Showing More Big Data

Towards visualisation of compute clusters in a mixed-purpose ELK Analytics Cluster

HvA induction session September 2016

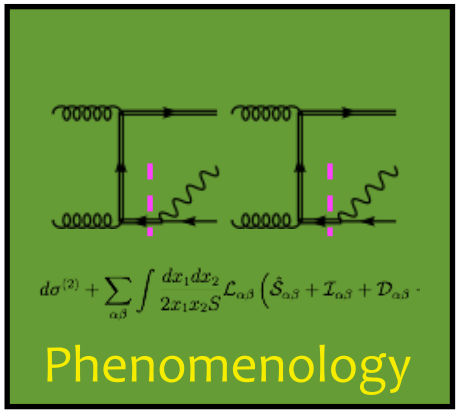
David Groep, Nikhef

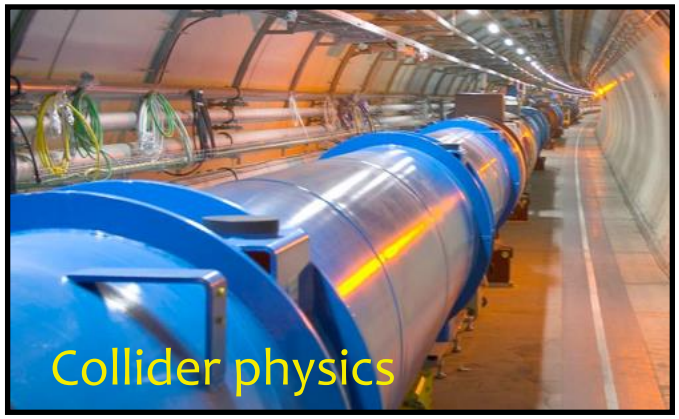
David Groep
Nikhef
Amsterdam
PDP programme

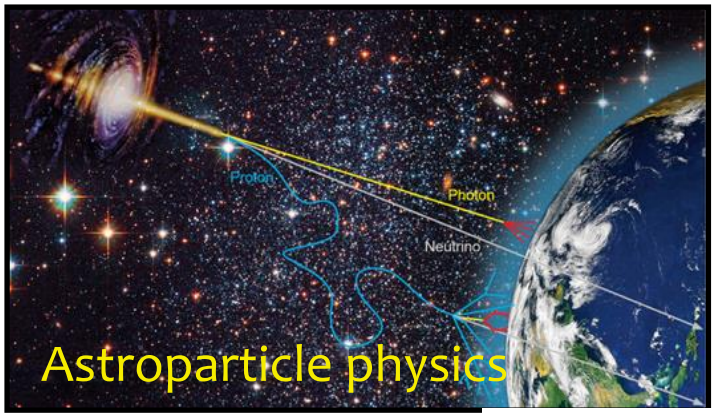
Nationaal instituut voor subatomaire fysica

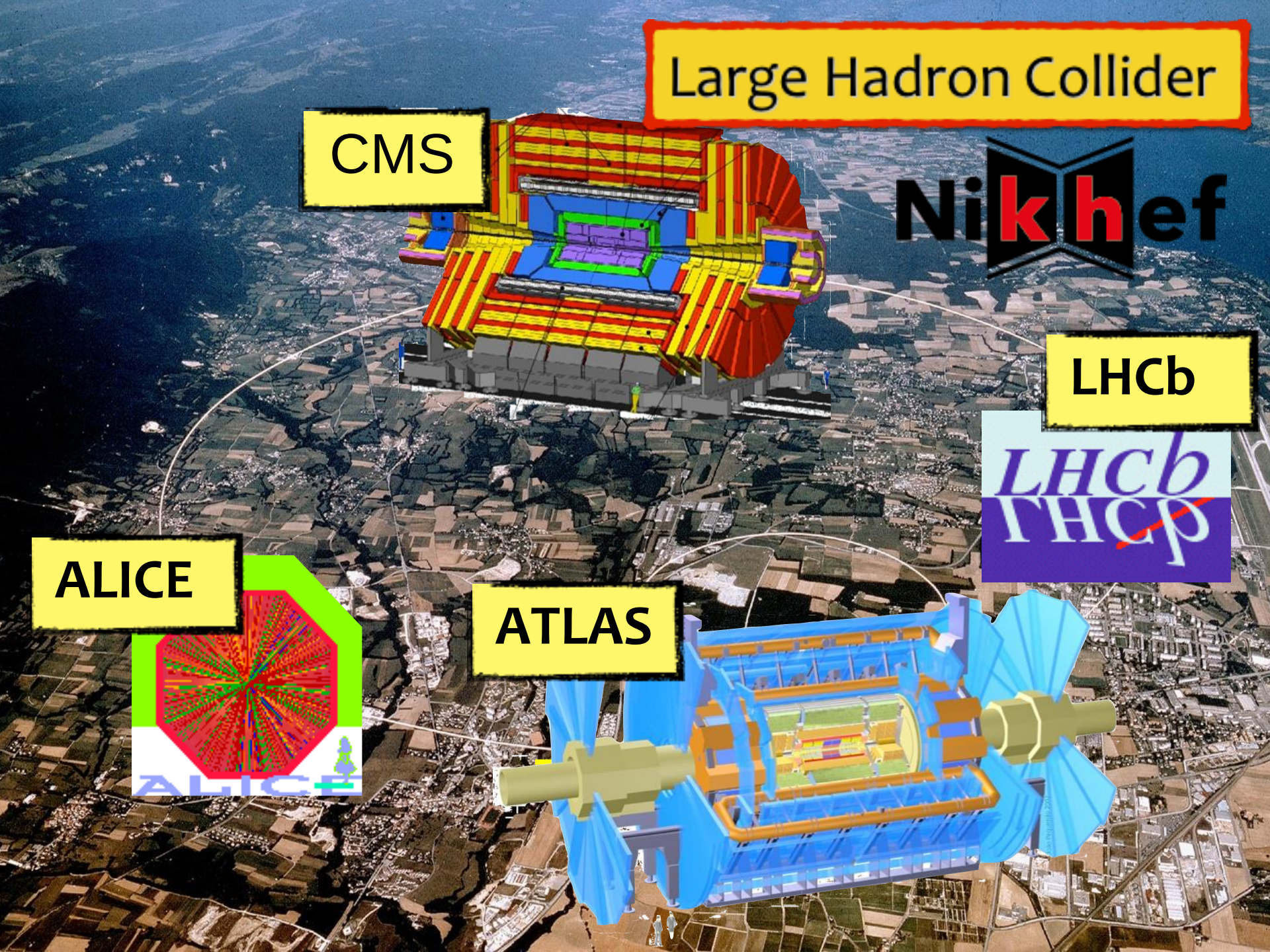
Verleggen van de grenzen van onze kennis

- **Accelerator-based particle physics**
  Experiments studying interactions in particle collision processes at particle accelerators, in particular at CERN;
- **Astroparticle physics**
  Experiments studying interactions of particles and radiation emanating from the Universe.

