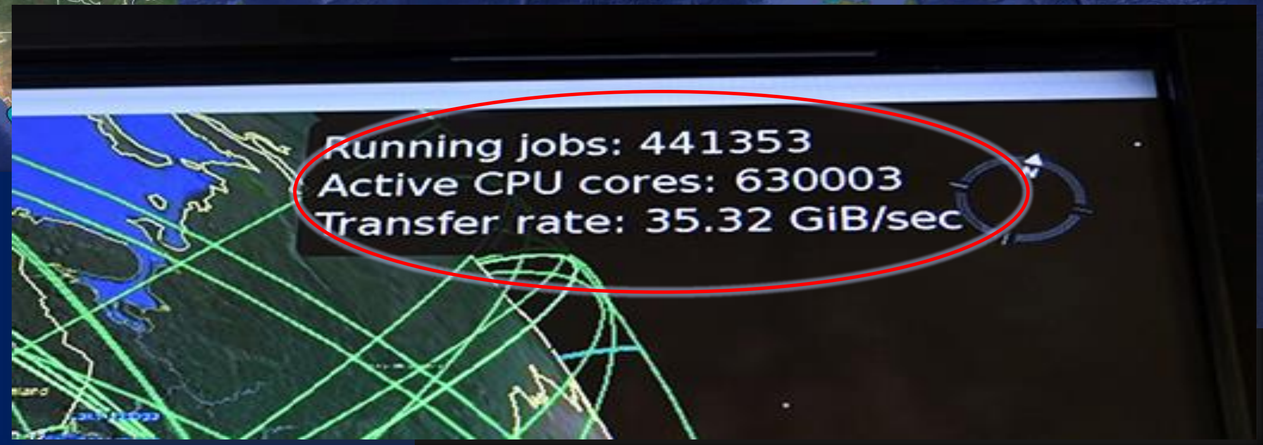
- dit is equivalent met zoeken van
1 persoon op 1000 wereldpopulaties
- oftewel één speld in 20 miljoen hooibergen

Netwerk: from 9600bps to 1.2Tbps



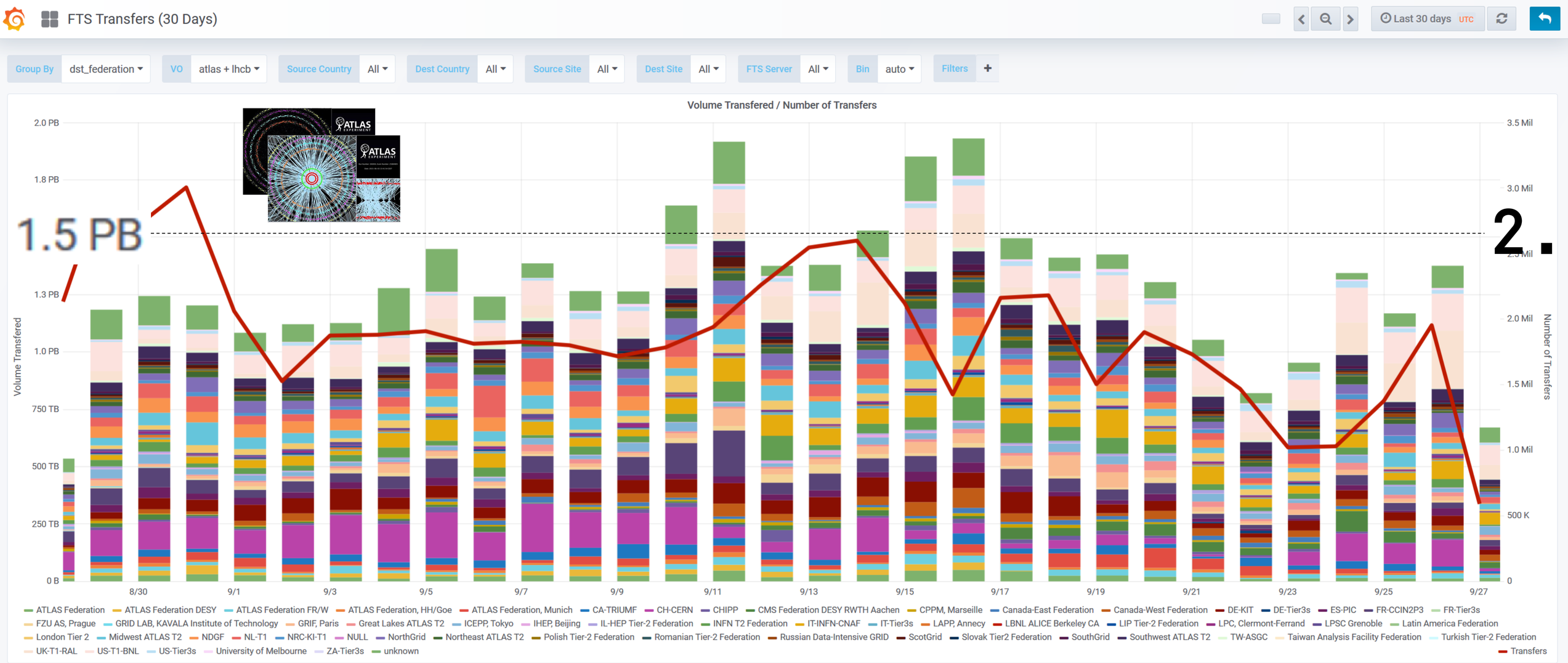
… globally distributed

meer dan 170 instituten in 42 landen en economiën



- *CPU: ~ 350,000 modern rekenkernen*
- *Disk 310 PB*
- *Tape 390 PB*

Data verdelen over de wereld T0>T1s



source: https://monit-grafana.cern.ch/d/000000420/fts-transfers-30-day on September 27, 2018

‘Ye olde compatibilitye’

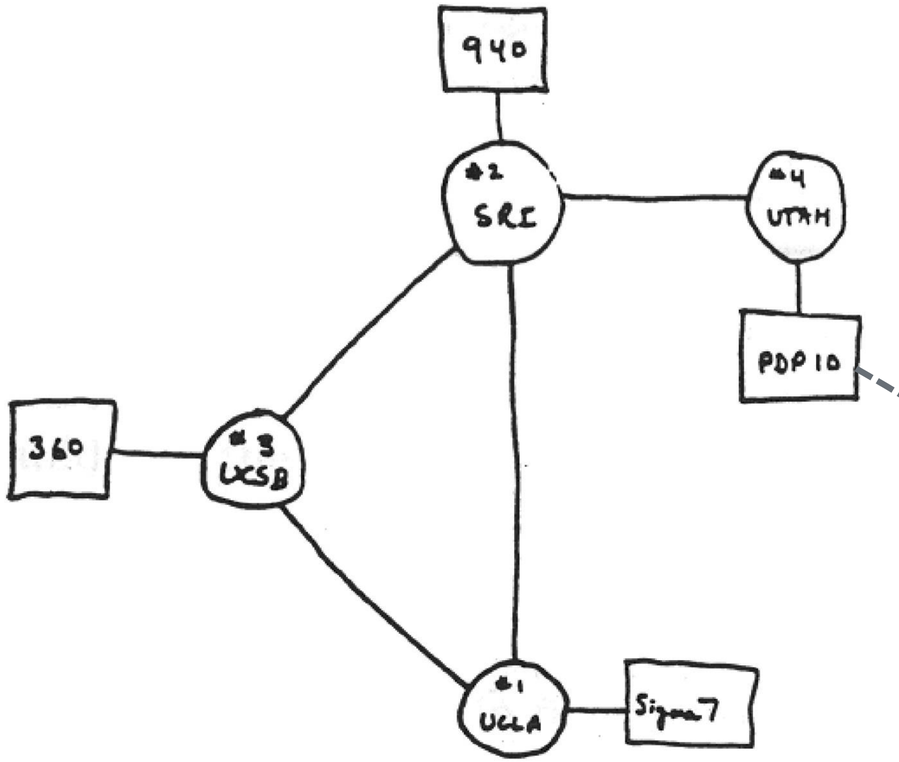
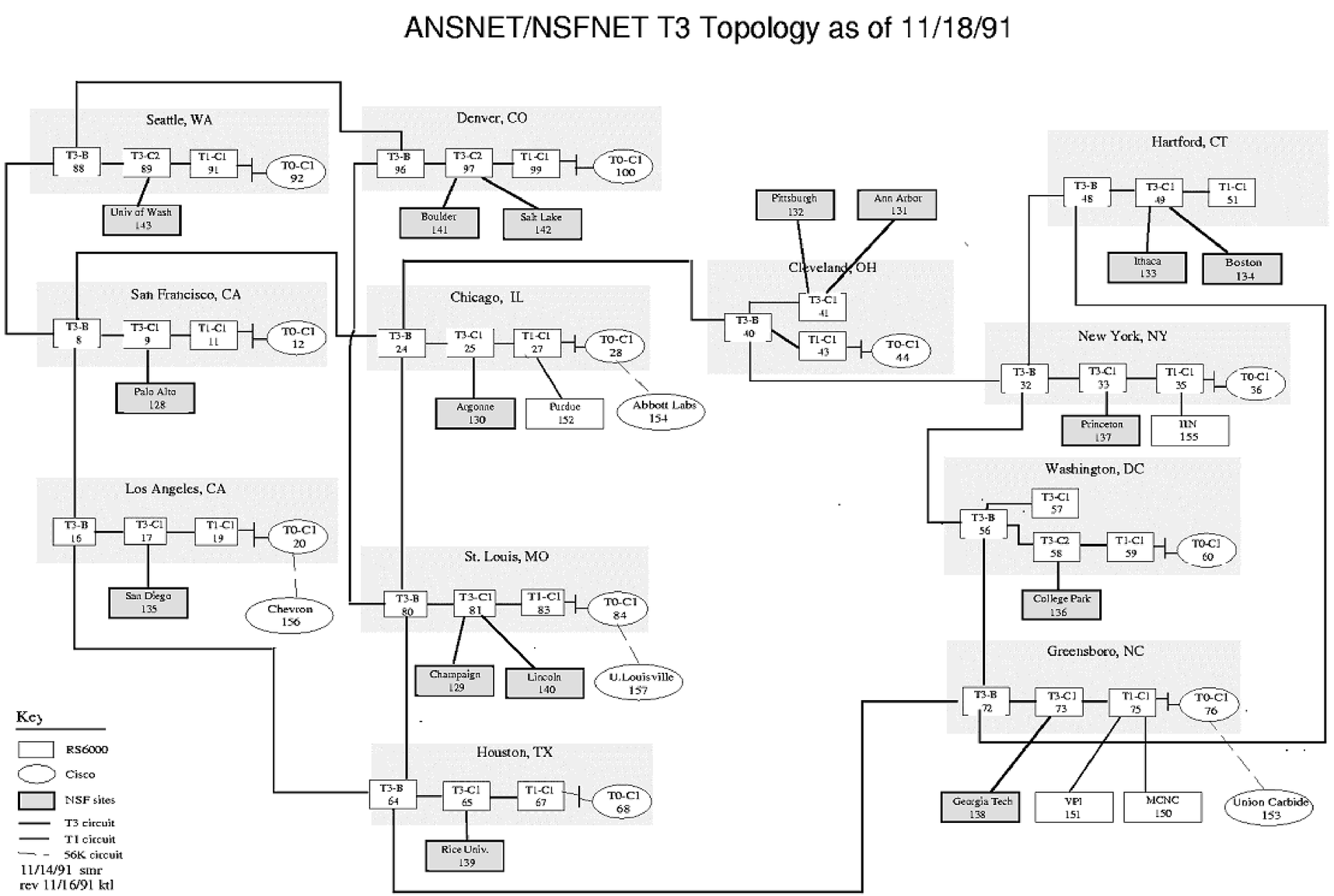
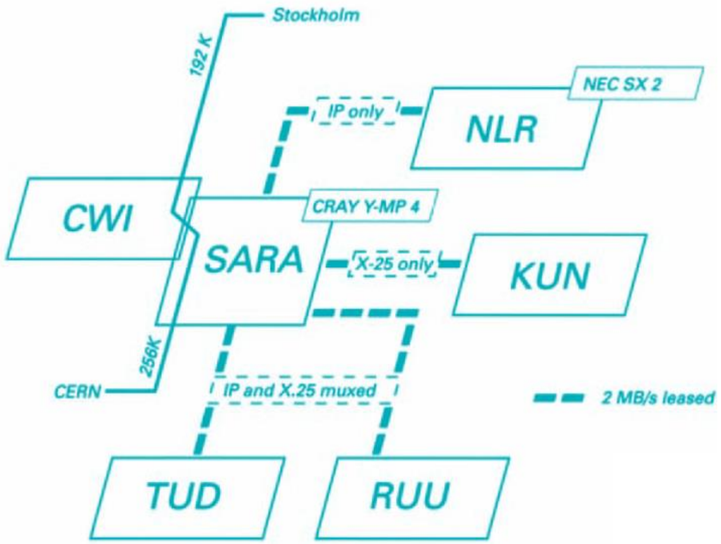


Image source: Alex McKenzie and "*Casting the Net*", page 56. See https://personalpages.manchester.ac.uk/staff/m.dodge/cybergeography/atlas/arpanet2.gif ; acoustocoupler: Wikimedia

NSFnet and early SURFnet





See https://personalpages.manchester.ac.uk/staff/m.dodge/cybergeography/atlas/historical.html for more historic maps

SURFnet 2, 1990

Collaboration & remote analysis

https:// /Welcome.html

Welcome to my ho

This page features, or should feature, r

The Web offers much, in:

 The Netherlands,  Eur

There are several other interesting plac

 NOS-'Teletekst' (received an hono

 WWW Virtual Library (via CERN)

 De Digitale Stad (the Digital City)

