

Worker Node Farm migration to IPv6

Author: Bruno Hoefft

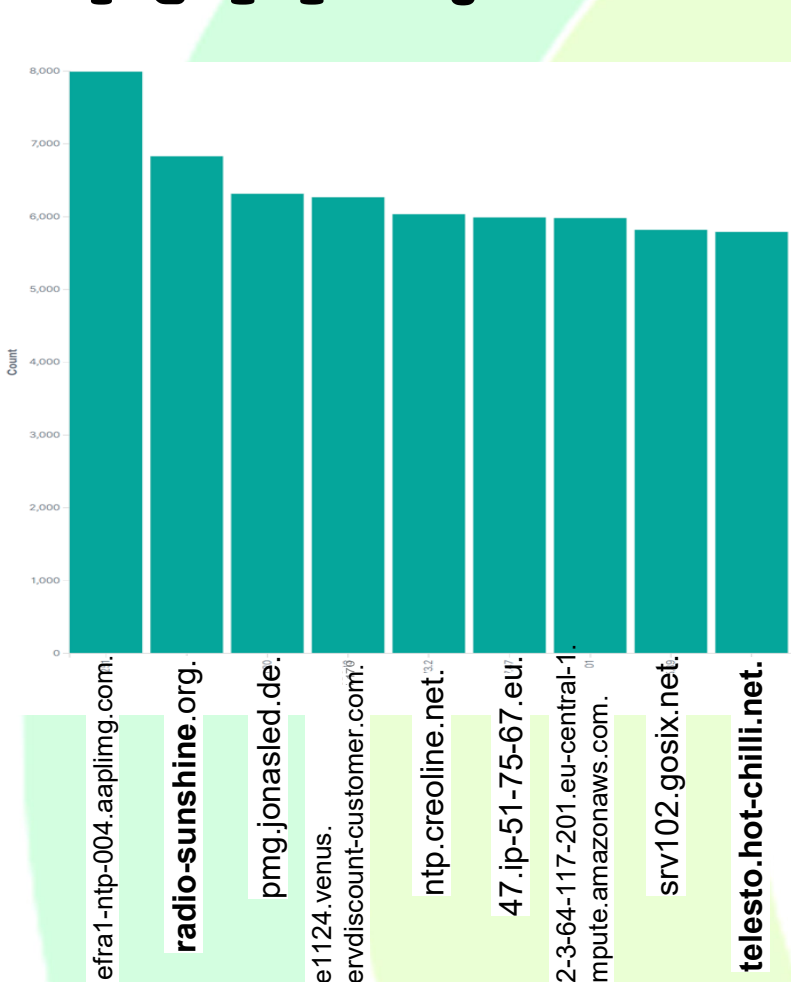
Co-Author: Andreas Petzold, Matthias Schnepf, Max Fischer

Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, {first.familyname}@kit.edu

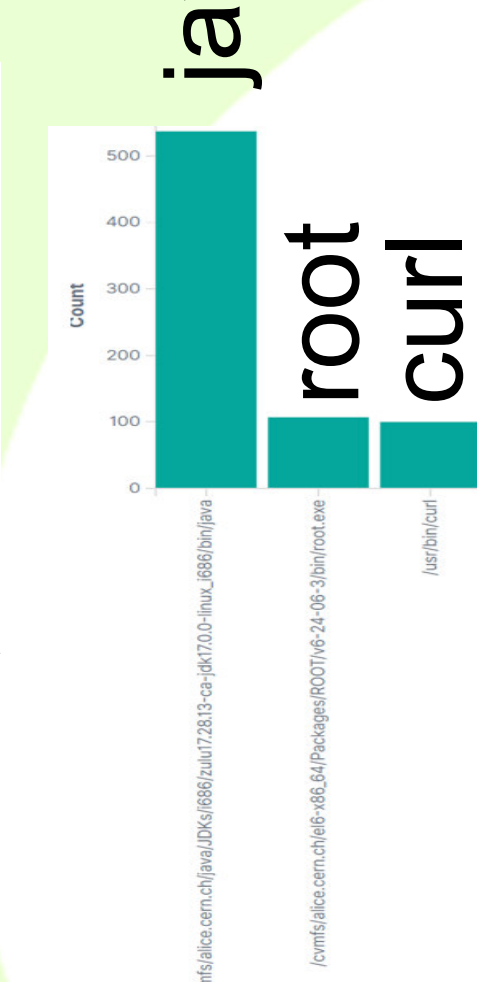
monitor process inter-communication capture at DE-KIT (GridKa) → identifying protocol usage

- packetbeat is collecting the network data and logstash is pushing the data to opensearch (former elastic search) for data storing
- kibana for visualizing
- the monitoring started with a small set of workernodes (storing the data „longterm“ → ~ 6 weeks)
- enlarging workernodes gradually - data keeping time had to be limited to less than one week only (storage size < 0,5 Tbyte)