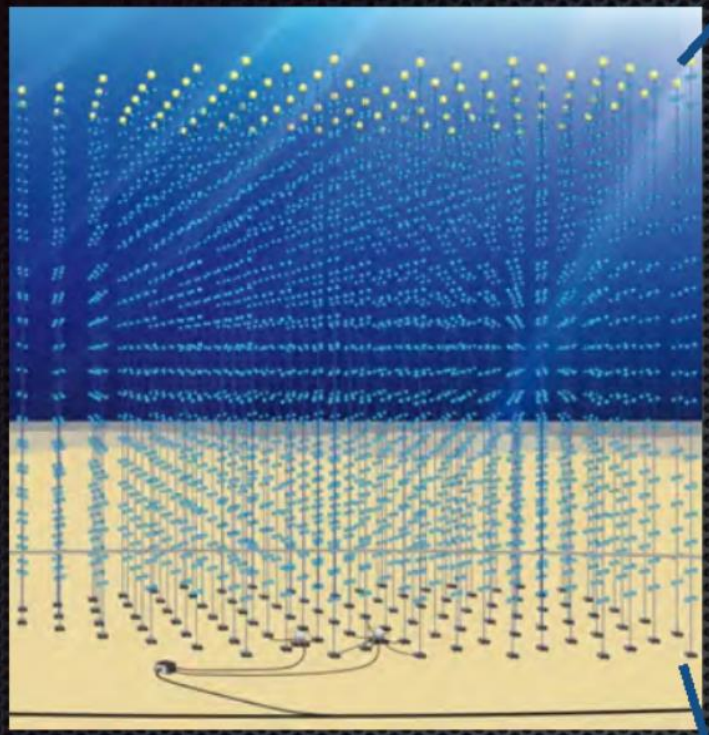


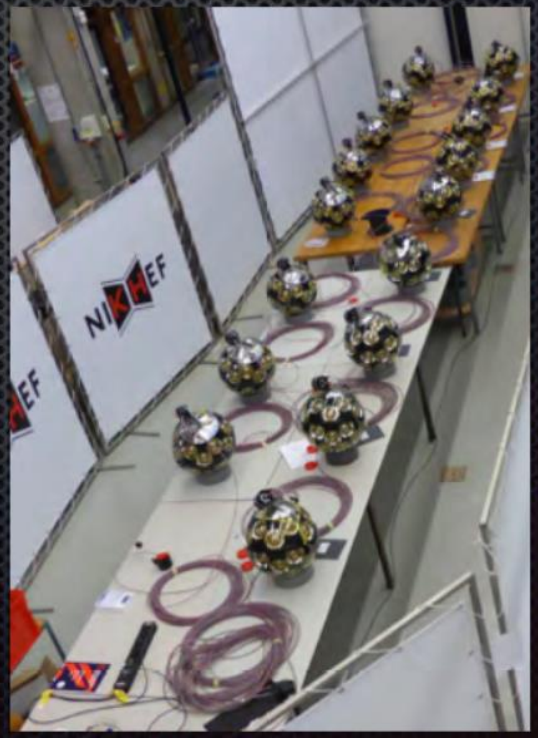





Large Hadron Collider
CMS
Nikhef
LHCb
LHCb
ALICE
ALICE
ATLAS

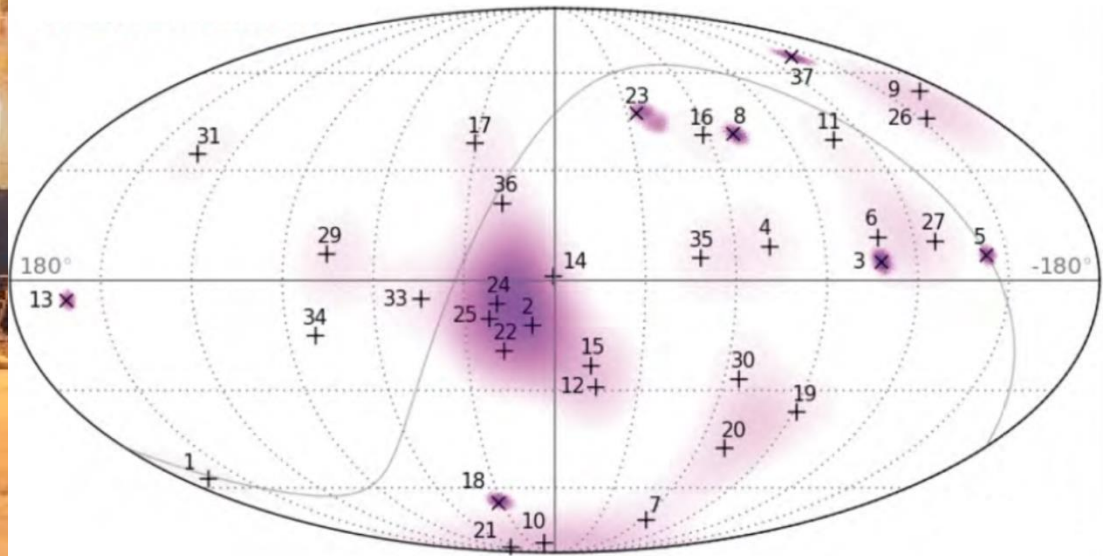
Nikhefs neutrino-detector: KM3NeT

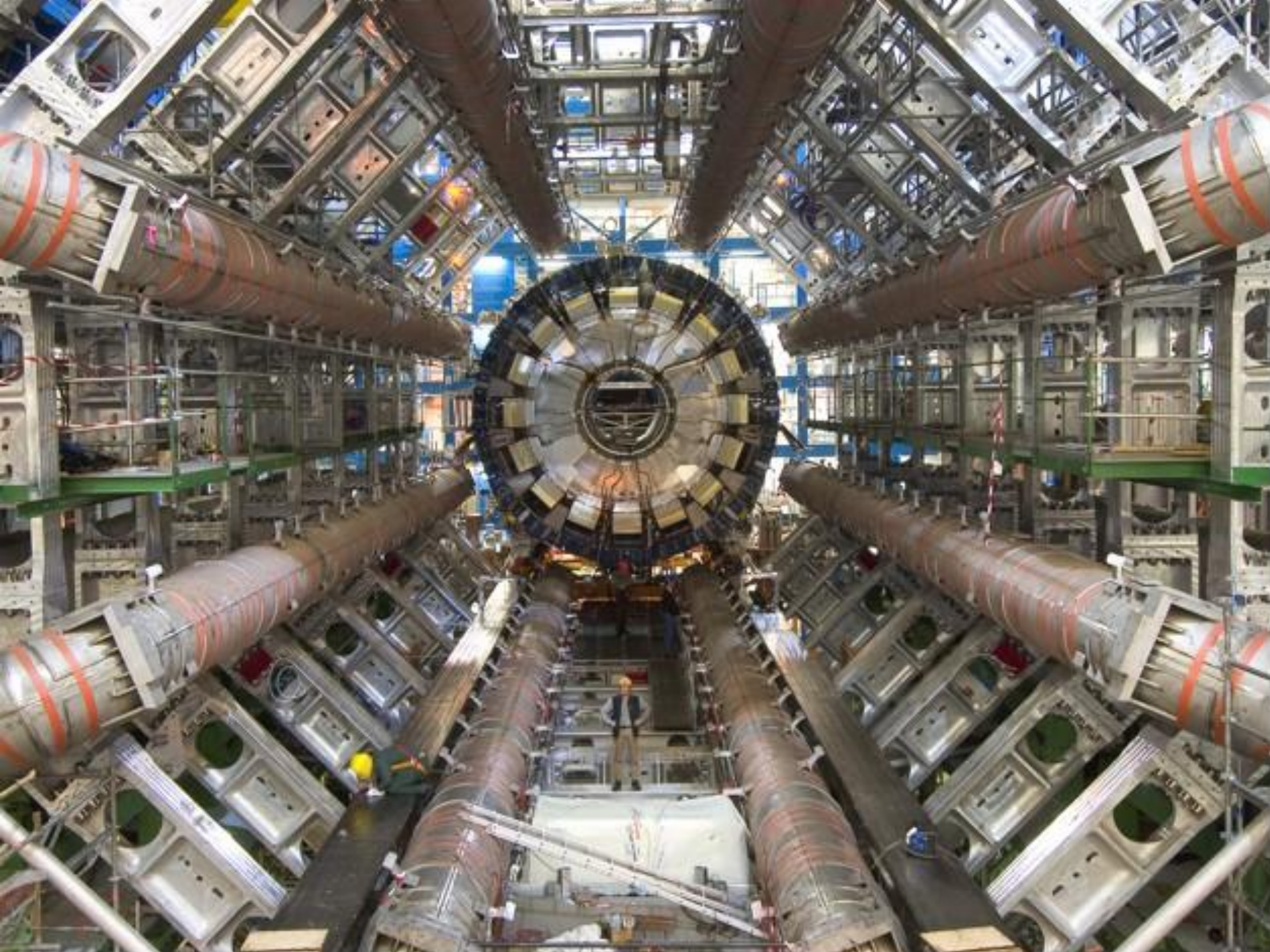




Little white structures prevent the HV bases and cables to touch each other

De Melkweg

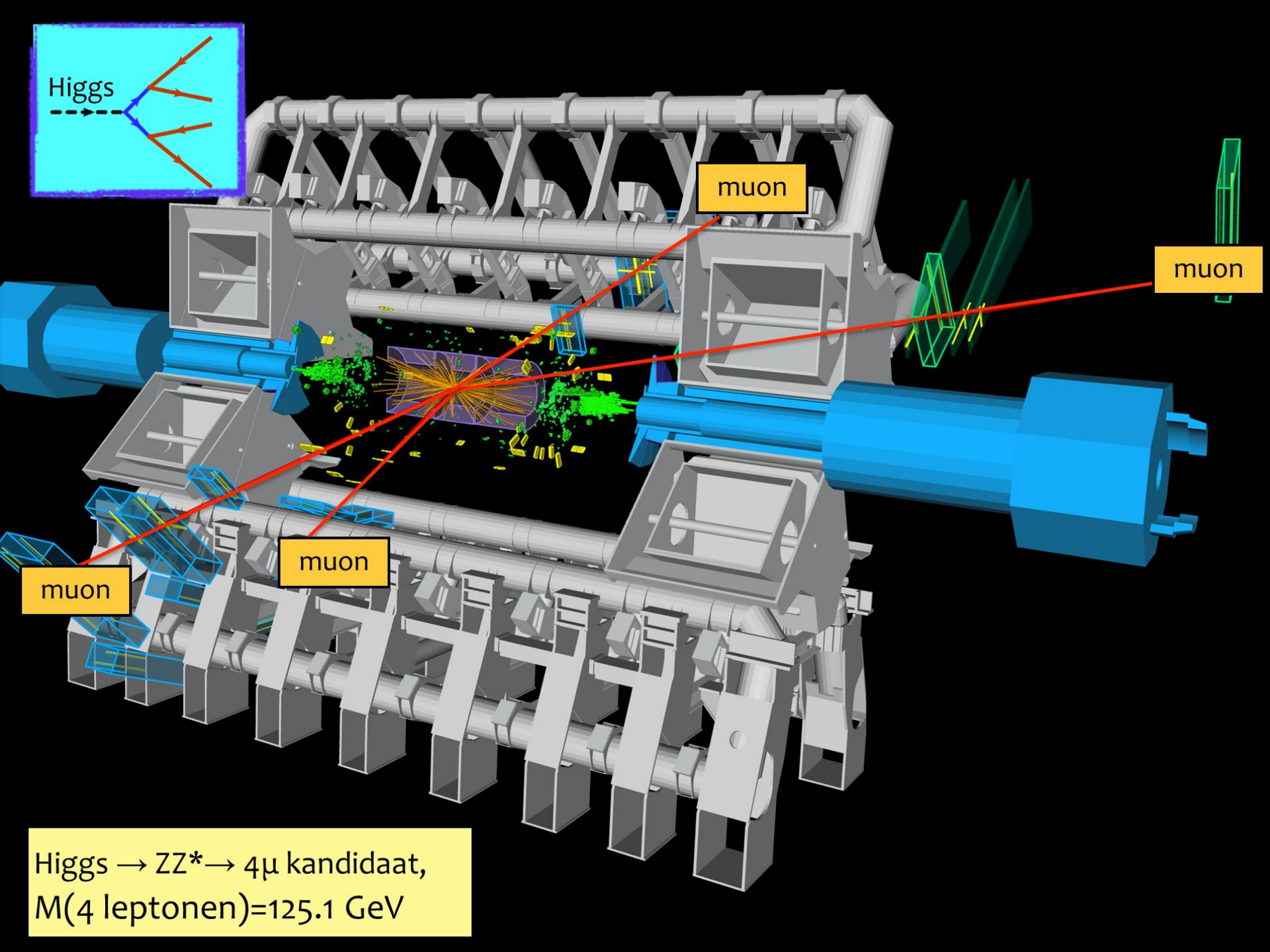




Kans Higgs deeltje:

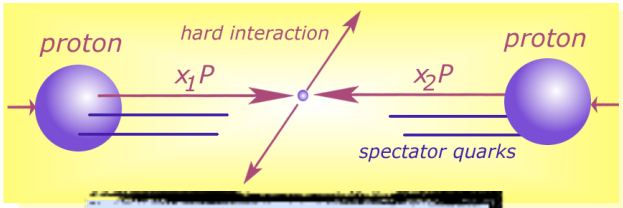
1 op de 1.000.000.000.000 bostingen

- Dit is equivalent met zoeken van 1 persoon op 1000 wereldpopulaties
- Oftewel één naald in 20 miljoen hooibergen

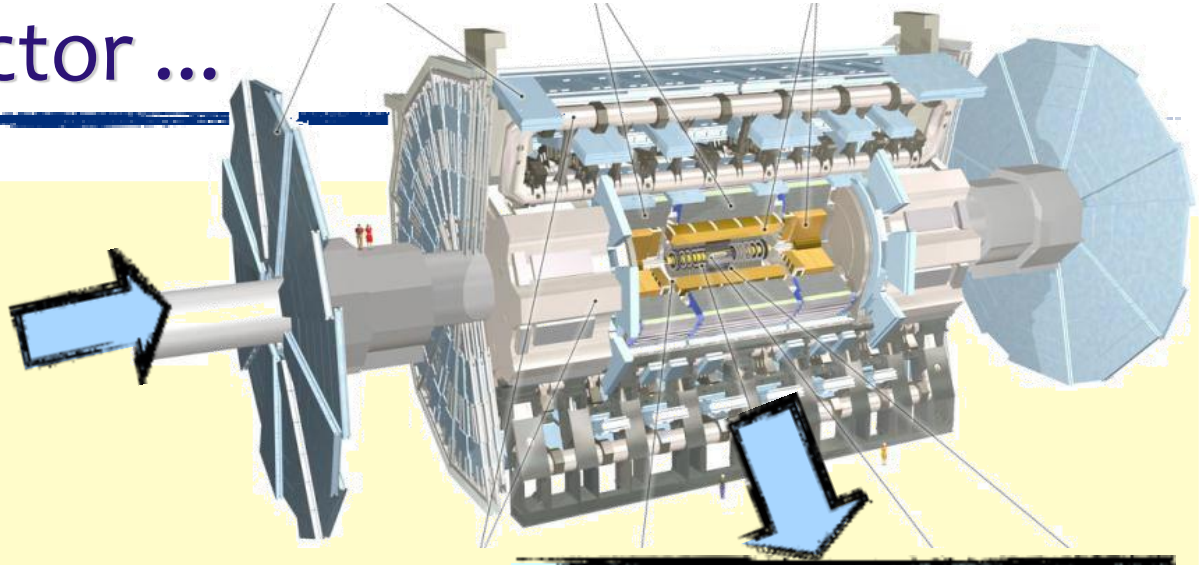


Higgs → ZZ*→ 4μ kandidaat,
M(4 leptonen)=125.1 GeV

Detector to doctor ...

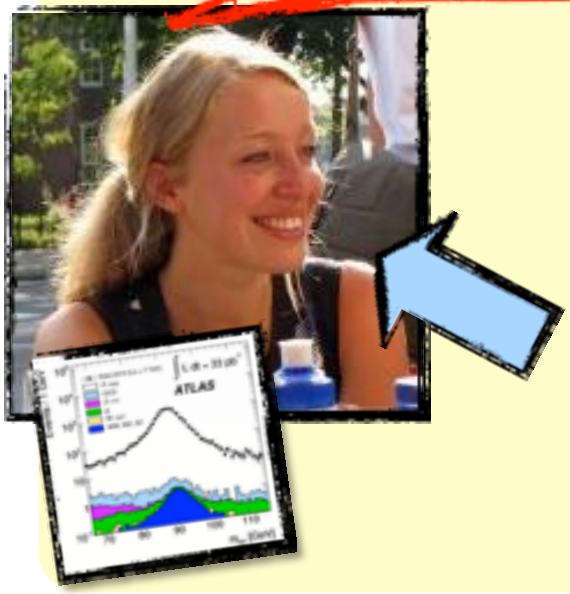


40 miljoen / seconde



Trigger systeem selecteert 600 Hz
~ 1 GB/s data

Analyse van botsingen door promovendi



and processing

Data distributie met GRID computers

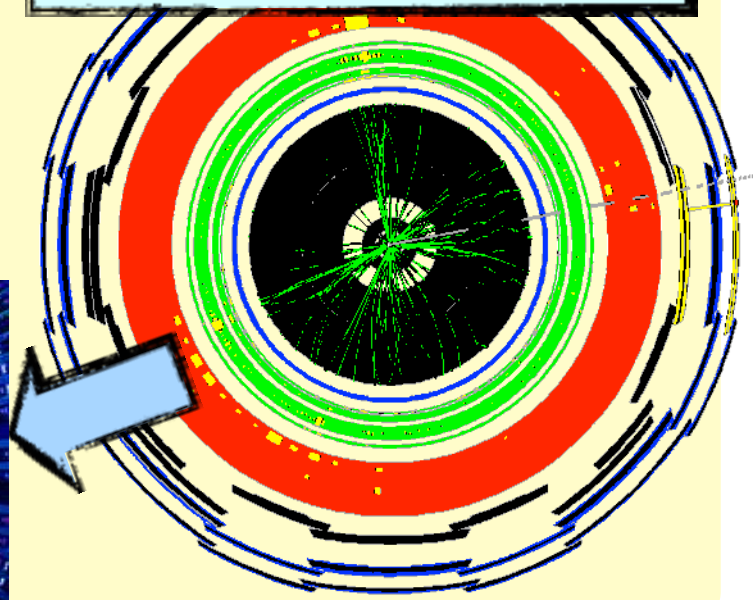


Image source: joint CERN (wLCG) and EGI

Organisations participating in the global collaboration of e-Infrastructures

Even just for wLCG, supporting the CERN LHC programme

More than 200 independent institutes with end-users
More than 50 countries & regions
More than 300 service centres
Handful regional 'service coordination organisations'
500 000 CPU cores, 200+PByte storage
One independent 'policy-bridge' PKI