 VU Campus Informatie Systeem

 SURFNET and NLnet

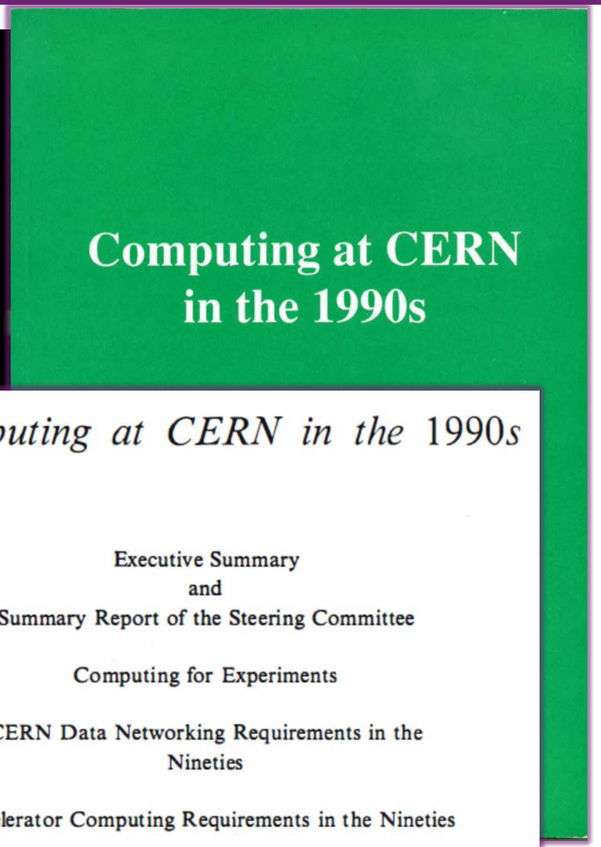
 Meteorological satelite pictures (vi

I work at NIKHEF, a research institute l

A short summary of the daily news (in c

What about my homepage at NIKHEF?

```
WWW - World Wide Web at NIKHEF

   Help [1]     About this program, and the World-Wide Web

   NIKHEF [2]   NIKHEF information

   CERN [3]     CERN Information, ftp access to file server asis01.cern.ch

   HEP [5]      High Energy Physics

   WWW [6]      The home page of WWW on info.cern.ch

   NETFIND [7]  World Wide User Lookup of email addresses

        [End]

1-7, Quit, or Help: 2
```



Computing at CERN in the 1990s

Executive Summary
and
Summary Report of the Steering Committee

Computing for Experiments

CERN Data Networking Requirements in the Nineties

Accelerator Computing Requirements in the Nineties

Computing for Engineering at CERN in the 1990s

MIS Computing in the 90s

Theory Computing in the 1990s

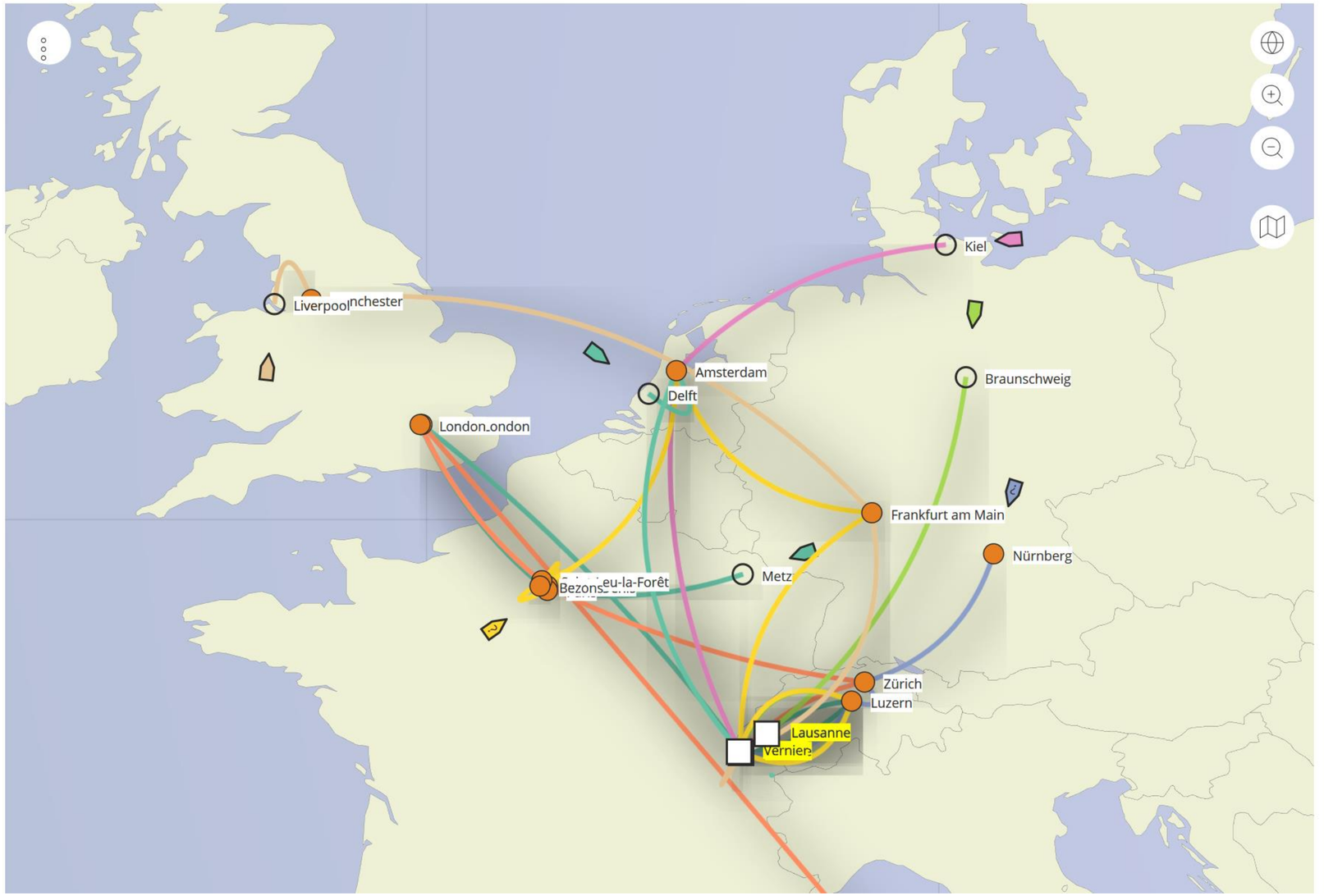
July 1989

Sources:
nic.nikhef.nl, trivia.nat.vu.nl:8000, and cds.cern.ch/record/206085/

how would *you* get to CERN?

Traceroute measurement to linuxsoft.cern.ch (multihomed)

General Information | Probes | Map | TraceMON | IPmap (beta) | Results



Data: TraceMON IPmap
from RIPE NCC Atlas
atlas.ripe.net
measurement 9249079

BGP, and Policy Based Routing

Ik ben CERN, AS513, en …
ik wil m'n pakketje kwijt …

188.184.38.9

Ik ben AortaNet, AS6830,
(maar stiekem toch LibertyGlobal)

Ik ben LibertyGlobal, AS9141,
en breng je wel ergens heen

Ik ben KPN, AS286,
en breng je wel ergens heen

Ik ben Sunrise CH, AS6730,
en breng je wel ergens heen

Source

Ik ben ook GEANT, AS20965,
en breng je naar AS1103 en 1104

Ik ben GEANT, AS21320,
en breng je naar AS1103 en 1104

Ik ben SURFsara, AS1162,
en breng je direct naar AS1104!

