DeiC

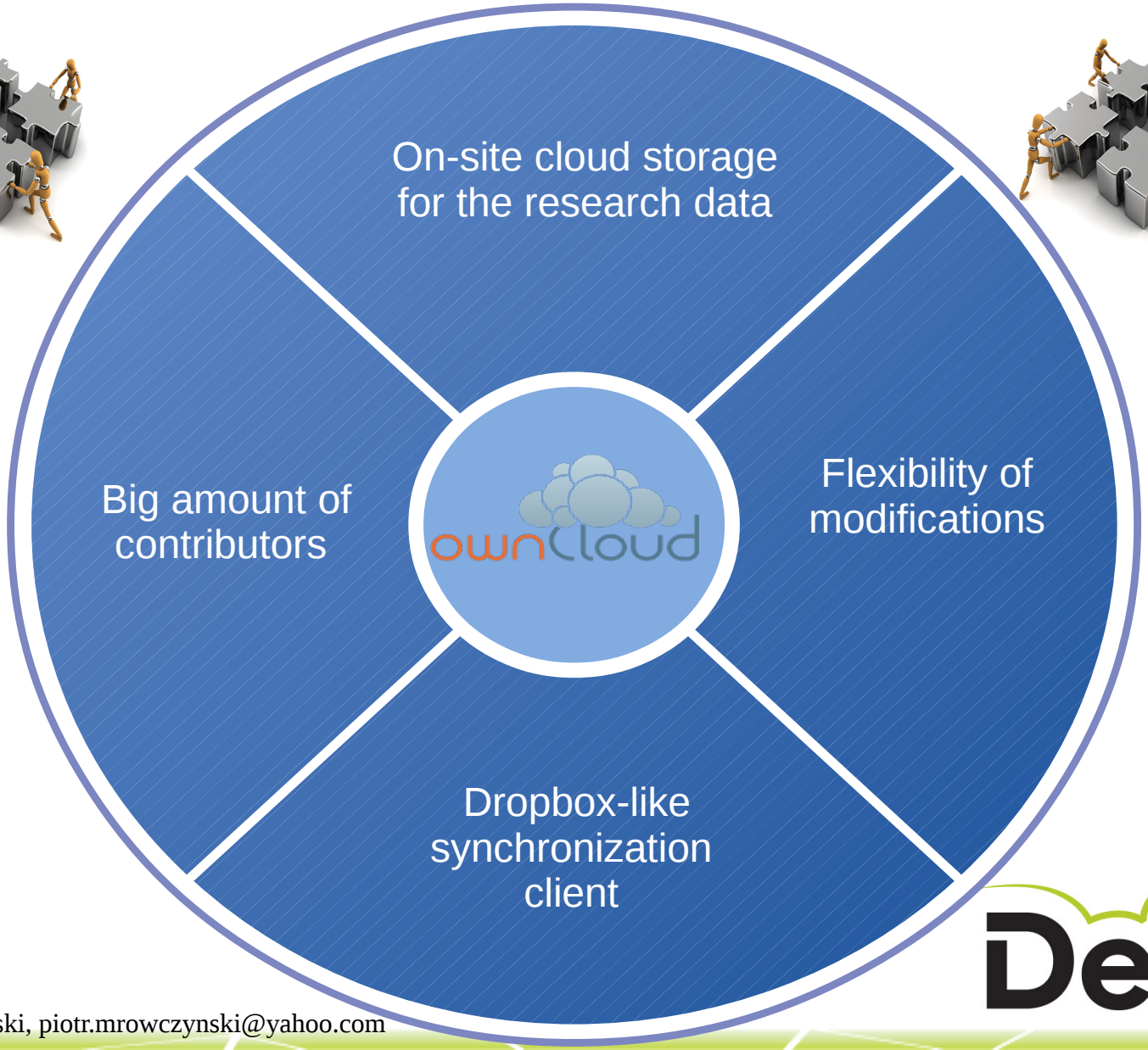
DANISH e-INFRASTRUCTURE COOPERATION

Benchmarking and testing ownCloud, CERNBox, Dropbox and Seafile

Piotr Mrówczyński
DeIC, Denmark, 11.01.2016
Bachelor's Diploma

DeiC

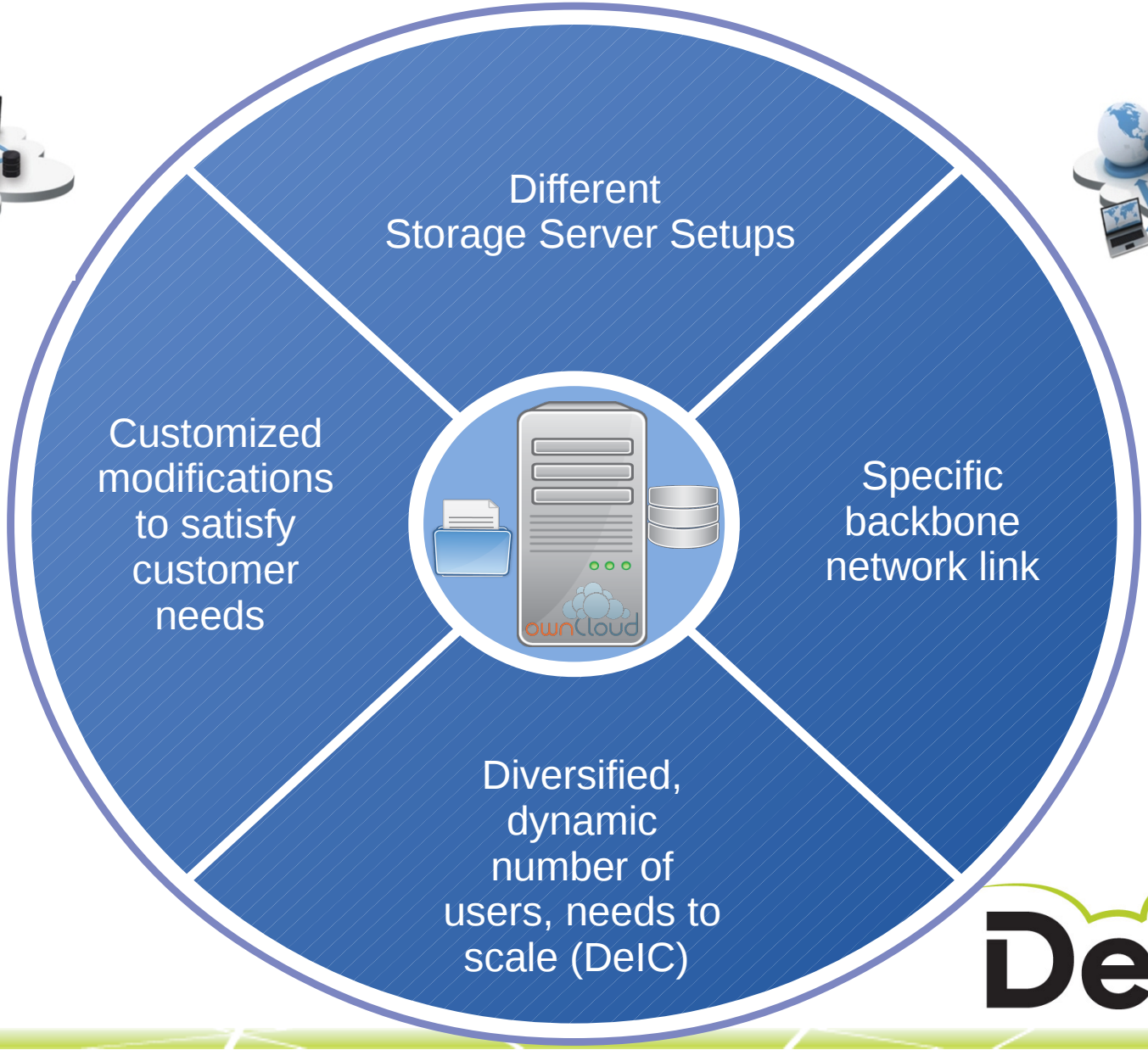
ownCloud as a framework



DeiC

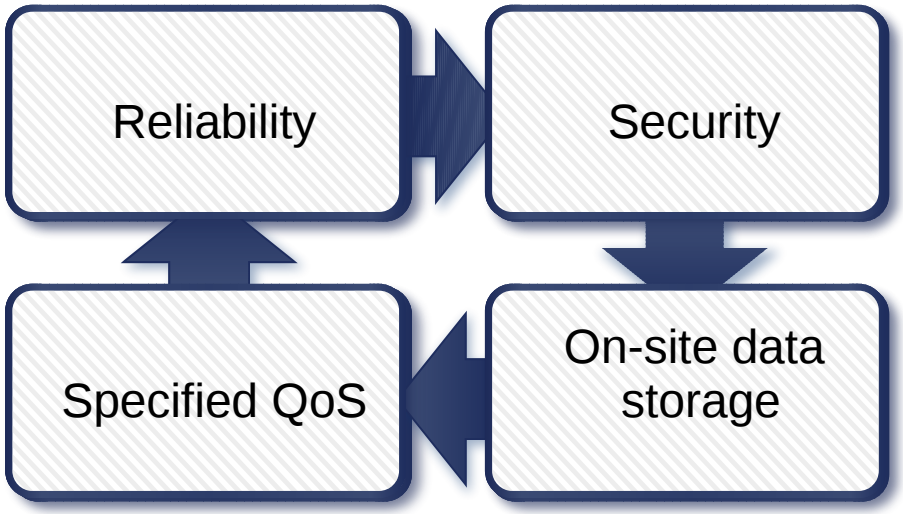
Piotr Mrówczyński, piotr.mrowczynski@yahoo.com