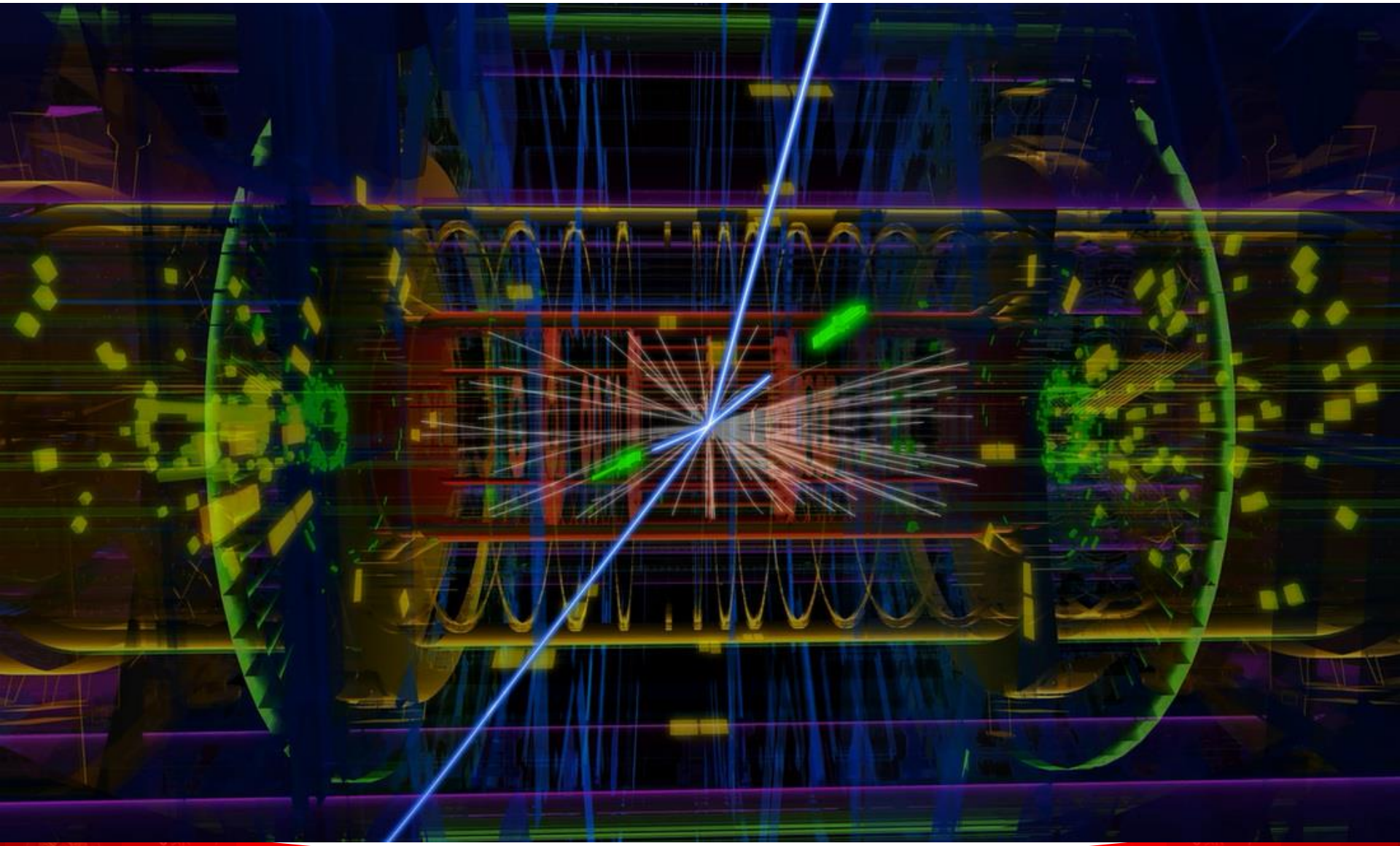
PRACE

Open Science Grid

EGI

XSEDE
Extreme Science and Engineering Discovery Environment

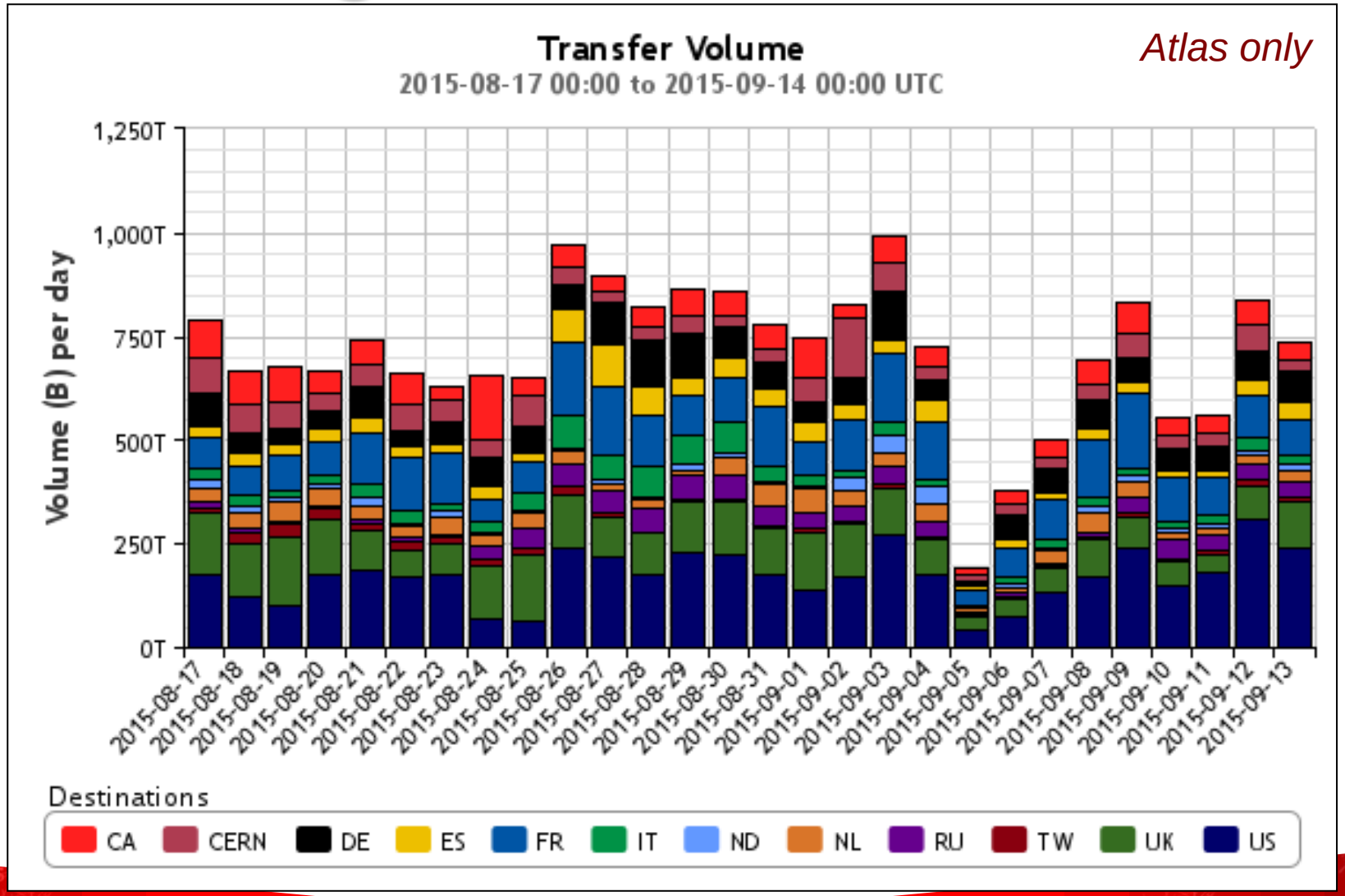
Atlas: ~50 TByte/day raw data to tape; 1000 TByte/day processed data transfers



Nikhef

Image source: CERN, Atlas

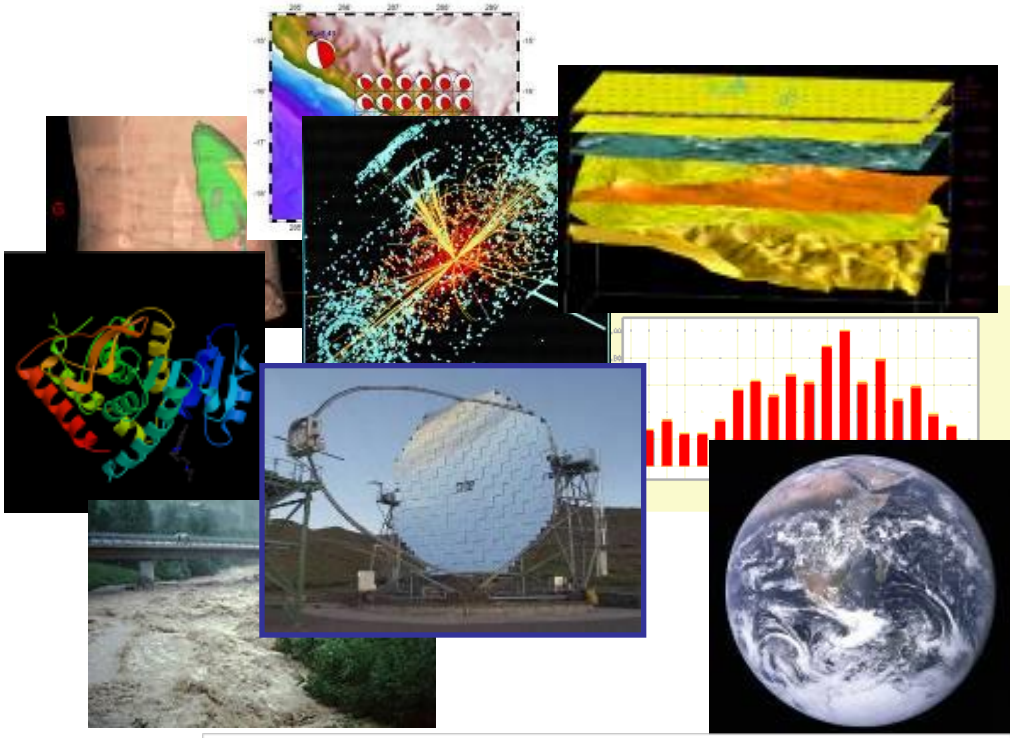
Big ‘as in Large’ Data

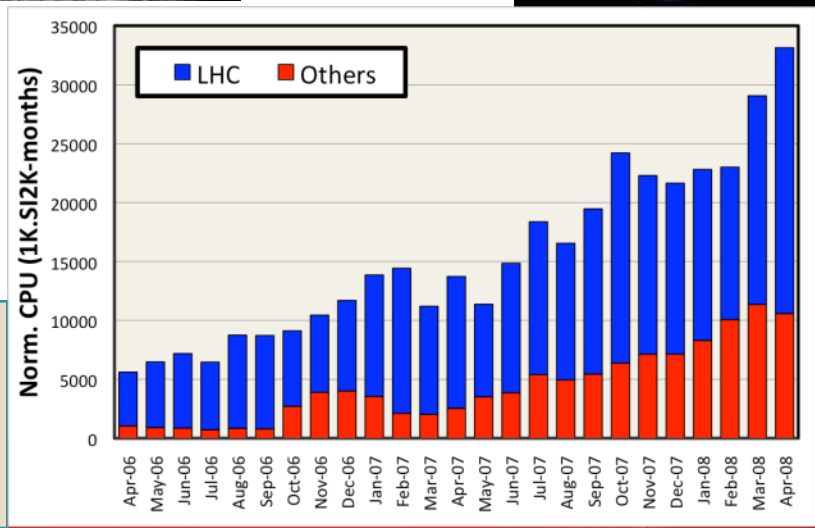


David Groep
Nikhef
Amsterdam
PDP programme

Shared e-Infrastructure

- >270 communities from many different domains
  - Astronomy & Astrophysics
  - Civil Protection
  - Computational Chemistry
  - Comp. Fluid Dynamics
  - Computer Science/Tools
  - Condensed Matter Physics
  - Earth Sciences
  - Fusion
  - High Energy Physics
  - Life Sciences
  - …