Ik ben SURFnet, AS1103,
en breng je naar AS1104

Ik ben Nikhef, AS1104
en ik heb schijfruimte voor je

194.171.96.128/25

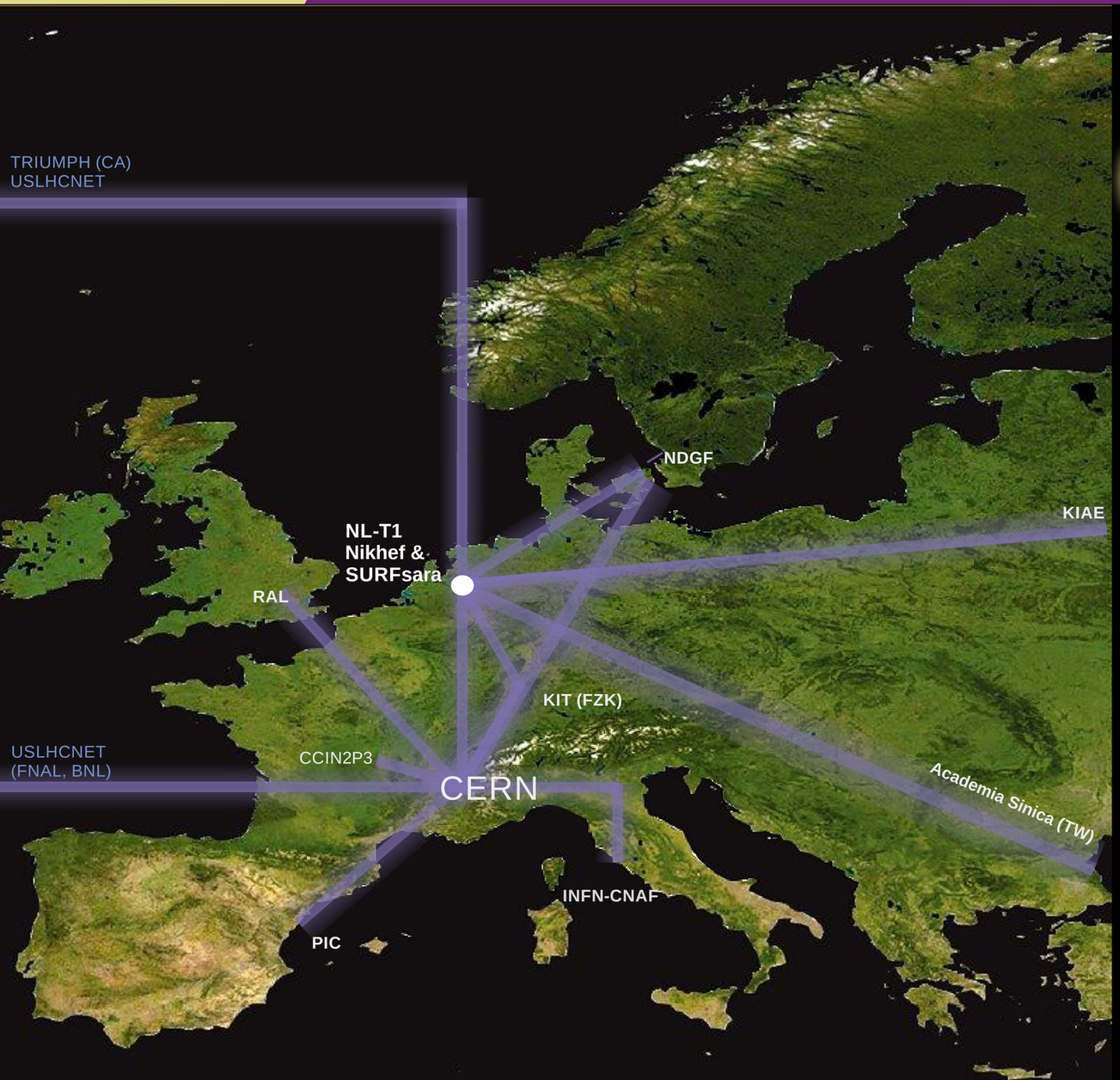
Path A
Path B
Path C

Ik ben SURFnet, AS1103,
en breng je naar AS1104

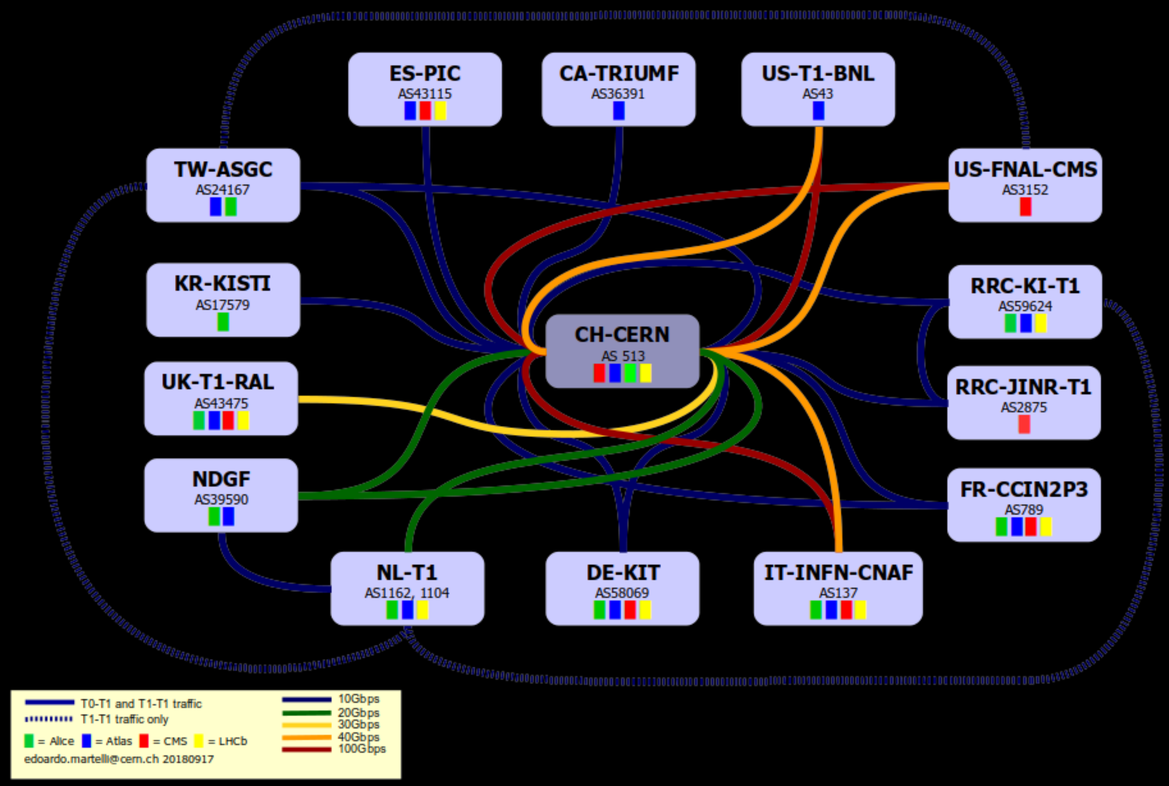
Destination

"segment routing"
image by David Penaloza, CISCO

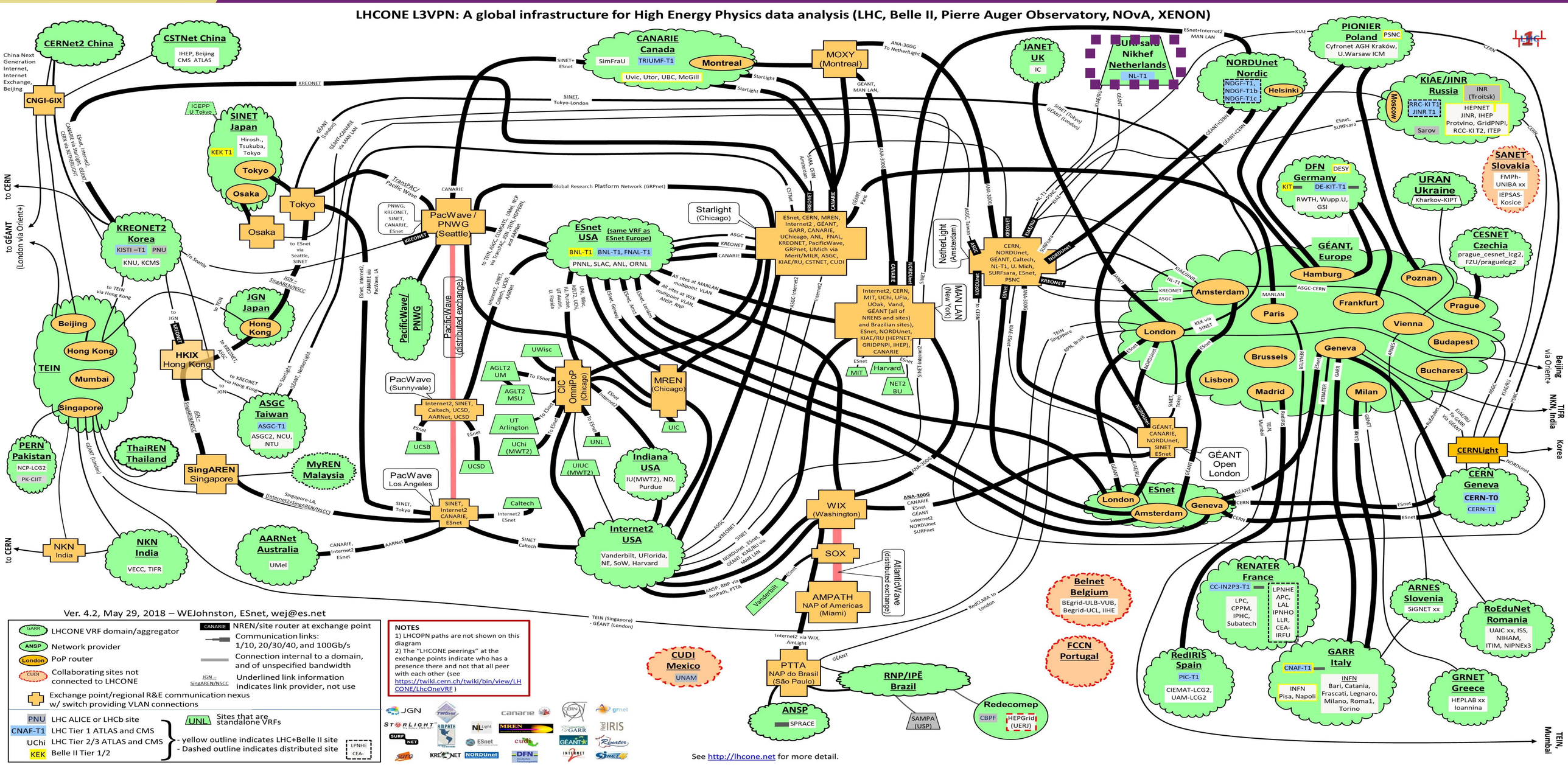
LHC ‘Optical Private Network’



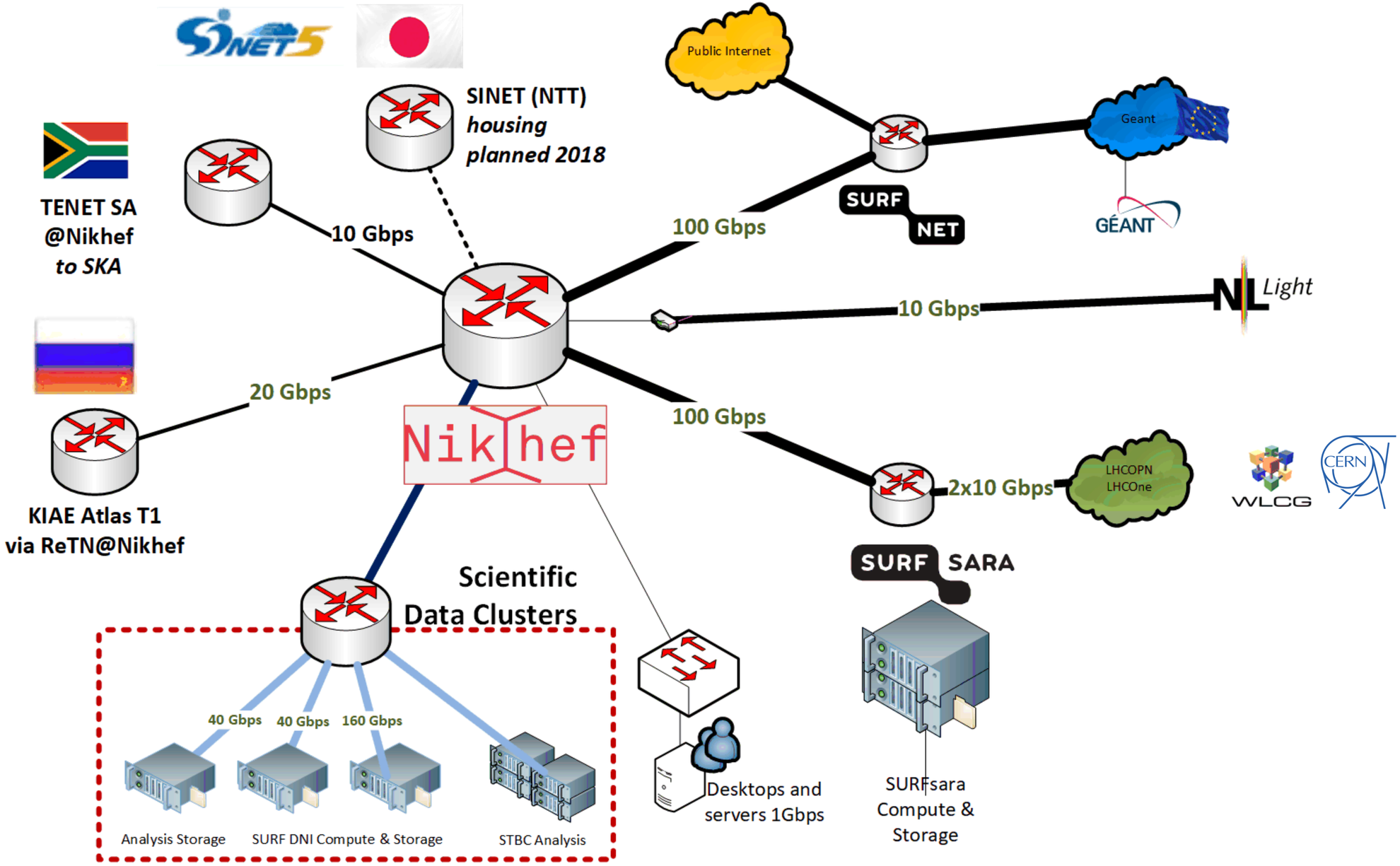
LHCOPN



‘LHC Open Network Environment’

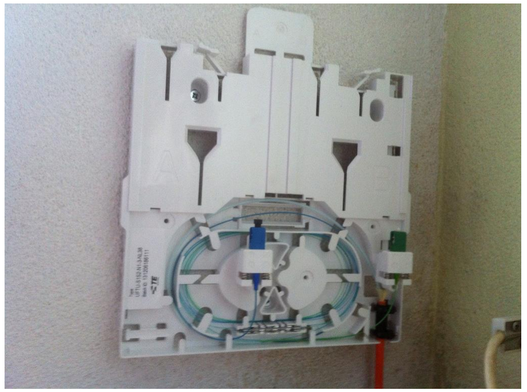


Direct peering infra for ‘big data’



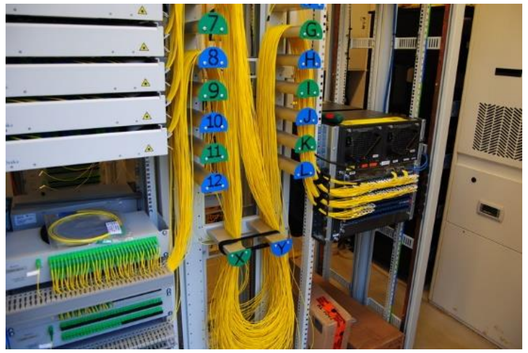
From 100 Gbps upwards ...

Thuis ‘FttH’
~1Gbps BX
single strand, SC



Nikhef
Data Processing
Facility
router ‘deel’

een KPN FttH
PoP in de wijk
100-1000Mbps



vergelijk:
VDSL BR straatkast
voor als je nog
op xDSL koper zit
20-180 Mbps

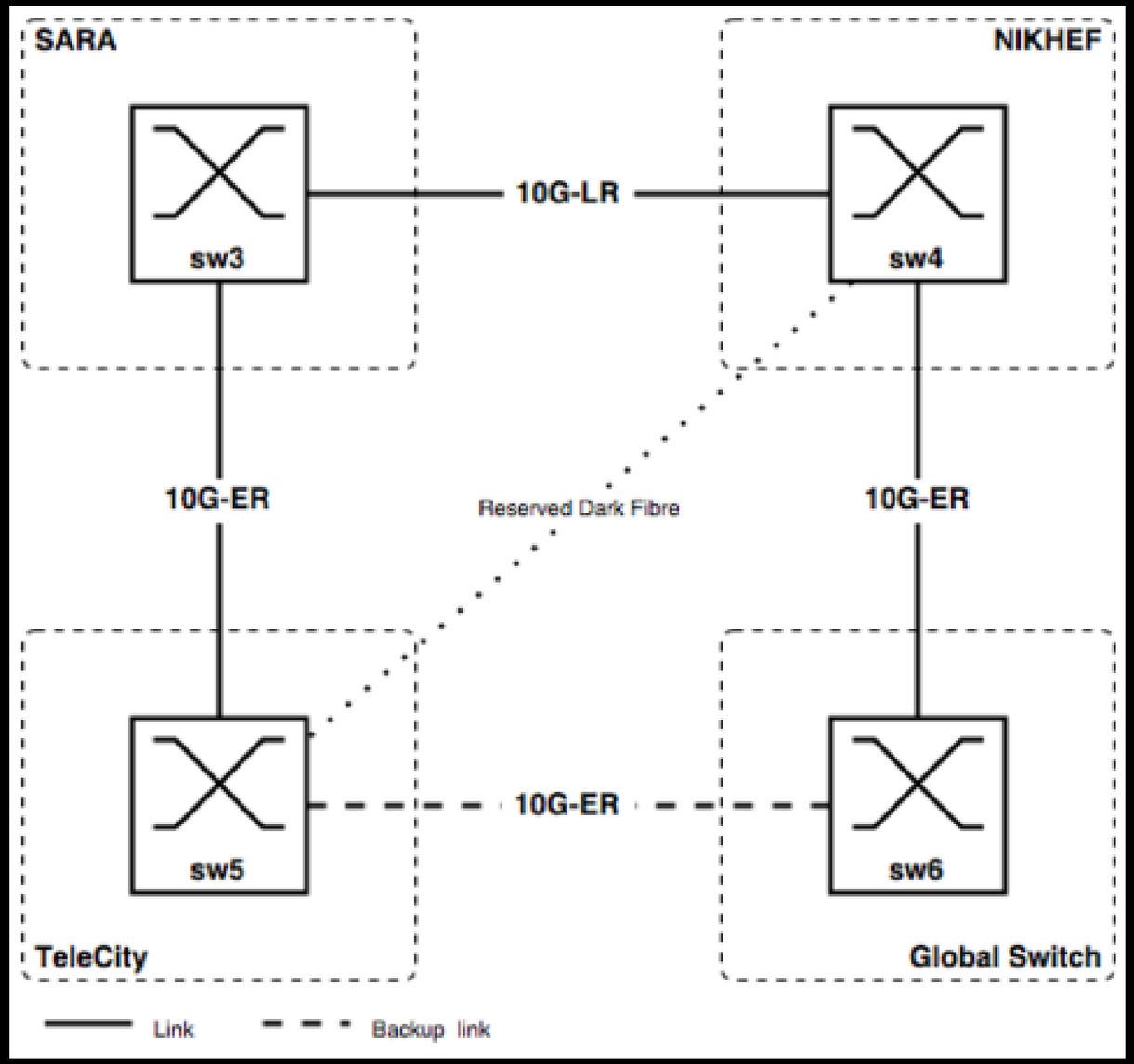
From the 80’s to ’18

this is how H1.40 could have looked ...
it is actually the Nikhef-K computer room
on the Oosterringdijk

