

BioQuant's Large Scale Data Facility

Usage Patterns, Problems, Solutions

Christian Thiemann Marc Hemberger

BioQuant IT, Ruprecht-Karls-Universität Heidelberg

LSDMA 2013 Spring Meeting

BioQuant LSDF
IBM SONAS

BioQuant LSDF

IBM SONAS

240x
1 TB

480x
1 TB

480x
1 TB

300x
1 TB

480x
2 TB

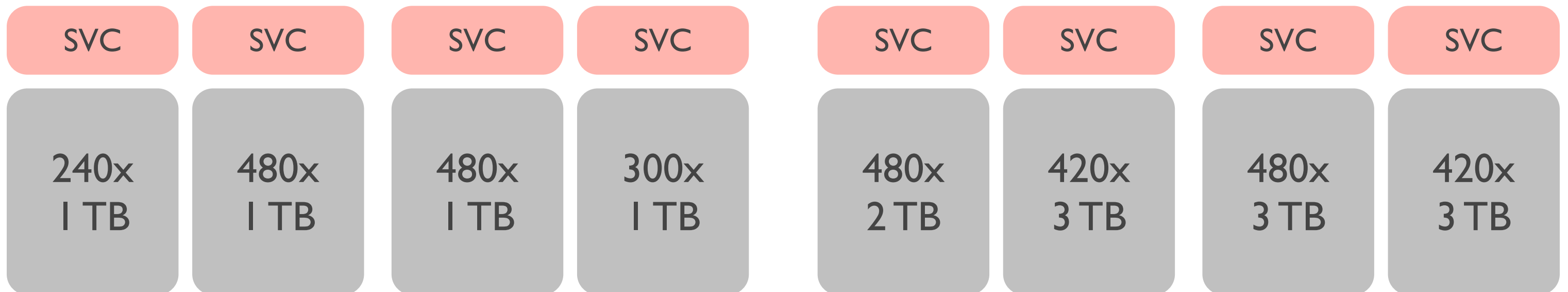
420x
3 TB

480x
3 TB

420x
3 TB

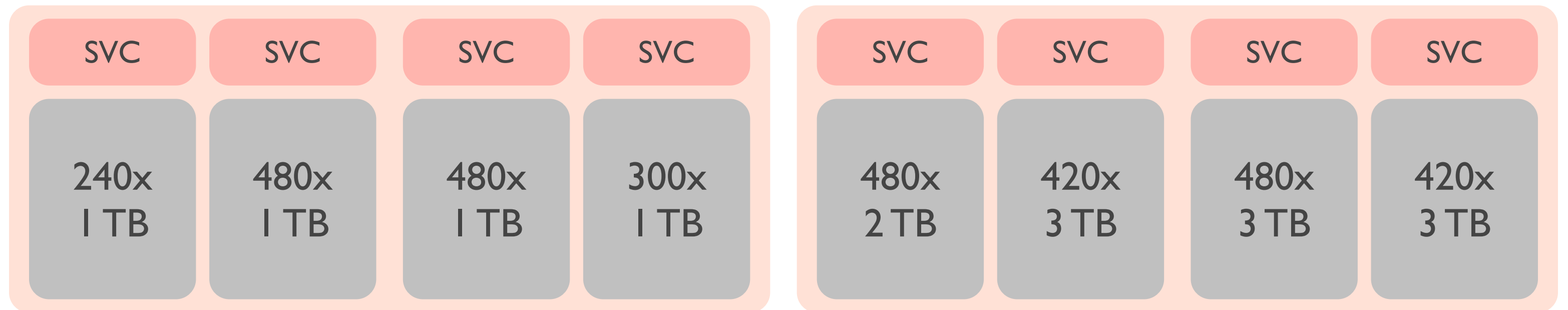
BioQuant LSDF

IBM SONAS



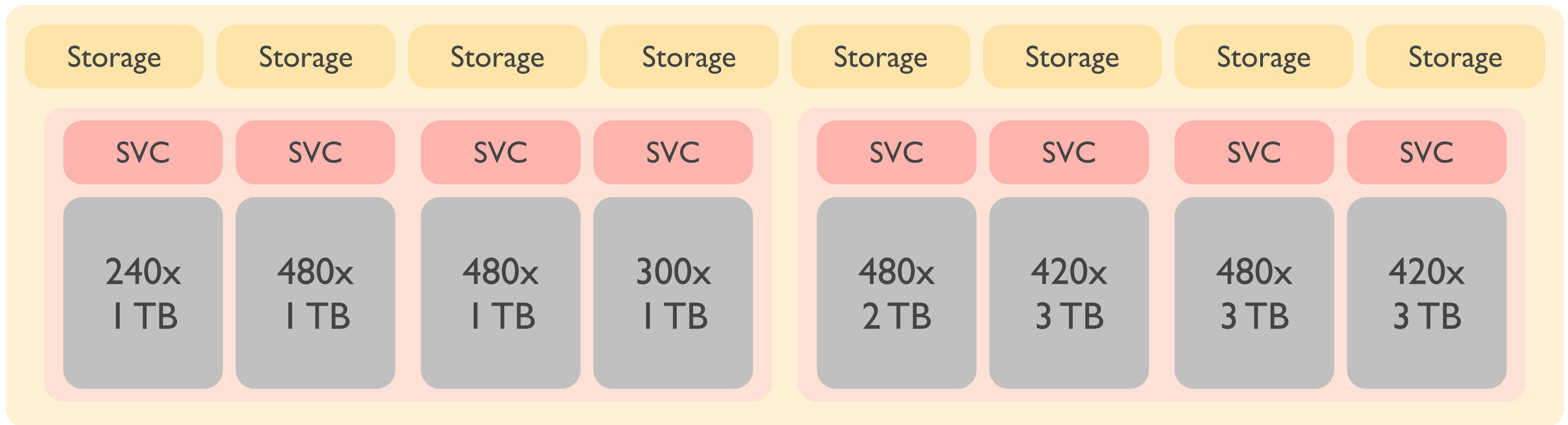
BioQuant LSDF

IBM SONAS



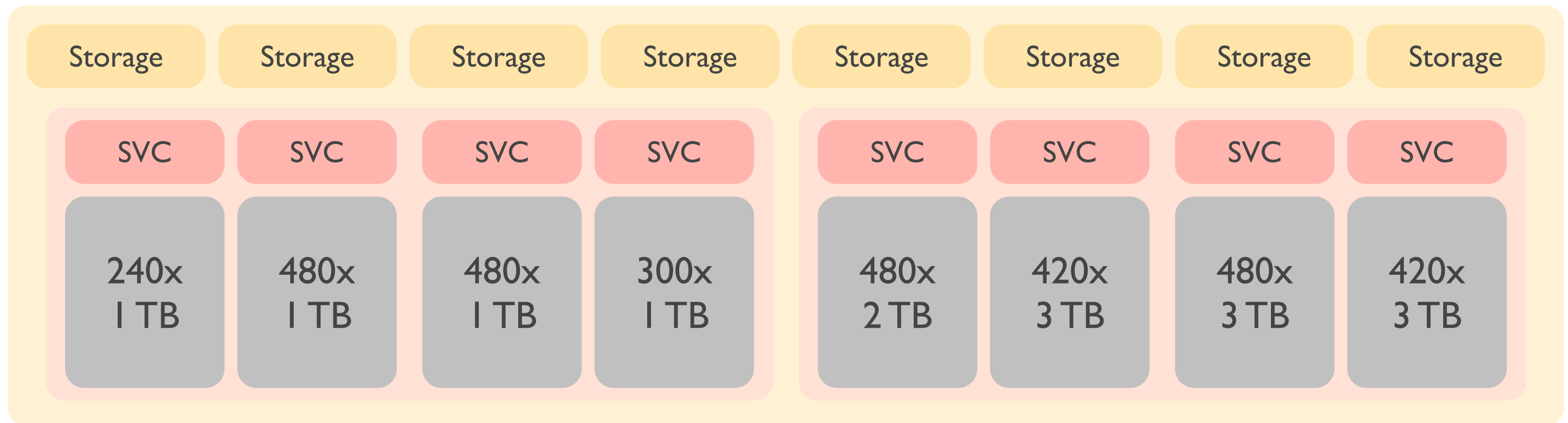
BioQuant LSDF

IBM SONAS



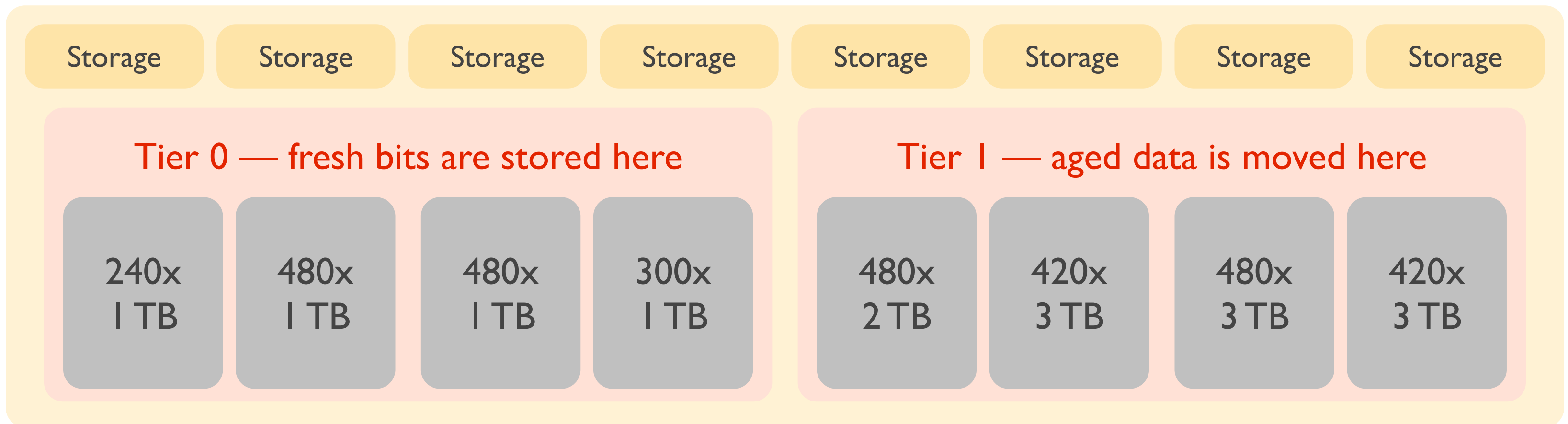
BioQuant LSDF

IBM SONAS



BioQuant LSDF

IBM SONAS



BioQuant LSDF IBM SONAS

NFS & CIFS

2x 10 GBit/s — 1.1 GB/s (read) — 700 MB/s (write)

3,300 HDDs in RAID6 (8+2) — 6.1 PB (physical) — **4.3 PB** (logical)