ownCloud as a service



DeiC

Researcher – Cloud Service User



Researcher

User location very close to the server

Research data has different distributions then personal data

Stationary PC with stable internet connection

Computing VM connected to the cloud

Works on data – need for Research specific apps

Piotr Mrówczyński, piotr.mrowczynski@yahoo.com

SATISFACTION
=
REALITY – EXPECTATIONS

Could we really measure and monitor the expectations for the service, from the user perspective?

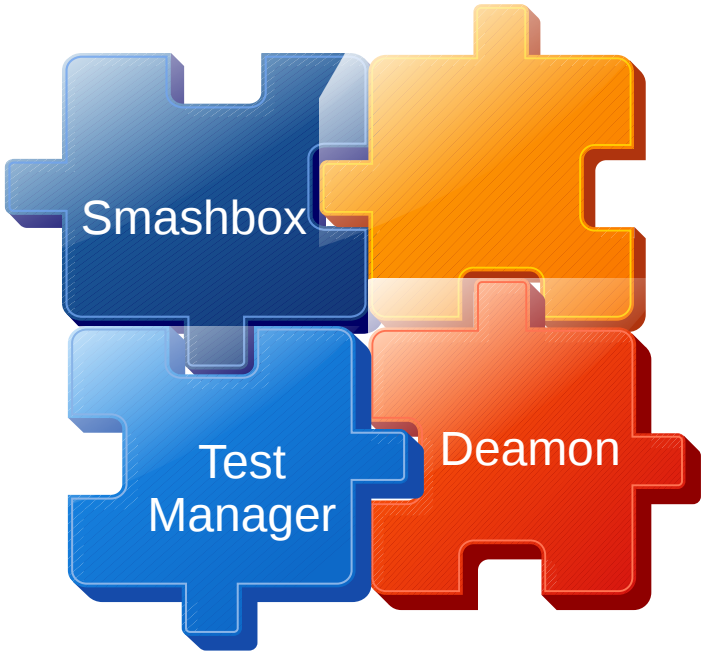
Piotr Mrówczyński, piotr.mrowczynski@yahoo.com

I tried to check it!

Smashbox

Developed by
Jakub Moscicki, CERN.

QA Testing Tool for
the ownCloud clients

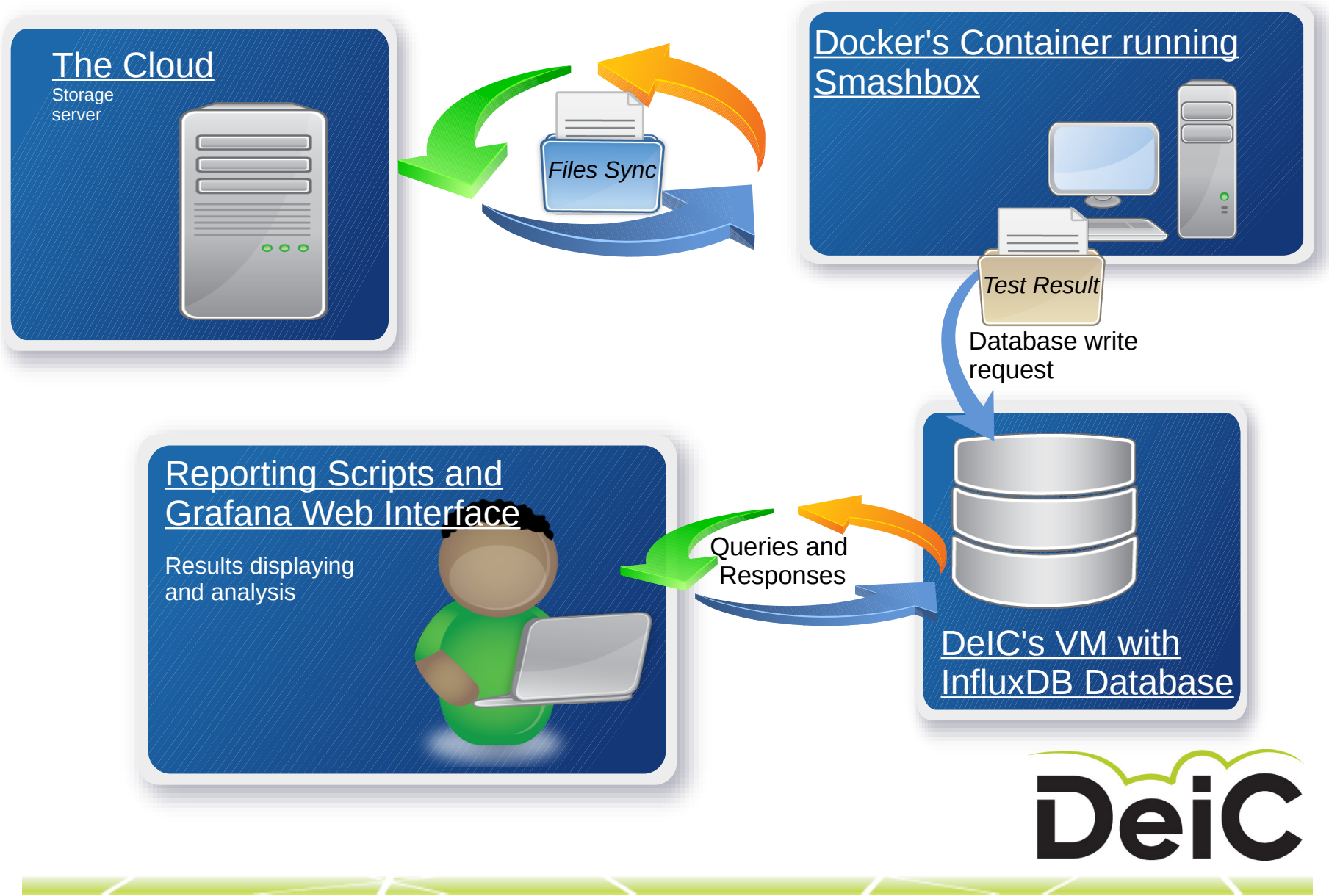


Test Manager

Extended for performance testing, more clients, detailed analysis of the test and remote result storage