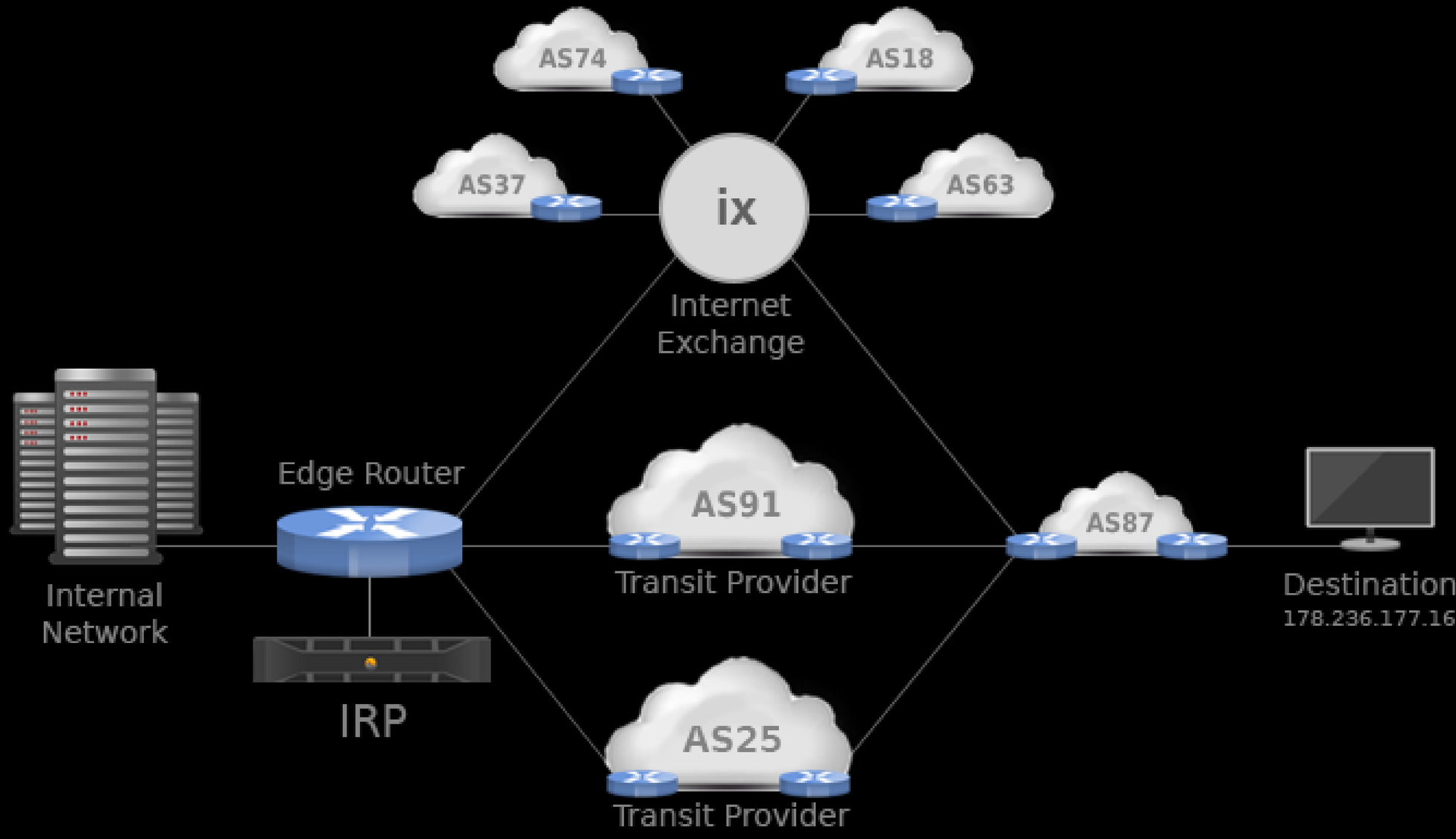
IBR LAN @Nikhef: the first exchange

1996, Nikhef H1.40

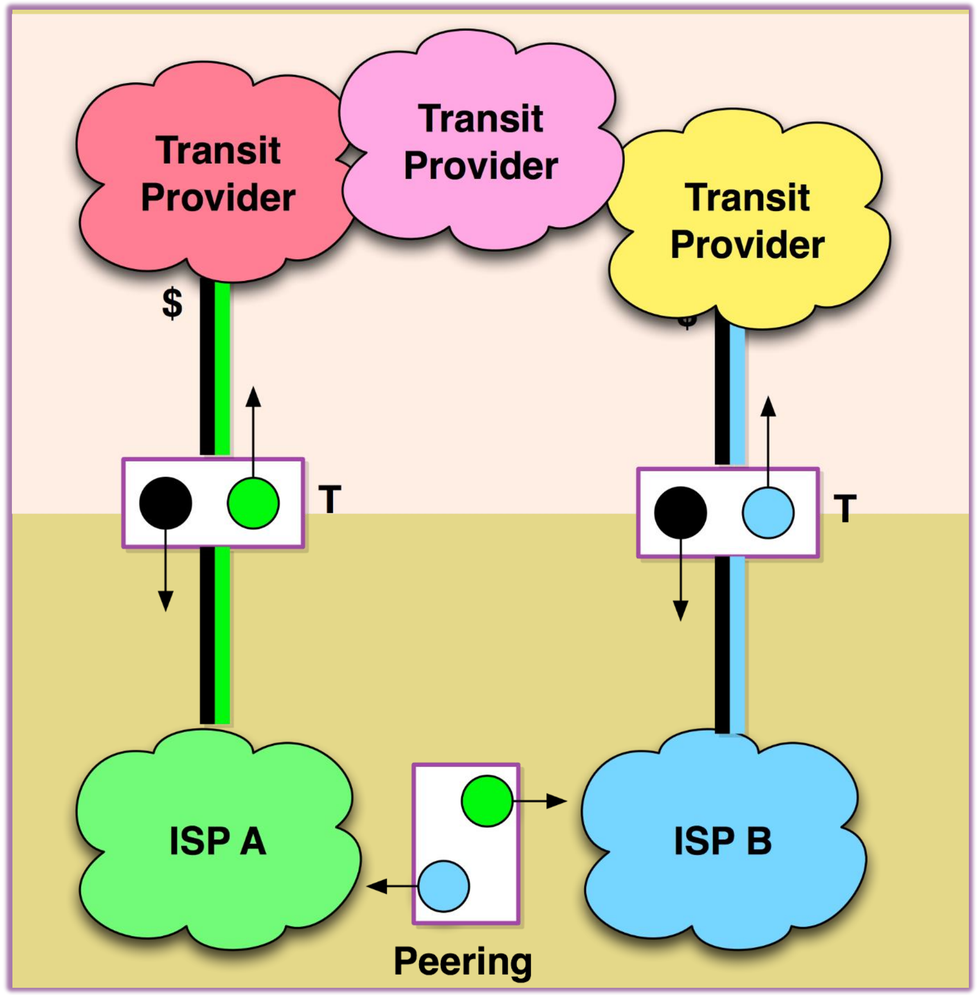
AMS-IX in ~2002



Beyond IX-es: direct transit



and private direct peering



Network Data Centres and Housing

- the more ISPs are located together, the easier private peerings become
- parties ‘pay their own way’ to the housing location, and have their own equipment – so it’s not ‘free’
- peerings can be settlement-free or paid
- *model takes traffic away from IX-es*

PeeringDB

For a quick summary on peering, see e.g. http://drpeering.net/core/ch4-Internet-Peering.html

Friends ... the more the merrier

PeeringDB

Search here for a network, IX, or facility.

Advanced Search

NIKHEF Amsterdam

Organization	Nikhef
Website	http://www.nikhef.nl
Address 1	Science Park 105
Address 2	
Location	Amsterdam, NH, 1098 XG
Country Code	NL
Geocode	52.356394, 4.950837
CLLI Code	AMSTNL
NPA-NXX	
Notes	

Local Exchanges

Filter

Exchange	Long Name	Networks
AMS-IX	Amsterdam Internet Exchange	803
Asteroid Amsterdam	Asteroid Amsterdam IX	28
DATAIX	Global Network Managment Ltd	177
DF-IX	Data Facilities Internet Exchange	24
GE-CIX	Global Internet Exchange & Peering Network	2
Global-IX	GlobalNet	61
GN-IX	Groningen Internet Exchange	3
Hopus	The HOPUS - the routed exchange	20
NetIX	NetIX Communications Ltd.	45
NL-ix	Neutral Internet Exchange	331
Speed-IX	Speed Internet Exchange	82

Networks

Peer Name	
KPN	
Apple Inc.	
Amsterdam Internet Exchange BV	
TELE2	
next layer Telekommunikationsdie	
GTT Communications (AS3257)	
TDC A/S	
Deutsche Telekom	
BT	
Daisy Communications	
Breedband Nederland B.V.	
Zayo (Abovenet Communications	
Elisa Corporation	
STRATO AG	
BICS	6774
Liberty Global	6830
Six Degrees - Datahop (6DG)	6908
Hurricane Electric	6939
Colt	8220
atom86	8455
OBIT	8492
1&1 Internet	8560
Infracom Internet B.V.	8587
EspritTelecom	8608

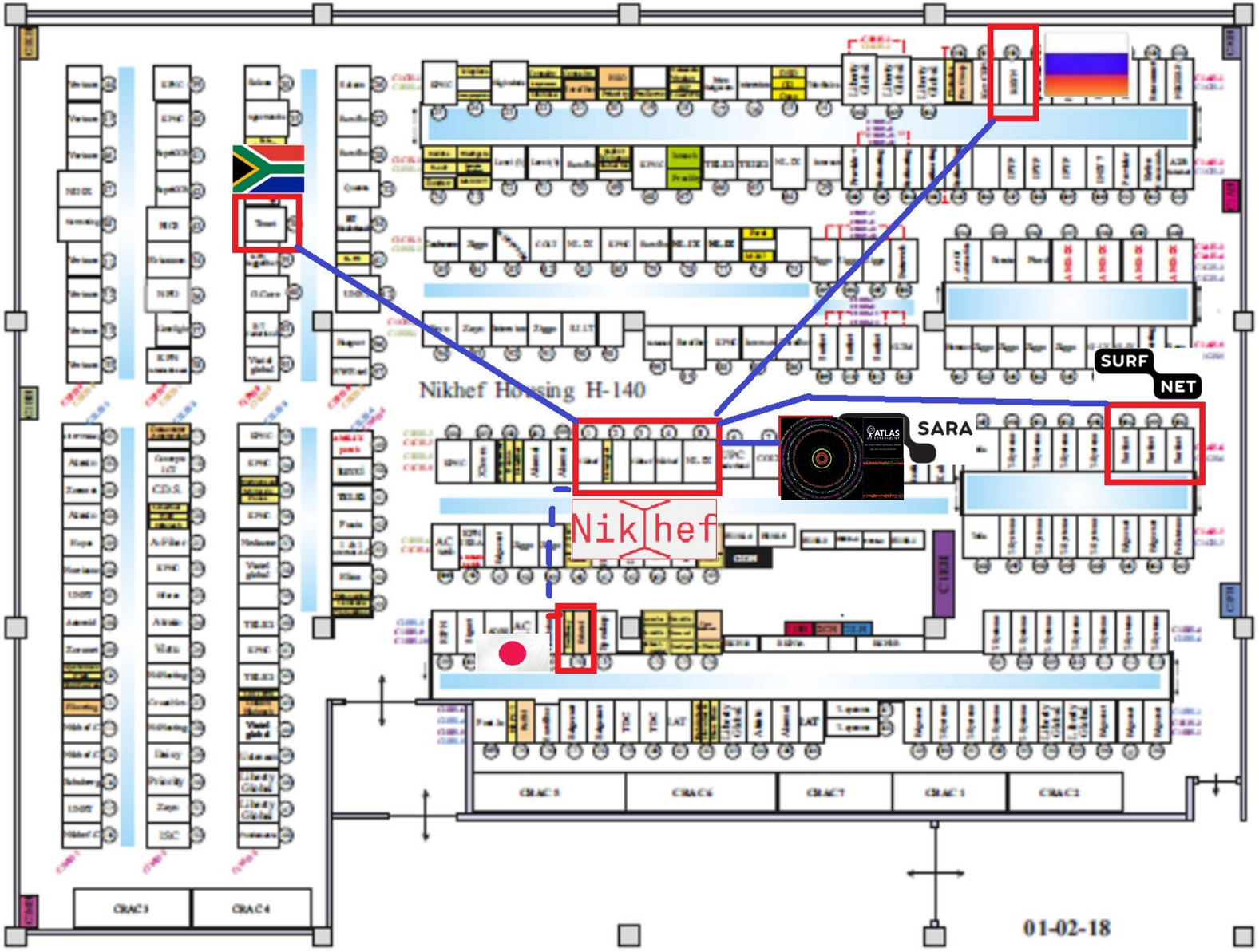
Networks

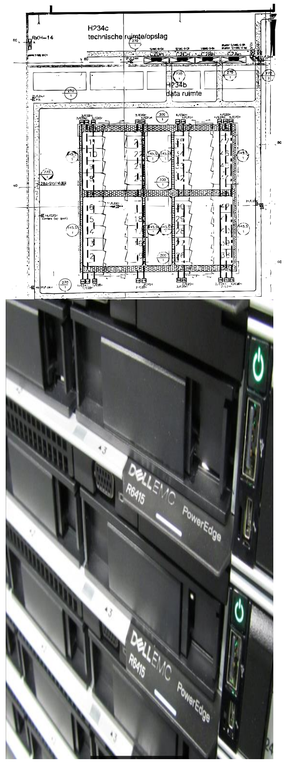
Filter

Peer Name	ASN
KPN	286
Apple Inc.	714
Amsterdam Internet Exchange BV	1200
TELE2	1257
next layer Telekommunikationsdienstleistungs- und BeratungsGmbH	1764
GTT Communications (AS3257)	3257
TDC A/S	3292
Deutsche Telekom	3320
BT	5400
Daisy Communications	5413
Breedband Nederland B.V.	5524
Zayo (Abovenet Communications Inc.)	6461
Elisa Corporation	6667
STRATO AG	6724
BICS	6774
Liberty Global	6830

PeeringDB

Touching the networks





Some quick housing facts

~180 networks, 33 carriers

2.5 MW redundant power, A+B feed

Data Processing Facility (400 kW)
high-power ~9kW/rack, 24 C

in Housing location ~4kW/rack, ~18 C

look for energy optimization methods ...

~21% additional power needed for cooling
but: with 3500 GJoule/yr of STES capacity
(~112 kW-yr) linked to student housing

first the tour
– please form two groups (but you will see everything)

then return for discussion and Q&A on your choice of topics

routing neutrality, DDoS, or open discussion

Discussion options

- ‘Back-office’ of net neutrality: peering wars and settlement
- Why do DDoS attacks keep persisting?
  How would carriers help, and why don’t they do it?
- *… and if you’re still not tired of GDPR:*
  would (or should) an ISP be classified as a data processor?