Tier 0 — fresh bits are stored here

240x
1 TB

480x
1 TB

480x
1 TB

300x
1 TB

Tier 1 — aged data is moved here

480x
2 TB

420x
3 TB

480x
3 TB

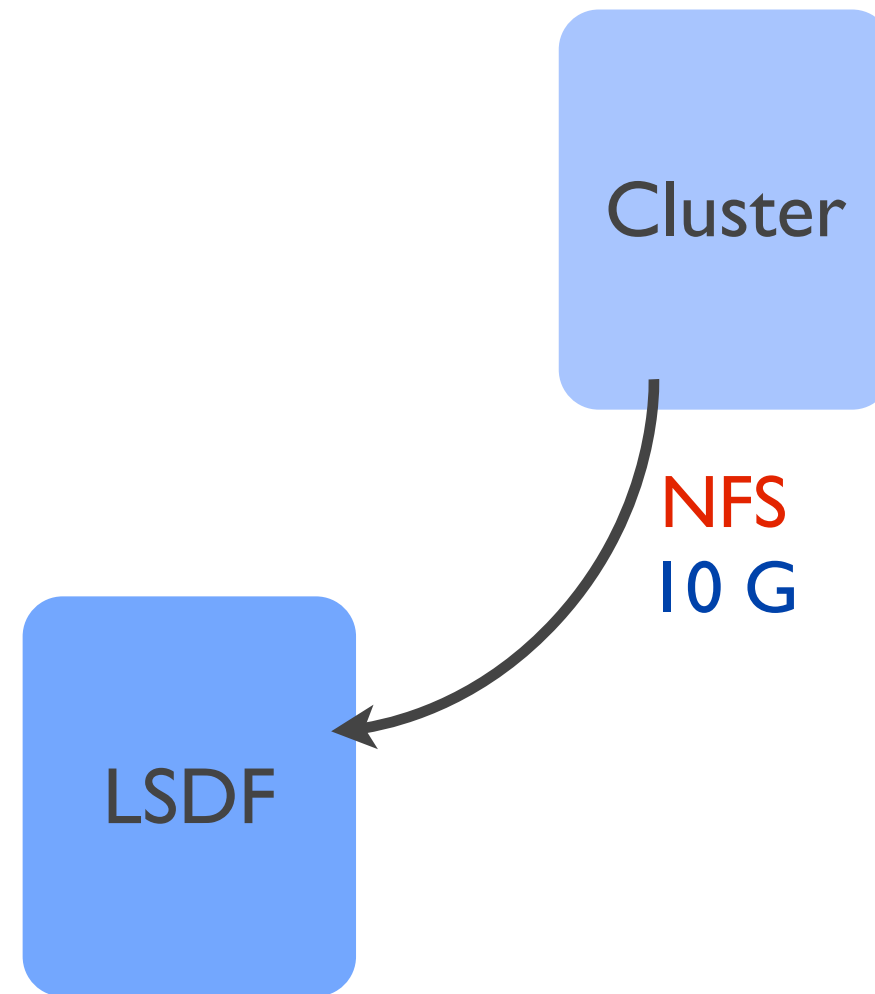
420x
3 TB

BioQuant LSDF

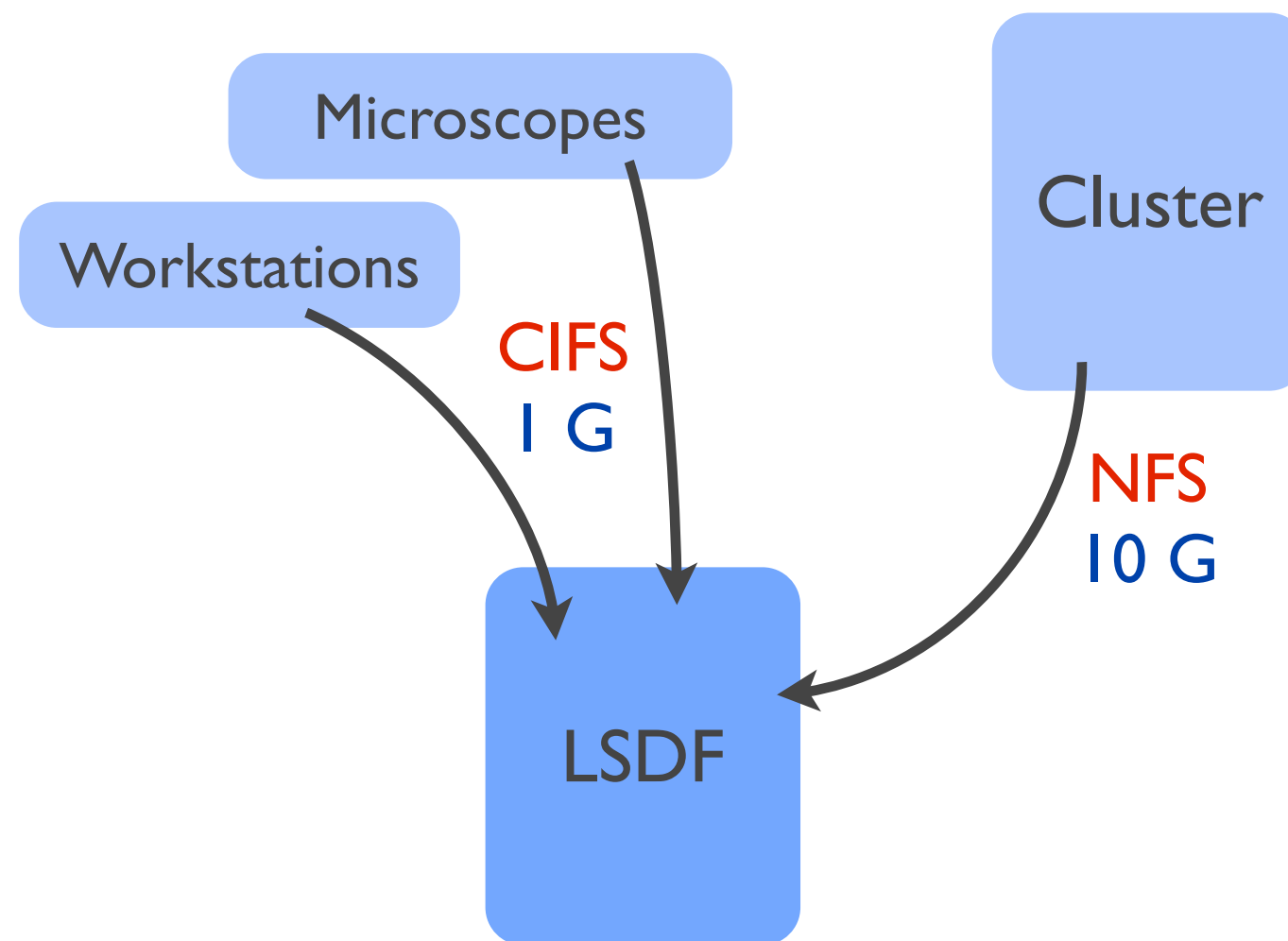


LSDF

BioQuant LSDF



BioQuant LSDF



BioQuant LSDF

DKFZ

Microscopes
Workstations

CIFS
1 G

Cluster

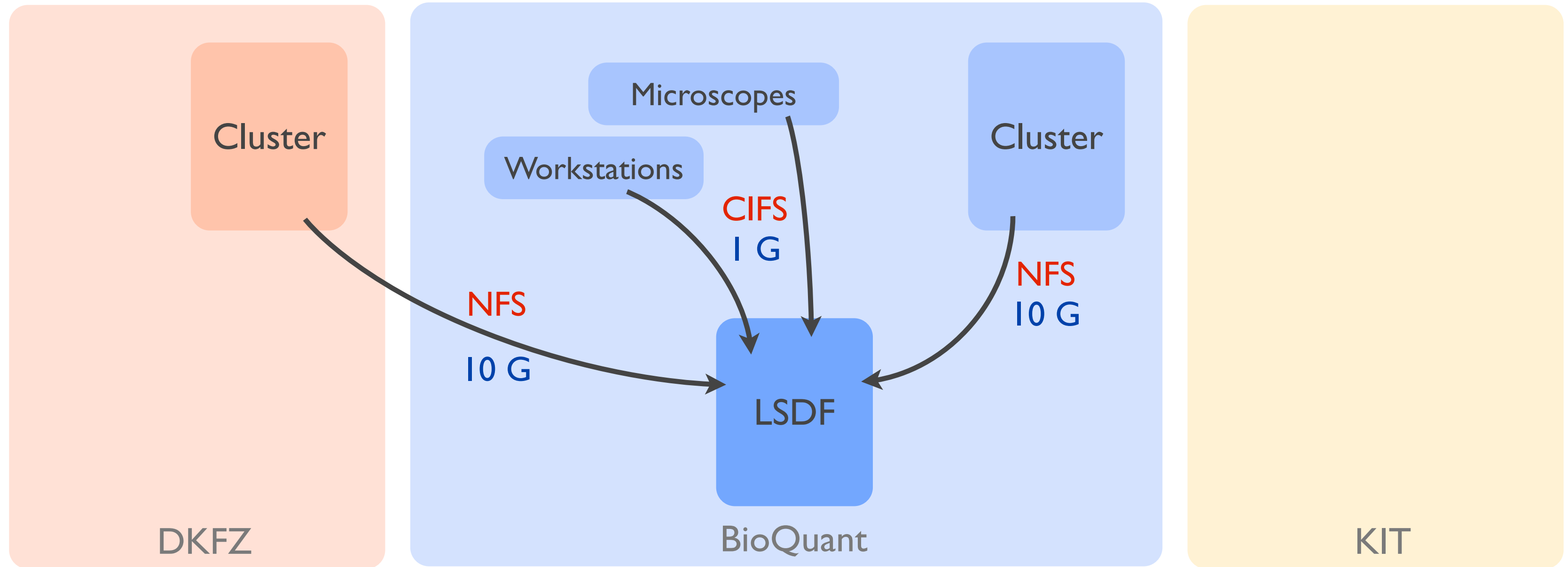
NFS
10 G

LSDF

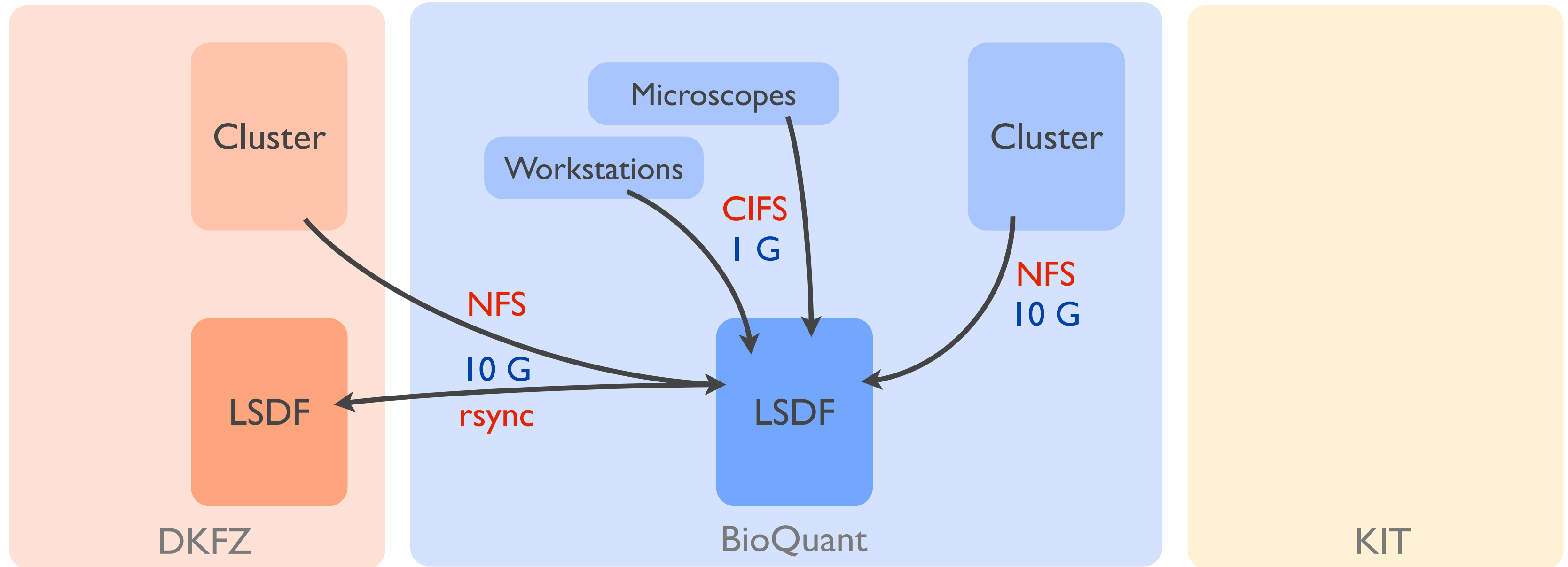
BioQuant

KIT

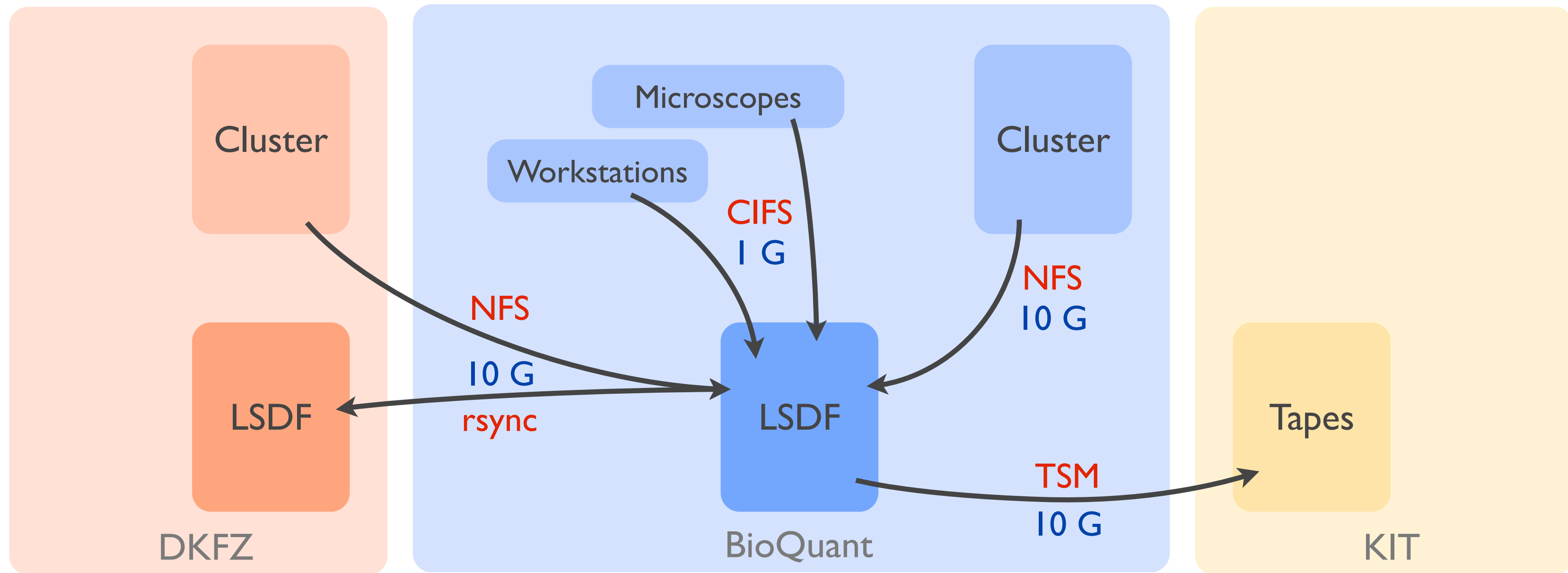
BioQuant LSDF



BioQuant LSDF



BioQuant LSDF



BioQuant LSDF

Group Shares

Microscopy

Sequencing

BioQuant LSDF

Group Shares ~ 80 TB

Microscopy ~ 70 TB

Sequencing 1.3 PB

BioQuant LSDF

Group Shares	~ 80 TB
Microscopy	~ 70 TB
Sequencing	1.3 PB
total	1.6 PB

BioQuant LSDF

Group Shares

Microscopy

Sequencing

BioQuant LSDF

Group Shares

RNA interference

Microscopy

Sequencing

BioQuant LSDF

Group Shares

RNA interference

Microscopy

image size ~ 2 MB

Sequencing

BioQuant LSDF

Group Shares

RNA interference

Microscopy

image size ~ 2 MB

Sequencing

~ 2,000 images per sample / run
future runs up to 70,000 images

BioQuant LSDF

Group Shares

