International Industry-Academia Workshop on Cloud Reliability and Resilience

Berlin, 7-8 November 2016

Dependable Storage and Computing using Multiple Cloud Providers

Alysson Bessani

Ciências ULisboa | Informática

www.di.fc.ul.pt/~bessani

Abstract & Bio

Despite the widespread success of public cloud services, issues such as outages, data loss, and vendor lock-in have prevented many organizations to embrace the cloud. An intuitive way to avoid these problems would be to replicate data and services in different clouds, avoiding thus the need to completely trust any single cloud provider. In this talk I'll discuss the benefits and challenges of this approach, with special focus on the practical insights gained from building several multi-cloud storage systems.

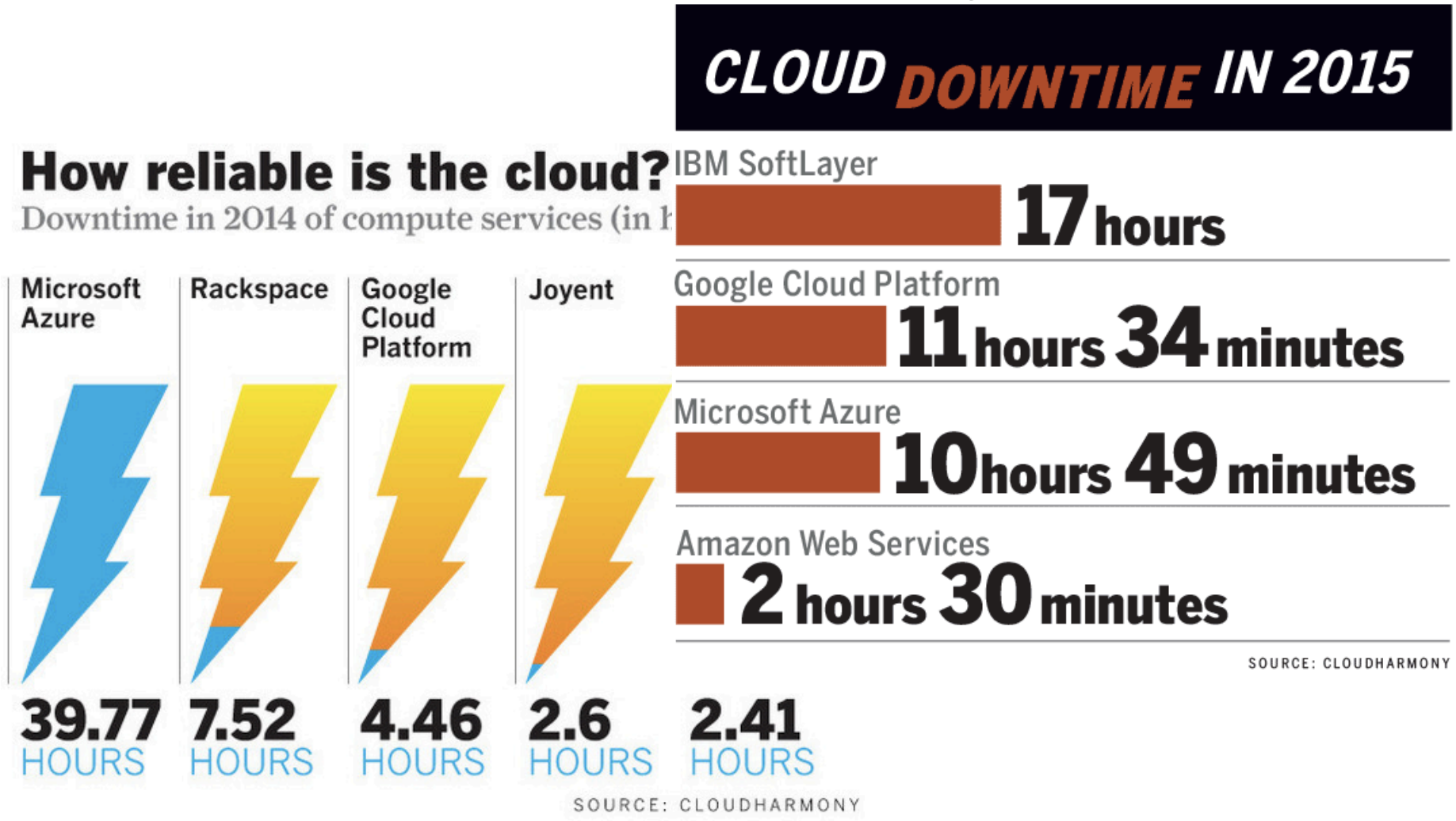
Alysson Bessani is an Associate Professor of the Faculty of Sciences of the University of Lisboa, Portugal, and a member of LaSIGE research unit. He holds a PhD in Electrical Engineering from Federal University of Santa Catarina, Brazil (2006), was a visiting professor at Carnegie Mellow University (2010), and a visiting researcher at Microsoft Research Cambridge (2014). Alysson co-authored more than 100 papers, participated in several FP7 and H2020 projects and is coordinating the DiSIEM H2020 project. His current interests lie in distributed systems, Byzantine fault tolerance, cloud storage, blockchain design and applications, and security monitoring. More information about him can be found at http://www.di.fc.ul.pt/~bessani.

Cloud Sec. & Dep.

- Secure and dependable services are a necessary condition for the long-term existence of a cloud provider
  - Clients need to trust providers
  - Providers need to justify this trust

- To stay in the market, a provider needs to invest on Sec. & Dep., however…

Unavailability



Number of nines

99,9% = 8,76 hours 99,99% = 53 minutes

Unavailability

www.crn.com/slide-shows/cloud/300081477/the-10-biggest-cloud-outages-of-2016-so-far.htm

The 10 Biggest Cloud Outages Of 2016 (So Far)

by **Joseph Tsidulko** on July 27, 2016, 3:04 pm EDT

Email this CRN article

1 2 3 4 5 6 7 8 9 10 ... next ›

Cloud Outages Are Business Outages

Every business is becoming a software business, the maxim goes, which means cloud outages translate to business outages.

As enterprises migrate more mission-critical workloads into production cloud environments, mere minutes of downtime from a provider can significantly impact profits, damage relations with customers, and cause IT administrators to prematurely age.

But while the global economy increasingly hinges on the ability of cloud services providers -- especially those operating at hyper-scale proportions -- to guarantee uptime and maintain service, outages are still common.

The causes can range from power outages to faulty software updates to overloaded servers to database errors. And far too often, we never learn the true nature and scope of the service failure.

Here are some of the cloud outages that have grabbed headlines in the first half of 2016.