Net neutrality and peering policy

From the recitals in Regulation (EU) 2015/2120

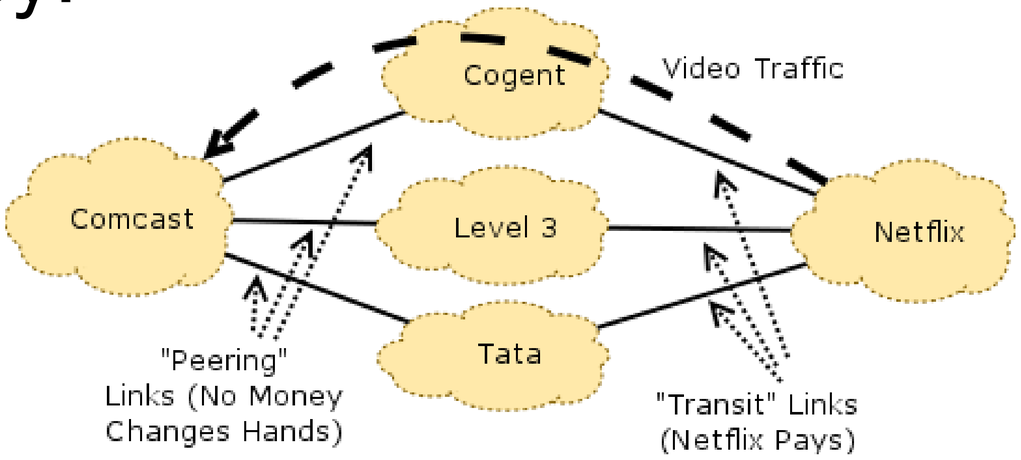
(6) End-users should have the right to access and distribute information and content, and to use and provide applications and services without discrimination, via their internet access service. The exercise of this right should be without prejudice to Union law, or national law that complies with Union law, regarding the lawfulness

(7) In order to exercise their rights to access and distribute information and content and to use and provide applications and services of their choice, end-users should be free to agree with providers of internet access services on tariffs for specific data volumes and speeds of the internet access service. Such agreements, as well as any commercial practices of providers of internet access services, should not limit the exercise of those rights and thus circumvent provisions of this Regulation safeguarding open internet access. National regulatory and other competent authorities should be empowered to intervene against agreements or commercial practices which, by reason of their scale, lead to situations where end-users' choice is materially reduced in practice. To this end, the assessment of agreements and commercial practices should, inter alia, take into account the respective market positions of those providers of internet access services, and of the providers of content, applications and services, that are involved. National regulatory and other competent authorities should be required, as part of their

requirements of specific categories of traffic, and thus of the content, applications and services transmitted. Reasonable traffic management measures applied by providers of internet access services should be transparent, non-discriminatory and proportionate, and should not be based on commercial considerations. The requirement for traffic management measures to be non-discriminatory does not preclude providers of internet access services

In the US: ‘traffic wars’

- All major providers claim they’re a ‘Tier-1 carrier’
- refuse settlement-free peering with equally large (or larger) partners to pressure them for money:
  - Cogent – Google
  - Comcast – Cogent
  - Hurricane Electric – Cogent
  - AT&T – Comcast
- Yet all have their own back-haul networks …



See e.g.
http://www.dslreports.com/forum/r31868626-Speed-Experiencing-site-slowdowns-Comcast-and-Cogent-are-at-it-again
https://qz.com/256586/the-inside-story-of-how-netflix-came-to-pay-comcast-for-internet-traffic/

yet Europe is not entirely immune

E.g. Liberty Global:

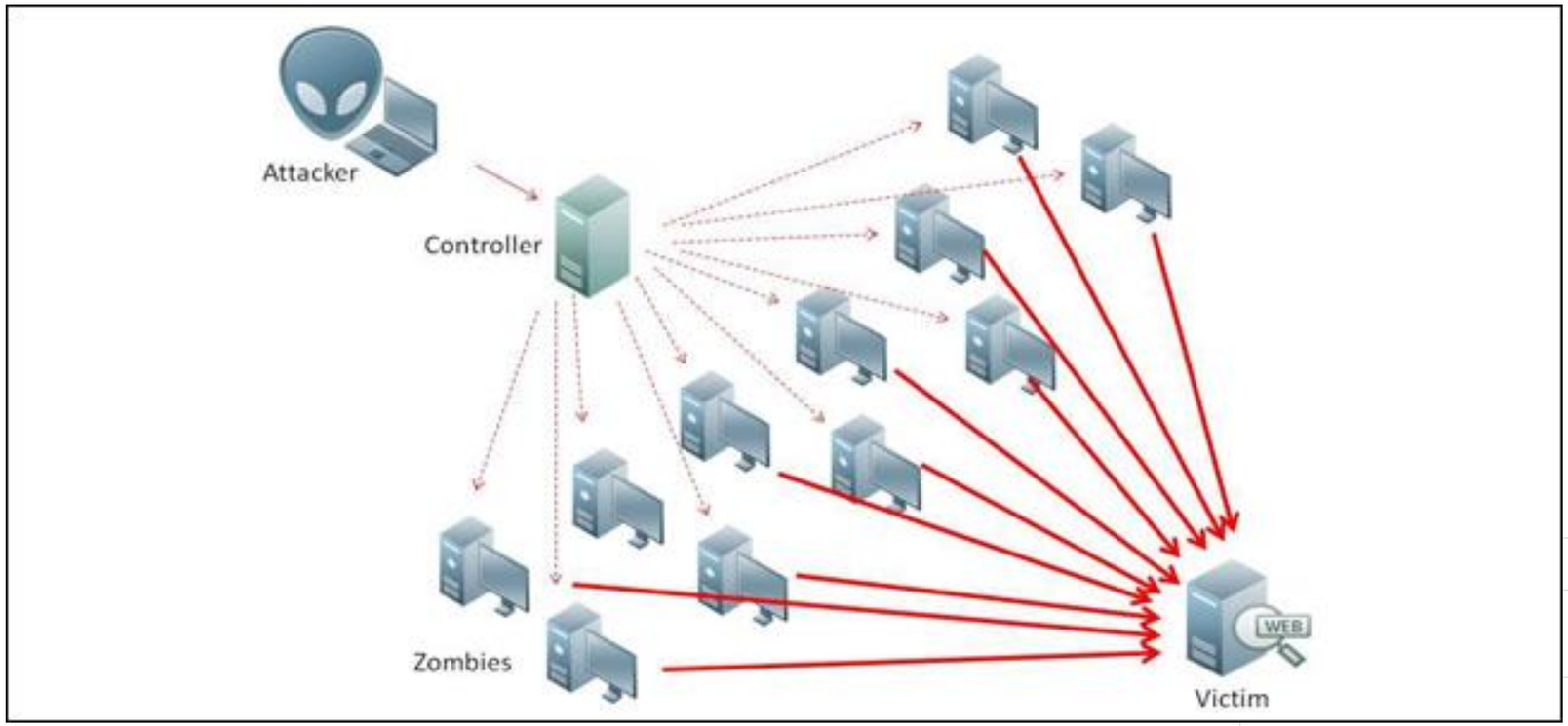
even if its own ‘Ziggo’ customers want to view your content and themselves already pay for that via their subscription,

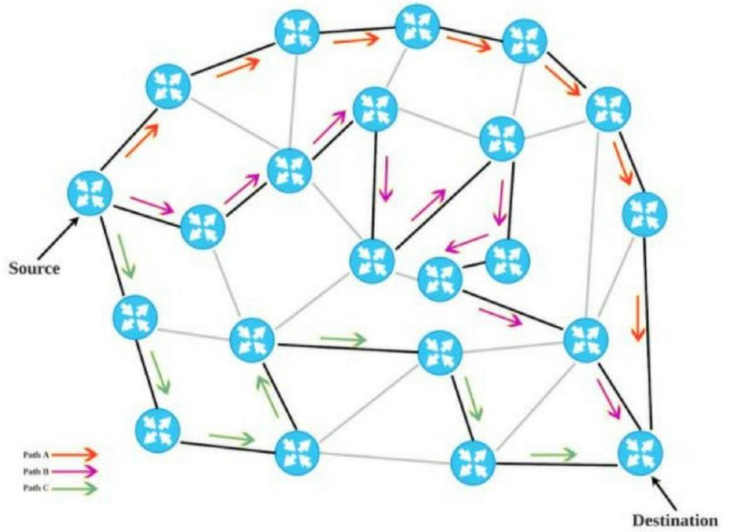
… *Liberty Global still wants you as a content provider to pay up as well!*

(getting payed twice for the traffic is quite profitable …)

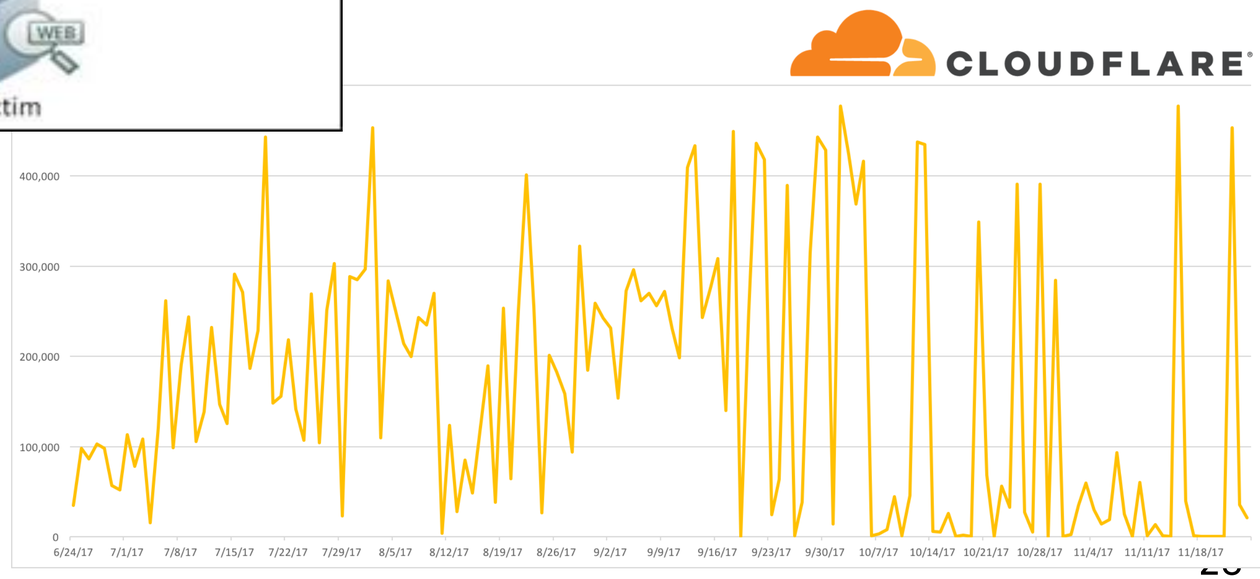
https://www.libertyglobal.com/operations/business-services/global-peering-principles/
https://www.libertyglobal.com/pdf/countries/Settled-Peering-Policy-March-2015.pdf

‘Did you order your DDoS today?’





Max SYN flood DDoS size per day in Mbps



Traffic graph source: https://blog.cloudflare.com/the-new-ddos-landscape/

Some common DDoS reasons

From Cyber Warfare: Techniques, Tactics and Tools for Security Practitioners; Jason Andress, Steve Winterfeld

Organised groups

- Extortion
- Anti-competitive practices (inside and outside)
- Hacktivism
- APTs

Individuals also *very* common

- attacking your game opponent
- attacking your school to get your exam deferred

‘Styn’ is clearly a student …

```
INSERT INTO `users` VALUES
(102199,'Styn','john1968', '¤¤¤¤¤¤@¤¤¤¤¤¤.¤¤¤¤¤',0,1,1422654207,0,0,'');
```

from home with Ziggo as the ISP, and one in a while using a Vodafone mobile browser

```
INSERT INTO `checkip` (`id`, `ip`, `user`) VALUES (87,'¤¤.75.183.125','Styn');
INSERT INTO `checkip` (`id`, `ip`, `user`) VALUES (138,'¤¤.217.77.43','Styn');
```

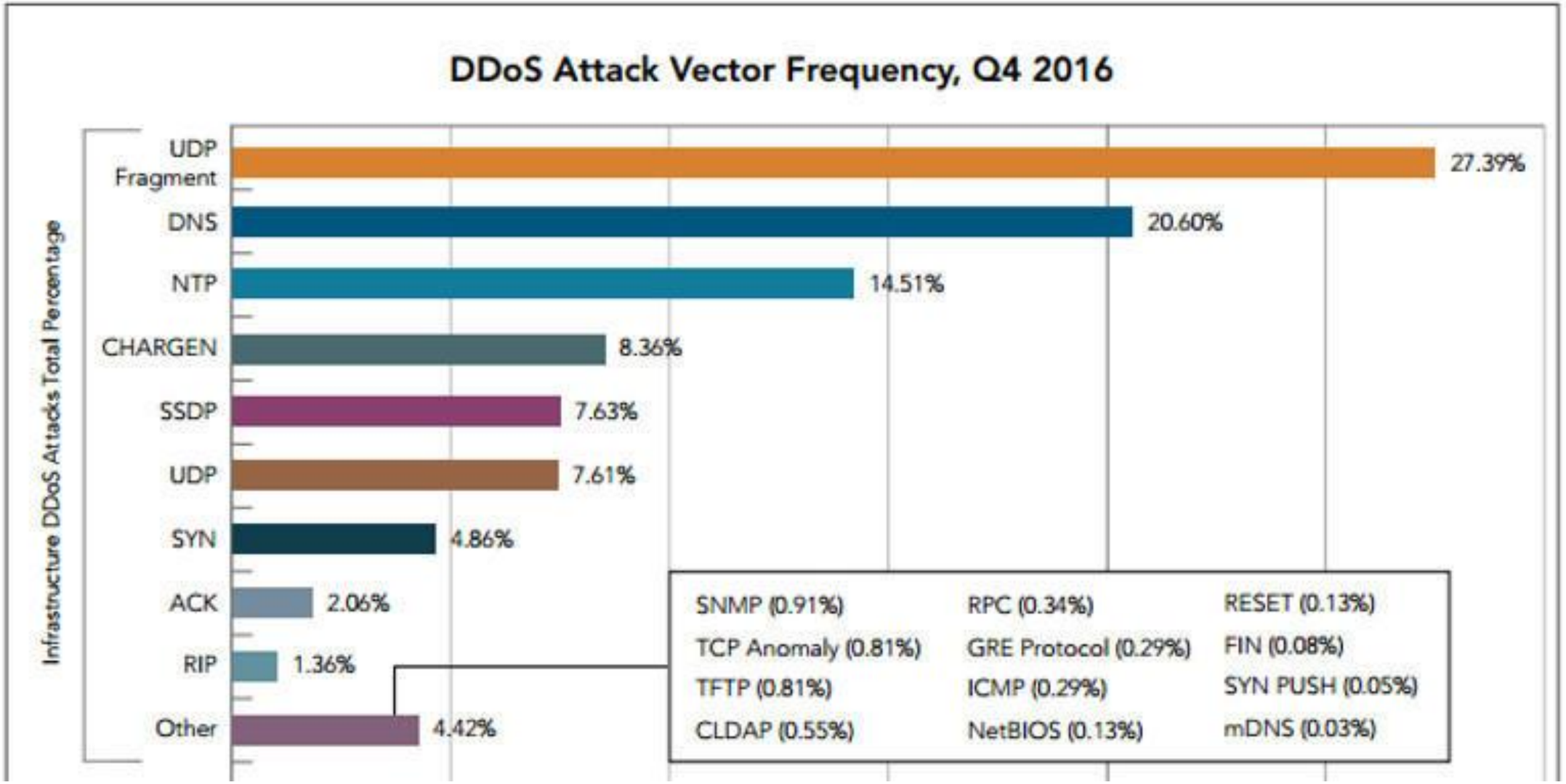
and 95 attacks ordered, against a Dutch ROC, and later a Dutch university:

```
INSERT INTO `logs` VALUES (14071,'Styn','¤¤¤.97.64.171',80,95,'¤¤¤¤',...);
```