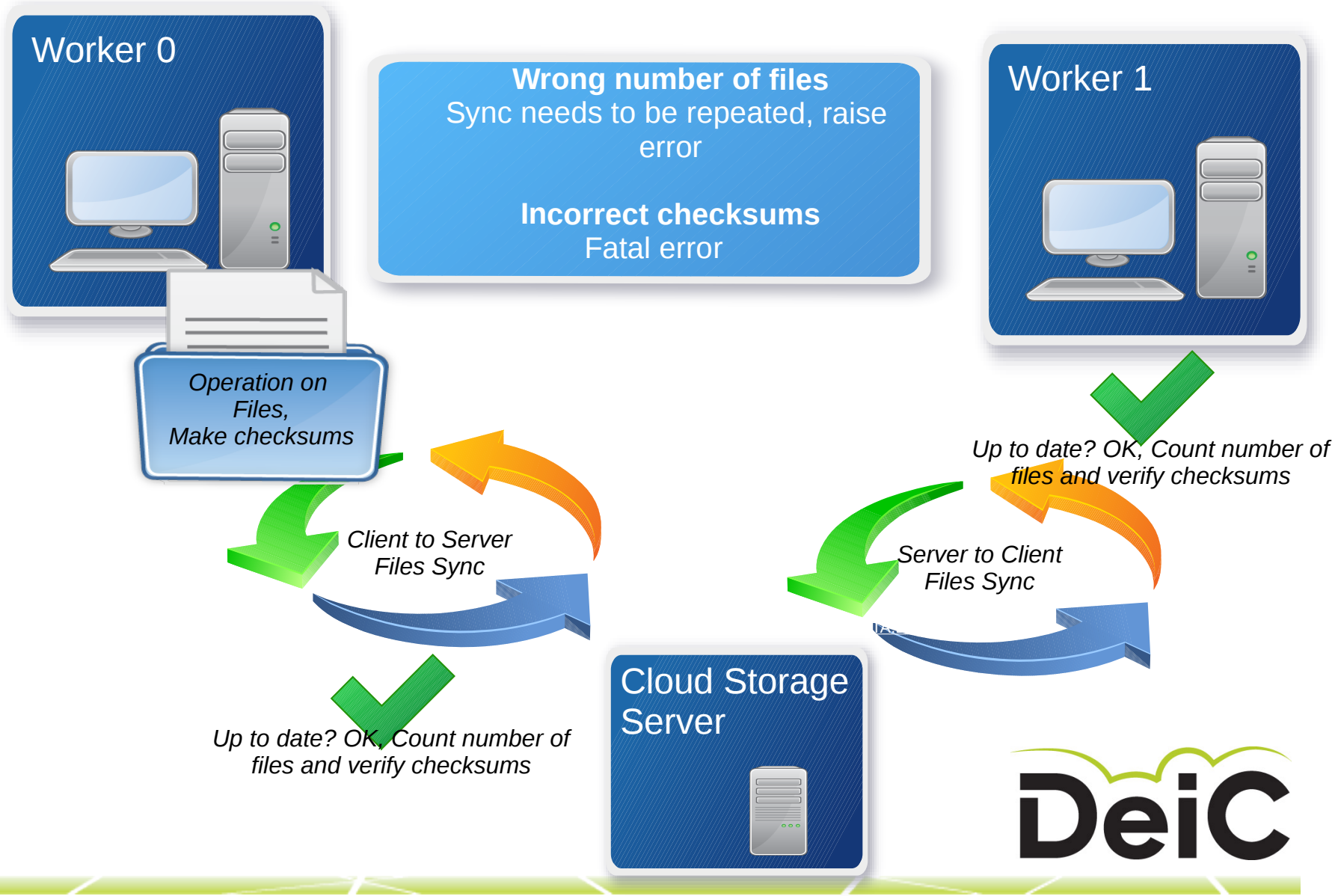
DeiC

Smashbox Analytics



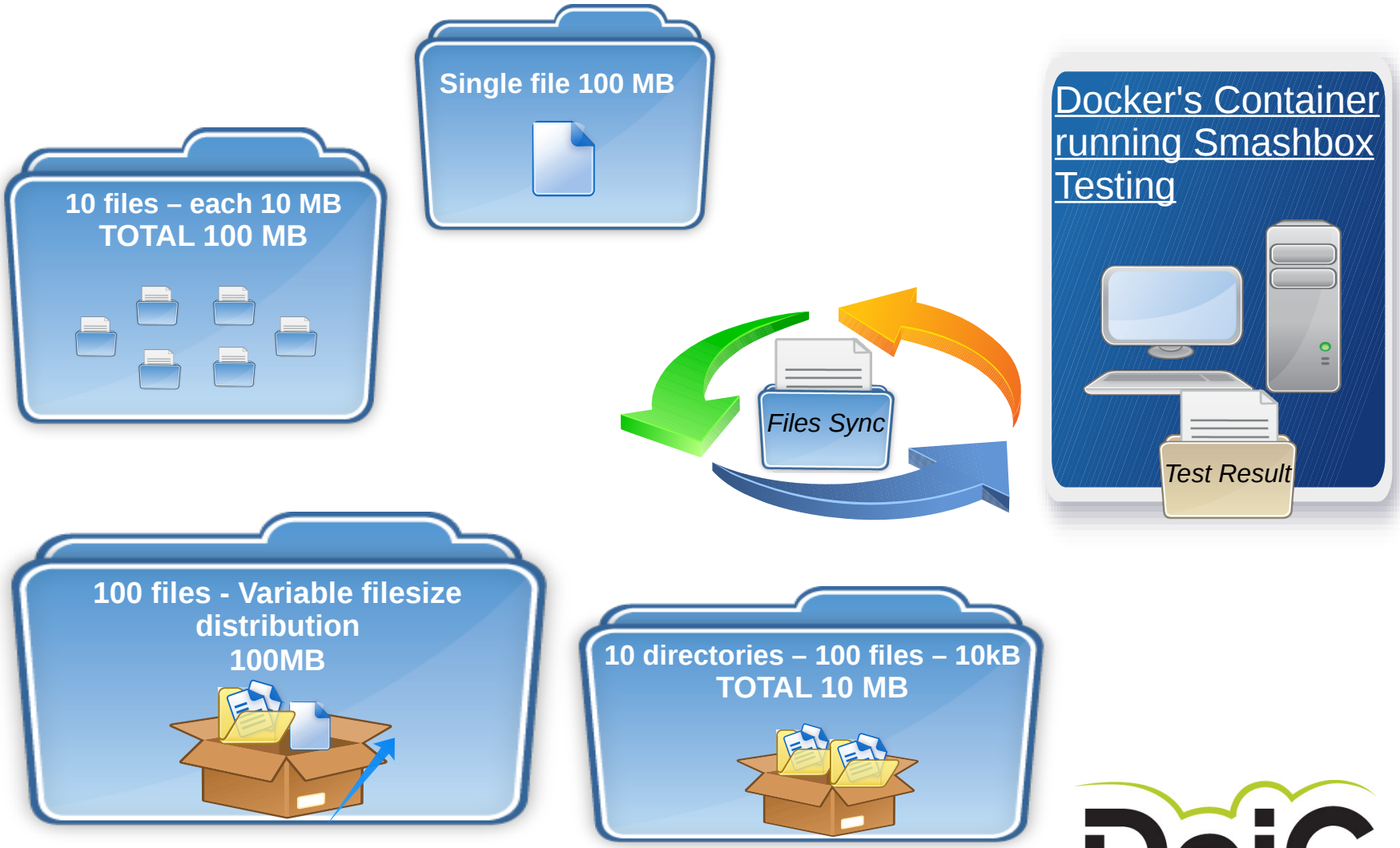
How does it work?