Data Loss

2011 **Amazon's Cloud Crash Disaster Permanently Destroyed Many Customers' Data**

2015

NEWS ANALYSIS

OOPS: Google "loses" your cloud data (sky falling; film at 11)

Lightning fell from sky -- from cloud to cloud, as't were

2016

The Register®

Biting the hand that feeds IT

DATA CENTRE SOFTWARE SECURITY TRANSFORMATION DEVOPS BUSINESS PERSONAL TECH SCIENCE EMERGENT TECH BOOTNOTES

Data Centre ▸ Cloud

Salesforce.com crash caused DATA LOSS

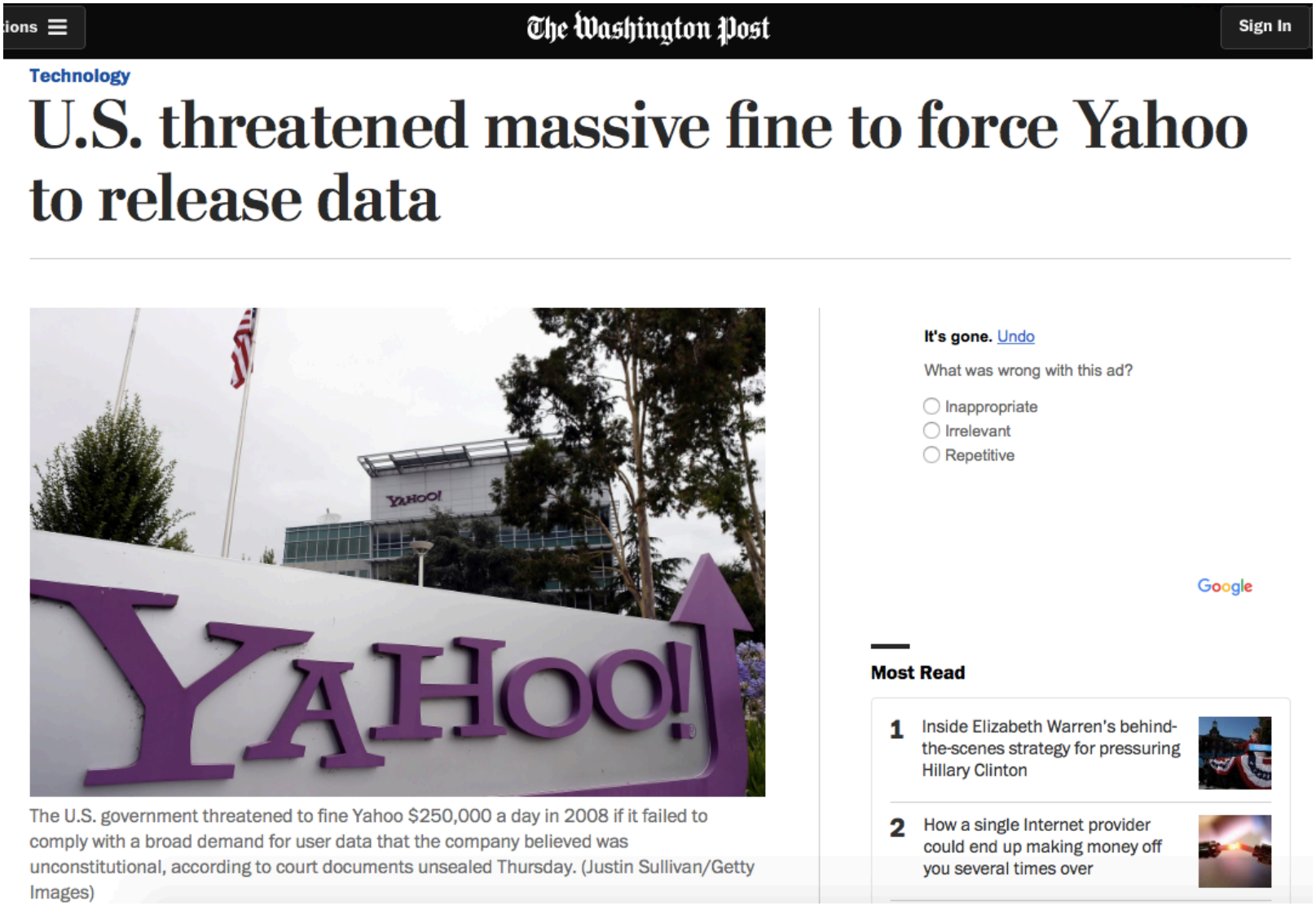
Three-and-half-hours of data has evaporated. Maybe forever

13 May 2016 at 00:00, Simon Sharwood 493

More like this

Outages

Privacy Issues…



ions ☰ The Washington Post Sign In

Technology

U.S. threatened massive fine to force Yahoo to release data

The U.S. government threatened to fine Yahoo $250,000 a day in 2008 if it failed to comply with a broad demand for user data that the company believed was unconstitutional, according to court documents unsealed Thursday. (Justin Sullivan/Getty Images)

It's gone. Undo

What was wrong with this ad?

- Inappropriate
- Irrelevant
- Repetitive

Google

Most Read

1 Inside Elizabeth Warren's behind-the-scenes strategy for pressuring Hillary Clinton

2 How a single Internet provider could end up making money off you several times over

InfoWorld FROM IDG

The dirty dozen: 12 cloud security threats

Introducing the 'Treacherous 12,' the top security threats organizatio when using cloud services

MORE

MORE LIKE THIS

The mc damagi

9 top threats to clo security

How to Service

on IDG Answers

What is 'Google do security threat?

Ciências ULisboa

Navigators

Important!

To the best of my knowledge, there is no work on the existing technical literature saying that using the cloud is less secure than having everything on premises