David Groep
Nikhef
Amsterdam

Applications have moved from testing to routine and daily usage

~80-95% efficiency

Global data flows

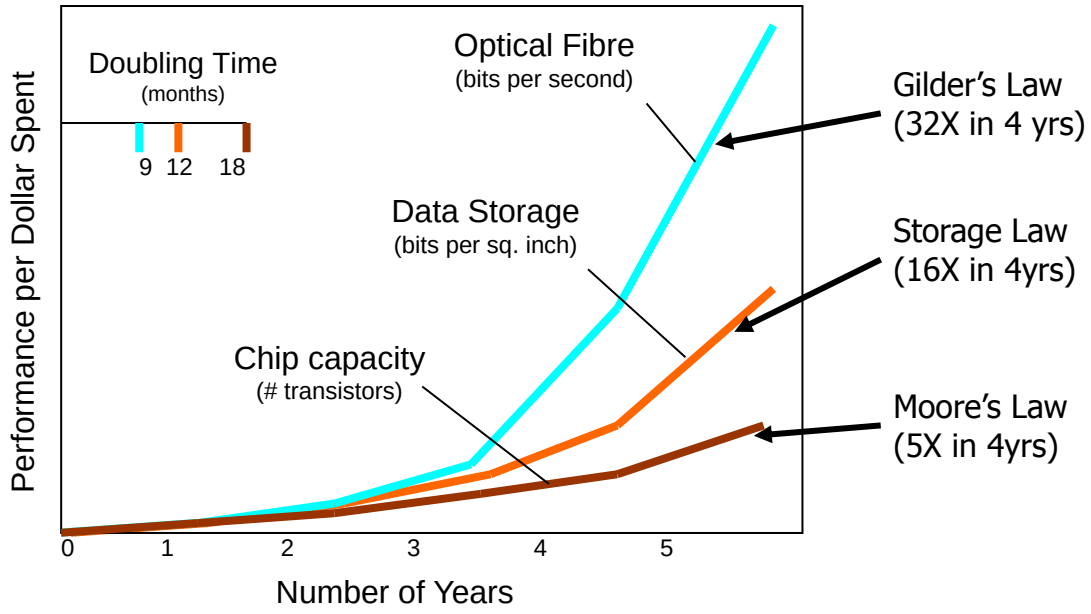
~150GByte, 12hrs
per (human) genome
per sequencer

But 1000+ sequencers…

Genome sequencing at the Beijing Genomics Institute BGI
Photo: Scotted400, CC-BY-3.0

David Groep
Nikhef
Amsterdam
PDP programme

There's always a network close to you



NL Light

SURF NET

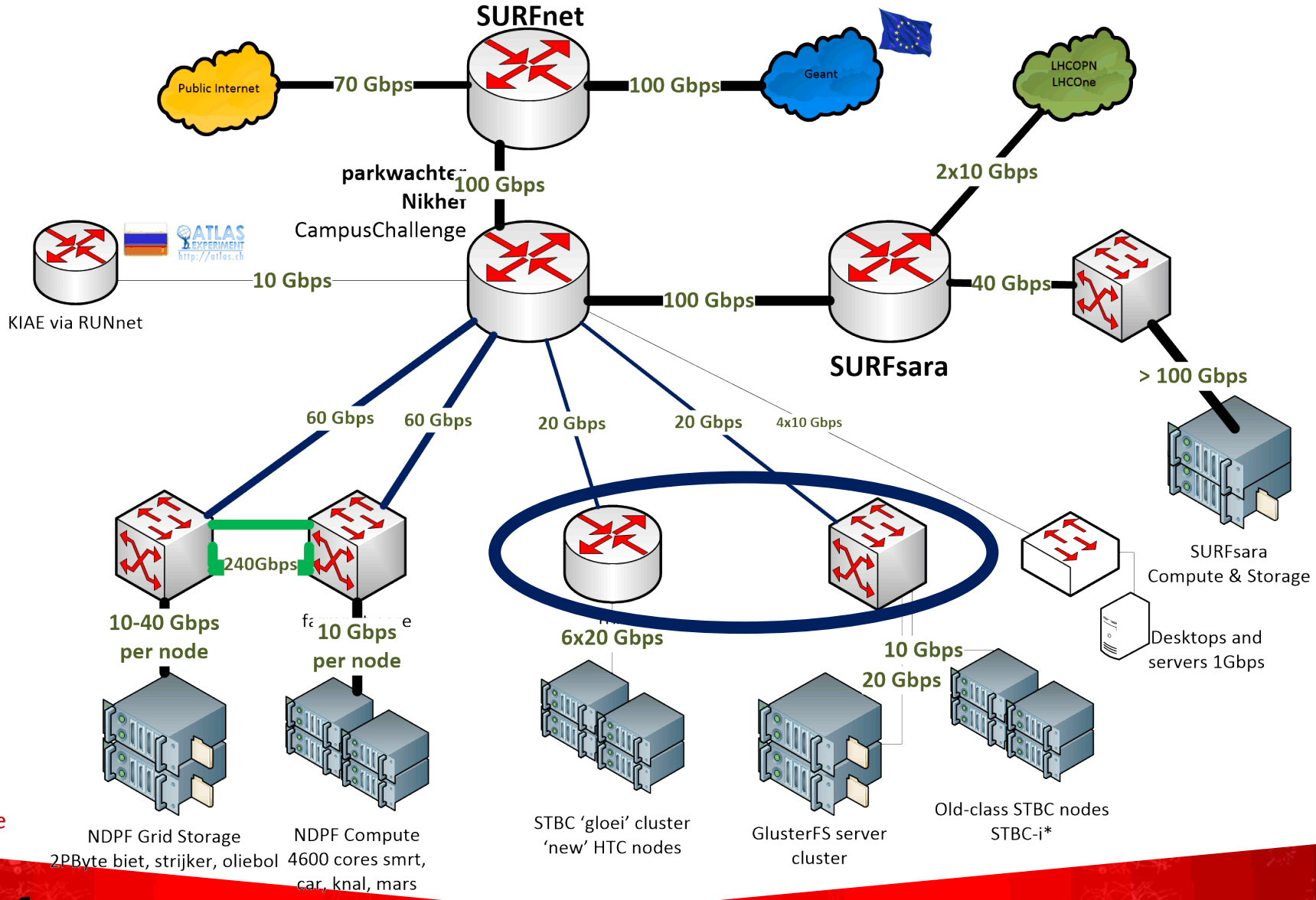
SURFnet pioneered 'lambda' and hybrid networks in the world

- and likely contributed to the creation of a market for 'dark fibre' in the Netherlands

There's always fibre within 2 miles from you – where ever you are!

(it's just that last mile to your home that's missing – and the business model of your telecom provider…)

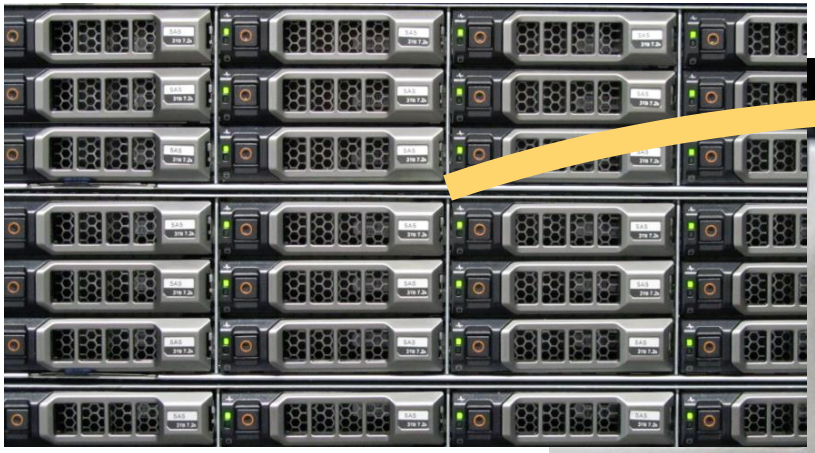
Nikhef Data Processing network



David Groep
Nikhef
Amsterdam
PDP programme

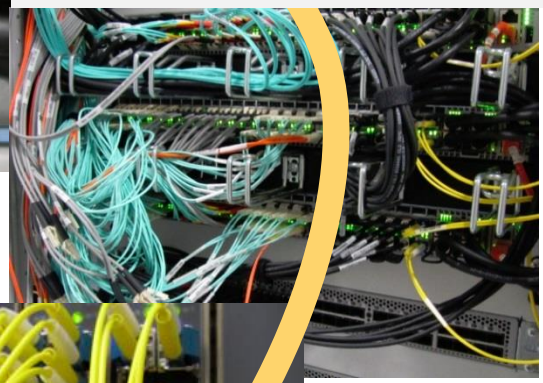
Nikhef

The Flow of Data at Nikhef



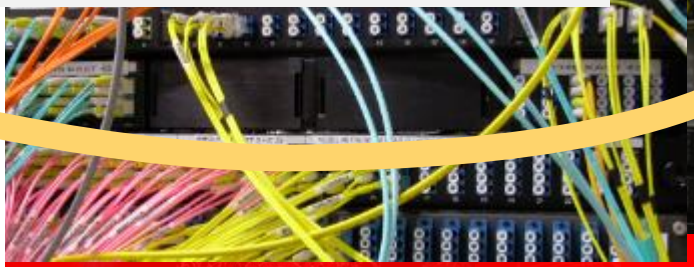
10 – 40 Gbps per server

240 Gbps interconnect



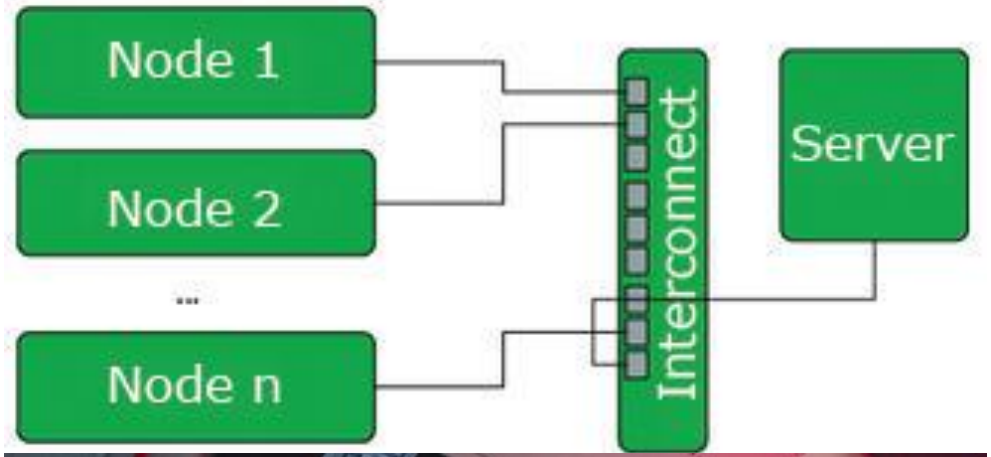
44 disk servers
~3 PiB (~3000 TByte)
2 control & DB nodes