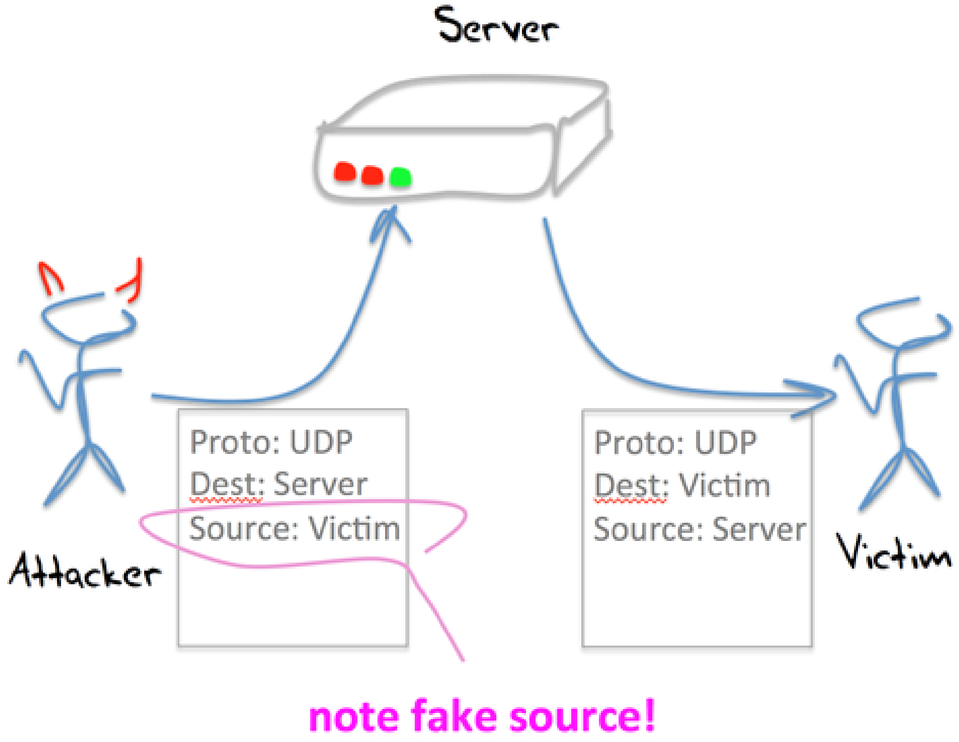
but meanwhile also on some Ziggo addresses – assumed his game opponents in a dorm

data thanks to a nice data leak of the LizardSquad DDoS-for-hire service in 2015 ☺

DDoS usage of IP protocols



Reflection & amplification attacks



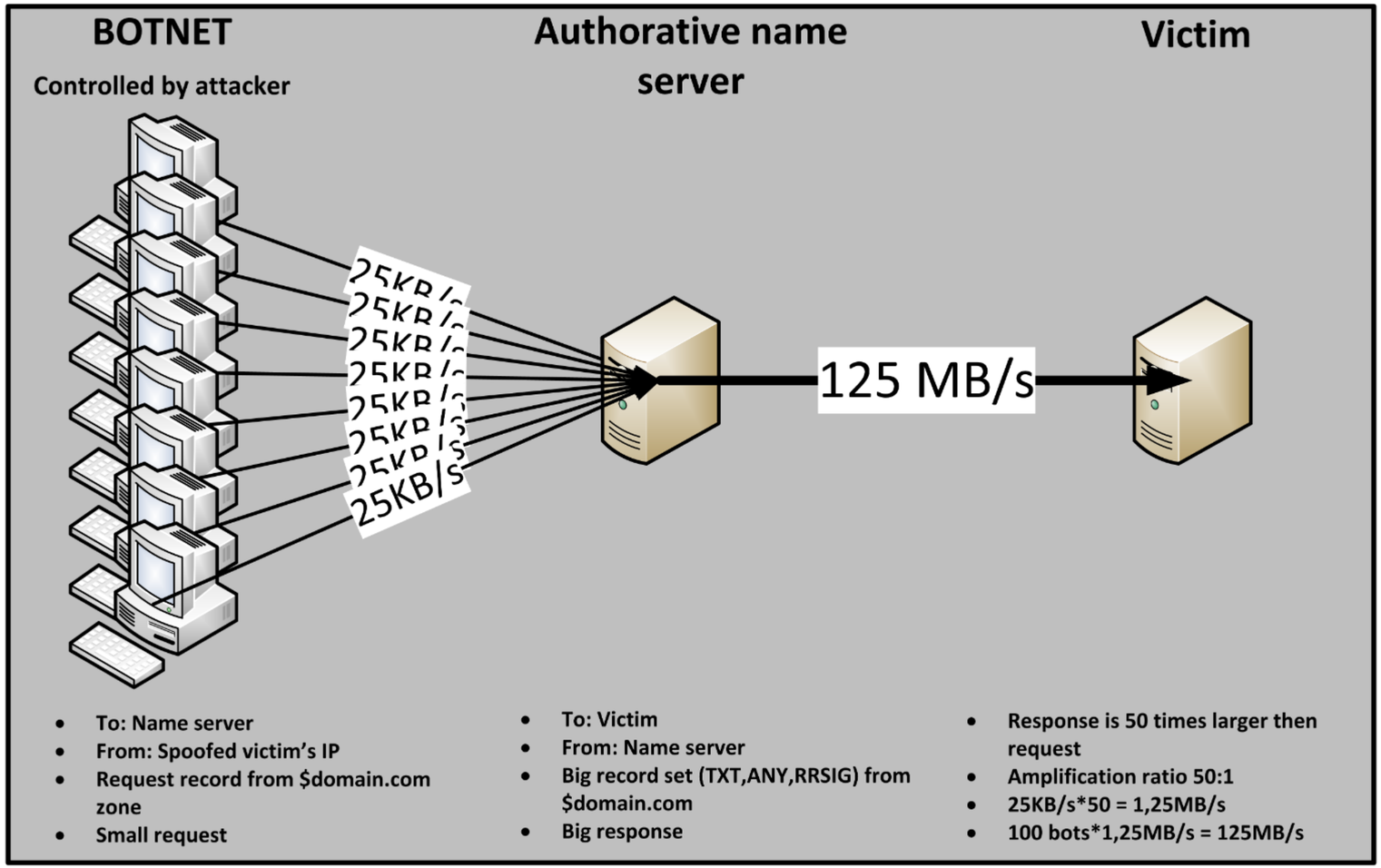
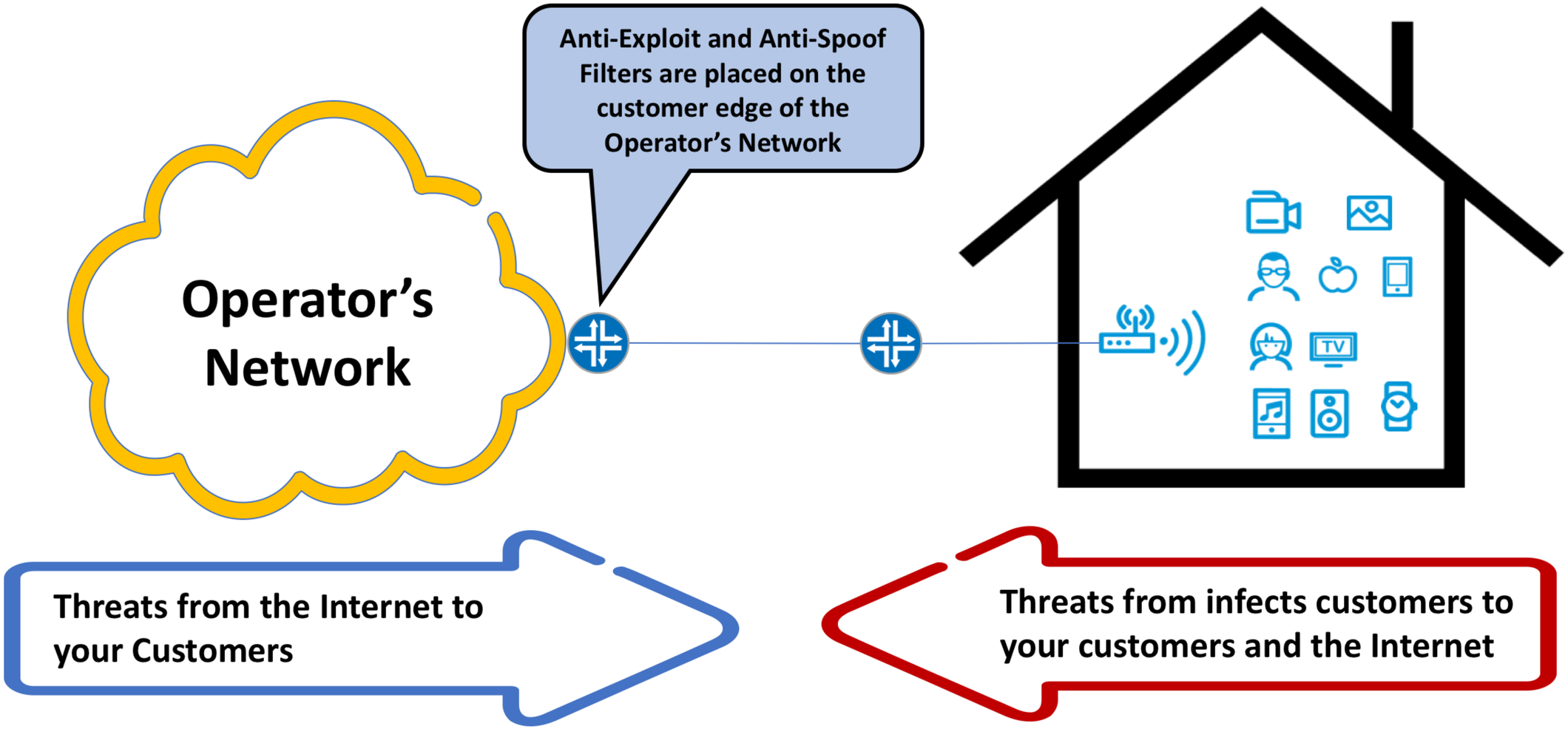


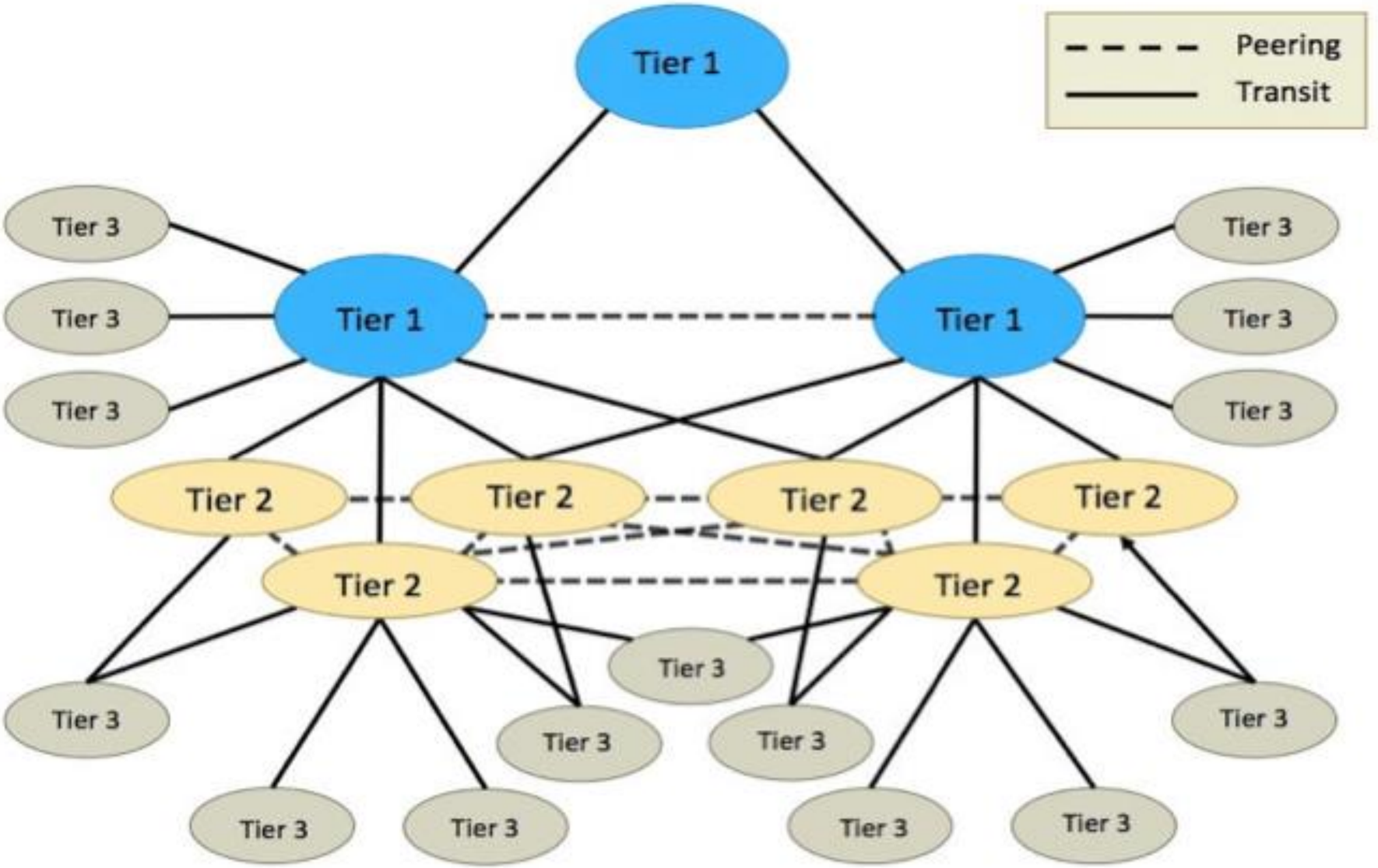
Image source top-left : https://labs.ripe.net/Members/gih/ntp-for-evil
NLNet labs: https://www.nlnetlabs.nl/downloads/publications/report-rrl-dekoning-rozekrans.pdf

‘BCP-38’



Read up on Barry Greene’s pages at SENKI: http://www.senki.org/exploitable-port-filtering/

At the T1's it's too late ...