Cloud services work pretty well…

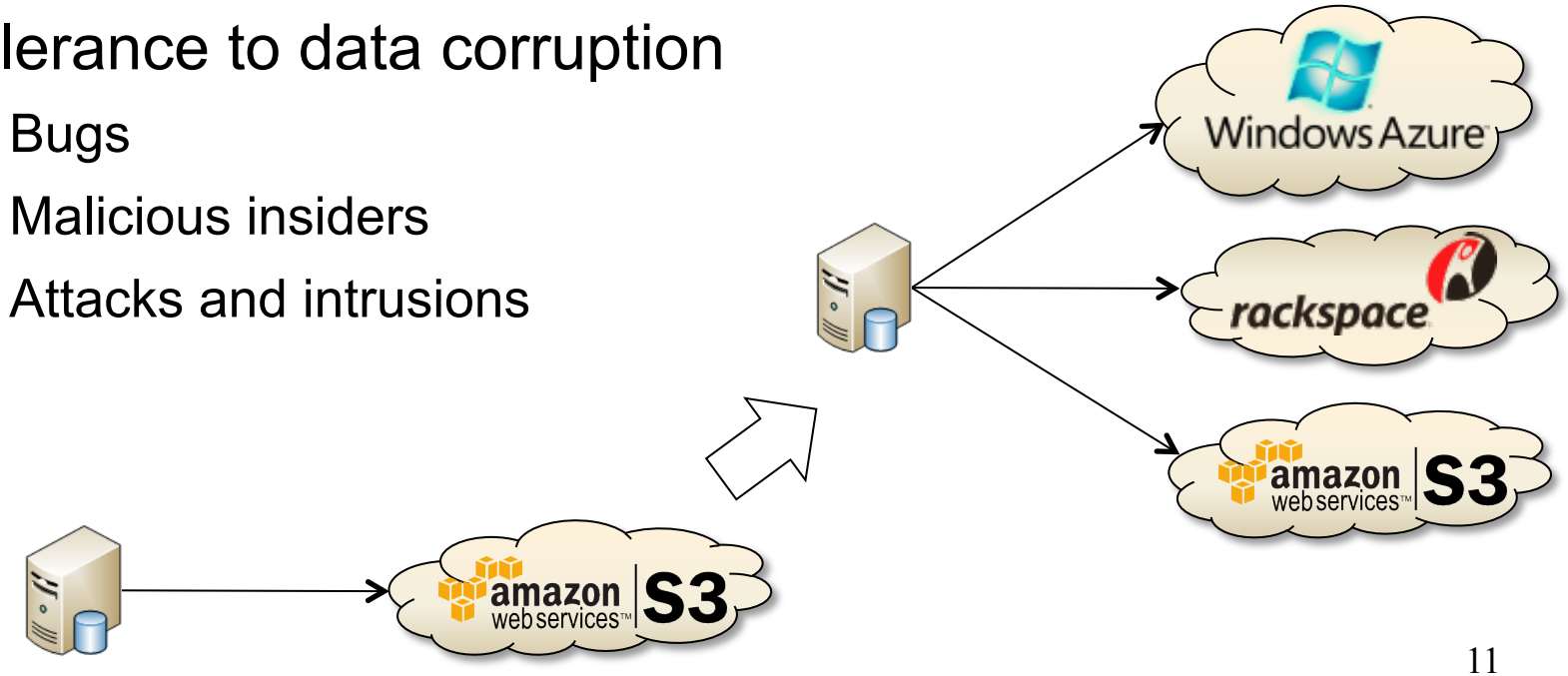
but they are not perfect!

Our view

- The best way to use the cloud in a dependable and secure way is to **not rely on a single provider**
  - Don't base your service on a single provider, i.e., *an application works if the provider where it is running is correct*
  - Instead, consider **distributed trust**: *an application is correct if at most f out-of n services/providers are faulty*

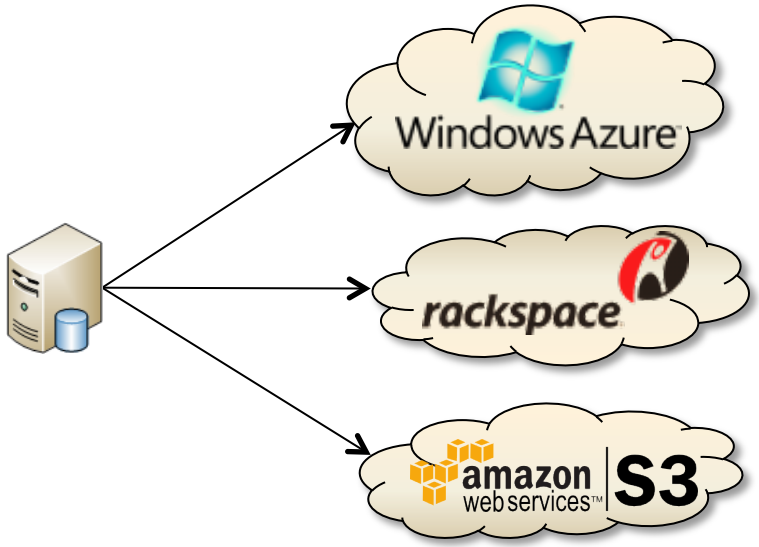
Why a Cloud of Clouds?

- Survive datacenter and cloud outages
- Avoid vendor lock-in
- Achieve better read performance
- Tolerance to data corruption
  - Bugs
  - Malicious insiders
  - Attacks and intrusions



Cloud-of-clouds Service

- Two fundamental characteristics
  1. No modification on existing cloud services
  2. No collaboration between providers



Cloud-of-Clouds Storage

Cloud Storage Diversity

- 24+ cloud object storage services
  - Locations: 68
  - Model: Standard - Archival
  - Cost per GB/month stored: $0,001 - $0,13
  - Cost per GB downloaded: $0,01 - $0,25
  - Availability SLA: undefined, 99,9%-100%
  - Access control: inexistent, URLs, …, ACLs,
  - Standards compliance: 0-33+

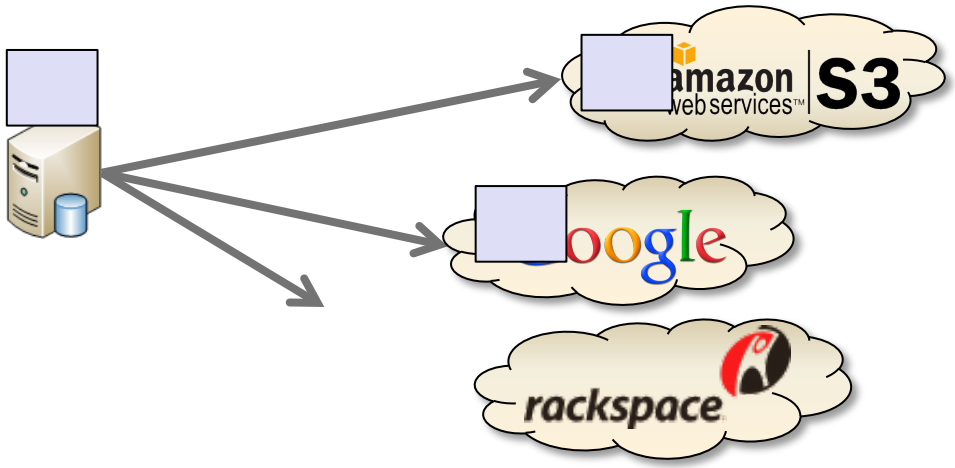
Fundamentals

- Let's consider the problem of implementing a cloud-of-clouds object storage tolerating **one** cloud failure
- Assume you cannot know for sure if a unresponsive cloud is faulty or not

Writing Data

One needs to write at least two copies of the data!

What if one of the clouds take too much to acknowledge the write?