RNA interference

Microscopy

image size ~ 2 MB

Sequencing

~ 2,000 images per sample / run
future runs up to 70,000 images

many samples per study / experiment
“typical” throughput: 20,000 images per day

peak throughput: 3,000,000 images (3 TB) in three weeks

BioQuant LSDF

Group Shares

RNA interference

Microscopy

image size ~ 2 MB

Sequencing

~ 2,000 images per sample / run
future runs up to 70,000 images

many samples per study / experiment
“typical” throughput: 20,000 images per day

peak throughput: 3,000,000 images (3 TB) in three weeks

many small files

BioQuant LSDF

Group Shares

Microscopy

Sequencing

many small files

BioQuant LSDF

Group Shares

Microscopy

Sequencing

Microscopes

raw data
write once

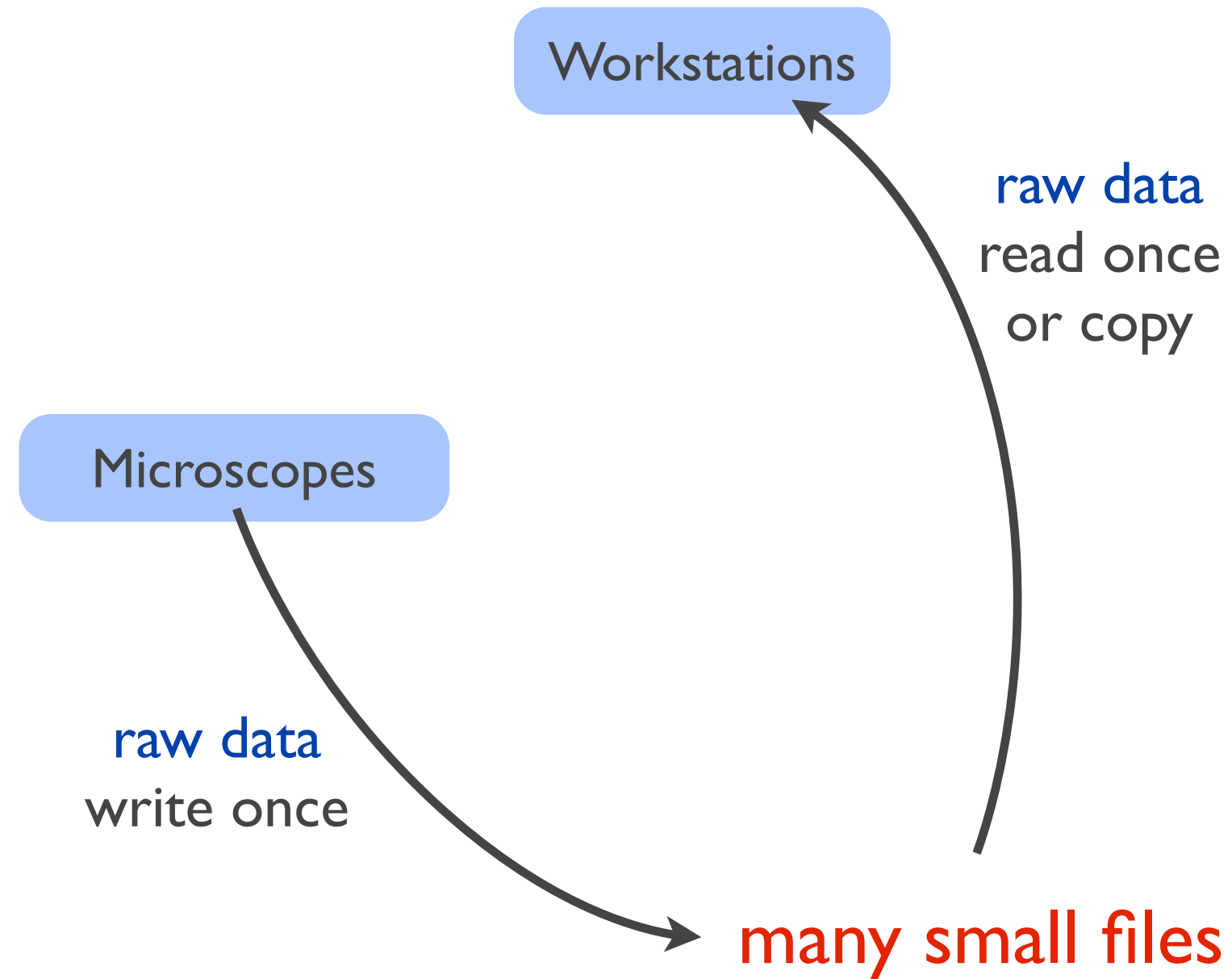
many small files

BioQuant LSDF

Group Shares

Microscopy

Sequencing

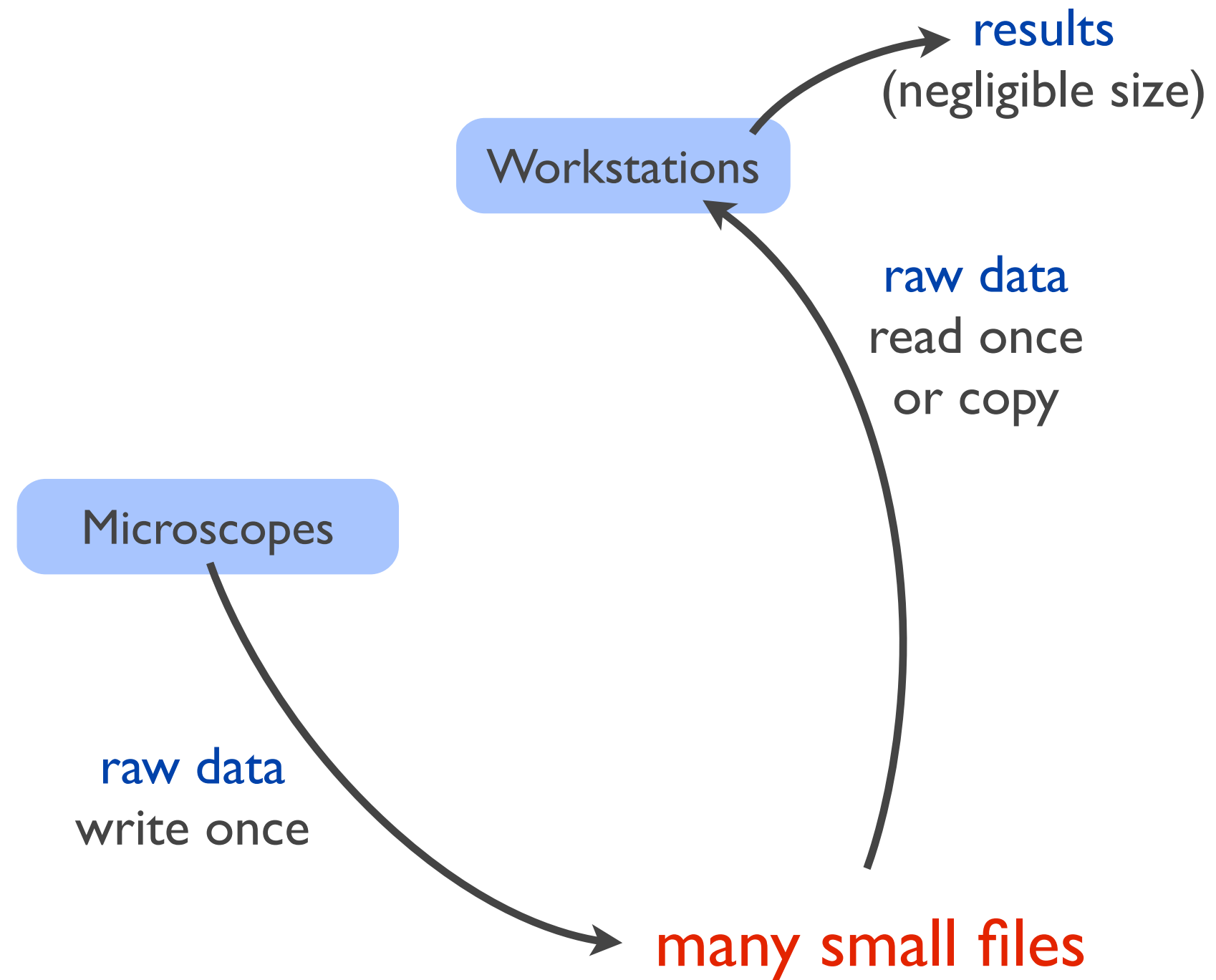


BioQuant LSDF

Group Shares

Microscopy

Sequencing

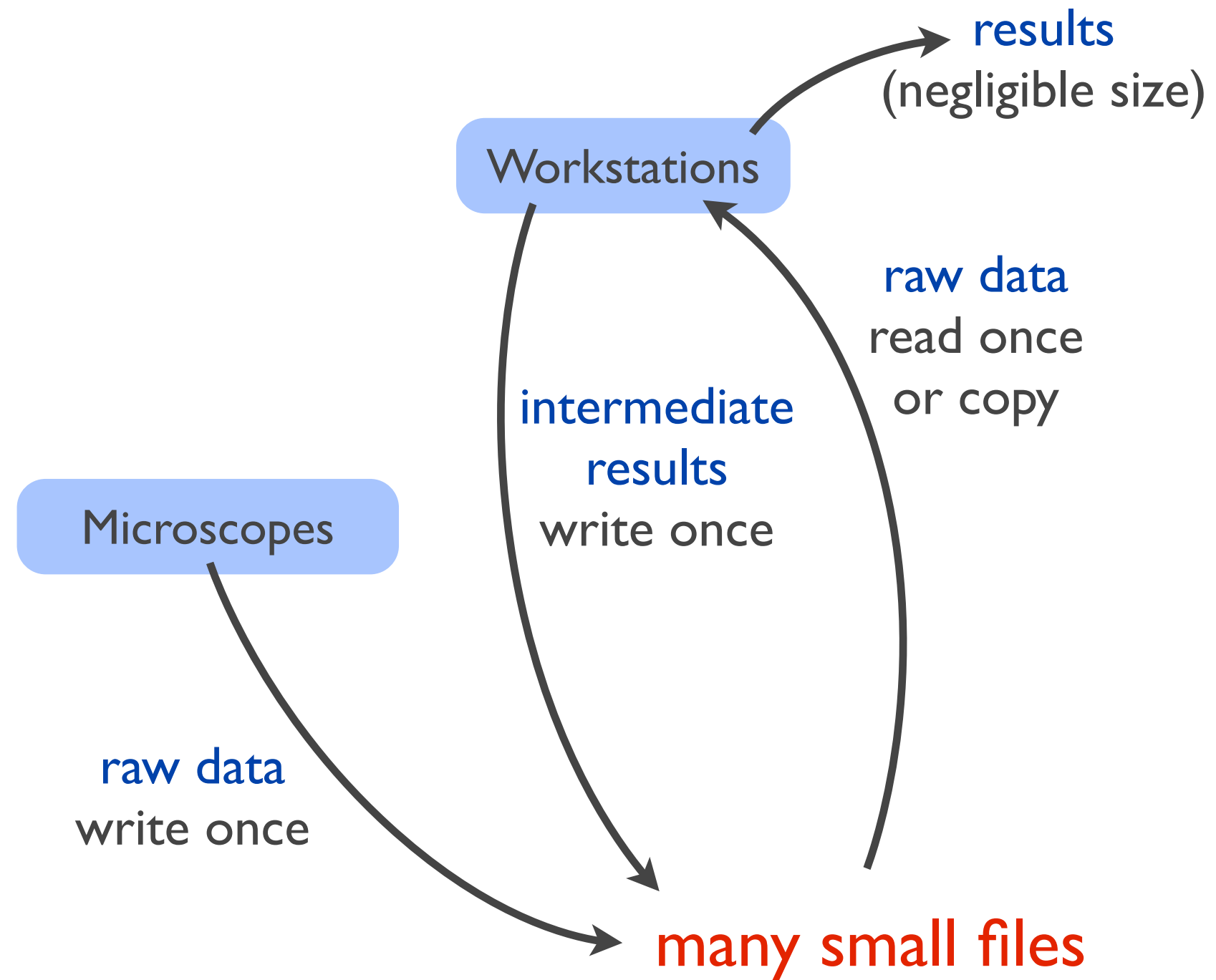


BioQuant LSDF

Group Shares

Microscopy

Sequencing



BioQuant LSDF

Group Shares

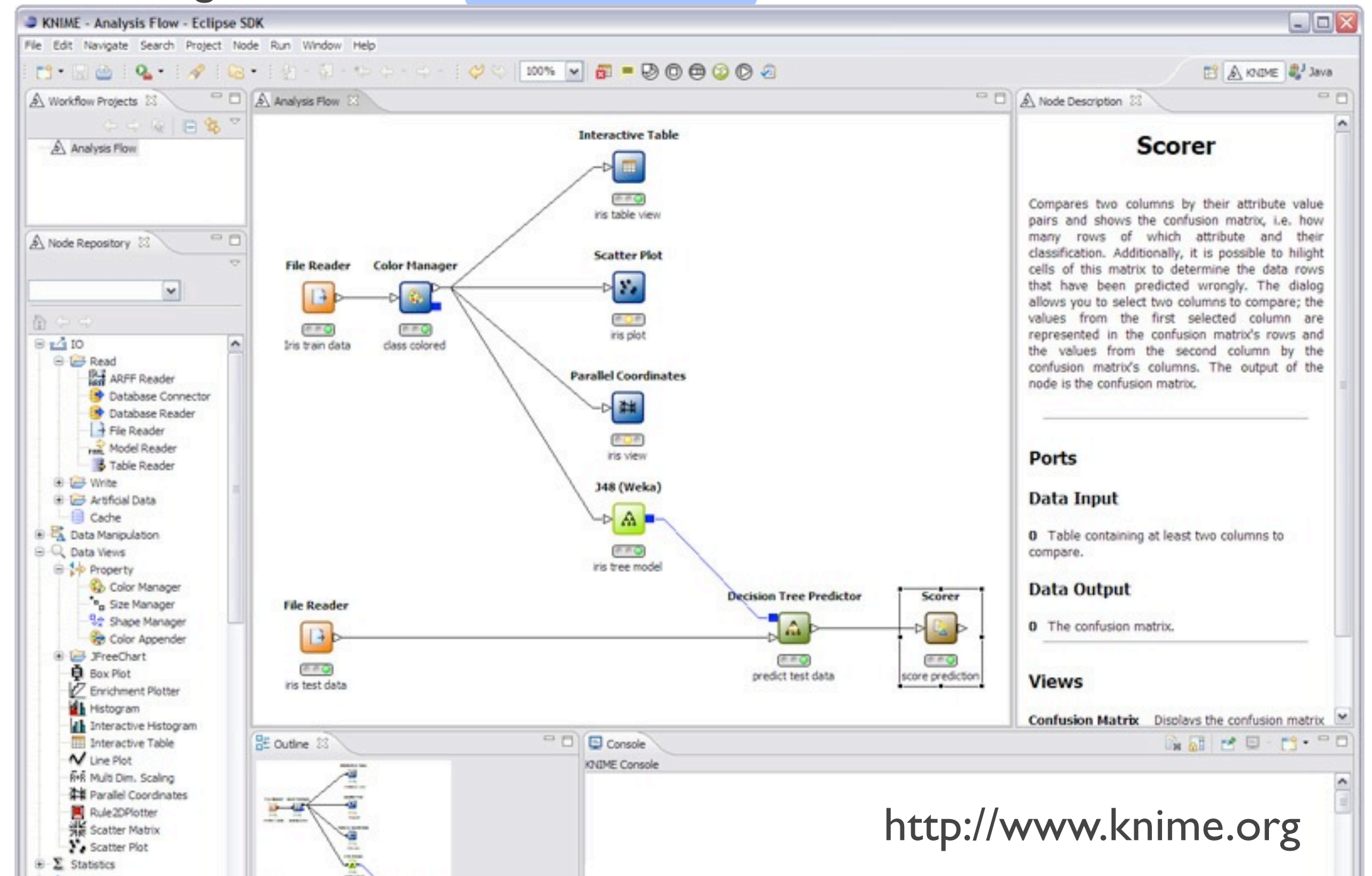
Microscopy

Sequencing

e.g. **KNIME**

Workstations

results
(negligible size)