The Test Structure



The Use Case

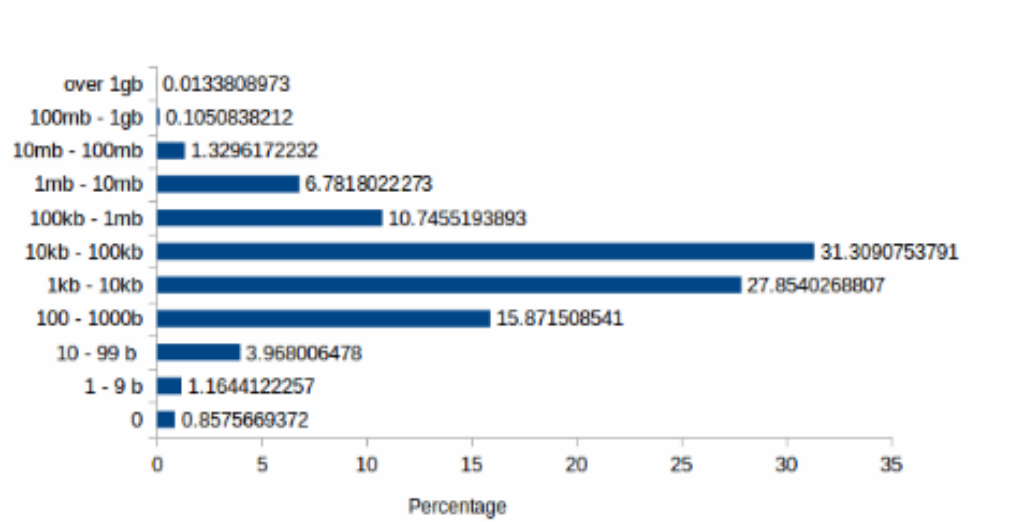
Four Different Scenarios



DeiC

Distribution of files in CERNBox

Building neutral scenario



Number of files: 100, Total size [MB] 98

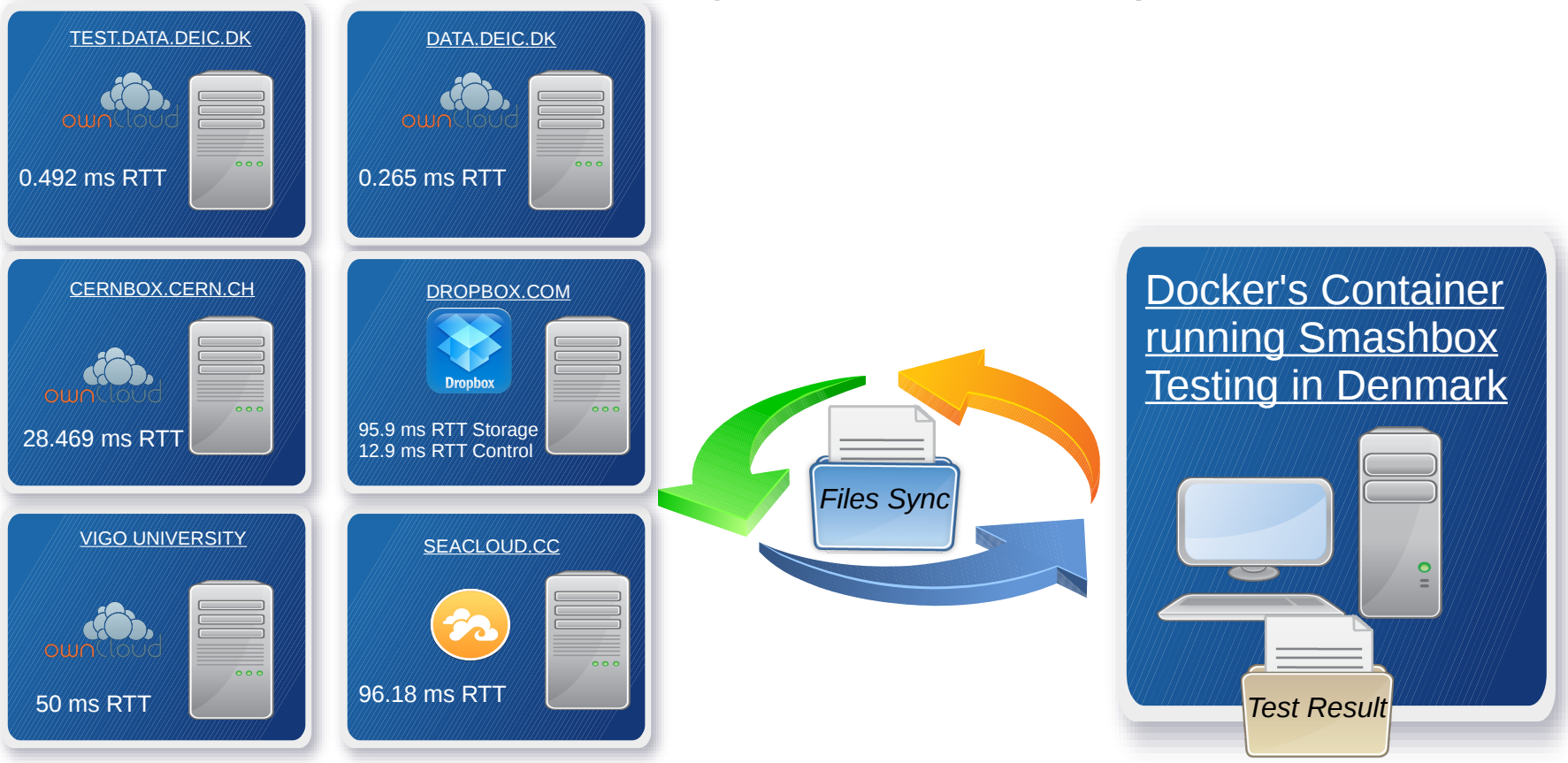
- 1 file[s] 50 MB, 1 file[s] 15 MB, 1 file[s] 1 MB, 1 file[s] 2 MB, 1 file[s] 3 MB, 1 file[s] 4 MB, 1 file[s] 5 MB, 1 file[s] 8 MB, 1 file[s] 10 MB
- 11 file[s] 500 kB
- 32 file[s] 50 kB
- 28 file[s] 5 kB
- 15 file[s] 1 kB
- 5 file[s] 100 Bytes

Statistics available thanks to Wojciech Jarosz, CERN

DeiC

The Use Case

Real-time Cloud Storage Services Comparison



Single Sync Completing Ratio

1 file - 100MB		
Server name	Failed	Passed
data.deic.dk	0	280
https://seacloud.cc/	2	297
cernbox.cern.ch	0	310
dropbox	0	329
ClawIO REDIS XATTR	0	42
test.data.deic.dk	0	330

10 files - 10MB		
Server name	Failed	Passed
data.deic.dk	0	279
https://seacloud.cc/	4	293
cernbox.cern.ch	0	311
dropbox	0	329
ClawIO REDIS XATTR	29	14
test.data.deic.dk	0	326

10 dir. Each 100 files - 10kB		
Server name	Failed	Passed
data.deic.dk	0	268
https://seacloud.cc/	5	293
cernbox.cern.ch	0	304
dropbox	0	320
ClawIO REDIS XATTR	0	44
test.data.deic.dk	0	320

100 files - file size distribution – total 100MB		
Server name	Failed	Passed
data.deic.dk	2	1123
https://seacloud.cc/	826	372
cernbox.cern.ch	0	1255
dropbox	0	1324
ClawIO REDIS XATTR	4	168
test.data.deic.dk	62	1259

Test machine at DeIC

* **LOCATION**: Office at Technical University of Denmark.

* **HARDWARE PARAMETRS**: SSD Disk, 8GB RAM, i5-4570 CPU @ 3.20GHz × 4

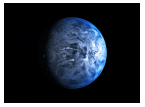
* **LINK**: University Network - network card 1Gbps. Low daily traffic. Shared switch.