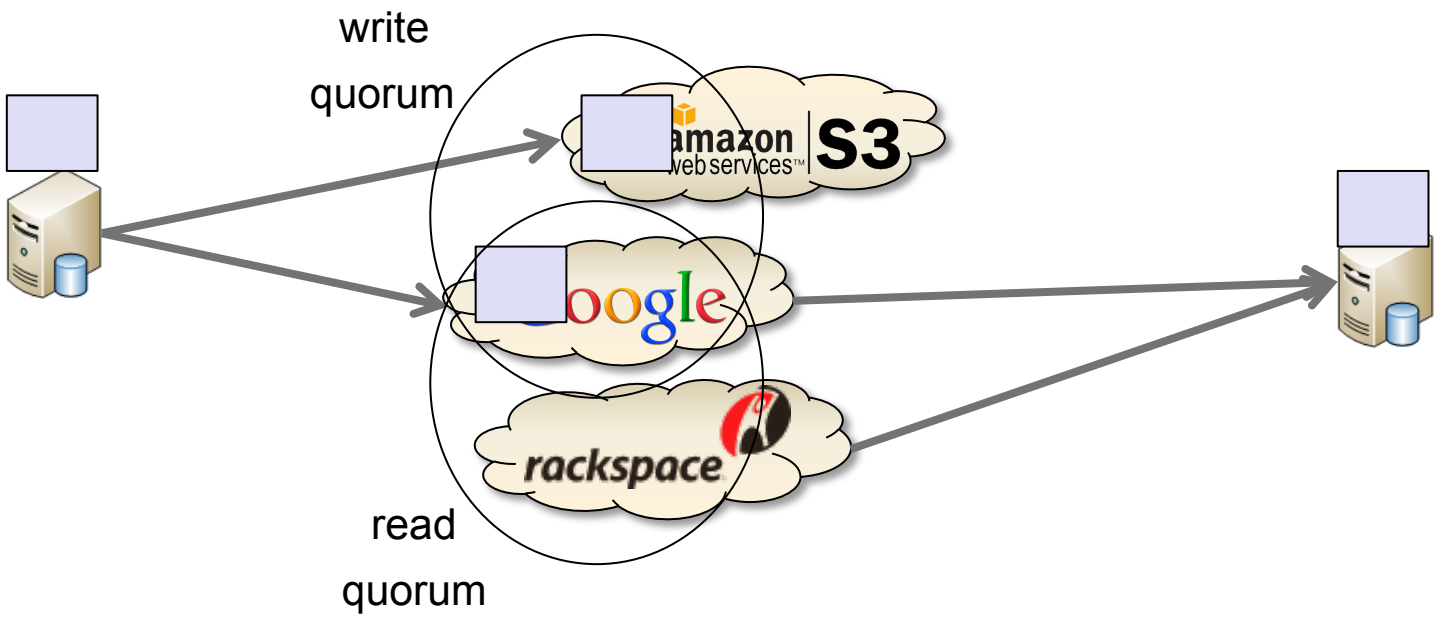
Use **three clouds**, and wait for at least two acks to complete the write

Reading Data

How many clouds you need to access for reading?

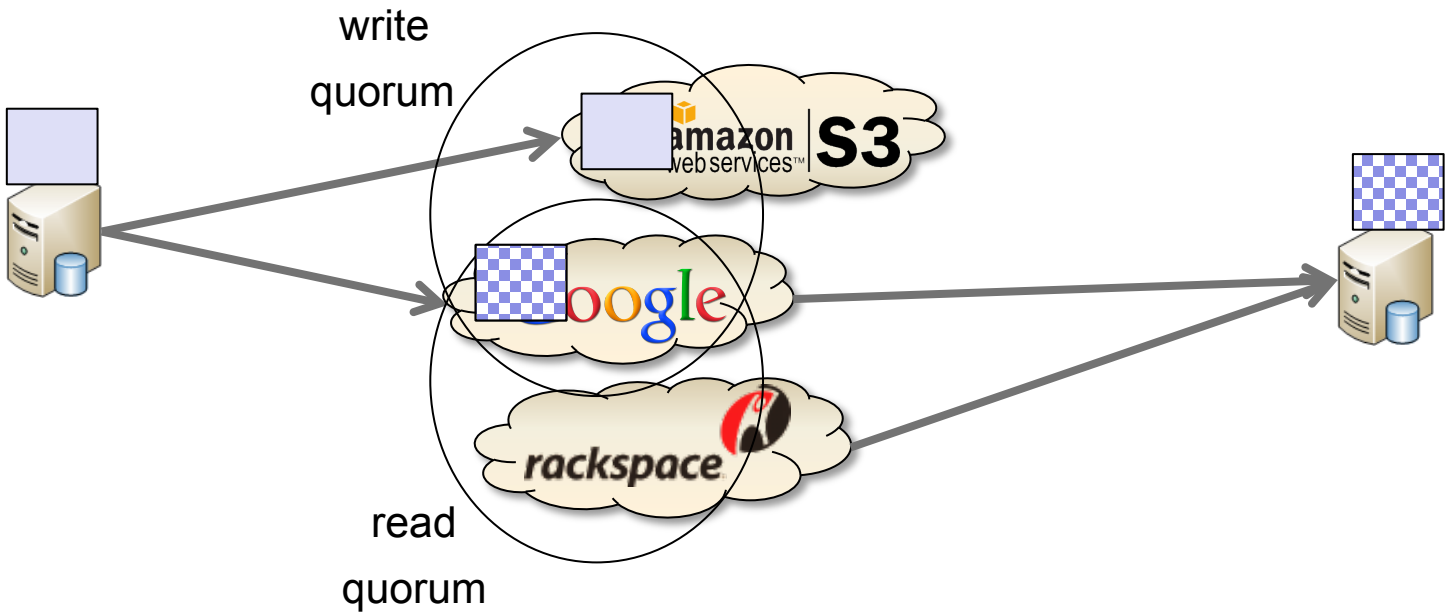
One cloud: might access an outdated/empty cloud