<http://www.knime.org>

BioQuant LSDF

Group Shares

Microscopy

Sequencing

BioQuant LSDF

Group Shares

International Cancer Genome Consortium

Microscopy

Sequencing

BioQuant LSDF

Group Shares

International Cancer Genome Consortium

Microscopy

~ 200 GB raw data per genome

Sequencing

BioQuant LSDF

Group Shares

International Cancer Genome Consortium

Microscopy

~ 200 GB raw data per genome

Sequencing

~ 2 TB after processing
intermediate results are retained

BioQuant LSDF

Group Shares

International Cancer Genome Consortium

Microscopy

~ 200 GB raw data per genome

Sequencing

~ 2 TB after processing
intermediate results are retained

throughput: 10–20 genomes per week

BioQuant LSDF

Group Shares

International Cancer Genome Consortium

Microscopy

~ 200 GB raw data per genome

Sequencing

~ 2 TB after processing
intermediate results are retained

throughput: 10–20 genomes per week

total ~ 2 PB of data

~ 1 PB at BioQuant LSDF

BioQuant LSDF

Group Shares

International Cancer Genome Consortium

Microscopy

~ 200 GB raw data per genome

Sequencing

~ 2 TB after processing
intermediate results are retained

throughput: 10–20 genomes per week

total ~ 2 PB of data

~ 1 PB at BioQuant LSDF

few large files

BioQuant LSDF

Group Shares

International Cancer Genome Consortium

Microscopy

~ 200 GB raw data per genome

Sequencing

~ 2 TB after processing
intermediate results are retained

throughput: 10–20 genomes per week

total ~ 2 PB of data

~ 1 PB at BioQuant LSDF

few large files

only one ICGC project stored at BQ-LSDF



BioQuant LSDF

Group Shares

International Cancer Genome Consortium

Microscopy

~ 200 GB raw data per genome

Sequencing

~ 2 TB after processing
intermediate results are retained

throughput: 10–20 genomes per week

total ~ 2 PB of data

~ 1 PB at BioQuant LSDF

few large files

different “views” of same data via symbolic links

BioQuant LSDF

Group Shares

Microscopy

Sequencing

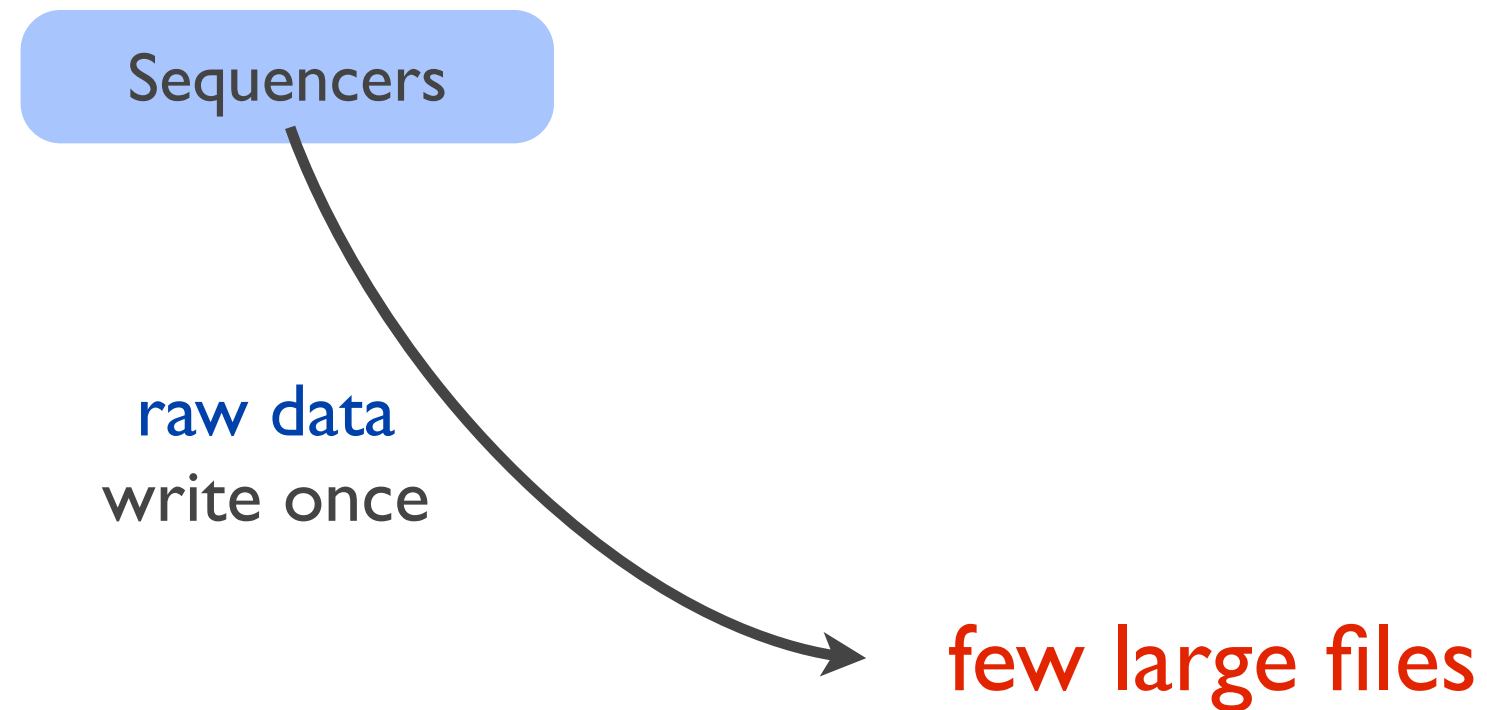
few large files

BioQuant LSDF

Group Shares

Microscopy

Sequencing

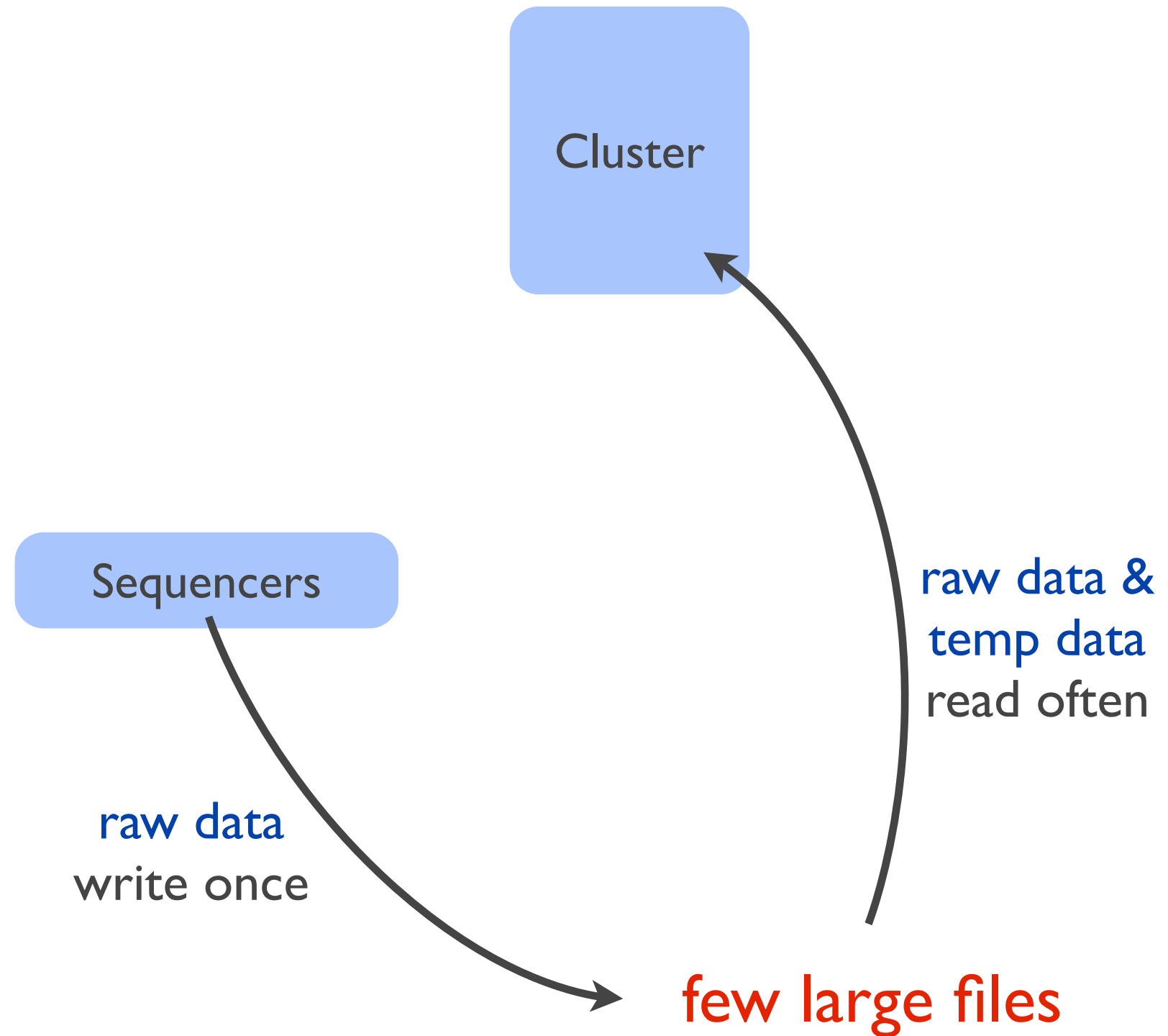


BioQuant LSDF

Group Shares

Microscopy

Sequencing

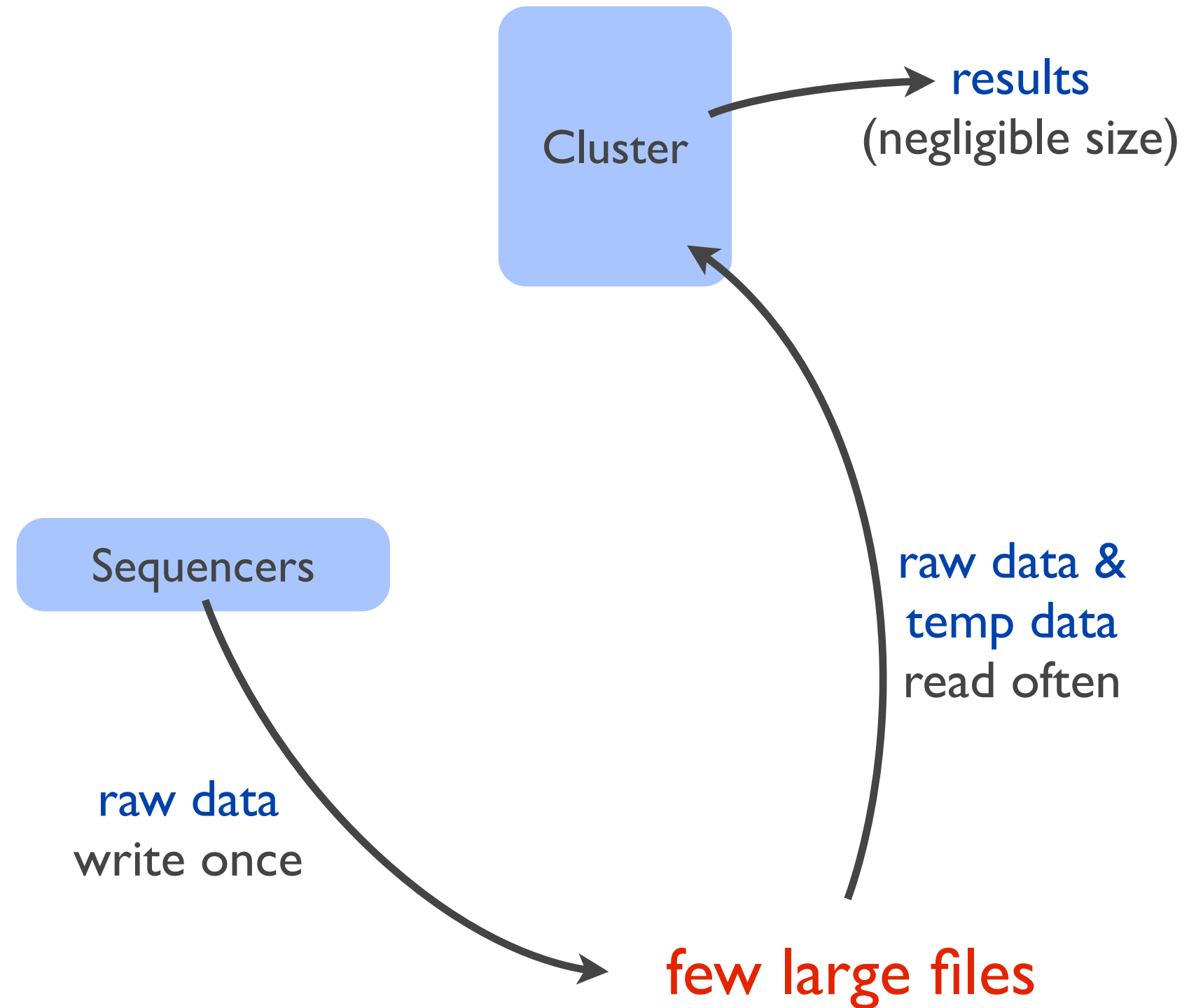


BioQuant LSDF

Group Shares

Microscopy

Sequencing

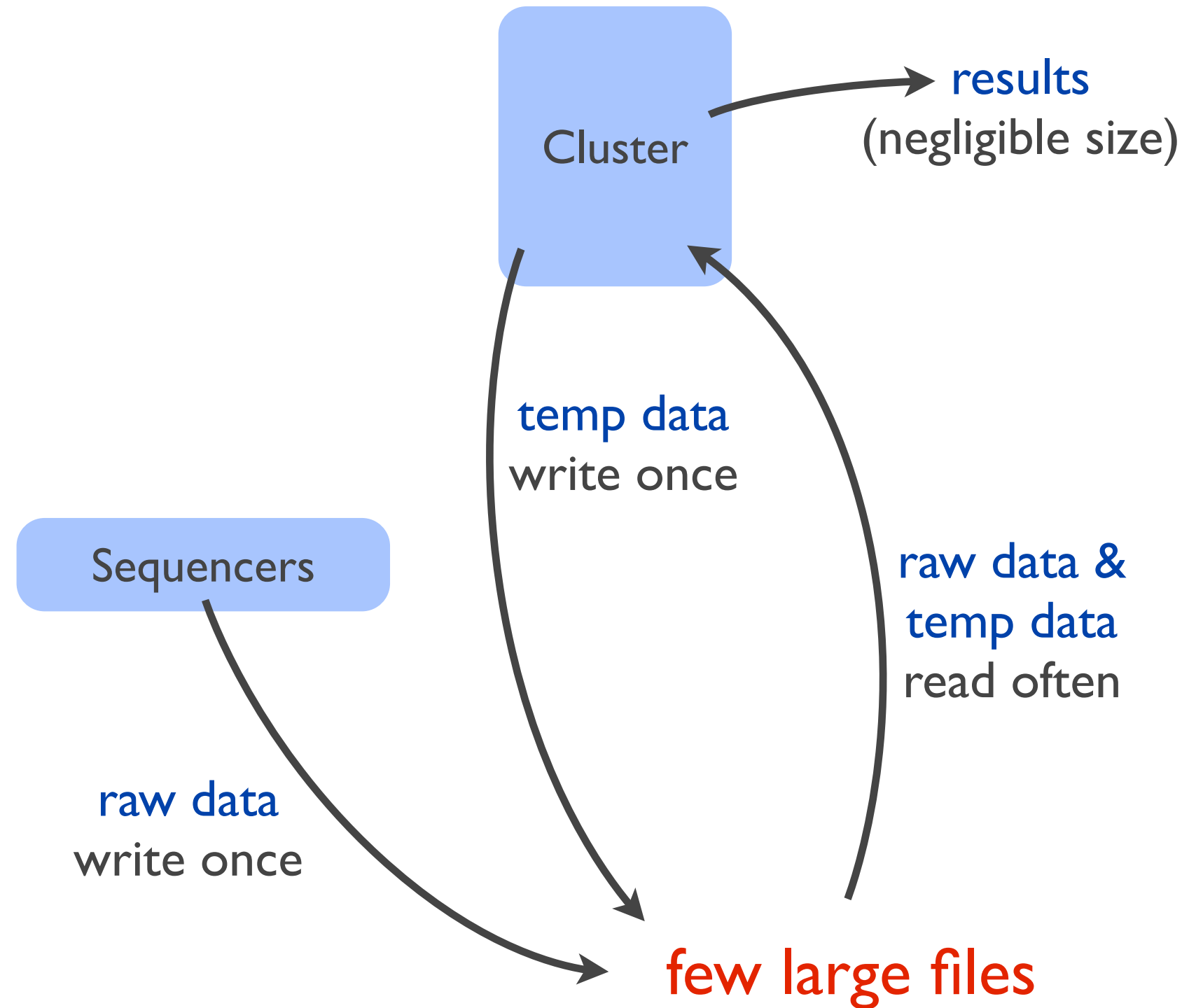


BioQuant LSDF

Group Shares

Microscopy

Sequencing



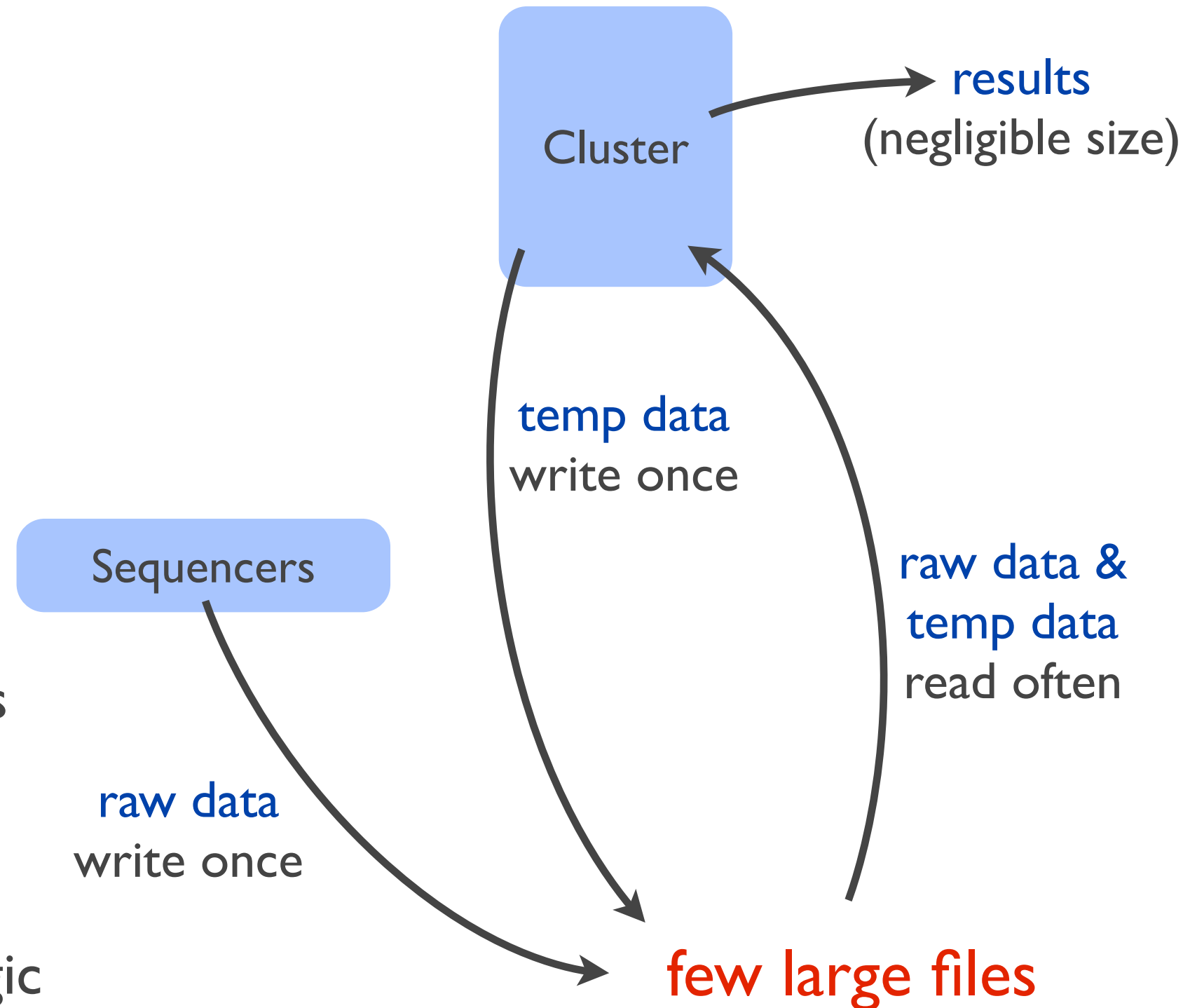
BioQuant LSDF

Group Shares

Microscopy

Sequencing

store as provided by sequencers
organize semantically (symlinks)
check for data corruption
perform genome alignment magic
compute standard measures



BioQuant LSDF

Cluster



















results
(negligible size)

Group Shares

Microscopy

Sequencing
One Touch Pipeline

Sylwester Radomski,
DKFZ

Home Overview Individuals Sequences Runs Progress Processes Run Submit Logout						
List of Workflows						
Enable auto refresh						
	Workflow	Count	Count of Failed	Last Success	Last Failure	Duration
-	mergedBamDiscoveryWorkflow	0		-	-	-
	 FileSystemConsistencyWorkflow	78		02:51:20 AM	-	1 h 51 min
	 BwaAlignmentWorkflow	100	52	Sat Dec 22 2012	Sat Dec 22 2012	1 day(s) 13 h
	 FastqcSummaryWorkflow	266		Wed Feb 27 2013	-	660 msec
	 FastqcWorkflow	2615	153	Wed Feb 27 2013	Tue Feb 05 2013	1 h 31 min
	 runUnpack	2		Thu Oct 18 2012	-	2 sec 539 msec
	 DataInstallationWorkflow	401	1	Wed Feb 27 2013	Thu Feb 21 2013	19 h 56 min
	 createSeqScan	3898	35	Tue Feb 26 2013	Tue Feb 19 2013	752 msec
	 DataLocationWorkflow	855	44	Fri Oct 12 2012	Fri Oct 12 2012	851 msec
	 loadMetaData	1062	7	Tue Feb 26 2013	04:42:02 PM	1 min 43 sec

BioQuant LSDF

Cluster

results
(negligible size)

Group Shares

Microscopy

Sequencing
One Touch Pipeline

Sylwester Radomski,
DKFZ

Home Overview Individuals Sequences Runs Progress Processes Run Submit Logout									
List of Runs									
Search: <input type="text"/>									
Name	Center	Storage Realm	Date Created	Date Executed	Blacklisted	Multiple Source	Fastqc		
130213 SN952 0141 BC1PB8ACXX	dkfz		2013-02-27 04:04:04		false	false	NaN		
130212 SN509 0208 BC1R00ACXX	dkfz	dkfz	2013-02-26 01:02:28	2013-02-12	false	false	0.82		
130212 SN509 0207 AD1U4CACXX	dkfz	dkfz	2013-02-26 12:24:40	2013-02-12	false	false	0.80		
130206 SN143 0612 BD1H8EACXX	gatk	bioquant	2013-02-26 10:56:42	2013-02-06	false	false	0.84		
130103 SN509 0201 AC1BVCACXX	dkfz	bioquant	2013-02-25 11:42:06	2013-01-03	false	false	0.83		
NG-6166 Exome Data batch22	gatk	dkfz	2013-02-22 11:48:39		false	false	NaN		
130201 SN486 0373 AC1C15ACXX	gatk	dkfz	2013-02-22 10:04:44	2013-02-01	false	false	NaN		
130205 SN7001397 0049 AC1PLNACXX	dkfz	bioquant	2013-02-21 12:53:22	2013-02-05	false	false	0.86		
130205 SN7001153 0110 BD1RYKACXX	dkfz	bioquant	2013-02-21 12:44:56	2013-02-05	false	false	0.78		
130205 SN1007 0135 AD1U5MACXX	dkfz	bioquant	2013-02-21 11:10:14	2013-02-05	false	false	0.83		
130129 SN486 0372 BD1GWAACXX	gatk	bioquant	2013-02-21 10:28:23	2013-01-29	false	false	0.75		
130205 SN1007 0136 BD1TBBACXX	dkfz	mixed	2013-02-19 11:51:28	2013-02-05	false	false	0.83		
130130 SN7001103R 0108 BD1N9DACXX	dkfz	bioquant	2013-02-19 11:13:45	2013-01-30	false	false	0.86		
130130 SN7001393 0039 BC1NMCACXX	dkfz	bioquant	2013-02-18 02:10:51	2013-01-30	false	false	0.81		
130130 SN7001103R 0107 AD1T8YACXX	dkfz	bioquant	2013-02-18 12:11:44	2013-01-30	false	false	0.86		
130129 SN7001153 0107 AC1K37ACXX	dkfz	bioquant	2013-02-18 12:07:34	2013-01-29	false	false	0.86		
130131 SN952 0138 AD1J8WACXX	dkfz	mixed	2013-02-18 11:06:32	2013-01-31	false	false	0.64		
130129 SN7001149 0123 BC1PMDACXX	dkfz	bioquant	2013-02-18 10:58:10	2013-01-29	false	false	0.80		
130103 SN952 0135 BD1H60ACXX	dkfz	dkfz	2013-02-15 04:00:00	2013-01-03	false	false	0.81		
130131 SN952 0139 BC1R14ACXX	dkfz	mixed	2013-02-14 10:13:00	2013-01-31	false	false	0.75		
130131 SN7001177 0109 AD1NUKACXX	dkfz	dkfz	2013-02-12 03:23:13	2013-01-31	false	false	0.62		
130129 SN509 0206 BC1K3GACXX	dkfz	dkfz	2013-02-12 03:19:10	2013-01-29	false	false	0.64		
130129 SN509 0205 AD1T51ACXX	dkfz	dkfz	2013-02-12 03:17:51	2013-01-29	false	false	0.61		
130122 SN7001397 0048 BC1DG4ACXX	dkfz	bioquant	2013-02-12 11:09:37	2013-01-22	false	false	0.86		
130122 SN7001397 0047 AD1U2BACXX	dkfz	bioquant	2013-02-12 11:02:25	2013-01-22	false	false	0.85		