Seacloud does not complete sync in one round

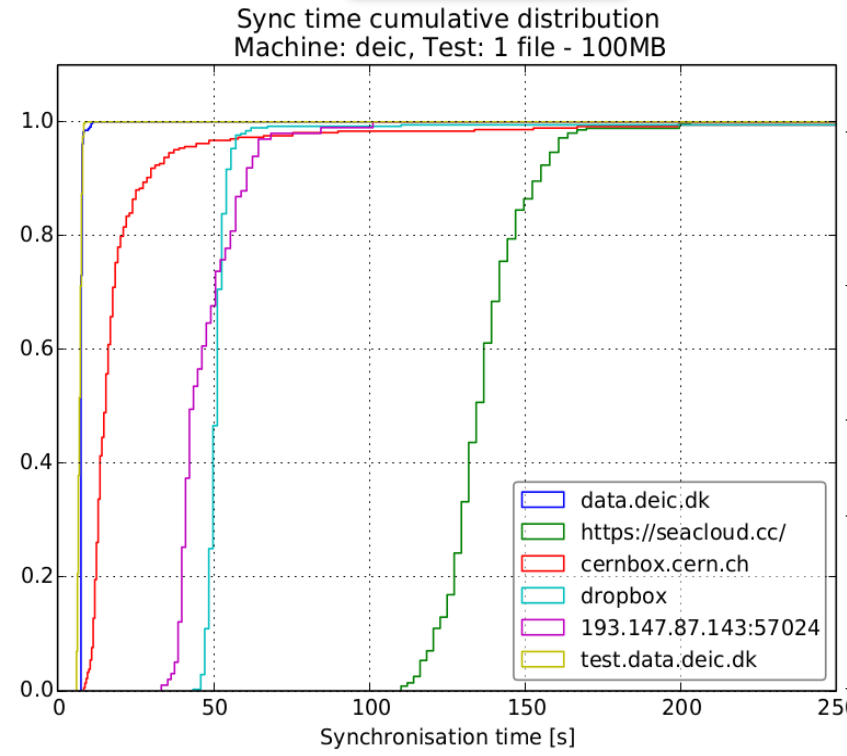
IT NEEDS >1 MORE ITERATIONS TO COMPLETE

Cloud Storage Services Benchmarking based on one parameter

Sync of few, big files



DeIC



99%

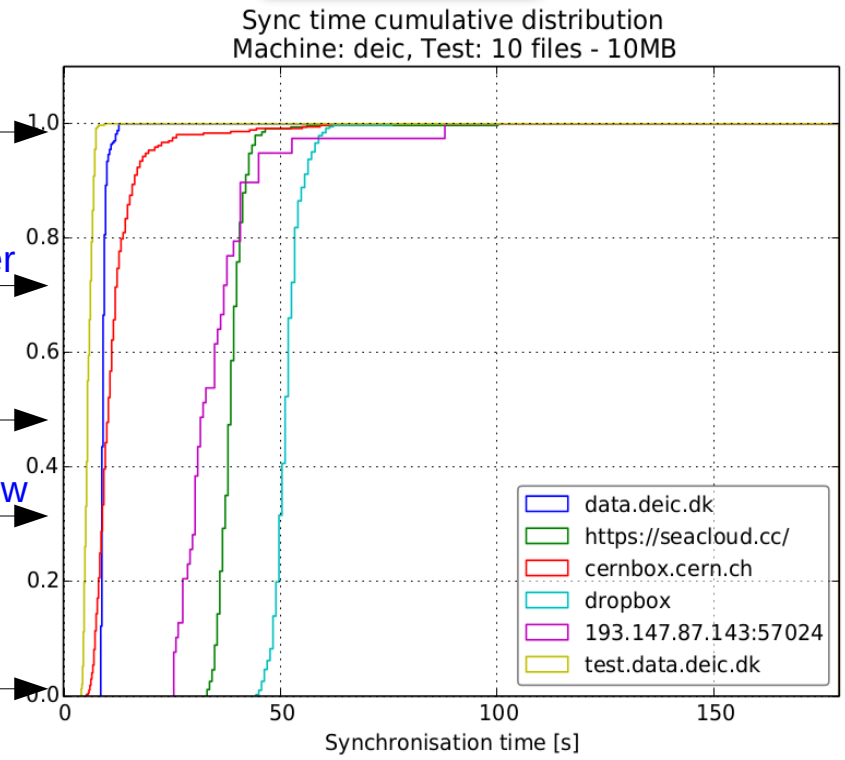
Values over median

Median

Values below median

1%

DeIC



DeiC

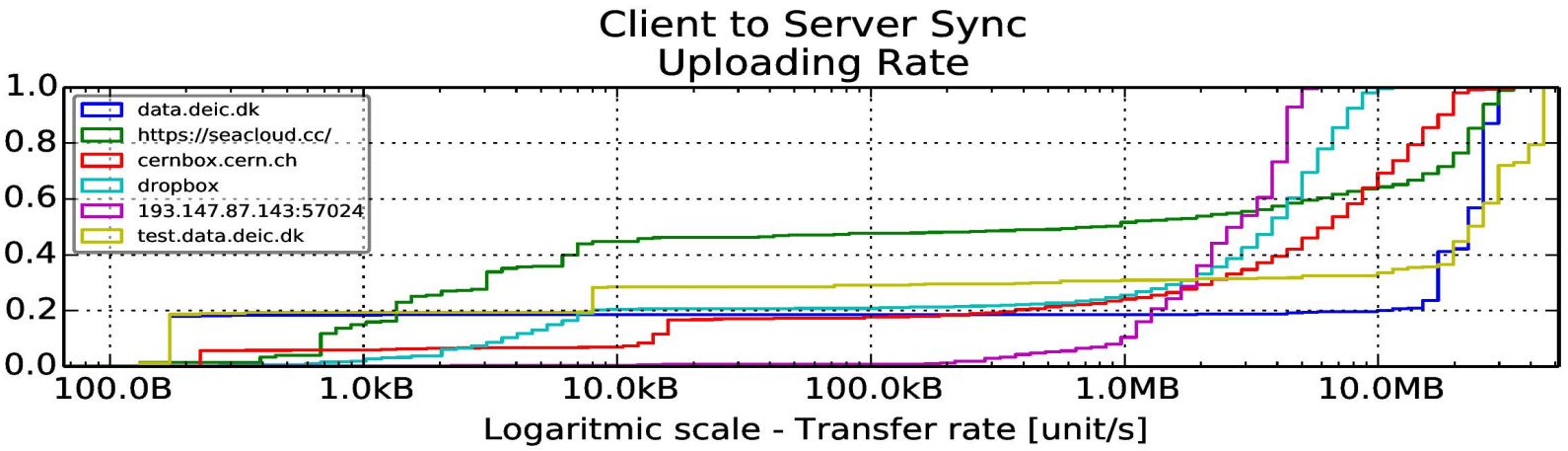
Why DeIC Data and CERNBox are so fast comparing to Dropbox,
for the researcher in Denmark?

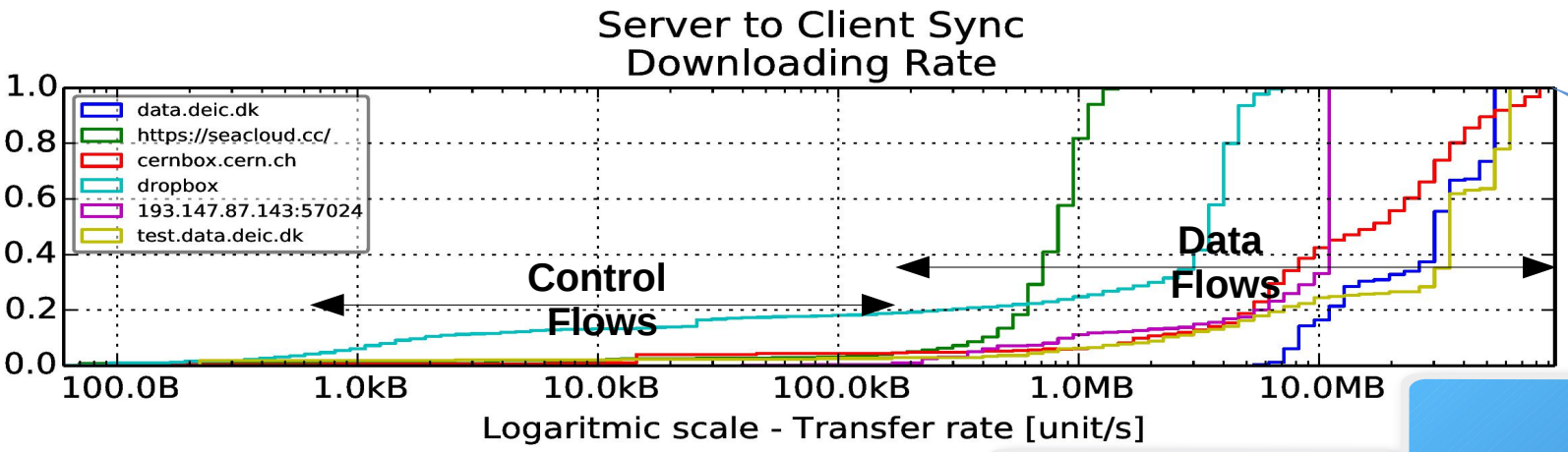
DeiC

Bandwidth characteristics – big files

Cumulative Distribution of transfer rates for all the measurements

Machine: deic, Test: 1 file - 100MB



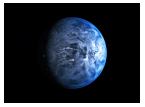


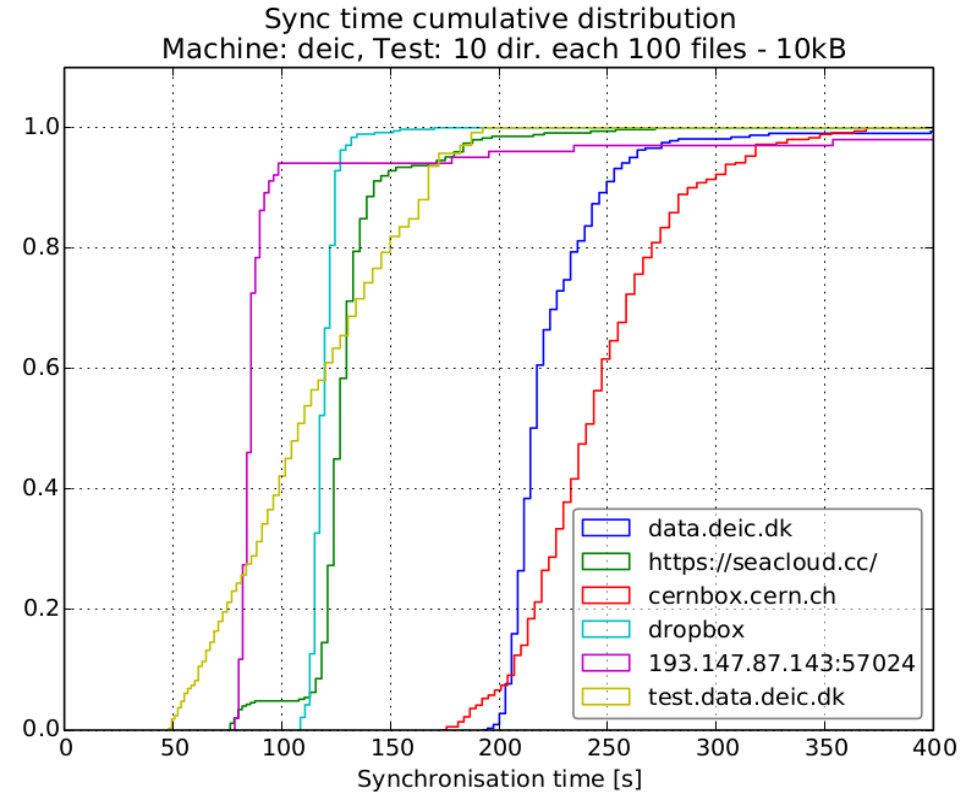
CERNBox could reach 90MB/s
Max for the test machine