Two clouds: always access the last complete write



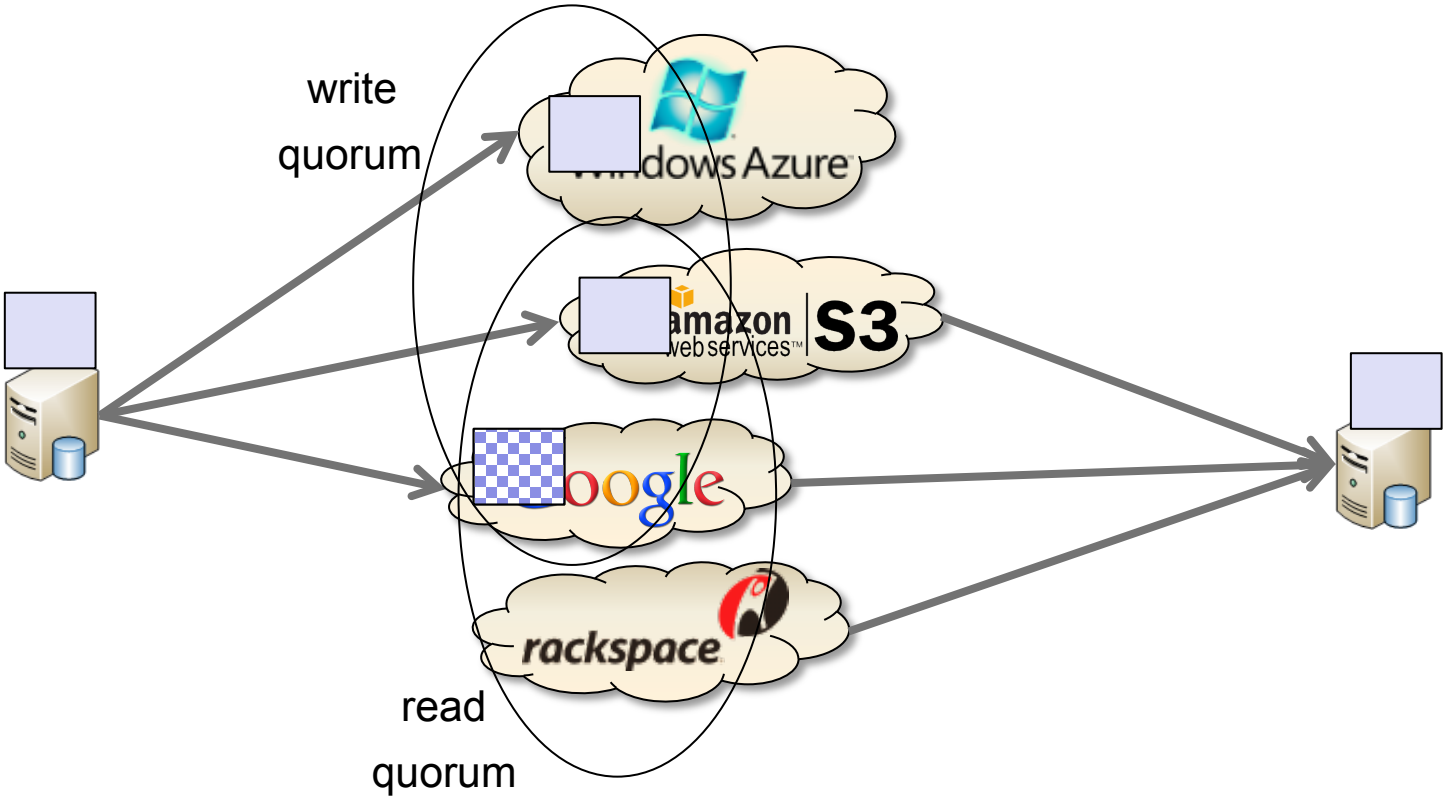
Reading Data II

What if data can be **modified or corrupted**?



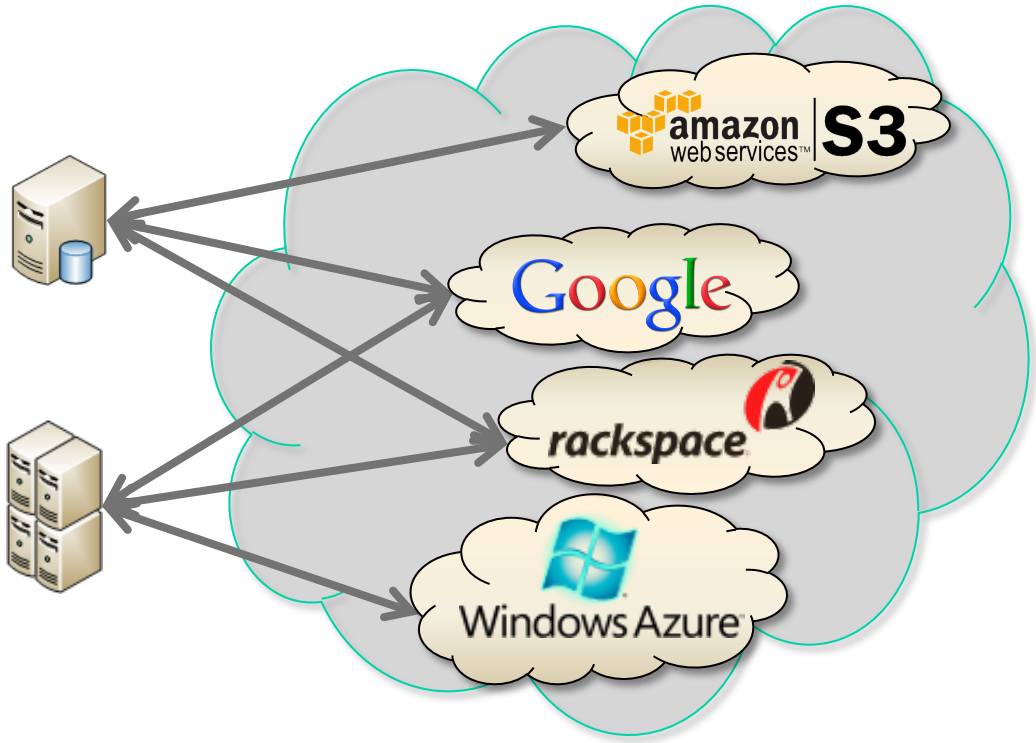
Reading Data II

To solve this we need more clouds and bigger quorums



DepSky: Dependable Cloud-of-Clouds Object Storage

[*EuroSys'11, ACM Trans. on Storage 2013*]