Peerings: SURFnet, SURFsara, Kurchatov, AMOLF, CWI, LHCOPN, LHCOne via SARA



>200 Gbps uplinks

Batch processing – cluster setup



David Groep
Nikhef
Amsterdam
PDP programme

Big Data Analytics for log analysis

- Analysis of log data is typical 'big data' problem
  - CERN tried Hadoop ('map-reduce')
  - RAL went with … ELK*
- For logs specifically, it's mostly efficient search
  - ElasticSearch (www.elastic.co)
  - LogStash (collect and parse logs, import to ES)
  - Kibana – analysis based on Apache Lucene + graphing
- Integrated into a single 'stack': ELK

David Groep
Nikhef
Amsterdam
PDP programme

*Appleyard et al., http://pos.sissa.it/archive/conferences/239/027/ISGC2015_027.pdf

LogStash

Data arrives in format A…
…process B occurs…
…data out in format C

e.g. convert syslog into json

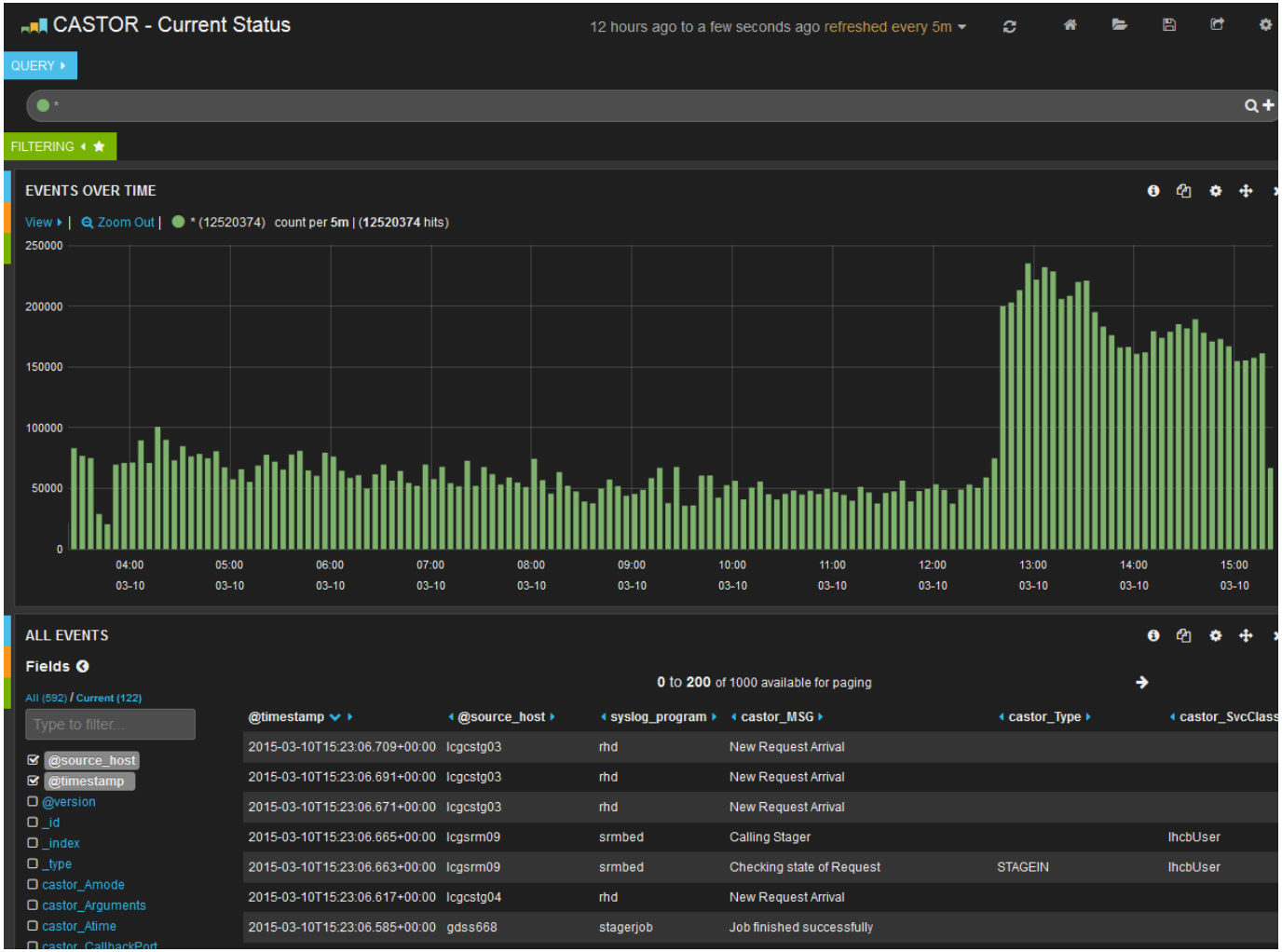
plugin documentation

inputs	codecs	filters	outputs
collectd	cloudtrail	advisor	boundary
drupal_dblog	collectd	alter	circonus
elasticsearch	compress_spooler	anonymize	cloudwatch
eventlog	dots	checksum	csv
exec	edn	cidr	datadog
file	edn_lines	cipher	datadog_metrics
ganglia	fluent	clone	elasticsearch
gelf	graphite	collate	elasticsearch_http
gemfire	json	csv	elasticsearch_river
generator	json_lines	date	email
graphite	json_spooler	dns	exec
heroku	line	drop	file
imap	msgpack	elapsed	ganglia
invalid_input	multiline	elasticsearch	gelf
irc	netflow	environment	gemfire
jmx	noop	extractnumbers	google_bigquery
log4j	oldlogstashjson	fingerprint	google_cloud_storage
lumberjack	plain	gelfify	graphite
pipe	rubydebug	geoip	graphtastic
puppet_facter	spool	grep	hipchat
rabbitmq		grok	http
rackspace		grokdiscovery	irc
redis		i18n	jira
relp		json	juggernaut
s3		json_encode	librato
snmptrap		kv	loggly
sqlite		metaevent	lumberjack

David Groep
Nikhef
Amsterdam
PDP programme

Nikhef

Analyse ‘by hand’: Kibana & Lucene



David Groep
Nikhef
Amsterdam
PDP programme

Graphic from Rob Appleyard, RTFC RAL – at ISGC 2015

Challenges ahead!

At the start

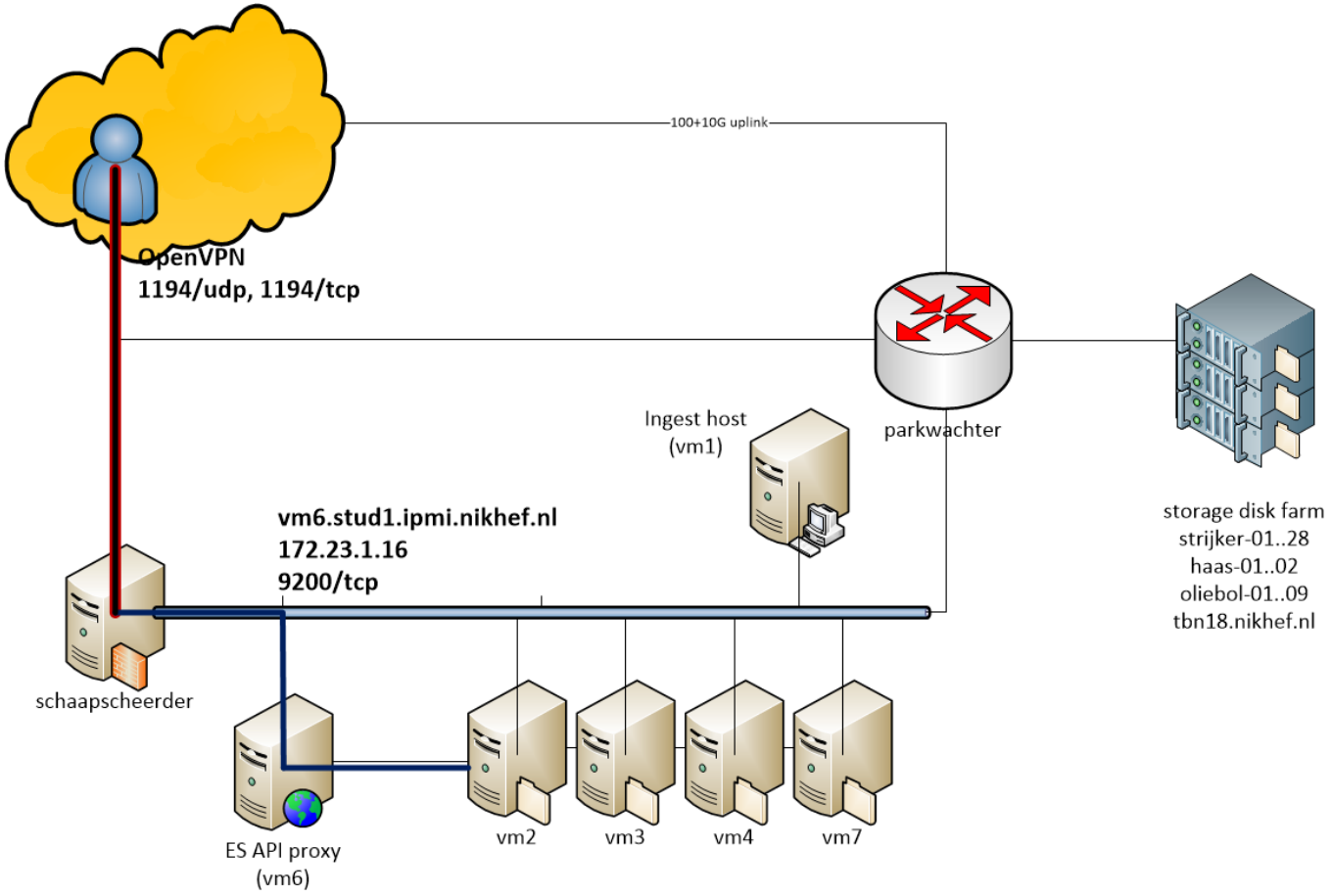
- ~500 difference data sources, 5500 CPU cores
- 5000+ users, working 24x7