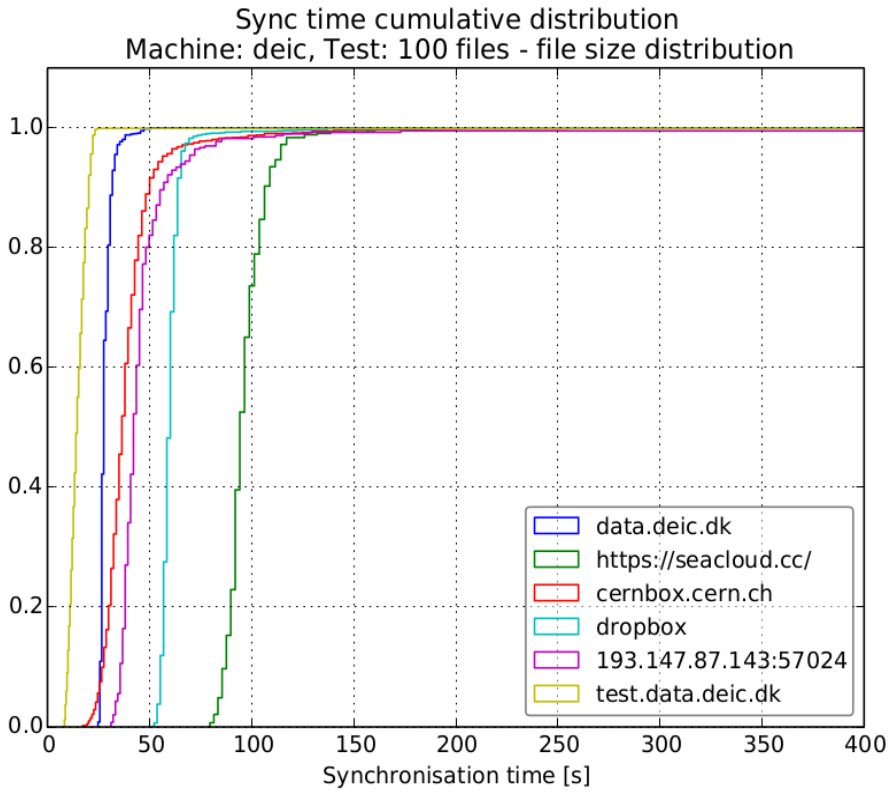
DeIC's services reach 60MB/s

Cloud Storage Services Benchmarking

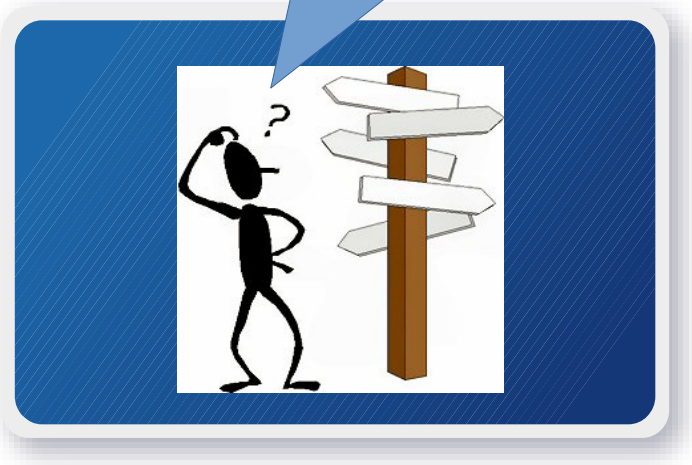
Sync of many files







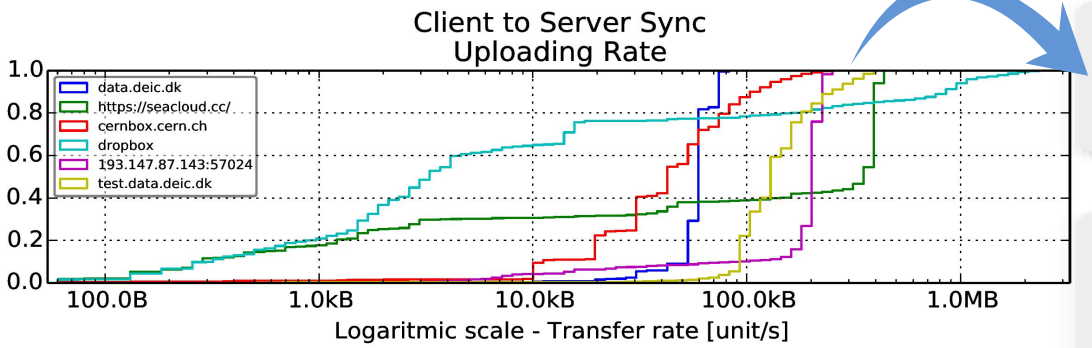
Why is the sync slower when we increase the number of small files?



OwnCloud's Bottleneck

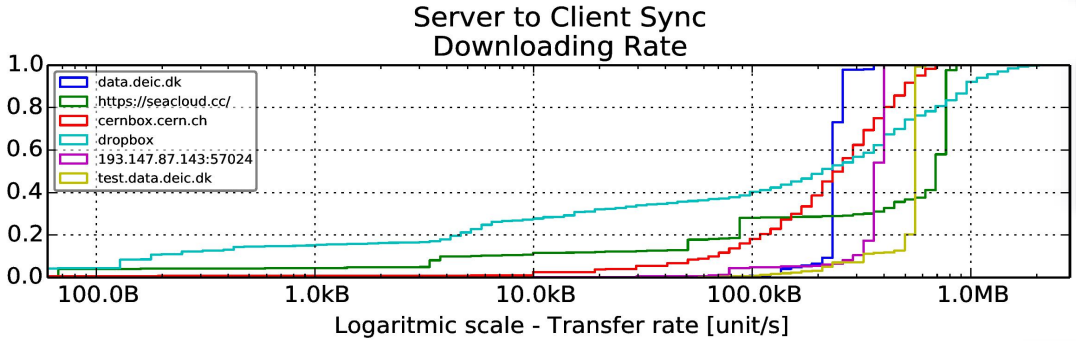
Big number of small files – 1000 files - 10kB

Machine: deic, Test: 10 dir. each 100 files - 10kB

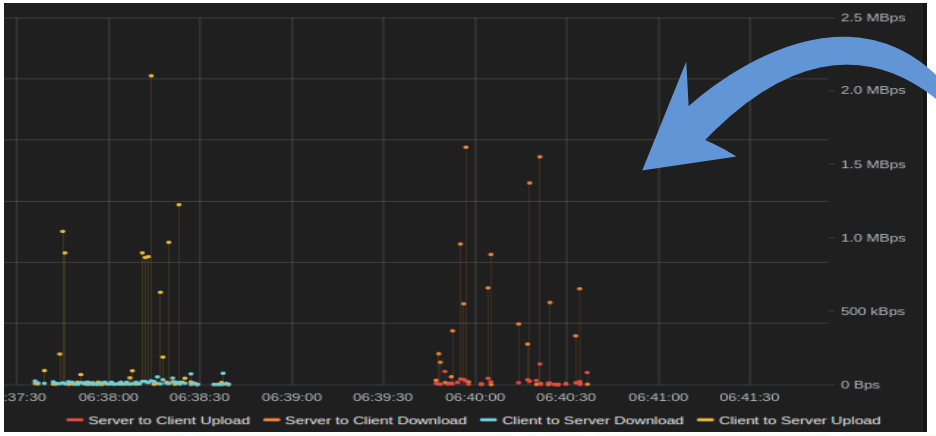


Owncloud is limited by specific amount of files/s

CERNBox users load causes changes in the amount of processed files/s



TEST DATA is increasing its sync time each iteration – bug detected – next slides



Client to Server: Dropbox is preparing package of files and sending

Server to client: Dropbox is receiving and processing in parallel

Let's move now to CERN, 30ms away from Denmark..

Latency to Dropbox and Seacloud stays the same

CERN TEST MACHINE

* **LOCATION**: Virtual Machine at the CERN ccmpute node

* **HARDWARE PARAMETRS**: 4GB RAM, Intel(R) Xeon(R) CPU E5-2650 v2 @ 1x 2.60GHz

* **LINK**:Inc Virtio network device

DEIC TEST MACHINE

* **LOCATION**: Office at Technical University of Denmark.

* **HARDWARE PARAMETRS**: SSD Disk, 8GB RAM, i5-4570 CPU @ 3.20GHz × 4

* **LINK**: University Network - network card 1Gbps. Low daily traffic. Shared switch.