NTP ?



java



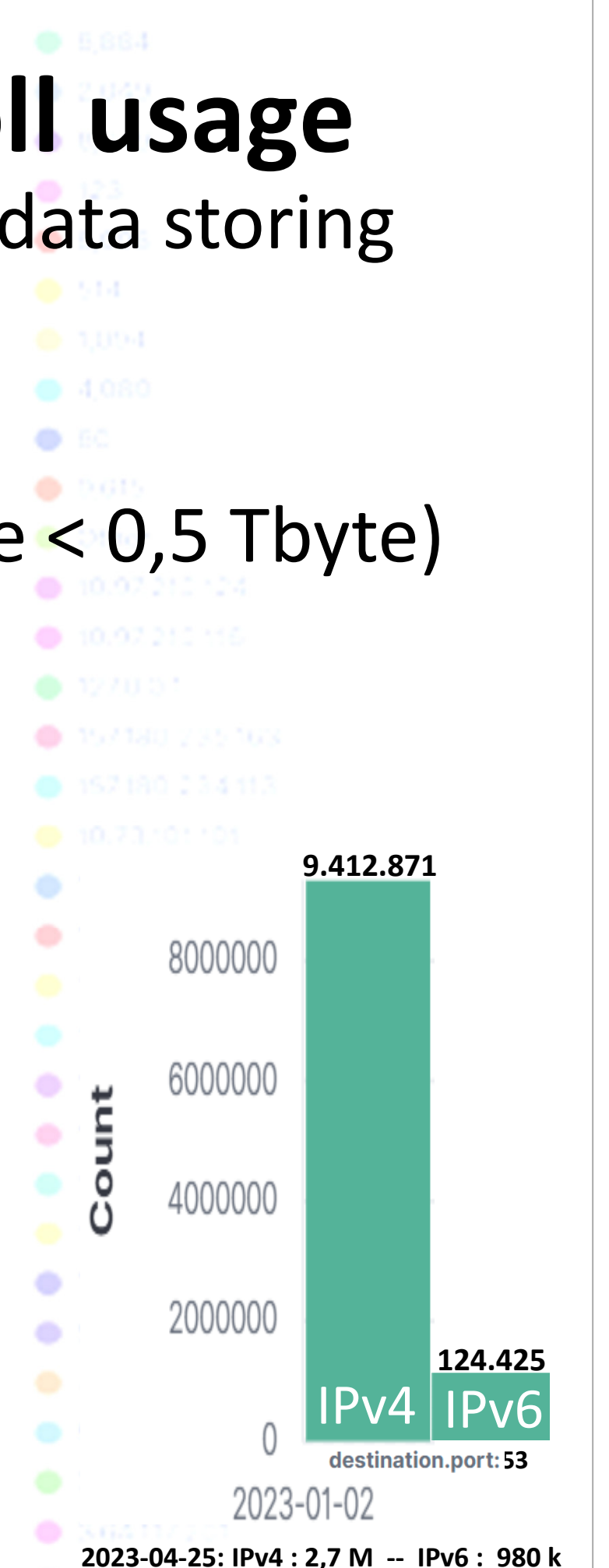
- many NTP / port 123 connections (IPv4 only)
- during 24 hours approx. 210.000
- NTP → IPv4 only (depending on dualstack enabling of rack-manager (40.000 internal))
- 160.000 external communications → some of the destination server have quite dubious „names“
- process-tracking
- the numbers of NTP communication process and matched process is not matching yet

S O L V E D

- NTP.ORG
- returns sometimes funny addresses

closer look at DNS

- GridKa DNS(Port 53):
- IPv4 only count : 9,412,871 (24 hours)
- DNS (Bind) server and WN are already dual-stack
- at WN resolve.conf first lines IPv4
- make sure first IPv6 DNS server addresses listed
- every new deployed host: the first lines are IPv6 resolver addresses of the resolve.conf file followed by the IPv4 addresses