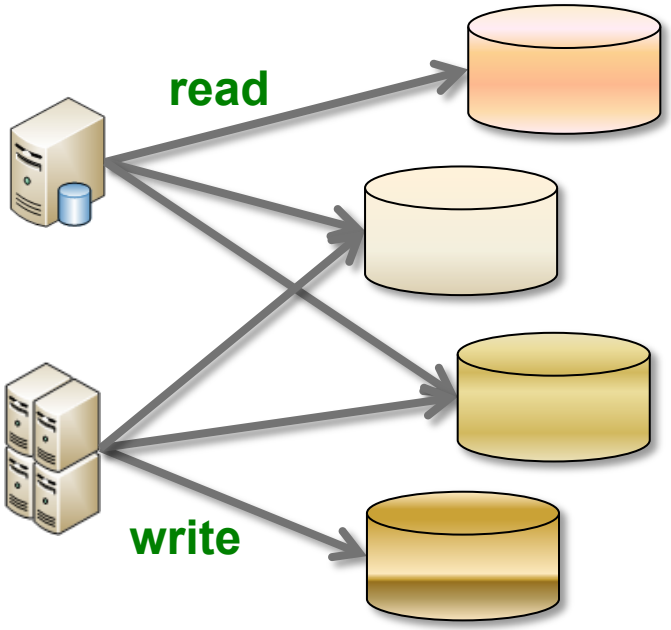


Availability, **Integrity** and **Confidentiality**

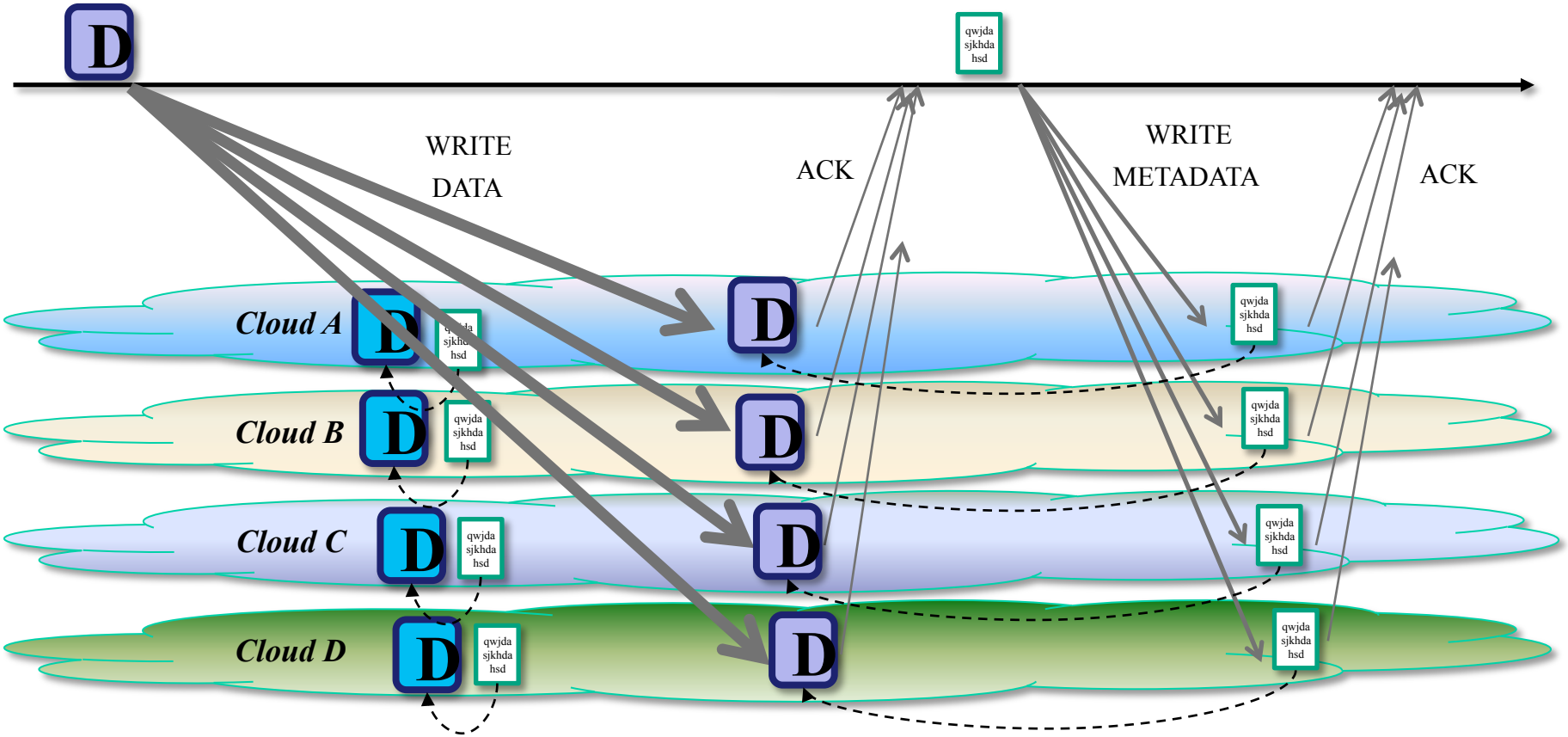
despite the failure of up to *f* clouds

Challenges for implementing Updatable Objects

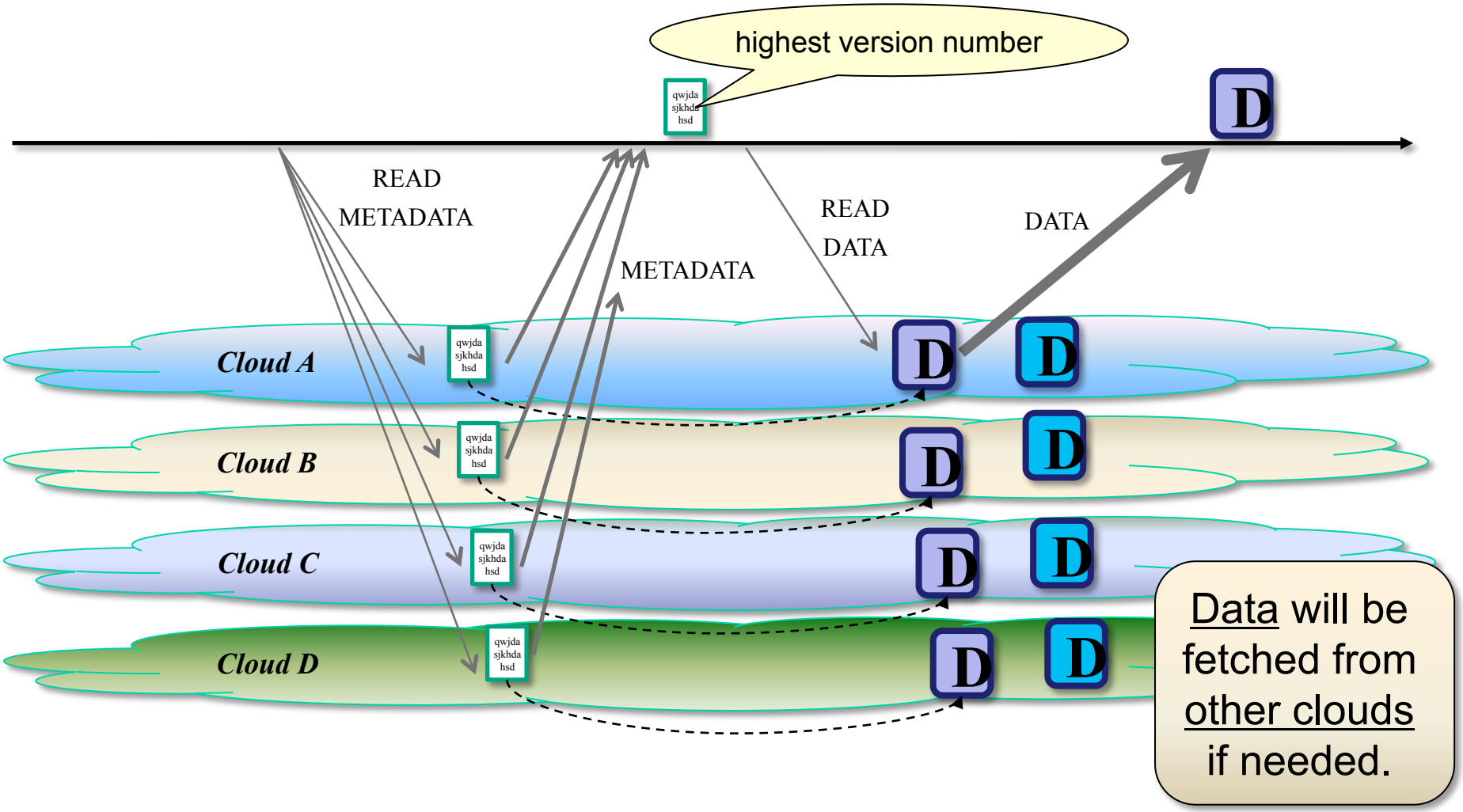
- How to implement an efficient replication protocol using only passive storage nodes?
- How to make it affordable?