Grafana debugging

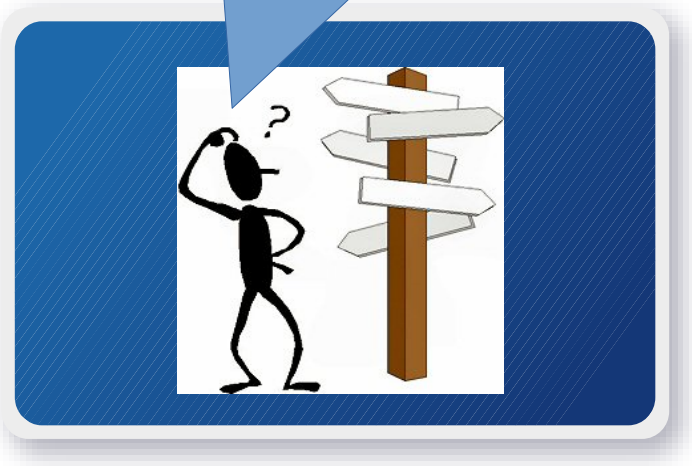
SLOWER MACHINE at CERN

The same moment of time at the different machines

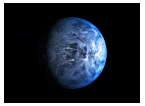
FASTER MACHINE at DeIC

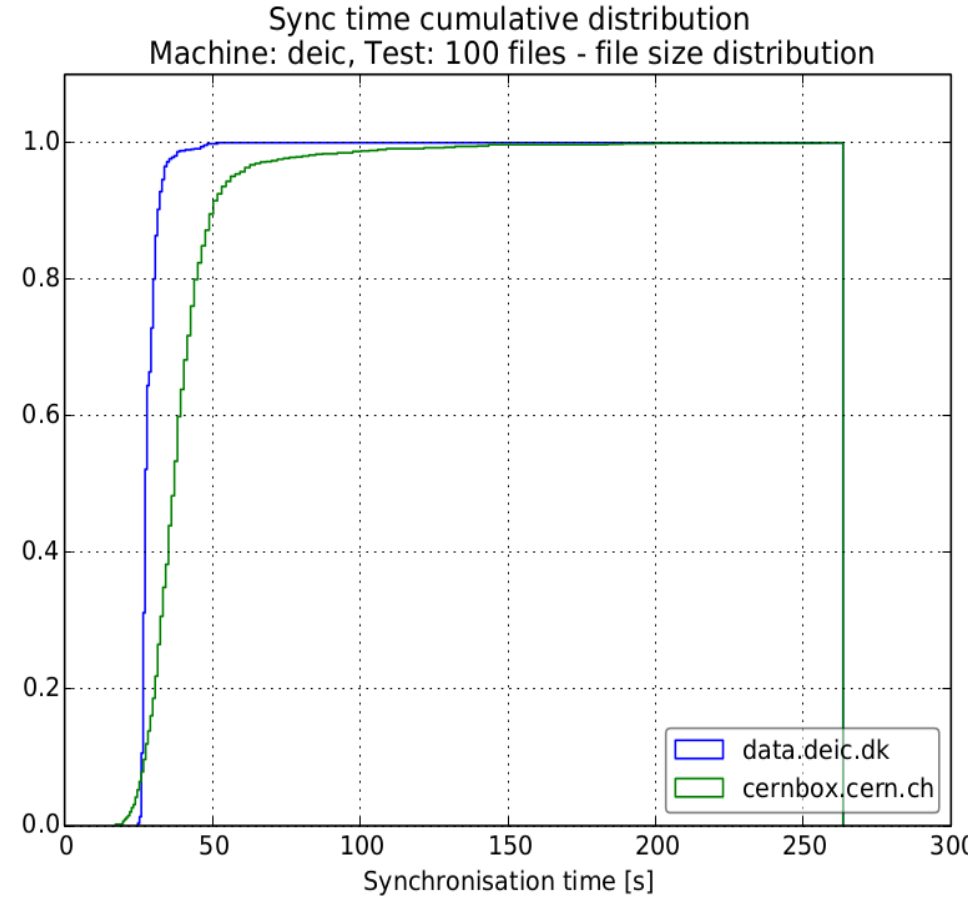
Cloud Storage Services Benchmarking

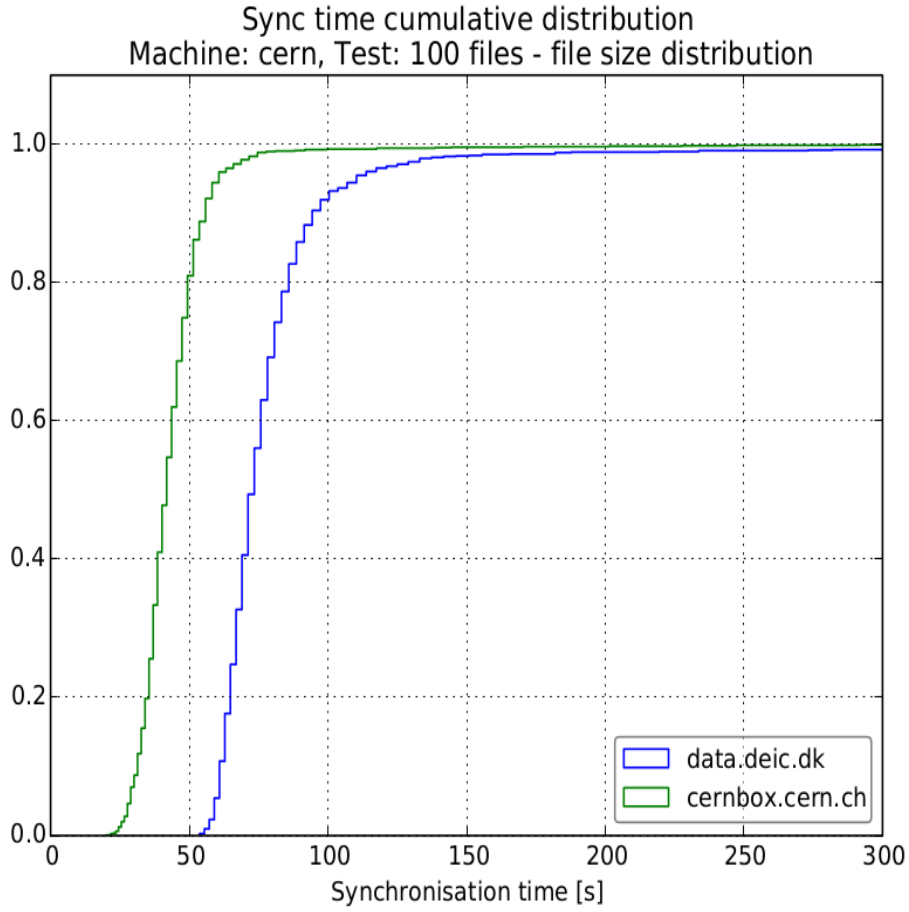
What about CERNBox and Data DeIC?