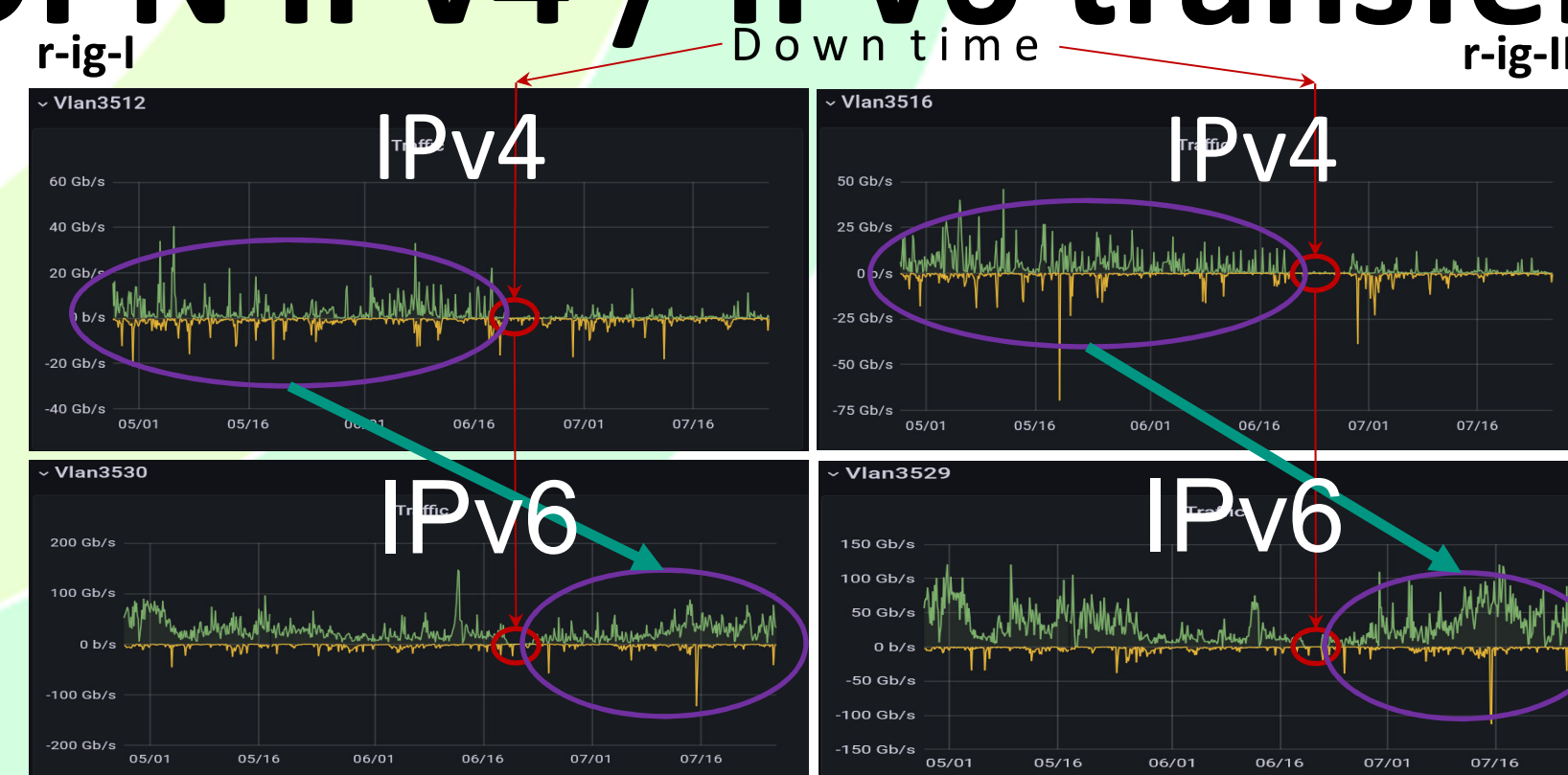


upgrade from dCache version 6.2.34 to 7.2.15

two day downtime at June 20th and 21st 2022

- HTTP-TPC transfers now prefer IPv6 address, if both endpoints support it.
- fixed handling of Storage Resource Reporting (SRR) requests over IPv6
- handle IPv6 address when running HTTP(s) Third Party Copy (TPC) with gridsite delegation
- Storage Resource Manager (SRM) : fix IPV6 logging for SRM

LHCOPN IPv4 / IPv6 transfer pattern after downtime



graph over 90 days traffic of LHCOPN moved from the IPv4 vlans after the downtime to the IPv6 Vlans

details of SQUID

- SQUIDS (proxyserver and Web-Cache):
- some SQUIDS still IPv4 only (migration to dualstack in process)
- significant part of connections via public IPv4
- => to check: if CVMFS can prefer IPv6?