BioQuant LSDF



results
(negligible size)

Group Shares

Microscopy

Sequencing
One Touch Pipeline

[Home](#) [Overview](#) [Individuals](#) [Sequences](#) [Runs](#) [Progress](#) [Processes](#) [Run Submit](#) | [Logout](#) |

Run details

General

[previous run](#) [next run](#)

Name	130206_SN143_0612_BD1H8EACXX
Sequencing Center	gatk
Sequencing Technology	HiSeq2000
Date Executed	2013-02-06
Date Created	2013-02-26 10:56:42.65
Meta Data Path	/icgc/lscdf/ftp/gatk/
Final Locations	/icgc/lscdf/project//pedbrain/pa/sequencing//whole_genome_sequencing/gatk/

Processing

DataInstallationWorkflow	Show Details
loadMetaData	Show Details
FastqcSummaryWorkflow	Show Details

Data Files

130206_SN143_0612_BD1H8EACXX_fastq.tsv	2013-02-26	true
--	------------	------

8 PABL94 CONTROL
WHOLE_GENOME PAIRED

Insert size: 278

number of base pairs: 37.7 G

- s s 8 1 sequence.txt.gz	astrocytome	✓ meta-data	✓ lscdf	✓ linked	16.5 GB	2013-02-27	FastQC (7/4/0)
- s s 8 2 sequence.txt.gz	astrocytome	✓ meta-data	✓ lscdf	✓ linked	16.8 GB	2013-02-27	FastQC (8/3/0)

Files not used:

OneTouchPipeline | Eils Labs | TBI DKFZ | Build: 19b3cc9 (production/1.8)

Sylwester Radomski,
DKFZ

Tbi-cluster

Logged in : Rolf Kabbe | [Log out](#)

Active jobs

Completed jobs

Nodes

Job categories

Projects

Preferences

Admin

About

Showing details for job : 929752

Use as MinJobId

Use as MaxJobId

Modify this job

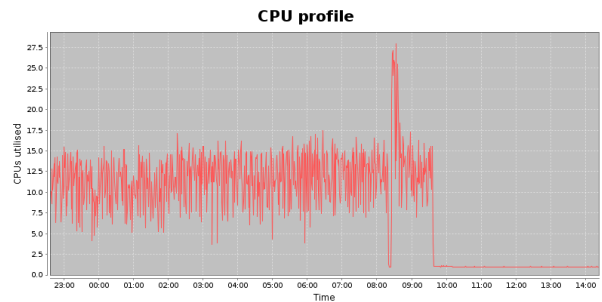
Delete Job

Purge Job

Extend Walltime

Requested and used resources

Attribute	Value
Owner	wangq
Name	LINA463K_LINA463K_merge_removedup
Submit arguments	-o /icgc/lsdf/mb/analysis/LINA/analysis/DKFZLSDF/results_per_pid/LINA463K_logs -j oe -M q.wang@dkfz.de -N LINA463K_LINA463K_merge_removedup -v CONFIG_FILE=/home/wangq/SVNrepository/ngs2/branches/pipelines/QCPipeline_BSseq/DKFZLSDF_LINA_CONFIG/pipelineParams_TEST.txt,pid=LINA463K,INPUT_BAMS=/icgc/lsdf/mb/analysis/LINA/analysis/DKFZLSDF/results_per_pid/LINA463K_temp/alignment /LINA463K_run121121_SN7001153_0093_BD1H88ACXX_2.rmdup.bam:/icgc/lsdf/mb/analysis/LINA/analysis/DKFZLSDF/results_per_pid/LINA463K_temp/alignment /LINA463K_run121127_SN952_0127_ADJHCEACXX_6.rmdup.bam:/icgc/lsdf/mb/analysis/LINA/analysis/DKFZLSDF/results_per_pid/LINA463K_temp/alignment /LINA463K_run130117_SN7001149_0121_BD1J20ACXX_2.rmdup.bam:/icgc/lsdf/mb/analysis/LINA/analysis/DKFZLSDF/results_per_pid/LINA463K_temp/alignment /LINA463K_run130117_SN7001149_0121_BD1J20ACXX_3.rmdup.bam:/icgc/lsdf/mb/analysis/LINA/analysis/DKFZLSDF/results_per_pid/LINA463K_temp/alignment /LINA463K_run130117_SN7001149_0121_BD1J20ACXX_4.rmdup.bam,OUTPUT_BAM=/icgc/lsdf/mb/analysis/LINA/analysis/DKFZLSDF/results_per_pid/LINA463K_temp/alignment /LINA463K_Slides_merged.rmdup.bam,METRICS_FILE=/icgc/lsdf/mb/analysis/LINA/analysis/DKFZLSDF/results_per_pid/LINA463K_temp/alignment/LINA463K_Slides_merged.rmdup.bam.dupmark.metrics.txt /home/wangq/SVNrepository/ngs2/branches/pipelines/QCPipeline_BSseq/analysis_tools/MergeBAM_MarkDup.sh
Queue	verylong
Job Type	BATCH
Status	RUNNING
Queued	Tue Feb 5 22:37:09 2013
Started	Tue Feb 5 22:37:09 2013
Ended	--:--:--
Exit Code	
Start count	1
Exec Host	tbi-dsx08
Requested walltime	50:00:00
Elapsed walltime	15:45:47
CPU time	135:19:48
CPU time / cores	135:19:48
CPUs avg. utilised	8.59
Cores	1
Memory Requested	5.0 GB
Maximum Memory Used	4.4 GB
Memory Efficiency	0.89
File Size	250.0 GB
Job Category	wangq.LINA463K-LINA463K-merge-removedup:cores=1:mem=5.0 GB:walltime=50:00:00

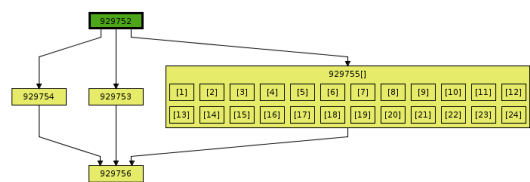


[Download profile](#)

Dependencies In	Dependencies Out
ACTIVE	BEFOREOK 929753
ACTIVE	BEFOREOK 929755[1]
ACTIVE	BEFOREOK 929754

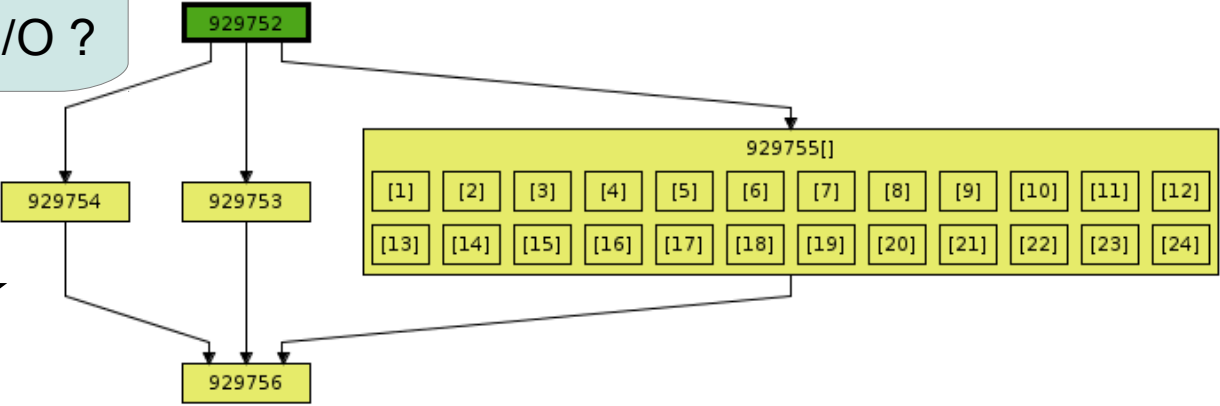
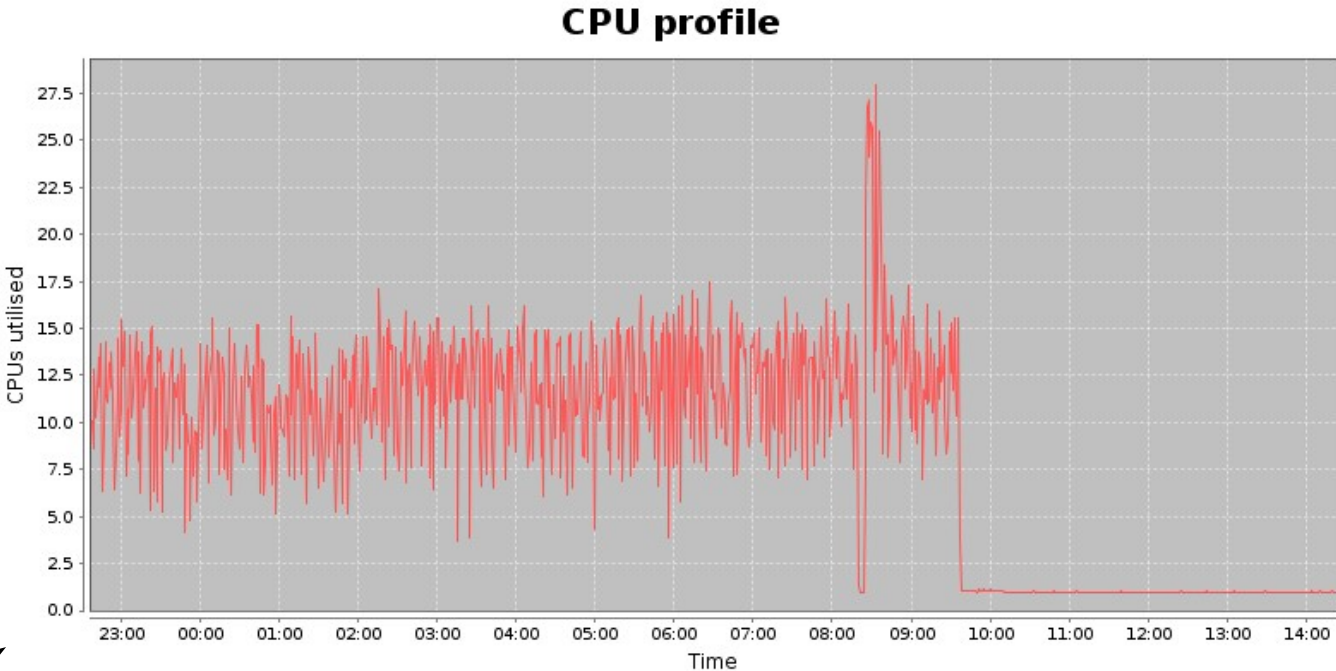
Warnings	
CPU	SEVERE More CPU-time used (135:19:48) than expected (15:45:47), ratio : 8.59
SPECIFICATION	INFO PBS will prevent that this job can create files larger than 250.0 GB

Dependency graph



Drop of CPU
Caused by I/O ?

Dependency-graph
Shows progress
of workflows



BioQuant

Group Shares

Microscopy

Sequencing

Results Database

Jürgen Eils, DKFZ

iCHIP Web - Image result for gene query - Mozilla Firefox

Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe

Welcome to www.iCHIP.de iCHIP Web - Displaying details iCHIP Web - Image result for gene qu...

https://ichip.bioquant.uni-heidelberg.de/iChipWeb2RMT/faces/secure/geneQuery.jsp

iCHIP

Home Data Study info News Results Exchange Feedback Image Admin

[Quick Navigation](#)

Use this navigation to navigate iCHIP quickly:

- Home
- News
- Query
- Search by gene
- Search for images
- Export study
- Study info

Image result

For the given search criteria the following gene was found: A2M

Please note: Every image is a **place holder** for all images of the spot / well. Click on an image to view all corresponding images of the spot / well.

channel

- ☒ dapi
- ☐ GFP

GENE: A2M RNA: 121211

[Download](#)

score	0.48
p-value	0.11
cell count median	667
gfp mean signal median	450.578
infection ratio median	0.201

RNA: 121212

[Download](#)

score	-0.85
p-value	0.19
cell count median	207
gfp mean signal median	683.922

[Download](#)

score	-0.85
p-value	0.19
cell count median	169
gfp mean signal median	674.393

© 2009 German Cancer Research Center, Heidelberg

BioQual

Group Shares

Microscopy

Sequencing

Results Database

Jürgen Eils, DKFZ

iCHIP Web - Displaying details - Mozilla Firefox

File Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe

Welcome to www.iCHIP.de iCHIP Web - Displaying details iCHIP Web - Image result for gene qu... +

https://services.ichip.de/web/faces/secure/entityDetail.jsp

Details

→ Displaying detailed information for: SK-MES-1 cell line

Hierarchy:

- cell line
 - SK-MES-1

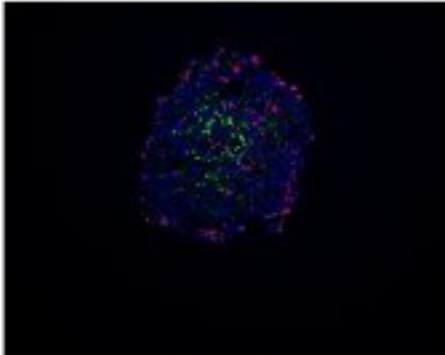
Values Experiments Images

☒ images per time point ☐ images per condition ☐ images per time point and condition ☐ images per slide ☐ images per section