And why DDoS persists …

Network Working Group
Request for Comments: 2827
Obsoletes: 2267
BCP: 38
Category: Best Current Practice

P. Ferguson
Cisco Systems, Inc.
D. Senie
Amaranth Networks Inc.
May 2000

Network Ingress Filtering:
Defeating Denial of Service Attacks which employ
IP Source Address Spoofing

Status of this Memo

This document specifies an Internet Best Current Practices for the Internet Community, and requests discussion and suggestions for

but ‘de cost gaet voor de baet uyt’ or ‘the tragedy of the commons’

- *Should we consider ISPs an accomplice for sending traffic not ‘their own’?*
- *Should we do the same for IP prefix hijacking?*
- *Or is it the responsibility of the IoT device’s owner? Or its manufacturer?*

Nikhef

David Groep
davidg@nikhef.nl
https://www.nikhef.nl/~davidg/presentations/
https://orcid.org/0000-0003-1026-6606

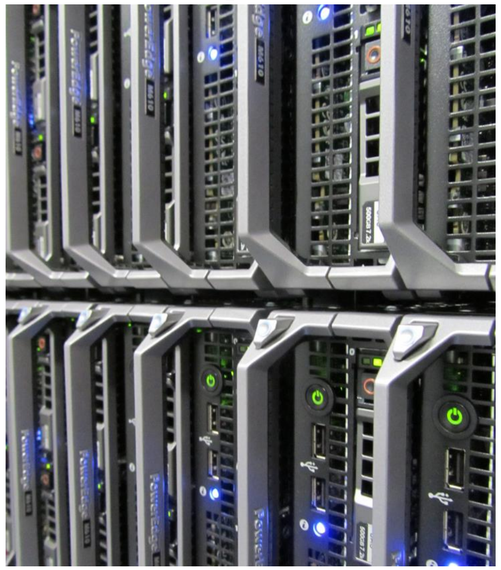
IT Infrastructures for Research

Secure Collaboration

we houden bij wie er mag rekenen, welk experiment wat gebruikt, en of onze rekenkracht goed gebruikt wordt ...

... en we niet misbruikt worden om het internet aan te vallen!

High Throughput Compute

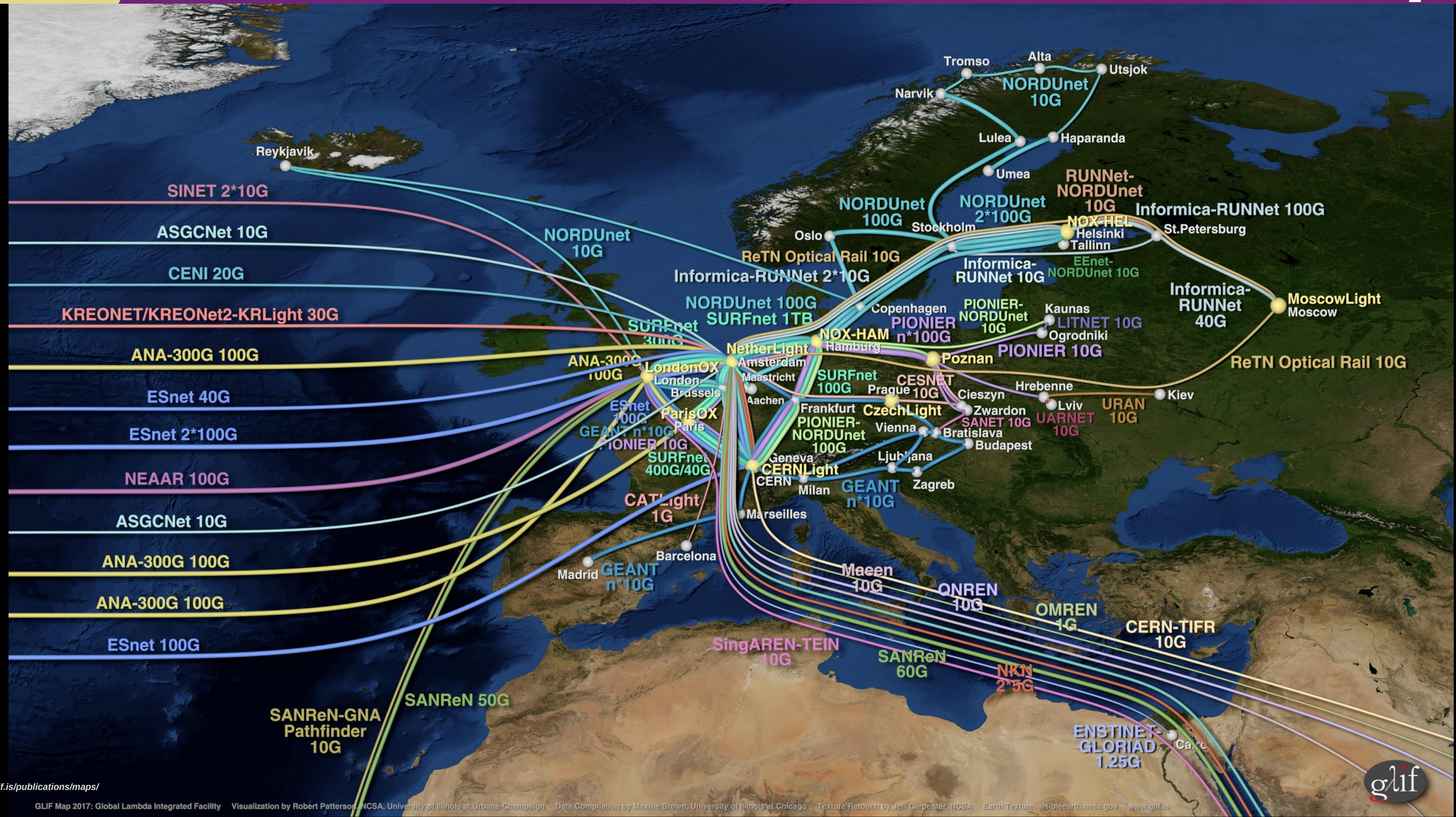


7000 rekenkernen (is ~300 servers) staan dag en nacht te rekenen aan onderzoeksdata

parallel data storage

de ~5 petabyte gegevens bij ons moeten we ook in één dag kunnen lezen, anders staan die 7000 rekenkernen te wachten ... die lezen ieder ~10MB per seconde!

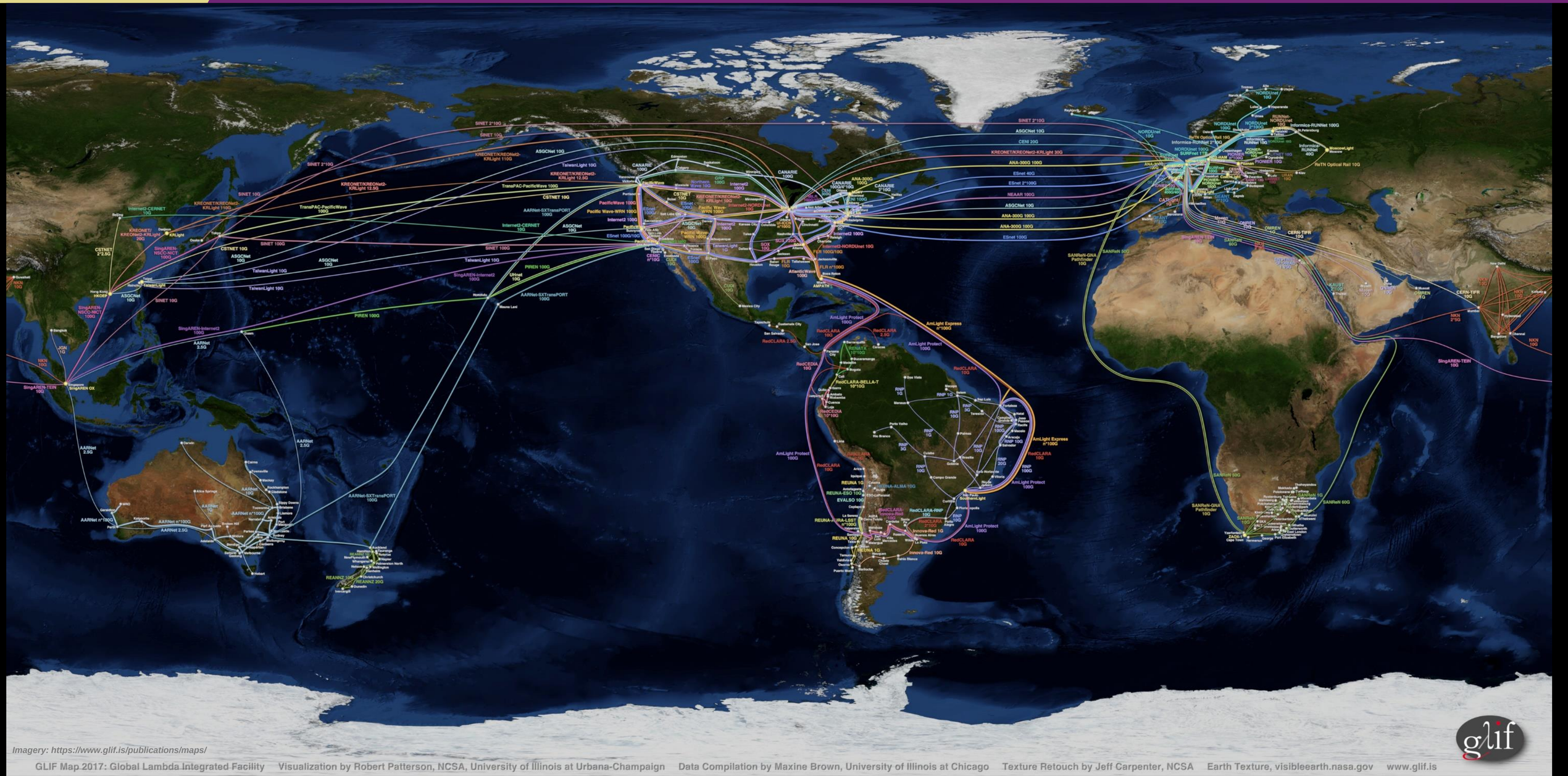
SURF and Netherlight



Imagery: https://www.glif.is/publications/maps/

Nikhef

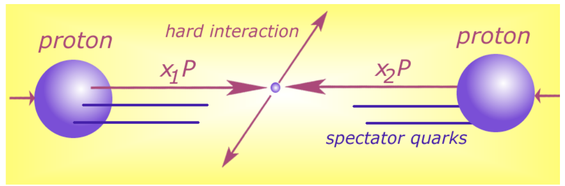
Global Lambda Integrated Facility

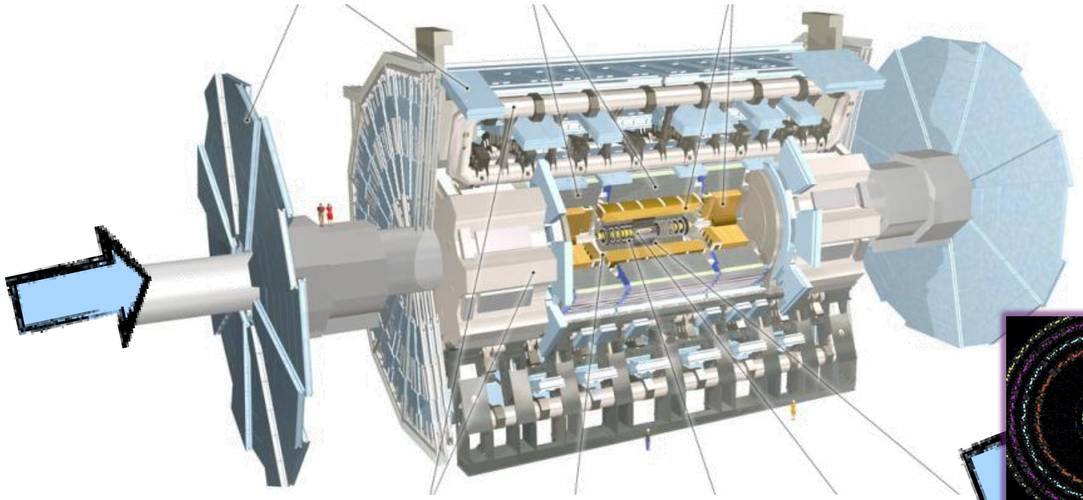


glif

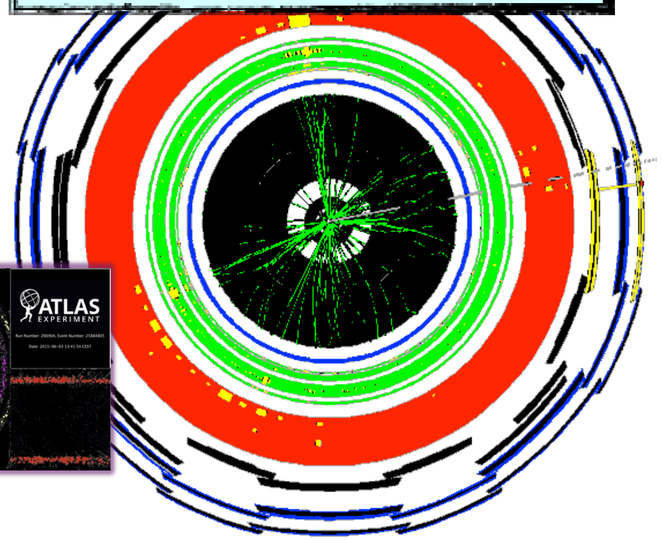
The LHC Analysis Chain

40 miljoen / seconde

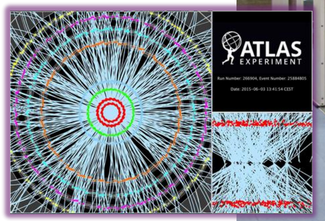




Trigger systeem selecteert 600 Hz ~ 1 GB/s data



Classificatie van deeltjes in de botsingen:
- *electronen*
- *muonen*
- *jets van hadronen*
- *...*