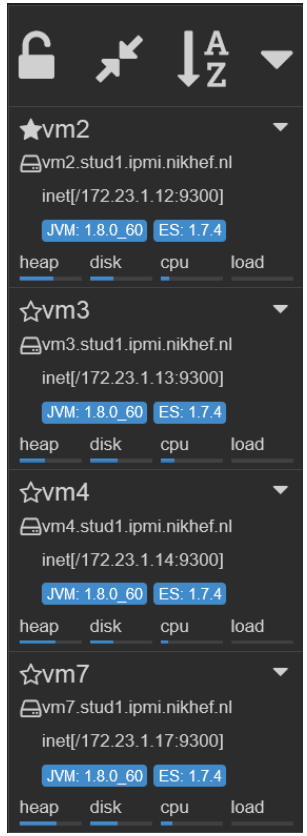
Your colleagues in 2015-2016 added

- setup of a big data analytics cluster (ELK)
- define queries and find some global anomalies ☺
- Focus on data transfers, but no compute …

David Groep
Nikhef
Amsterdam
PDP programme

ELK Cluster and Interface





David Groep
Nikhef
Amsterdam
PDP programme

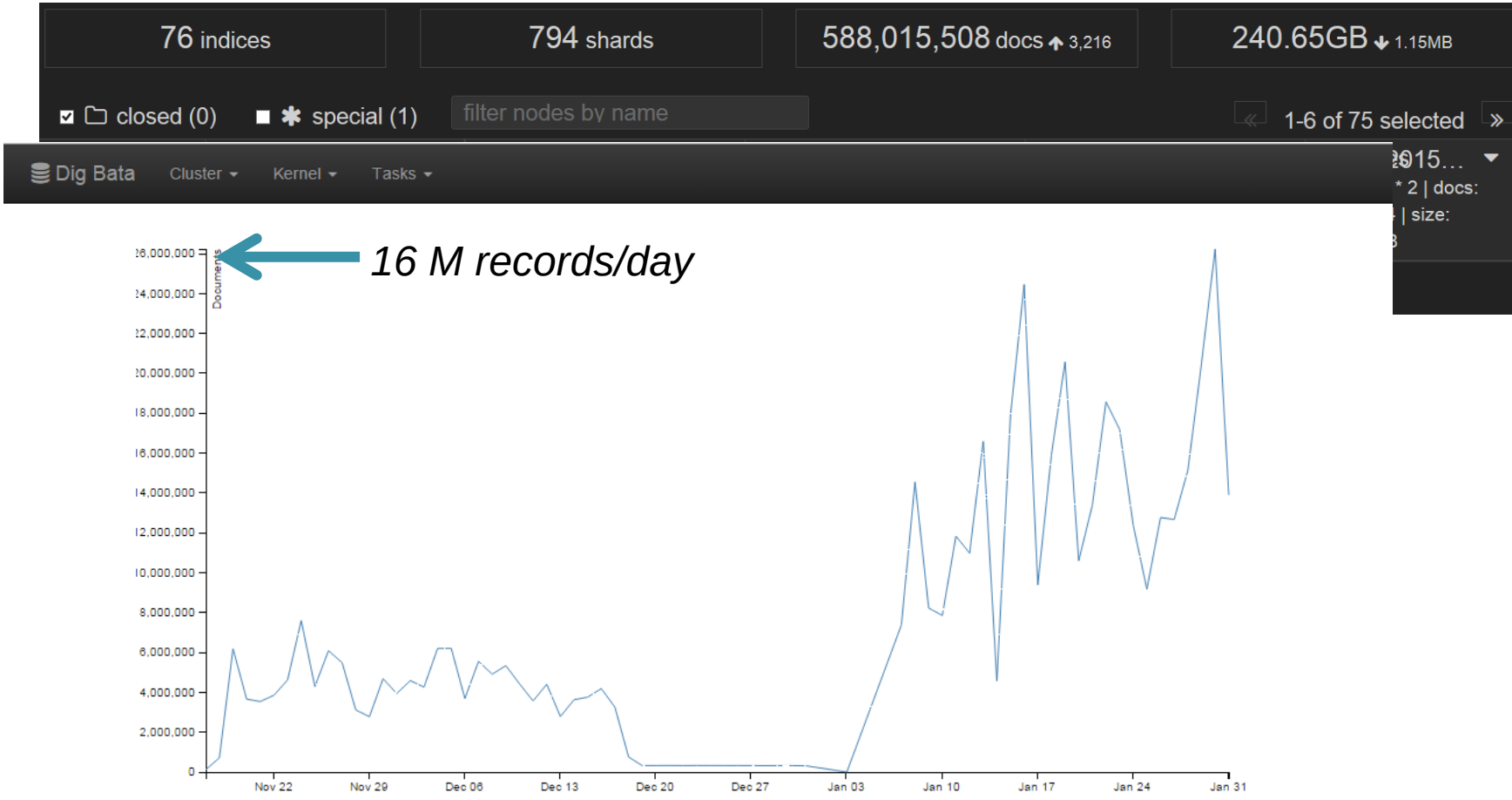
Filled with data

filter nodes by name | master | data | client

name	load average	cpu %	heap usage %	disk usage %	uptime
logstash-vm5.stud1.ipmi.nikhef.nl-7420-... vm5.stud1.ipmi.nikhef.nl inet[/172.23.1.15:9300] JVM: 1.8.0_60 ES: 1.7.0	N/A		25.0 used: 126.69MB max: 491.69MB	*no disk info for client nodes*	3d.
vm2 vm2.stud1.ipmi.nikhef.nl inet[/172.23.1.12:9300] JVM: 1.8.0_60 ES: 1.7.4		22.0	59.0 used: 2.36GB max: 3.98GB	38.0 free: 121.55GB total: 196.74GB	9d.
vm3 vm3.stud1.ipmi.nikhef.nl inet[/172.23.1.13:9300] JVM: 1.8.0_60 ES: 1.7.4		18.0	58.0 used: 2.33GB max: 3.98GB	45.0 free: 107.77GB total: 196.74GB	7d.
vm4 vm4.stud1.ipmi.nikhef.nl inet[/172.23.1.14:9300] JVM: 1.8.0_60 ES: 1.7.4		15.0	73.0 used: 2.94GB max: 3.98GB	37.0 free: 123.51GB total: 196.74GB	7d.
vm7 vm7.stud1.ipmi.nikhef.nl inet[/172.23.1.17:9300] JVM: 1.8.0_60 ES: 1.7.4		20.0	59.0 used: 2.36GB max: 3.98GB	44.0 free: 109.89GB total: 196.74GB	9d.