TIMEPOINT

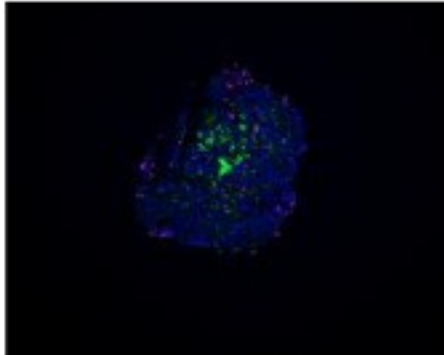
☒ 17 d ☐ 24 d ☐ 46 d

Images



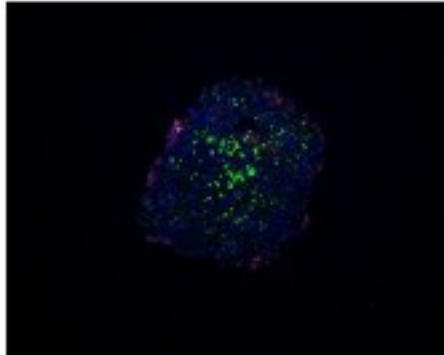
Download

SPHEROID	2
SECTION	2
SLIDE	06
CONDITION	1 mM Glucose, 0,28 mM O2
IF_STAINING	Ki67-ColIV-HOECHST



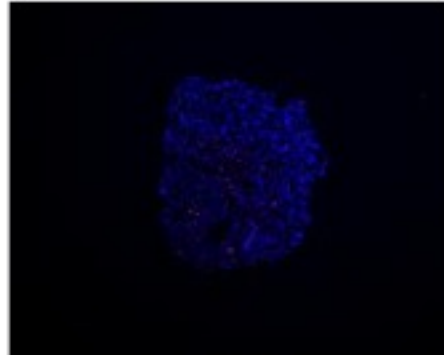
Download

SPHEROID	2
SECTION	3
SLIDE	06
CONDITION	1 mM Glucose, 0,28 mM O2
IF_STAINING	Ki67-ColIV-HOECHST




Download

SPHEROID	2
SECTION	4
SLIDE	06
CONDITION	1 mM Glucose, 0,28 mM O2
IF_STAINING	Ki67-ColIV-HOECHST



Download

SPHEROID	2
SECTION	5
SLIDE	06
CONDITION	1 mM Glucose, 0,28 mM O2
IF_STAINING	TUNEL-_-HOECHST

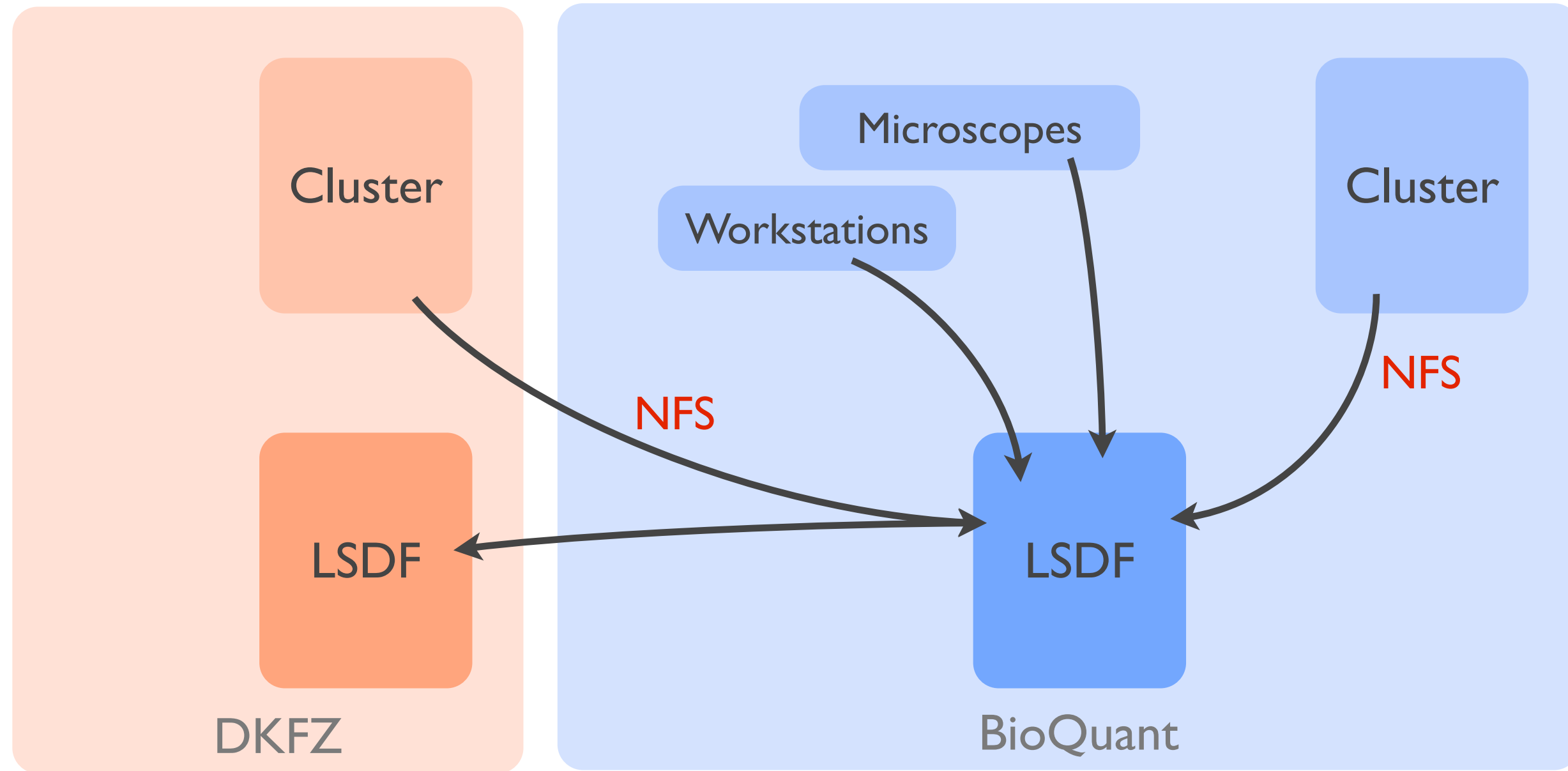


Download

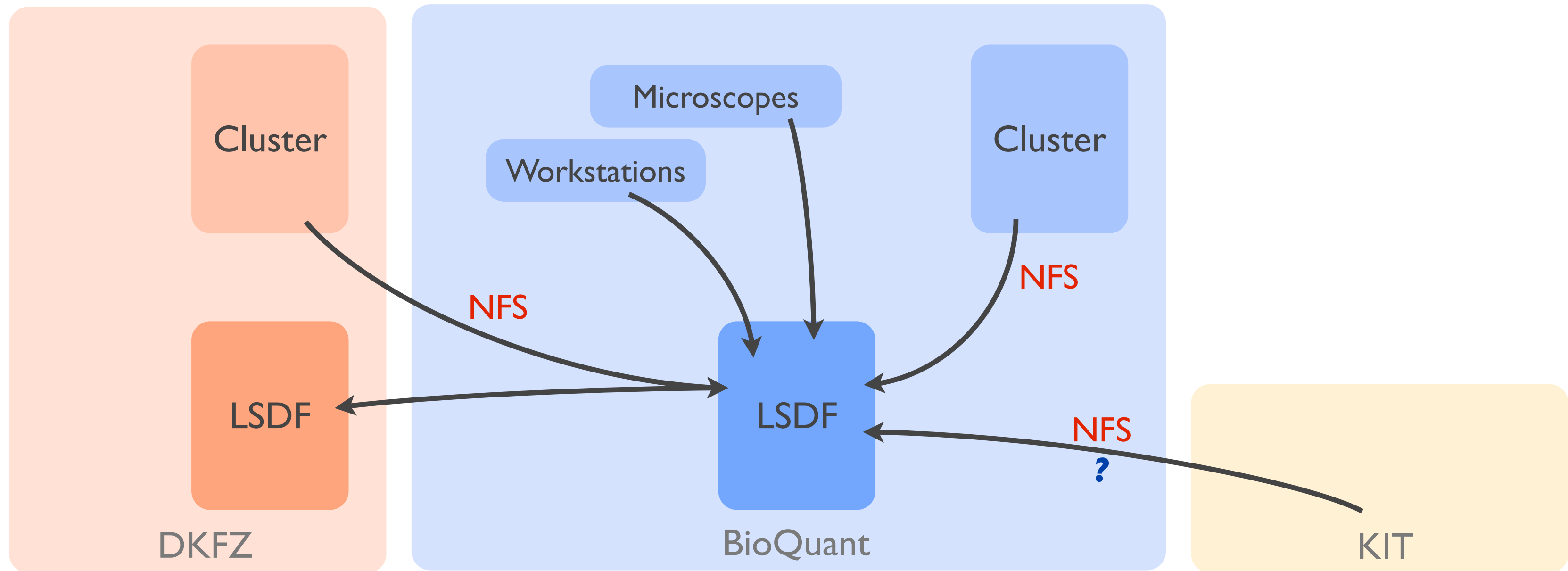
SPHEROID	2
SECTION	5
SLIDE	06
CONDITION	1 mM Glucose, 0,28 mM O2
IF_STAINING	TUNEL-_-HOECHST