(CVMFS → CernVM-File-System)

- CVMFS sending via http request to squid
- CVMFS has DN configuriert that needs to be resolved → default chooses IPv4 address

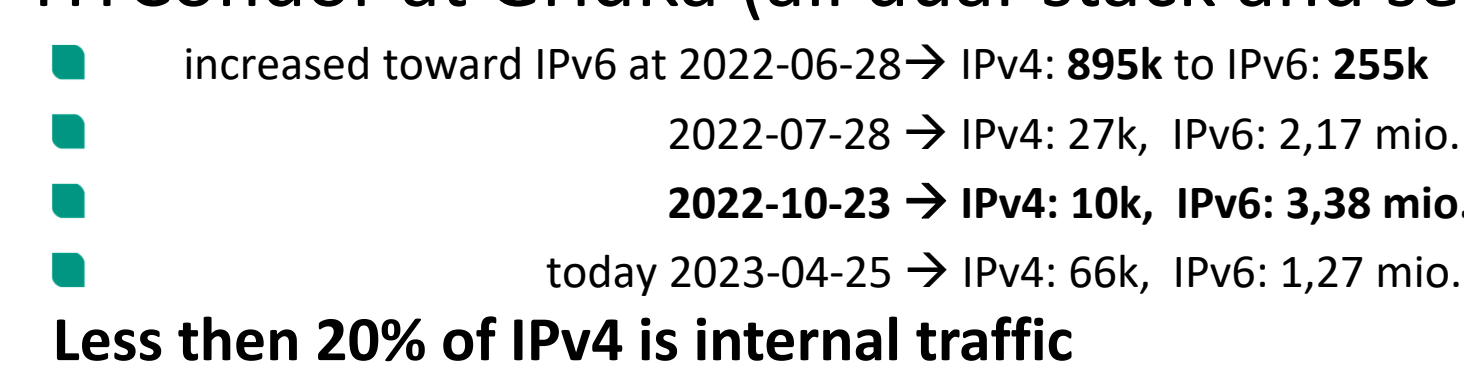
SQUIDS migrated all to dual-stack

During the second half of 2022 all SQUIDS migrated to dual-stack deployment CVMFS now

- manly IPv6 but:
- on WorkerNodes uses IPv6 (with deployed flag: CVMFS_IPFAMILY_PREFER=6)
- CVMFS frontier uses still IPv4 even while both systems dual-stack
- but switching of IPv4 → froniters will operate over IPv6
- statistic:
- July : IPv4 : 1,25 mio. IPv6: 9,6 mio.
- October : IPv4 : 4,44 mio. IPv6: 18 mio.
- December : IPv4 : 1,47 mio. IPv6 : 2,3 mio.

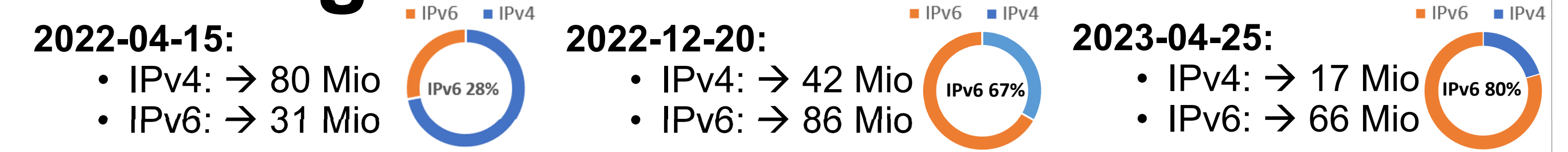
Batch-Processing -- LRMS (HT-Condor) all dual-stack

- LRMS (Local Resource Management System)
- HTCondor at GridKa (all dual-stack and set to prefer the protocol IPv6



Less than 20% of IPv4 is internal traffic

a view general statistics:



Logstash → is now IPv6 (dual-stack deployed)

Ratio 78% IPv6	2022-07-28 → IPv4 385k – IPv6 1,41M	} migration still in progress
Ratio 74% IPv6	2022-10-23 → IPv4 476k – IPv6 1,39M	
Ratio 66% IPv6	2022-12-23 → IPv4 227k – IPv6 450k	
Ratio 68% IPv6	today 2023-04-25 → IPv4 80k – IPv6 170k	

WN – deployment process

- Redhat Satellite Server (foreman)
- used for management of most GridKa hosts:
 - manages redhat subscriptions
 - controls kickstart installations (DHCP / PXE)
 - provides yum repos
 - provides CA (certificate authority) and ENC (encryptor) functionalities for puppet
- uses modular architecture - additional functionalities can be added via so called capsules
 - TFTP server (IPv6 ready - dual-stack)
 - Puppetmaster (IPv6 ready - dual-stack)
 - Pulp (software repository management (IPv6 ready - dual-stack))
 - DNS (IPv6 ready - dual-stack)
 - DHCP (currently DHCPv6 capsule not available)

Tasks ToDo:

- migration of Rack Manager – work in progress
- narrow down the still IPv4 communication
- packet monitoring configured
 - to list all unhandled IPv4 packets
 - 4080 – Condor rooster Monitor daemon → solved
 - 8884 – ALICE: operation report
 - 2049 – NFS
 - 8649 – Ganglia gmond
 - 1094 – XrootD
 - 961[89] – LRMS (20% only internal to WN-Farm)
- PXE – Boot + DHCPv6 (first boot addr. Distribution)
- identify the next service for IPv6 migration tasks