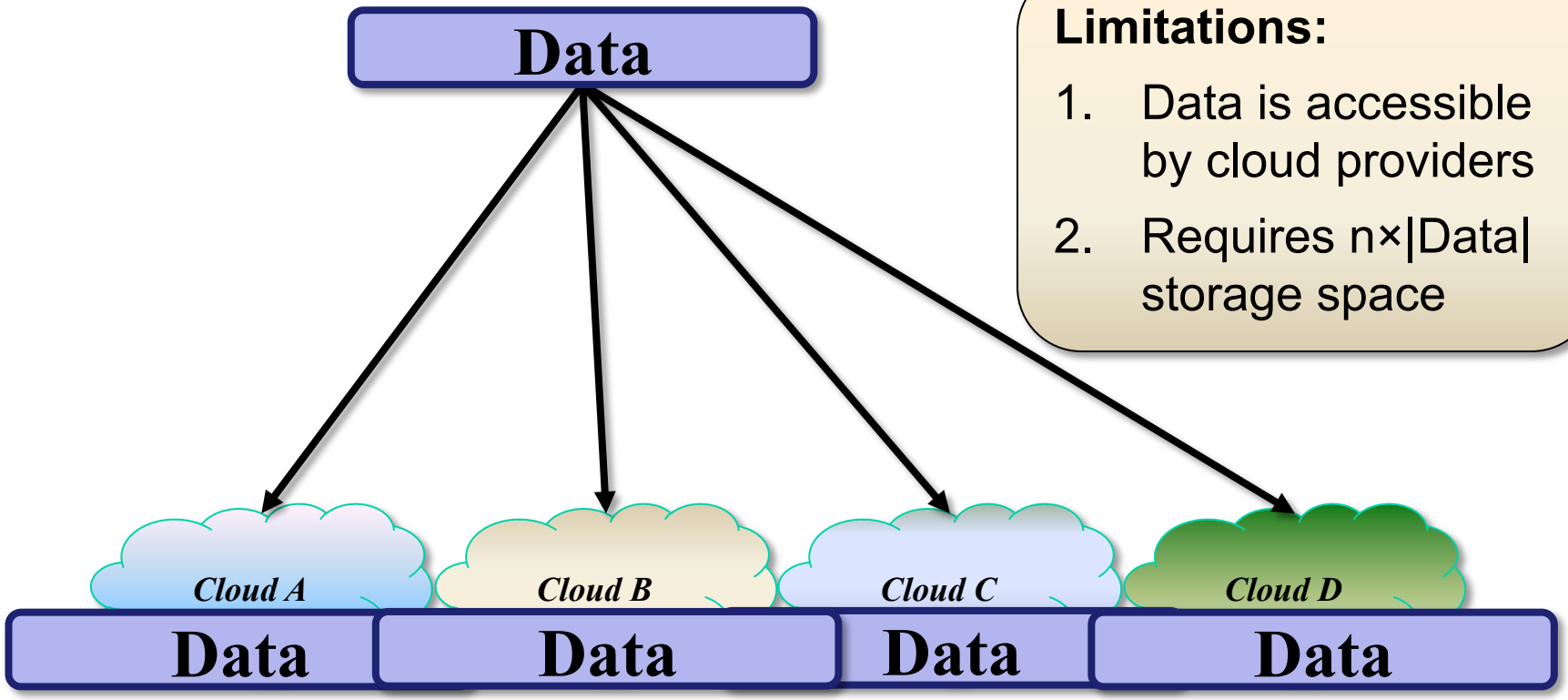
DepSky Write



DepSky Read


highest version number
qwjda sjkhda hsd
D
READ
METADATA
READ
DATA
METADATA
DATA
Cloud A
Cloud B
Cloud C
Cloud D
Data will be fetched from other clouds if needed.