© 2009 German Cancer Research Center, Heidelberg

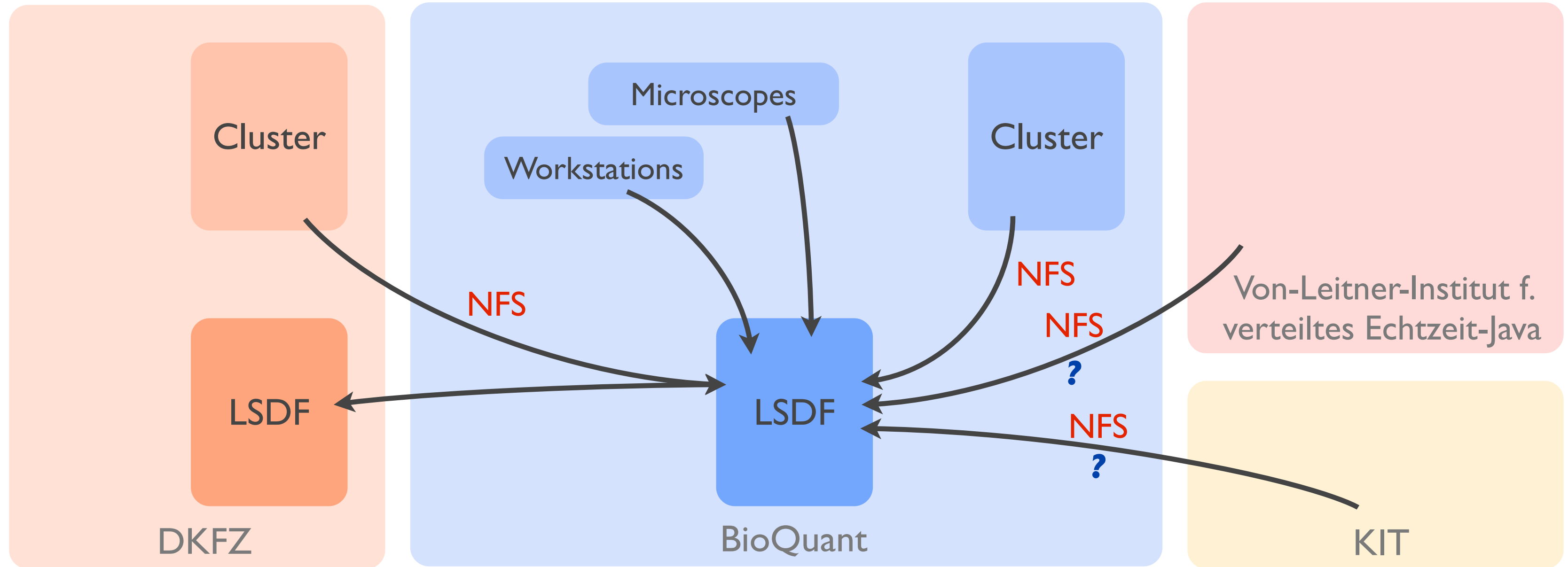
BioQuant LSDF



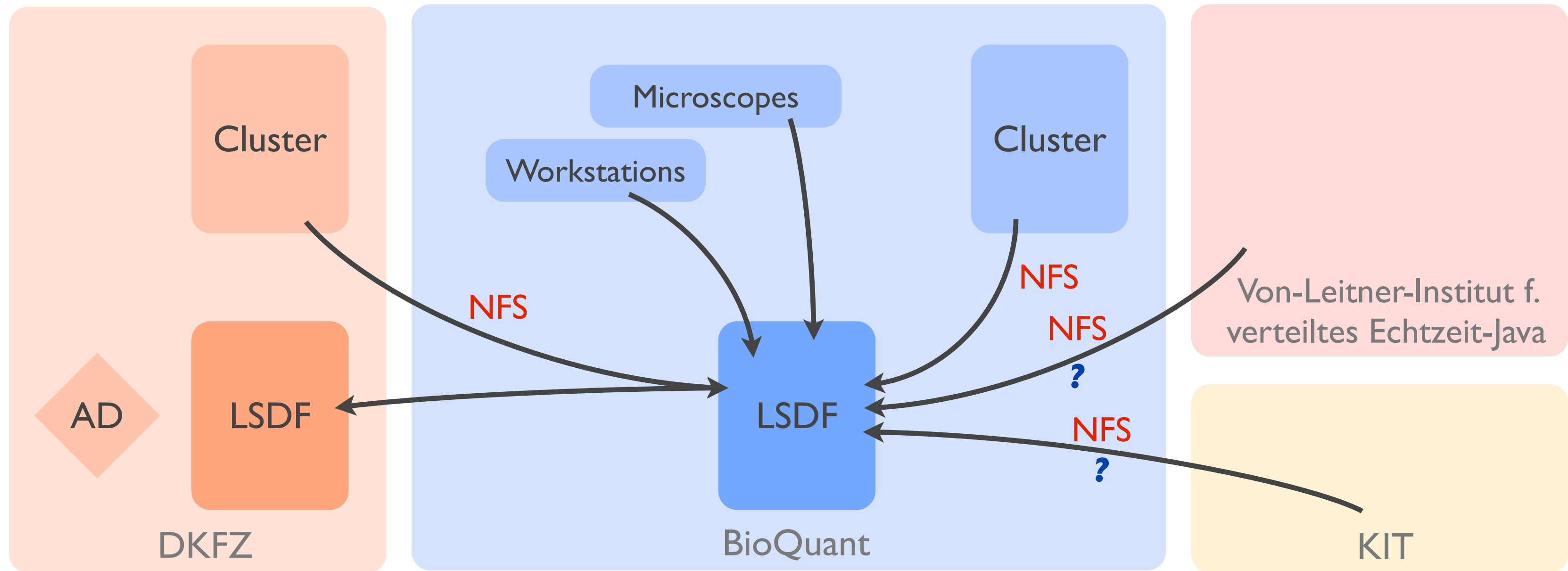
BioQuant LSDF



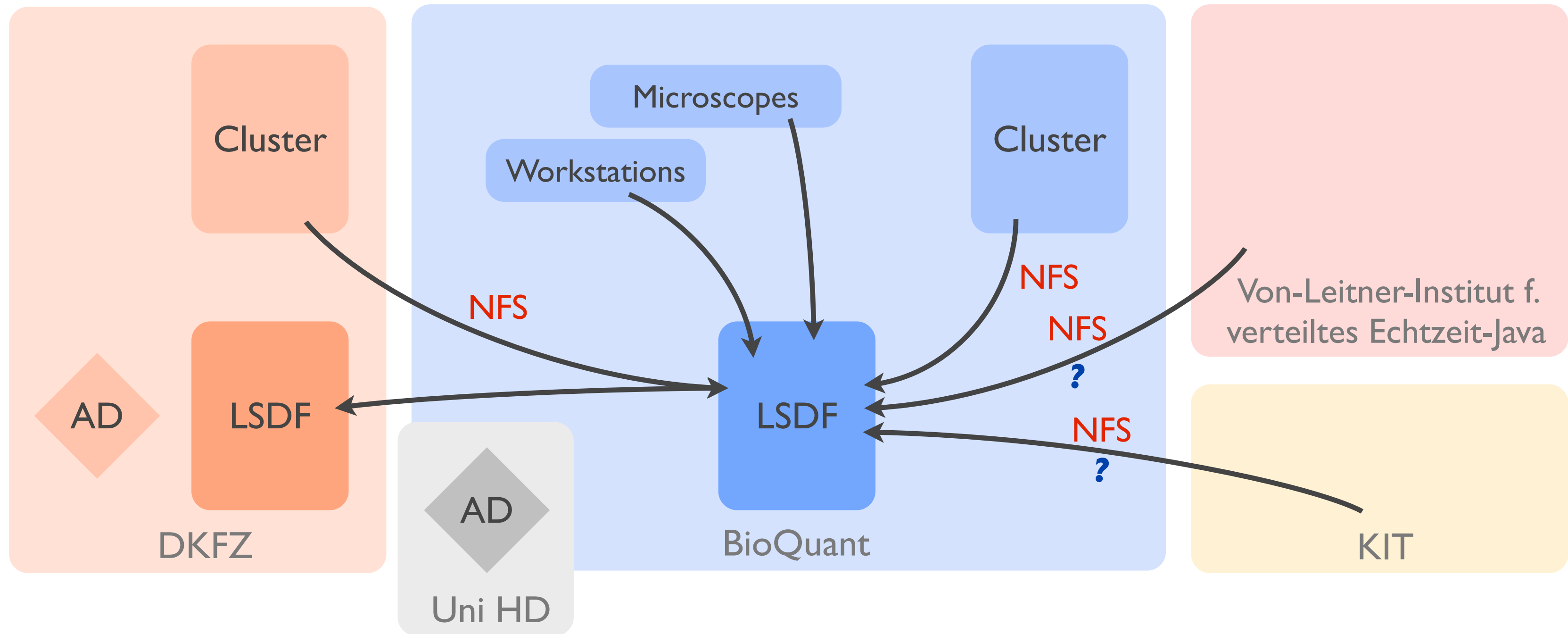
BioQuant LSDF



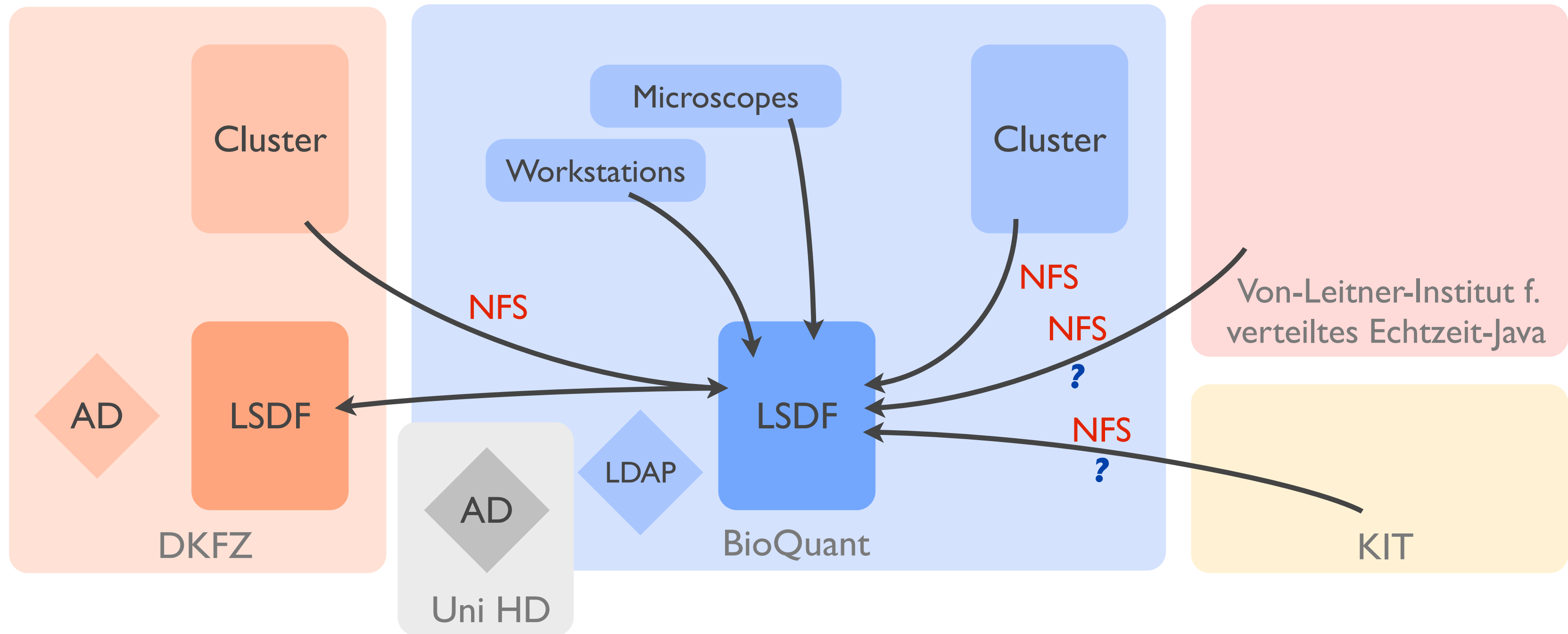
BioQuant LSDF



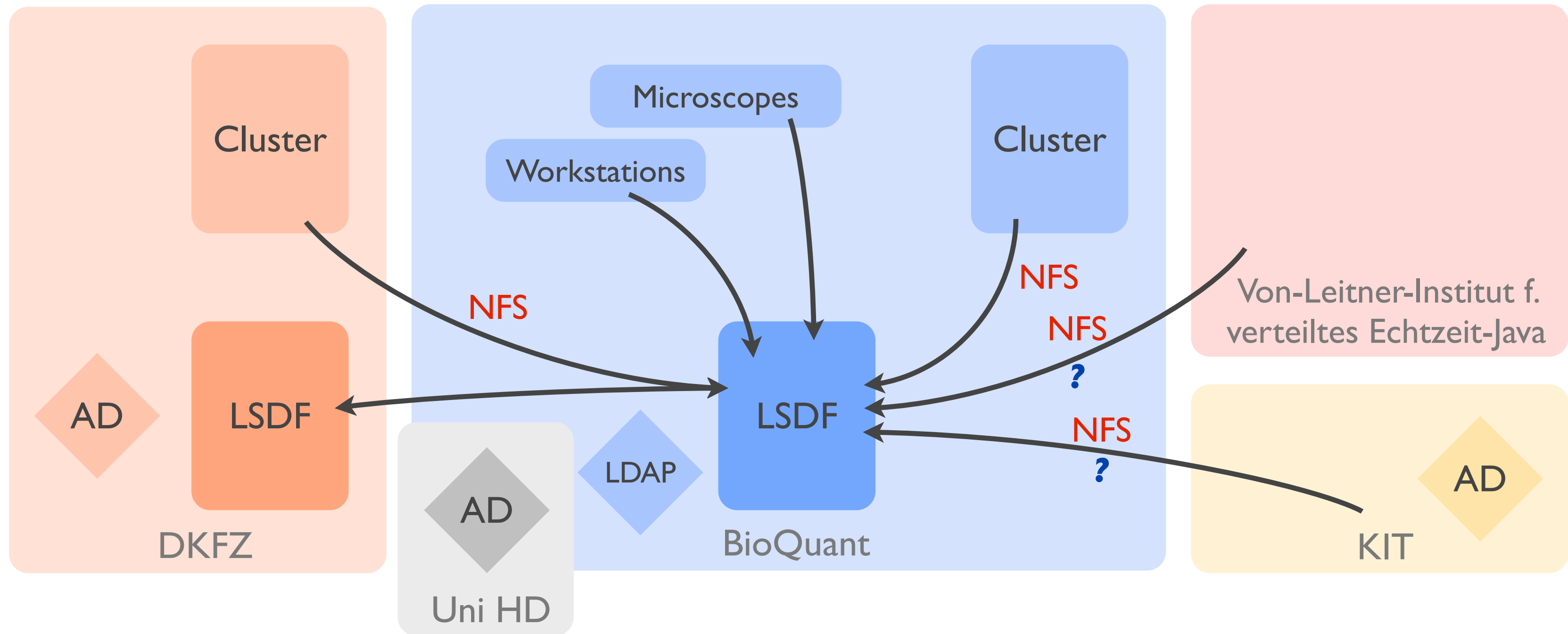
BioQuant LSDF



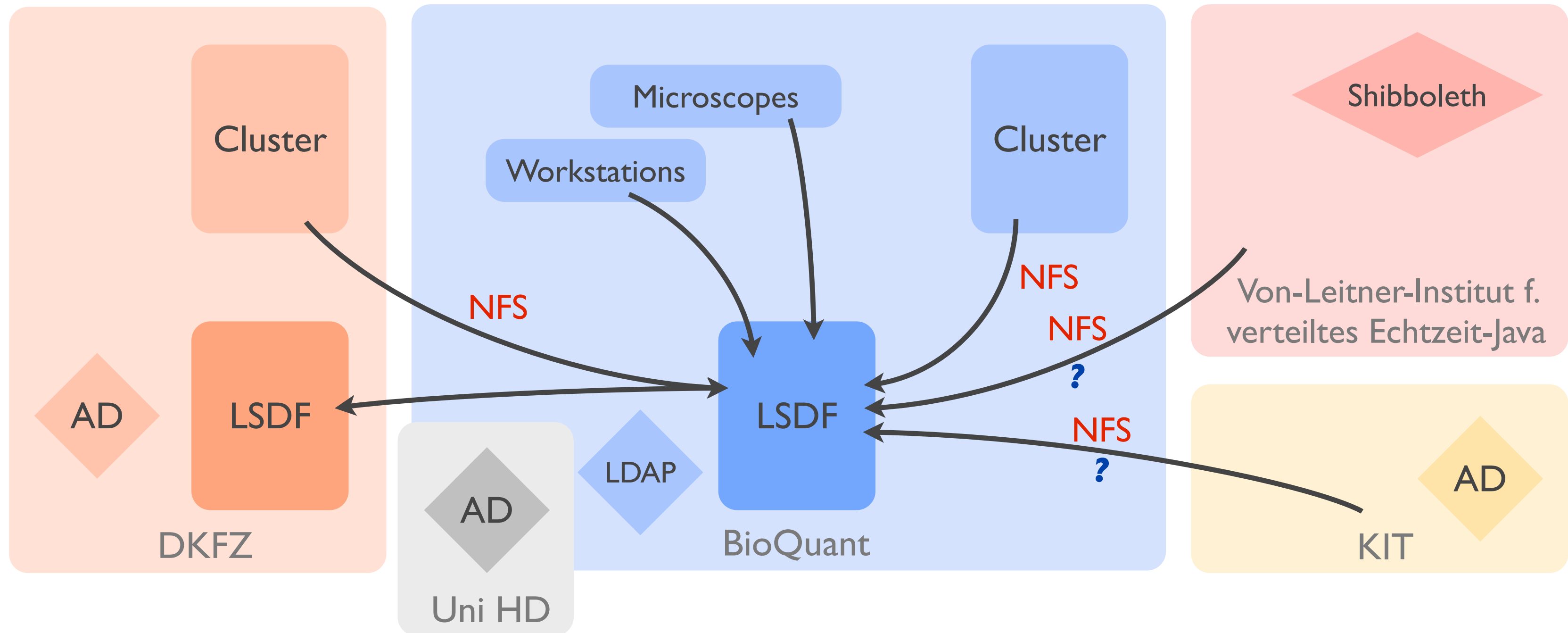
BioQuant LSDF



BioQuant LSDF



BioQuant LSDF



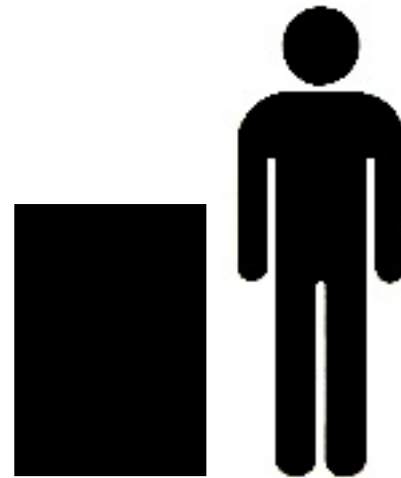
The Problem



The Problem Authentication

NFS via RPC AUTH_UNIX

COMMUNITY BANK
SAFE DEPOSIT BOXES



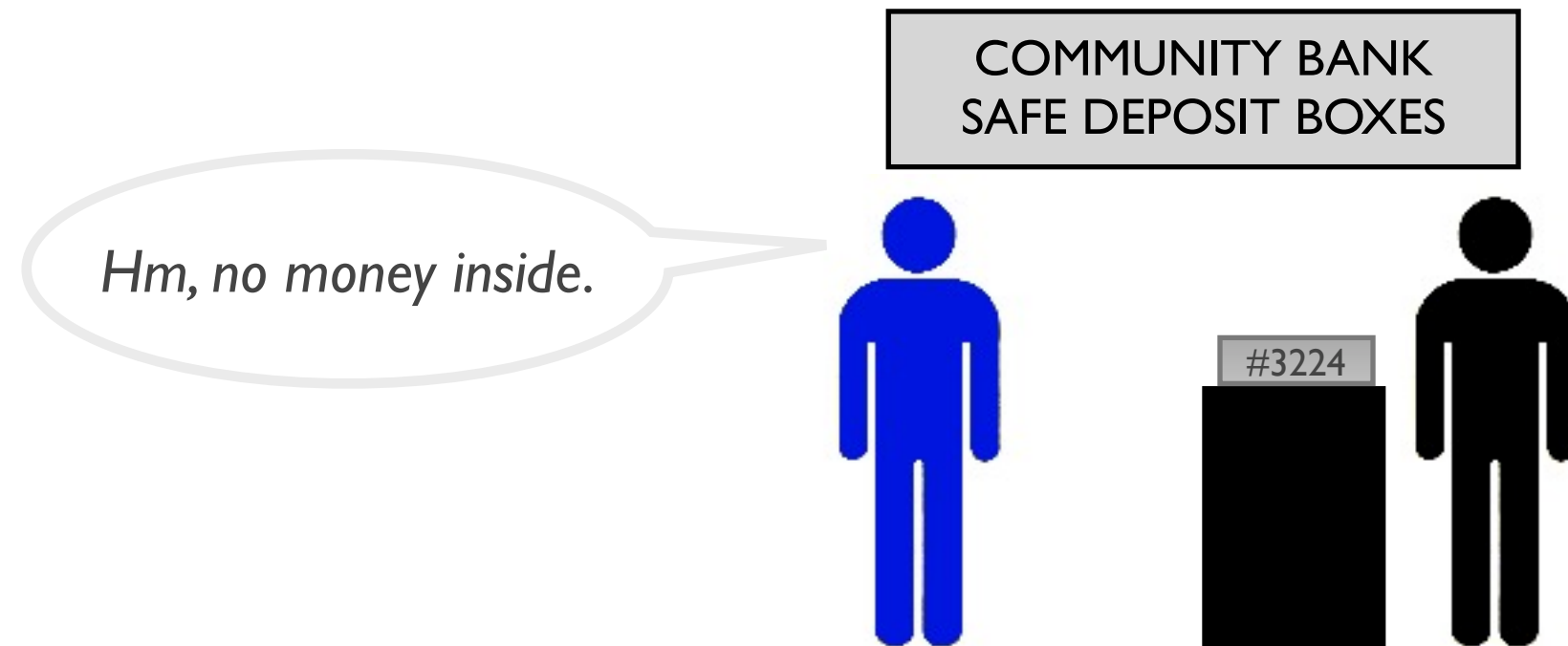
The Problem Authentication

NFS via RPC AUTH_UNIX



The Problem Authentication

NFS via RPC AUTH_UNIX



The Problem Authentication

NFS via RPC AUTH_UNIX



The Problem Authentication

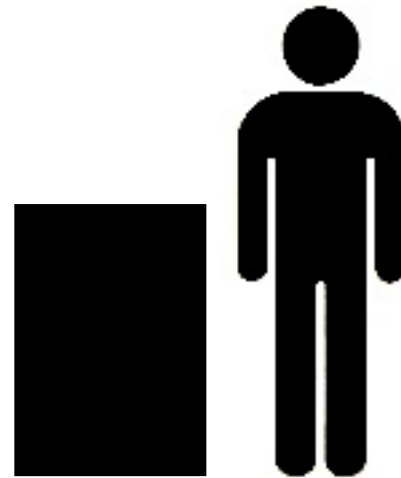
NFS via RPC AUTH_UNIX



The Problem Authentication

NFS via RPC AUTH_UNIX

COMMUNITY BANK
SAFE DEPOSIT BOXES



The Problem Authentication

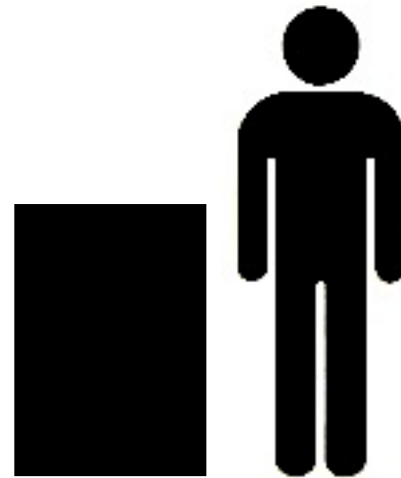
NFS via RPC AUTH_UNIX



The Problem Authentication

NFS via RPC AUTH_UNIX

COMMUNITY BANK
SAFE DEPOSIT BOXES



The Problem Authentication

NFS via RPC AUTH_UNIX



The Problem Authentication

NFS via RPC AUTH_UNIX



The Problem Authentication

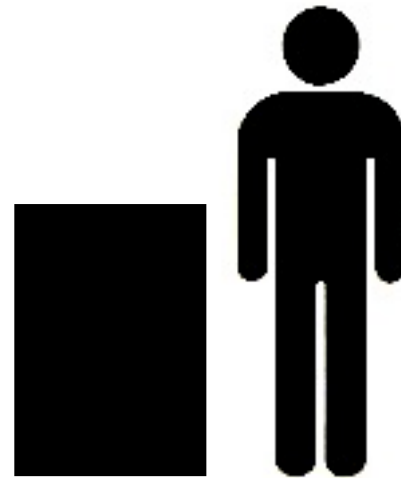
NFS via RPC AUTH_UNIX



The Solution Authentication

NFS via RPCSEC_GSS_KRB5

COMMUNITY BANK
SAFE DEPOSIT BOXES



The Solution Authentication

NFS via RPCSEC_GSS_KRB5



The Solution Authentication

NFS via RPCSEC_GSS_KRB5



BioQuant:
Christian Thiemann
is 3224

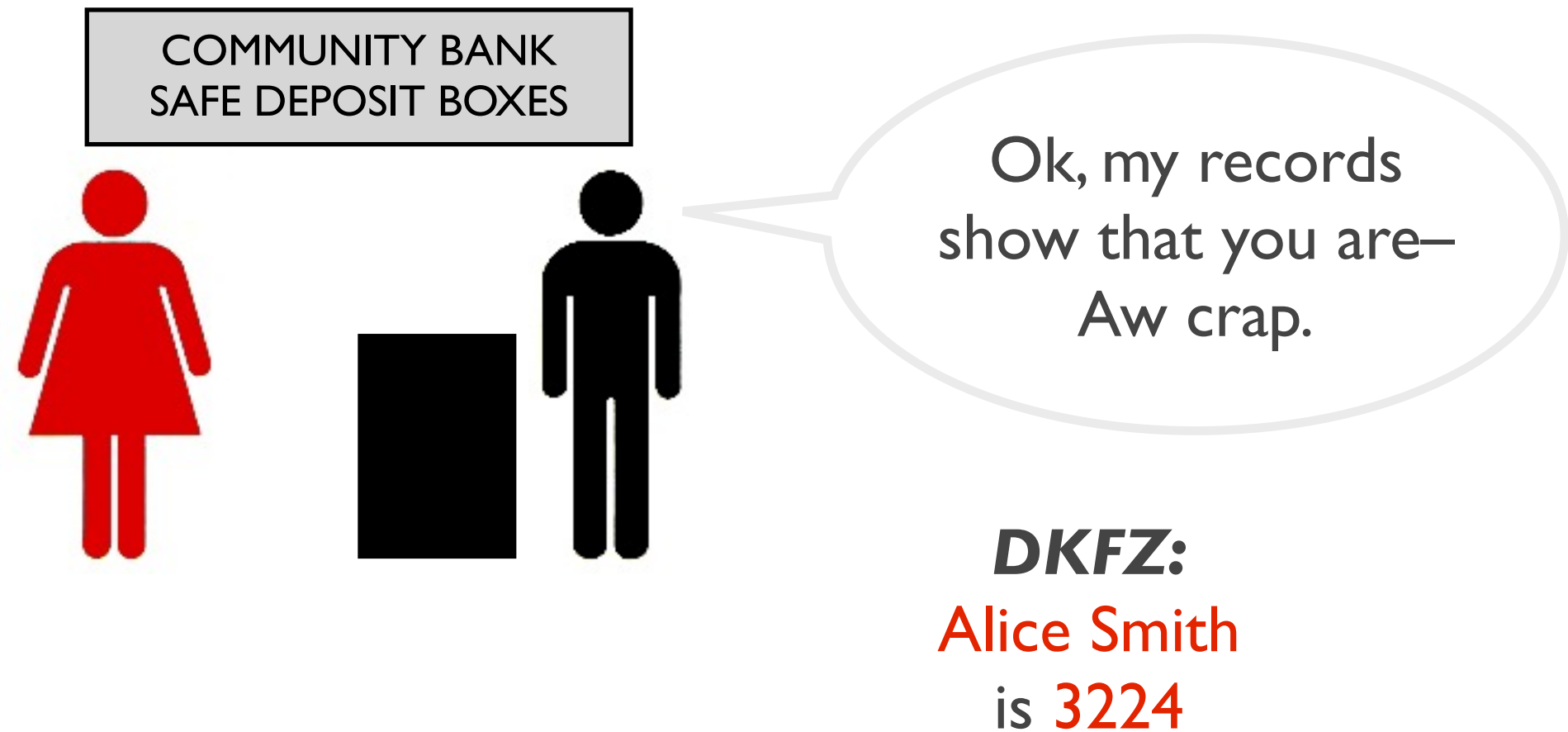
The Problem Differentiation

NFS via RPCSEC_GSS_KRB5 and multiple authorities



The Problem Differentiation

NFS via RPCSEC_GSS_KRB5 and multiple authorities



The Solution Differentiation

NFS via RPCSEC_GSS_KRB5 and multiple authorities and mapped ID spaces

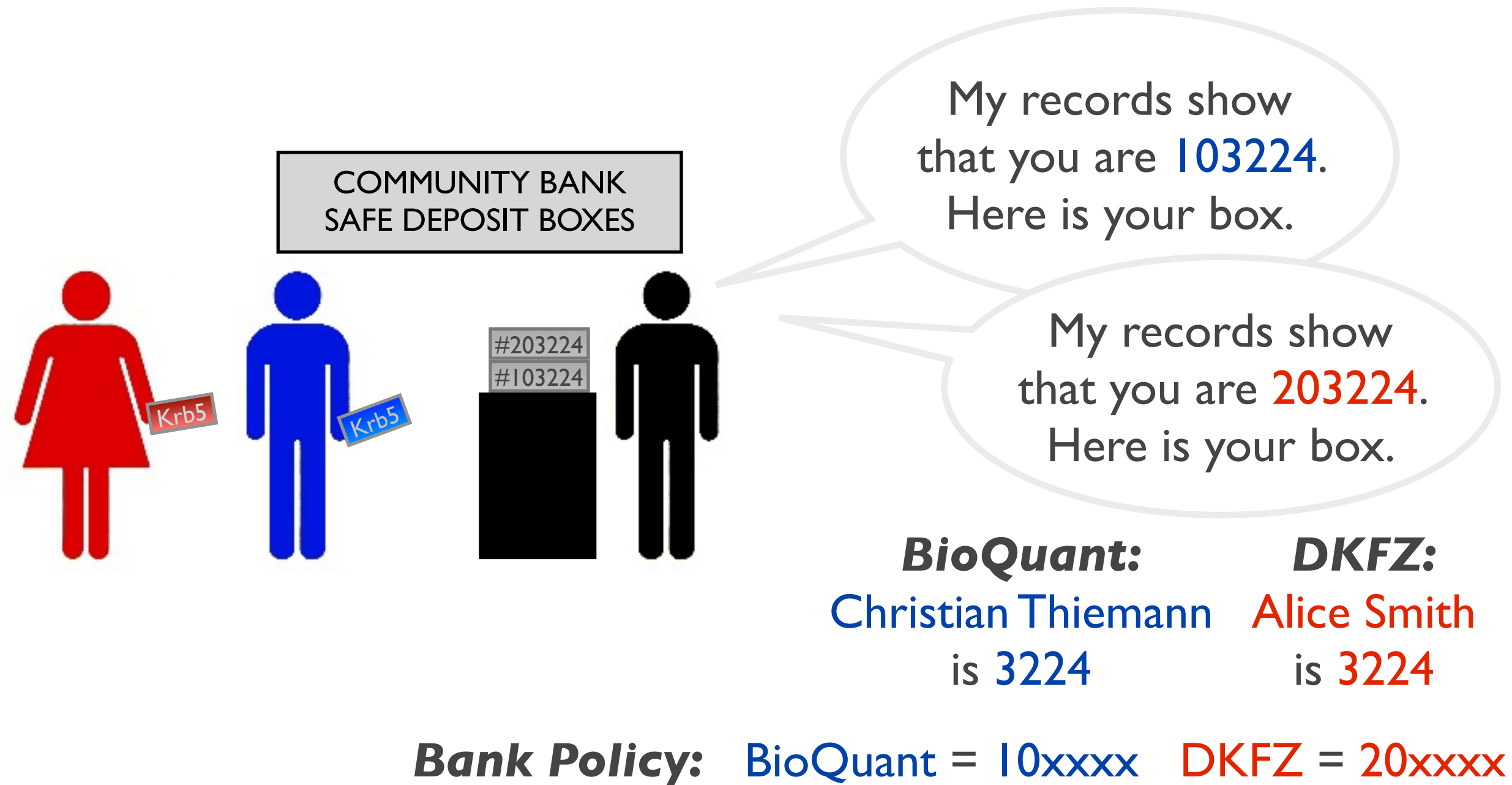


BioQuant:
Christian Thiemann
is 3224

DKFZ:
Alice Smith
is 3224

The Solution Differentiation

NFS via RPCSEC_GSS_KRB5 and multiple authorities and mapped ID spaces



The Problem

NFSv3 uses numeric UID/GIDs

BioQuant Client

LSDF

The Problem

NFSv3 uses numeric UID/GIDs

BioQuant Client

LSDF

Hi, I'm 3224. Who owns myfile.txt?

???

myfile.txt: user 103224 group 101000

```
graph LR; Client[BioQuant Client] -- "Hi, I'm 3224. Who owns myfile.txt?" --> LSDF[LSDF]; LSDF -- "myfile.txt: user 103224 group 101000" --> Client; Client --- Q[???];
```

The Problem

NFSv3 uses numeric UID/GIDs

BioQuant Client

LSDF

Hi, I'm 3224. Who owns myfile.txt?

myfile.txt: user 103224 group 101000

???

chgrp 1801 myfile.txt

What?

sudo chgrp 1801 myfile.txt

That only works on weak minds...
But here, have a sandwich instead.

The Workaround

NFSv3 on-route UID/GID translation

BioQuant Client

NFS Proxy

LSDF

The Workaround

NFSv3 on-route UID/GID translation

BioQuant Client

NFS Proxy

LSDF

Hi, I'm 3224. Who owns myfile.txt?

Hi, I'm 103224. Who owns myfile.txt?

Ok

myfile.txt: user 3224 group 1000

myfile.txt: user 103224 group 101000

The Workaround

NFSv3 on-route UID/GID translation

BioQuant Client

NFS Proxy

LSDF

Hi, I'm 3224. Who owns myfile.txt?

Hi, I'm 103224. Who owns myfile.txt?

myfile.txt: user 103224 group 101000

Ok

myfile.txt: user 3224 group 1000

chgrp 1801 myfile.txt

chgrp 101801 myfile.txt

Ok

sudo chgrp 1801 myfile.txt

sudo chgrp 101801 myfile.txt

Still ok

The Solution

NFSv4 usernames

BioQuant Client

LSDF

The Solution

NFSv4 usernames

BioQuant Client

LSDF

Hi, I'm 3224. Who owns myfile.txt?

Hi, I'm bq_cthiemann@bioquant. Who owns myfile.txt?