Latency influence







Daily distribution of synchronisation durations

100 files of different sizes

MAX OBSERVED

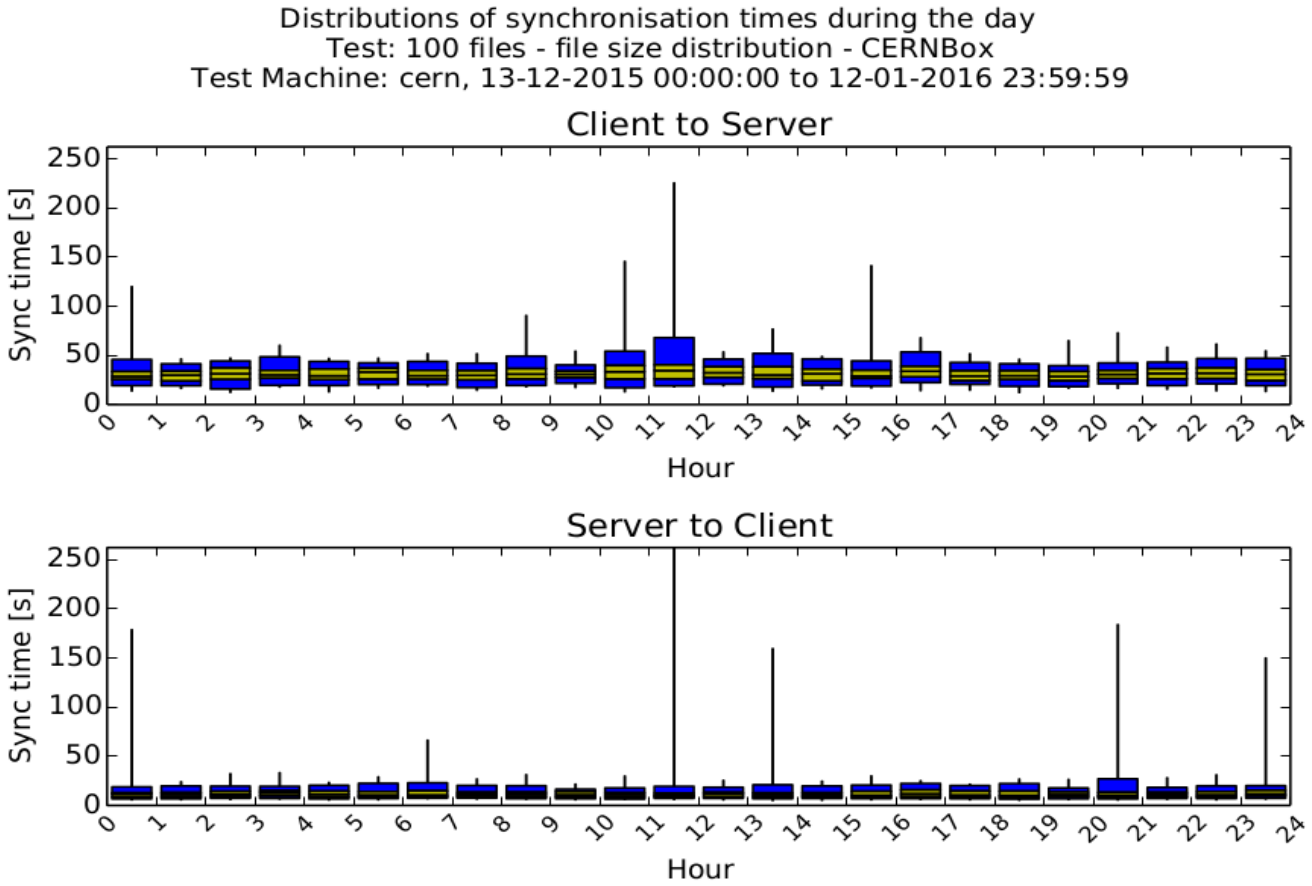
95-75%

75% - median
median - 25%

5-25%

MIN OBSERVED

CERN



Daily distribution of synchronisation durations

100 files of different sizes

MAX OBSERVED

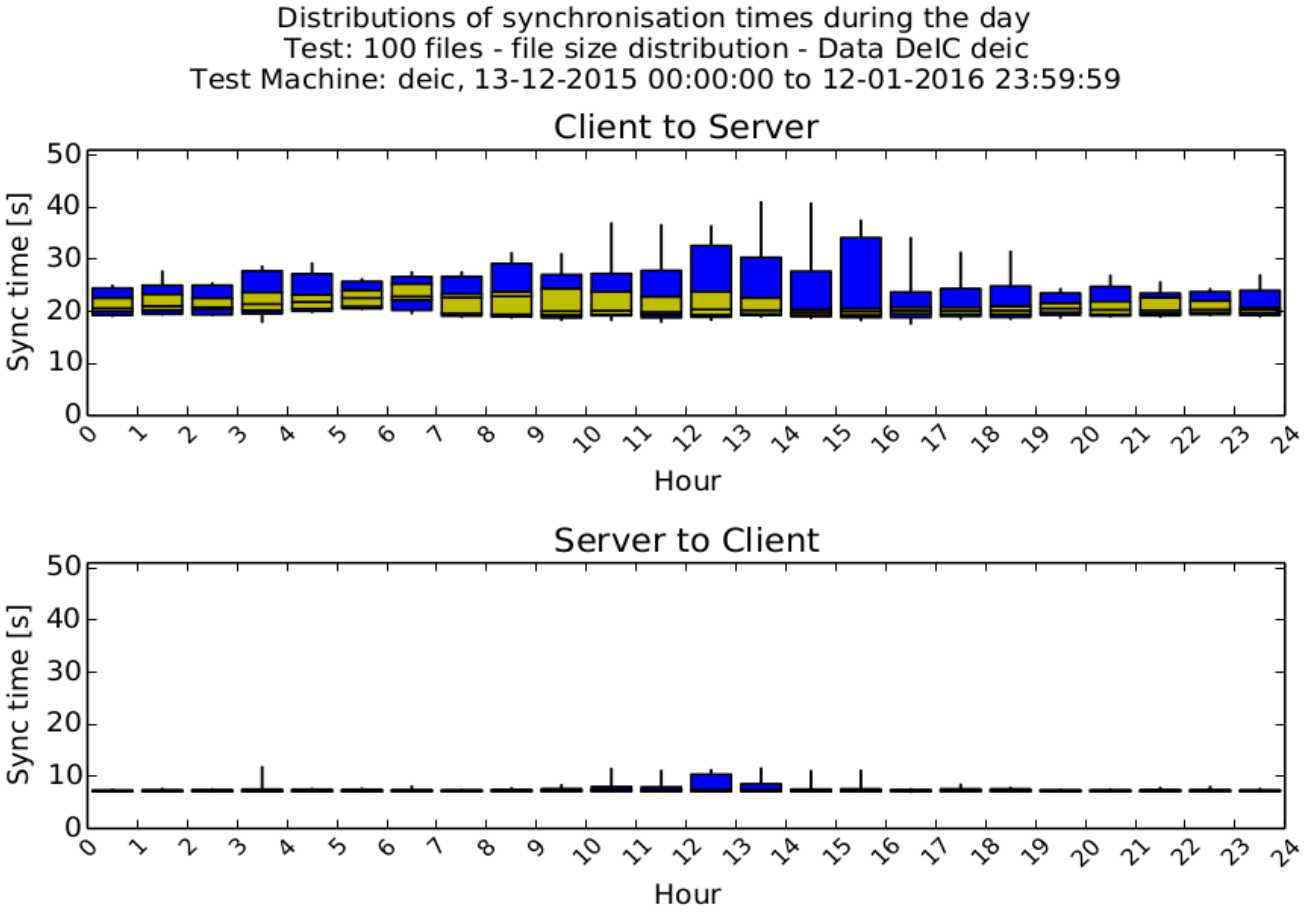
95-75%

75% - median
median - 25%

5-25%

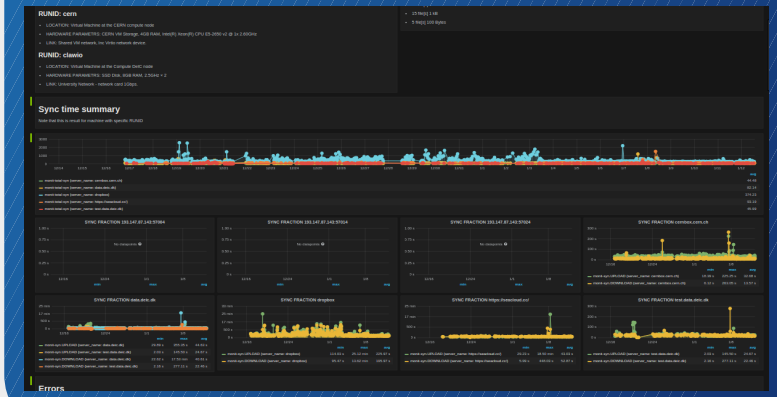
MIN OBSERVED

DEIC



Smashbox Analytics

Grafana Web Service

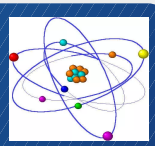


General reporting:

* benchmark instance according to scenario sync time

* monitor available bandwidth to the user during the test

Detailed reporting:



* Compare sync times during different times of the day

* Monitor number of completed single syncs and files synced

* Detailed analysis of sync in terms of stops in the transfer, its relative length and actual transfer being send

LIVE PRESENTATION
test.data.deic.dk probem

http://130.226.137.144:3000/

https://data.deic.dk/shared/72b556a0e3c3cd8502e5a0767558642f

Bottom Line

Smashbox as a monitoring node at your instance!

Easy installation by Docker Container

Detect difficult bugs faster

Cross-check services with your collaborators

Compare development stage services localy to validate expected changes

Thank you for your attention!

What more in my Bachelor's Thesis:

* Full analysis of the self-hosted cloud for scientists concept

* Dropbox deduplication in real-life

* Dropbox delta sync in real-life

* Detailed analysis of facts observed in the measurements – comparison of the services

Out of box
smashbox – DOCKER CONTAINER
https://hub.docker.com/r/mrow4a/smashbox/

Simulator
https://github.com/mrow4a/smashbox-sim

Analytical tools and demo
https://github.com/mrow4a/smashbox-tools

BROWSE DEMO RESULTS AT

GRAFANA DEMO
http://130.226.137.144:3000/
User demo:demo

Questions?

SOON AVAILABLE TO TRY AT YOUR INSTANCE

CERNBox Smashbox
https://github.com/cernbox/smashbox

DeiC