Analyse van botsingen door promovendi

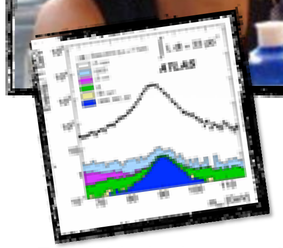
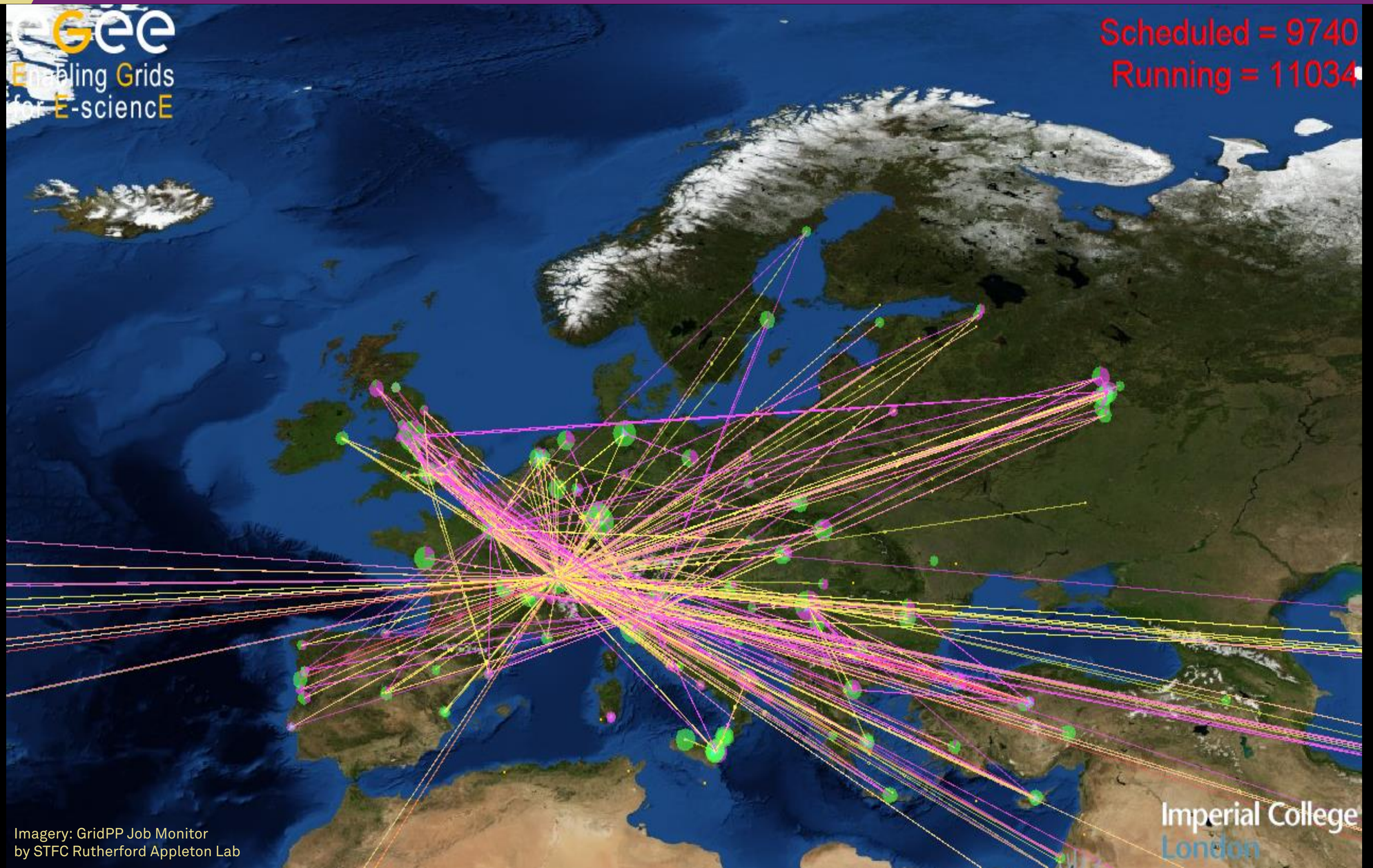


diagram: from Frank Linde; images: ATLAS collaboration, Nikhef. And sorry for the AVG-blur

Nikhef

Overal rekenen … ‘high-troughput’



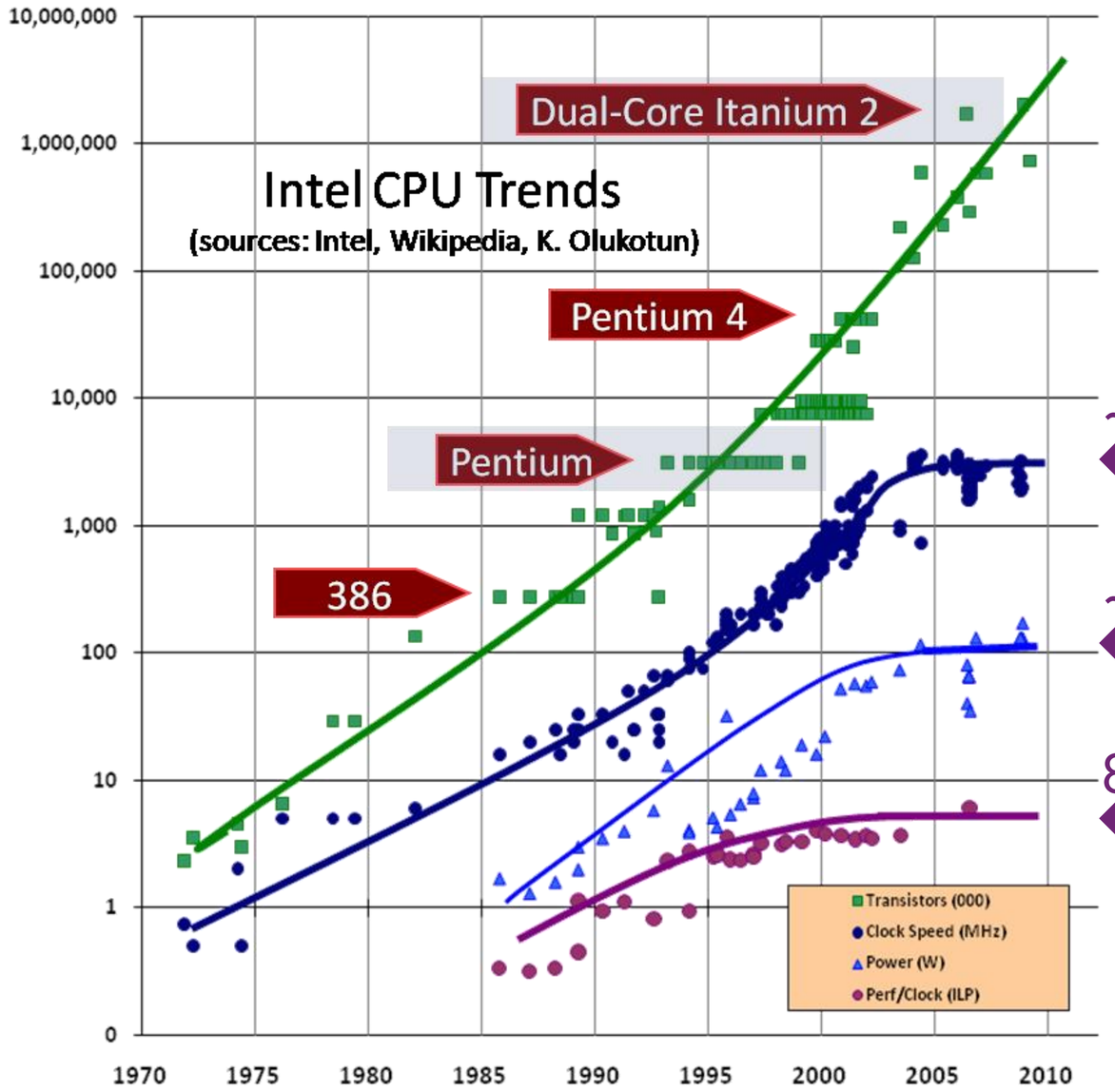
In Nederland:
De National
e-Infrastructuur
gecoordineerd
door SURF
www.surf.nl

SURF SARA

Europees
HTC Platform
door EGI
www.egi.eu

EGI

Verdeel en heers: we moeten wel



~ 4 GHz

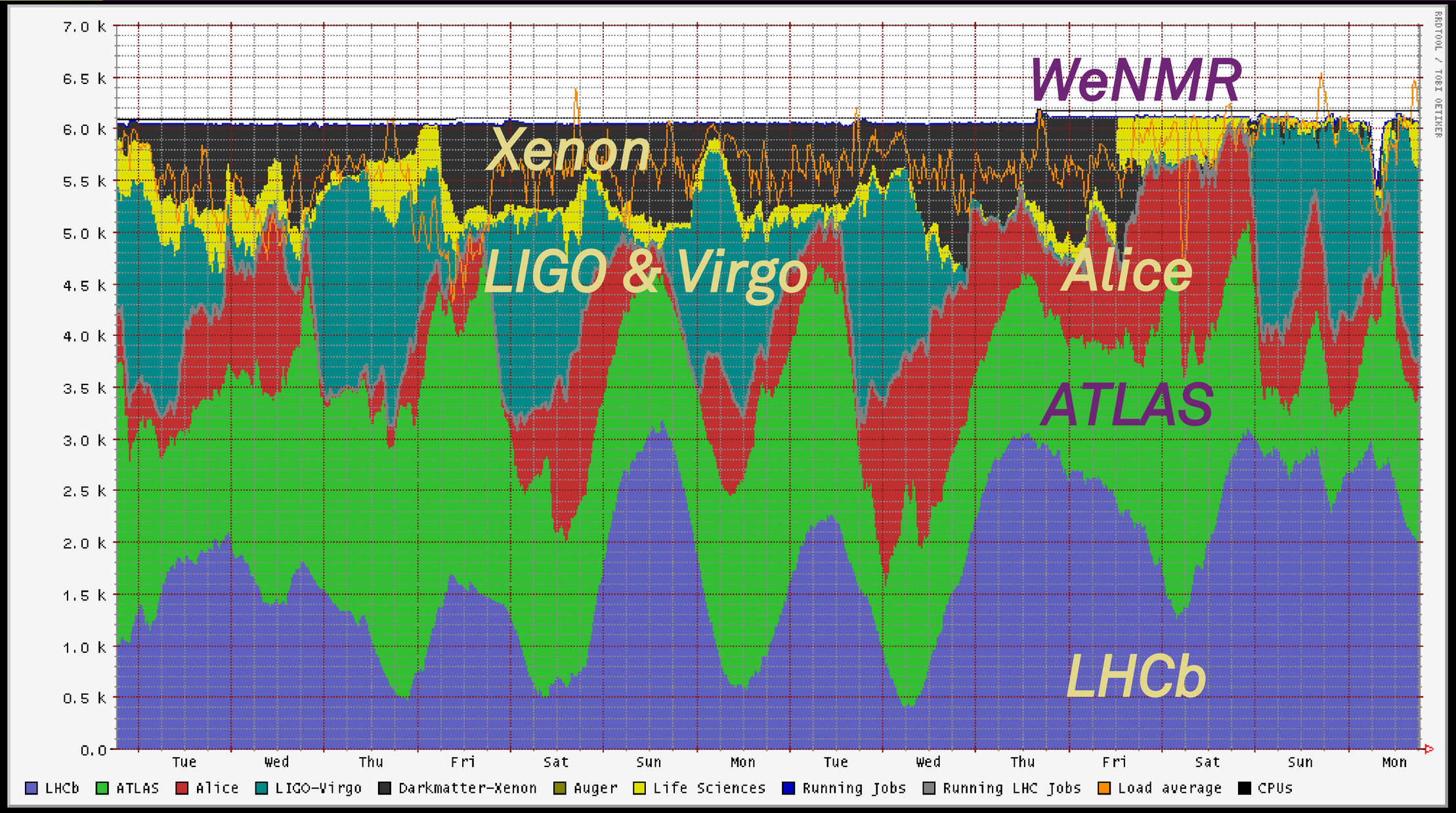
~ 40W

8-10 instructies per tick

Graphic: Herb Sutter
The Free Lunch Is Over: A Fundamental Turn Toward Concurrency in Software

Nikhef

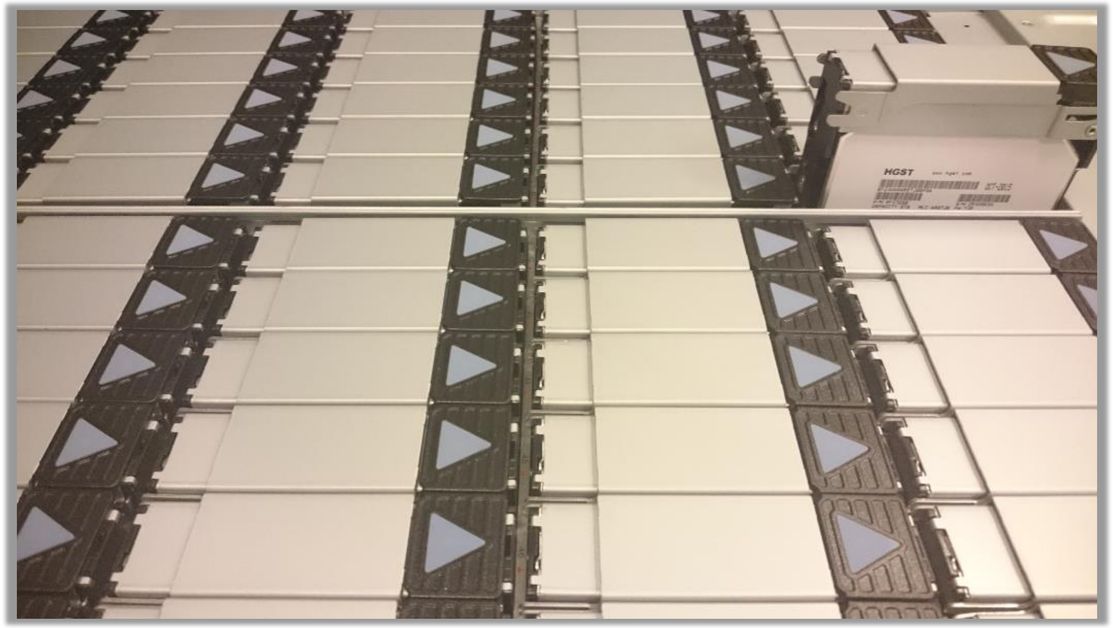
Zo hou je een rekencluster vol:



ew short 1 October 2018

Verdeel en heers

DotHill (HGST): 480 TByte gross capacity/4RU



Kengetallen van opslag

- capaciteit (‘terabytes’)
- bandbreedte (‘megabyte per seconde’)
- aantal ‘opdrachten’ per seconds (“IOPS”)

Daarom kan onze data niet op een USB stick – en doet je ‘thuis NAS’ oplossing het ook niet

… hoe leuk ik mijn eigen opslagdoosje ook vind van slechts 15 Watt met 16Terabyte voor € 915 …

5 Petabyte lezen in 1 dag?
Dat is dus 61 GByte per seconde!
(en dus ~500 Gbps)

Interception, exchanges, and peers

- Full mesh: no single predicable path for traffic
- Exchange point probably does not get you traffic you want
- Exfiltration at ‘the interesting’ edge more feasible:
  - traffic goes to the target
  - maybe fewer eyes on the equipment: attacks are easier
  - maybe your target does not have means to re-route?