myfile.txt: user 103224 group 101000

myfile.txt: user bq_cthiemann@BQ group bq_admins@BQ

Ok

The Solution

NFSv4 usernames

BioQuant Client

LSDF

Hi, I'm 3224. Who owns myfile.txt?

Hi, I'm bq_cthiemann@bioquant. Who owns myfile.txt?

myfile.txt: user 103224 group 101000

myfile.txt: user bq_cthiemann@BQ group bq_admins@BQ

Ok

chgrp 1801 myfile.txt

chgrp rattorturers@BQ myfile.txt

Ok

The Solution

NFSv4 usernames

BioQuant Client

LSDF

Hi, I'm 3224. Who owns myfile.txt?

NFSv4 development is slow:

“no economical incentive” (IBM / EMC)

Hi, I'm bq_cthiemann@bioquant. Who owns myfile.txt?

myfile.txt: user bq_cthiemann@BQ group bq_admins@BQ

Ok

chgrp 1801 myfile.txt

chgrp rattorturers@BQ myfile.txt

Ok

The Solution

NFSv4 usernames

BioQuant Client

LSDF

invisible hand



NFSv4 development is slow:
“no economical incentive” (IBM / EMC)

Hi, I'm 3224. Who owns myfile.txt?

Hi, I'm bq_cthiemann@bioquant. Who owns myfile.txt?

myfile.txt: user bq_cthiemann@BQ group bq_admins@BQ

Ok

chgrp 1801 myfile.txt

chgrp rattorturers@BQ myfile.txt

Ok

The Solution

NFSv4 usernames

BioQuant Client

LSDF
invisible hand



NFSv4 development is slow:
“no economical incentive” (IBM / EMC)

Ok IBM SONAS will be NFSv4-capable in 2014

hopefully...

chgrp 1801 myfile.txt

chgrp rattorturers@BQ myfile.txt

Ok

The Workaround

NFSv3 on-route UID/GID translation

BioQuant Client

NFS Proxy

LSDF

NFSv3 Ticket Agent: Kerberized Inter-Realm NFS Service

Hi, I'm 3224. Who owns myfile.txt?

Ok

chgrp 1801 myfile.txt

chgrp 101801 myfile.txt

Ok

sudo chgrp 1801 myfile.txt

sudo chgrp 101801 myfile.txt

Still ok

The Workaround

NFSv3 on-route UID/GID translation

BioQuant Client

NFS Proxy

LSDF

NFSv3 Ticket Agent: Kerberized Inter-Realm NFS Service

acts as a trusted client to the LSDF (using AUTH_SYS)

The Workaround

NFSv3 on-route UID/GID translation

BioQuant Client

NFS Proxy

LSDF

NFSv3 Ticket Agent: Kerberized Inter-Realm NFS Service

acts as a trusted client to the LSDF (using AUTH_SYS)

authenticates external clients using RPCSEC_GSS_KRB5

The Workaround

NFSv3 on-route UID/GID translation

BioQuant Client NFS Proxy LSDF

NFSv3 Ticket Agent: Kerberized Inter-Realm NFS Service

acts as a trusted client to the LSDF (using AUTH_SYS)
authenticates external clients using RPCSEC_GSS_KRB5
transparently translates UIDs & GIDs between LSDF and external ID spaces

The Workaround

NFSv3 on-route UID/GID translation

BioQuant Client

NFS Proxy

LSDF

NFSv3 Ticket Agent: Kerberized Inter-Realm NFS Service

Hi, I'm 3224. Who owns myfile.txt?

Ok

chgrp 1801 myfile.txt

chgrp 101801 myfile.txt

Ok

sudo chgrp 1801 myfile.txt

sudo chgrp 101801 myfile.txt

Still ok

The Workaround

NFSv3 on-route UID/GID translation

BioQuant Client

NFS Proxy

LSDF

NFSv3 Ticket Agent: Kerberized Inter-Realm NFS Service

user-space daemon intercepting TCP traffic

The Workaround

NFSv3 on-route UID/GID translation

BioQuant Client

NFS Proxy

LSDF

NFSv3 Ticket Agent: Kerberized Inter-Realm NFS Service

user-space daemon intercepting TCP traffic

slows down filesystem operations (directory listings etc.) factor 2–3

sudo chgrp 1801 myfile.txt

sudo chgrp 101801 myfile.txt

Still ok

The Workaround

NFSv3 on-route UID/GID translation

BioQuant Client NFS Proxy LSDF

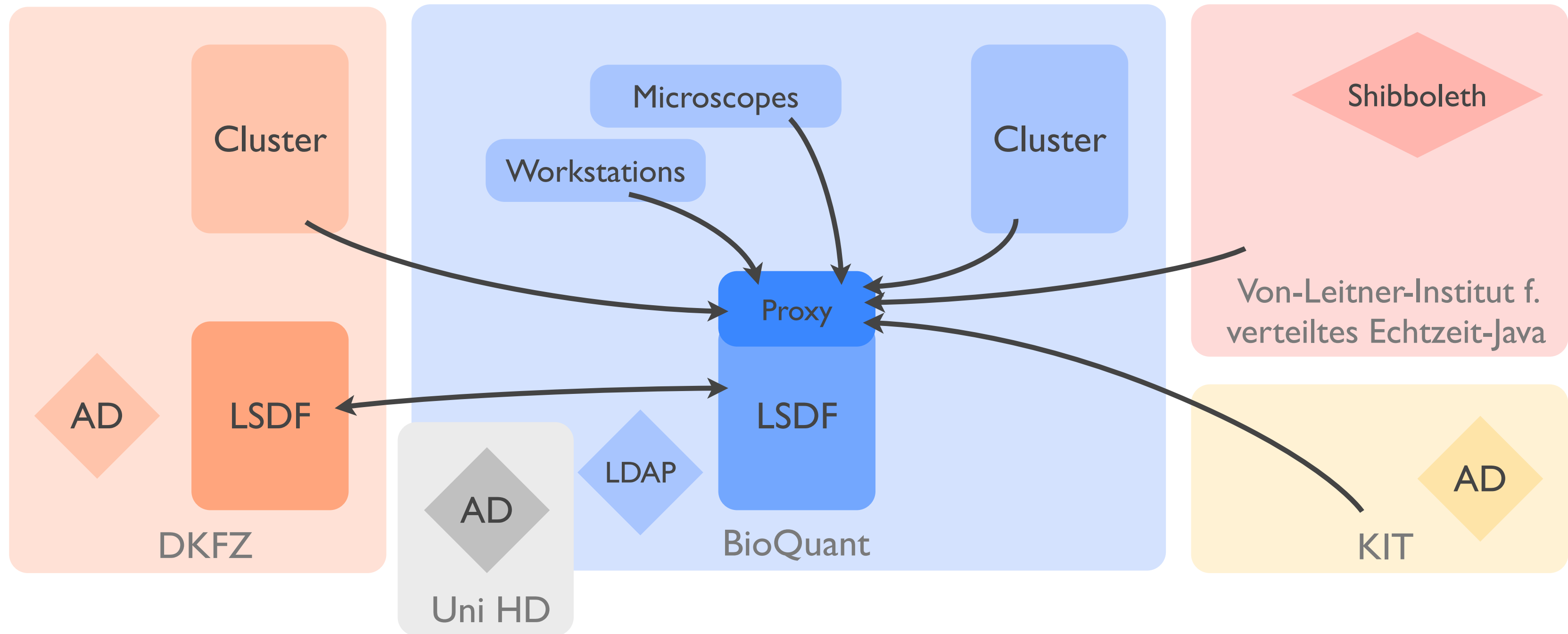
NFSv3 Ticket Agent: Kerberized Inter-Realm NFS Service

user-space daemon intercepting TCP traffic

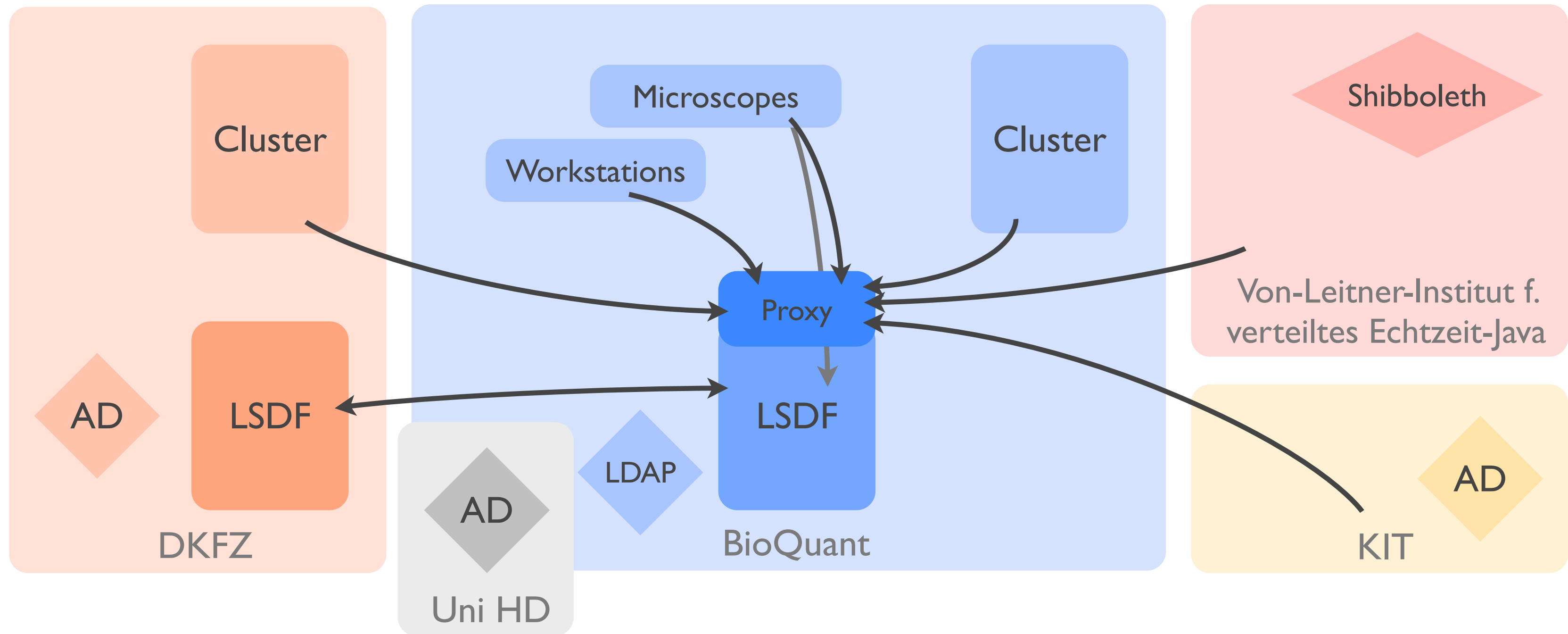
slows down filesystem operations (directory listings etc.) factor 2–3

no throughput penalty on read/write operations 10–20 GBit/s

BioQuant LSDF



BioQuant LSDF



BioQuant LSDF