David Groep
Nikhef
Amsterdam
PDP programme

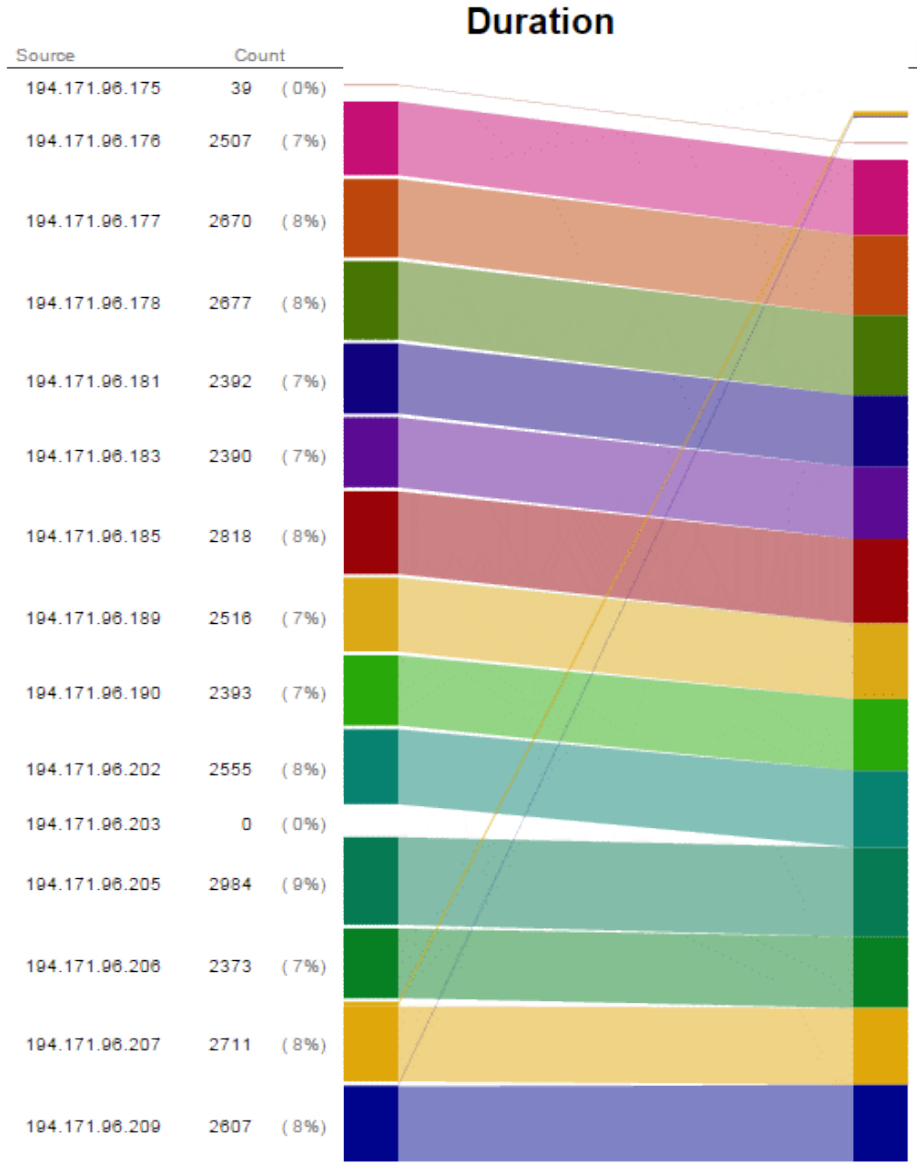
Filling it up: from 0 to ~ 500 million

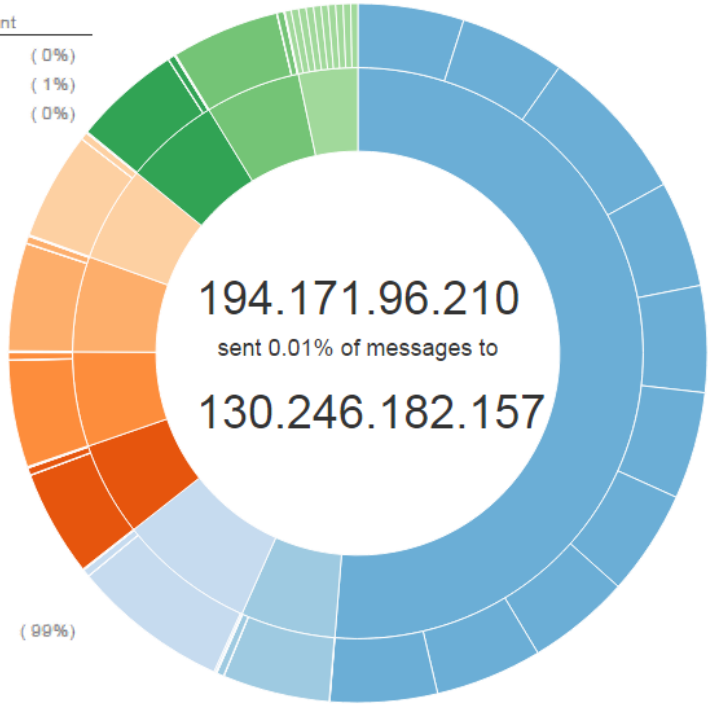


David Groep
Nikhef
Amsterdam
PDP programme

Graphics courtesy Jouke Roorda, Olivier Verbeek, Rens Visser

Nikhef



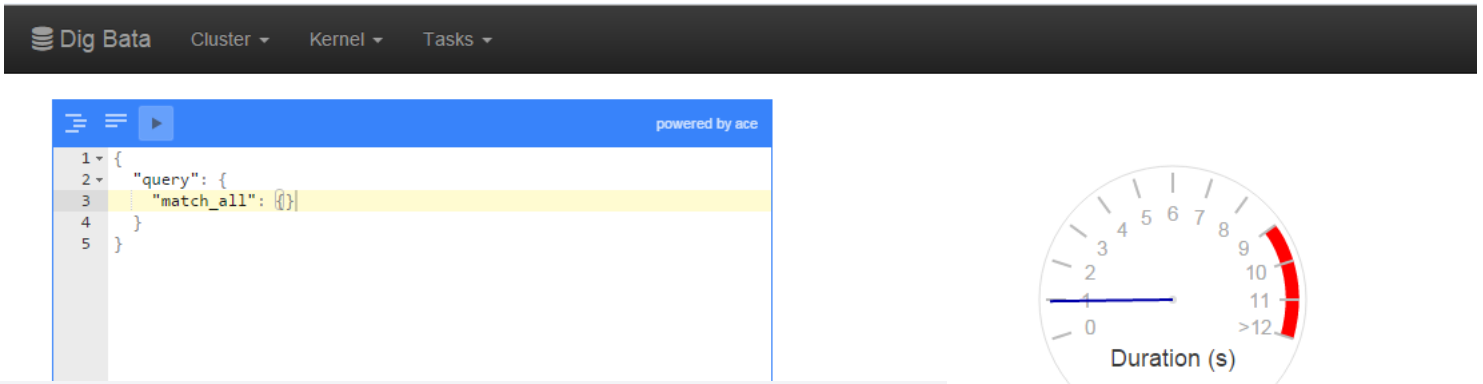


Data flows: a user at UC Irvine reading a data set from Nikhef, with some background from CERN and KIT Karlsruhe

Graphics courtesy Jouke Roorda, Olivier Verbeek, Rens Visser

Nikhef

Queries and responses




```
{
    "Query_begin_structuur": {
        "type_query": {
            "field": "naam field aan te spreken"
        }
    }
}
```

David Groep
Nikhef
Amsterdam
PDP programme

Graphics courtesy Jouke Roorda, Olivier Verbeek, Rens Visser

Nikhef

Processing batch logs

- Log file format is different (different service) so you'll need to parse, extract, and identify elements
- The 'grok' filter is used to process logs

- vm4.stud1.ipmi.nikhef.nl runs grok now '/grok/ls_config'

- If and when you need write access there, ask …

David Groep
Nikhef
Amsterdam
PDP programme

Access services

https://wiki.nikhef.nl/grid/HvABigDataVisualisation

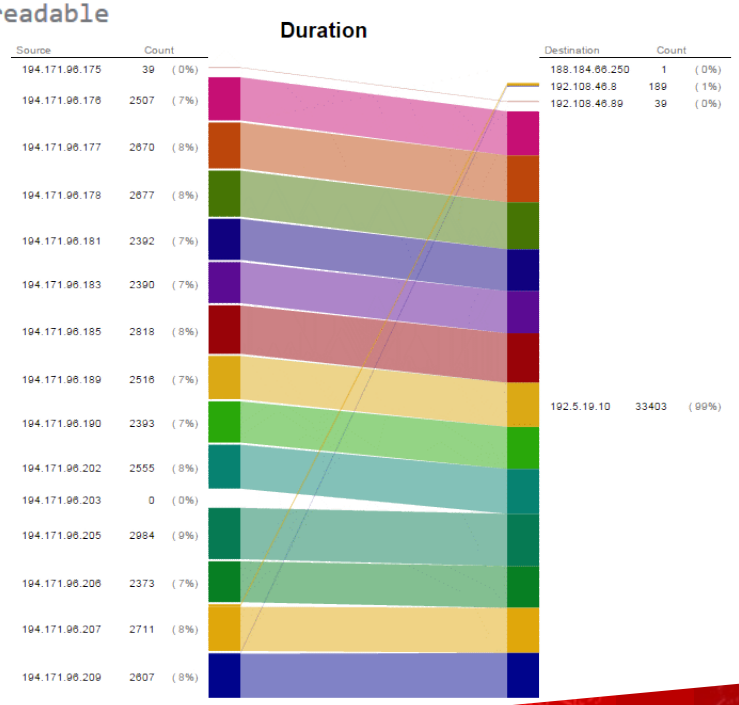
- OpenVPN configuration
- Access to full ES search API

- Don't kill the cluster just yet …
- Do not redistribute or copy records with PII
- Kindly observe the Nikhef AUP as well – and report any incidents to security@nikhef.nl

David Groep
Nikhef
Amsterdam
PDP programme

```
public List getDocsHistogram() throws IOException {
    return client.execute(
            new Search.Builder(new SearchSourceBuilder()
                    // This query searches for the right kind of log messages
                    .query(
                            QueryBuilders.boolQuery()
                                    .must(QueryBuilders.matchQuery("program", "dpm-gsiftp"))
                                    .must(QueryBuilders.matchQuery("msg_type", "transfer"))
                    )
                    // Newest results should be retrieved
                    .sort(SortBuilders.fieldSort("@timestamp").order(SortOrder.DESC))
                    // Get 25 entries to keep the graph readable
                    .size(25)
                    .toString()
            ).addIndex(logsIndex).build())
            .getHits(Map.class)
            // Java 8 streams because they're awesome :)
            .stream()
```

Your call!
Questions?



David Groep
Nikhef
Amsterdam
PDP programme

Nikhef

Code sample: Jouke Roorda, PDP-sample-app