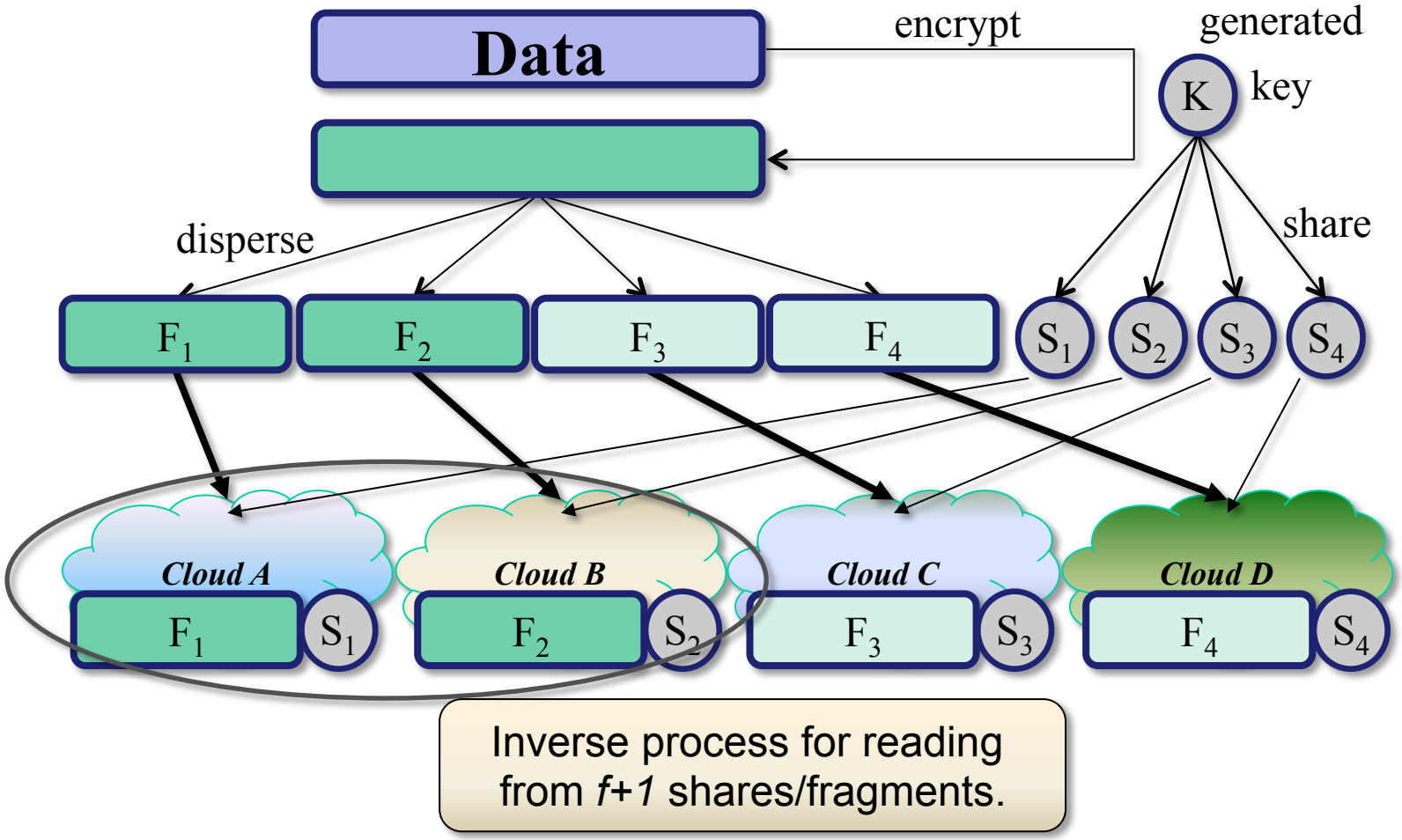
DepSky Confidentiality and Storage Efficiency



Limitations:

1. Data is accessible by cloud providers
2. Requires n×|Data| storage space

DepSky Confidentiality and Storage Efficiency



Ciências ULisboa

Navigators

Consistency Proportionality

- The consistency provided by **DepSky** is the same as the base storage clouds
  - If the weakest consistency cloud provides eventual consistency, DepSky provides eventual consistency
  - If the weakest consistency cloud provides “read your writes”, DepSky provides “read your writes”
  - If the weakest consistency cloud provides regular storage, DepSky provides regular storage
- This notion may be useful for other systems

Ciências ULisboa

Navigators

Practical Considerations

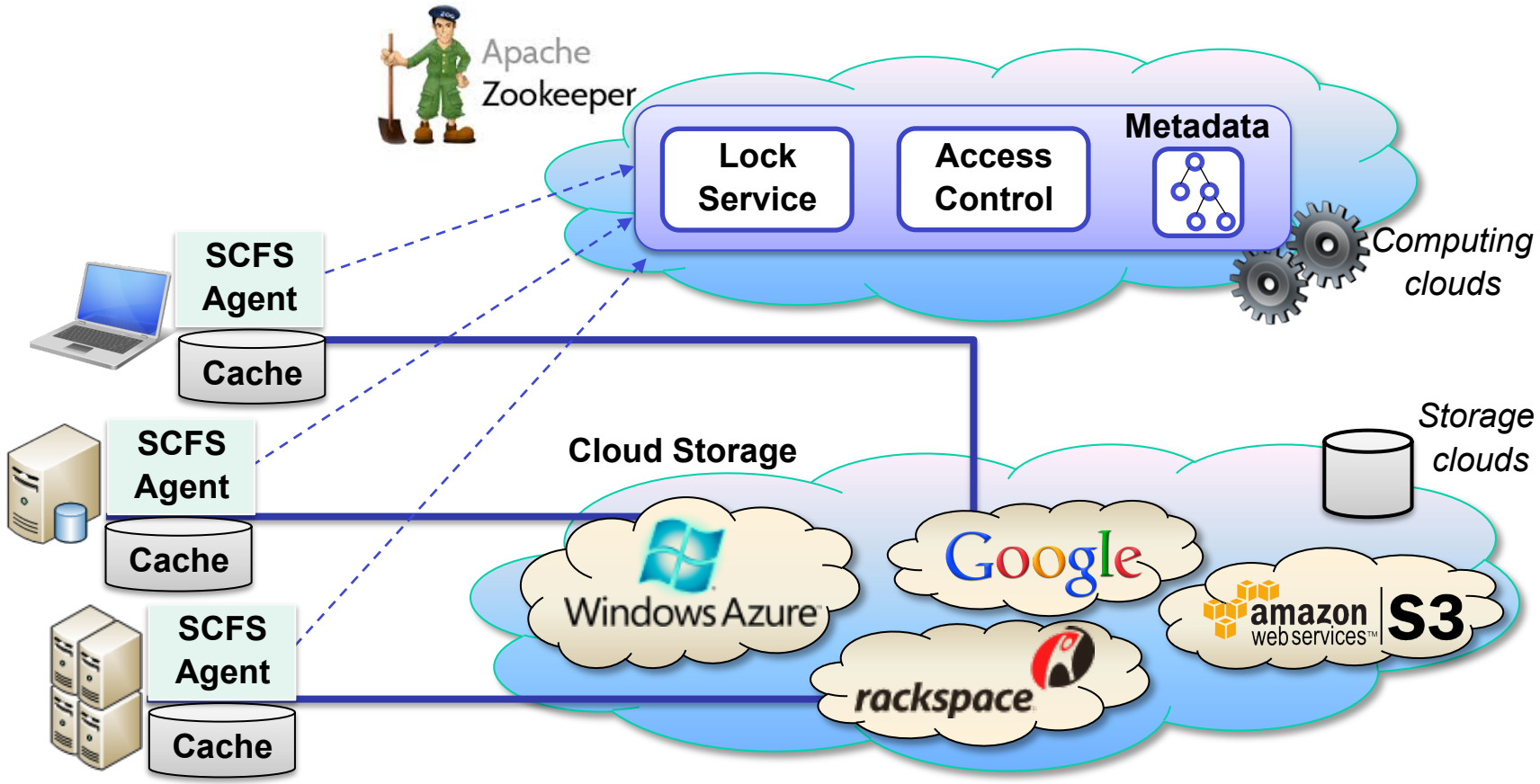
- Read/write latency ≈ accessing a single cloud
- Storage costs roughly 50% higher
- DepSky is a Java programming library

  http://cloud-of-clouds.github.io/depsky
- It does not support concurrent writers to the same object without using locks
  - Recent extensions for multi-writer storage (2016)

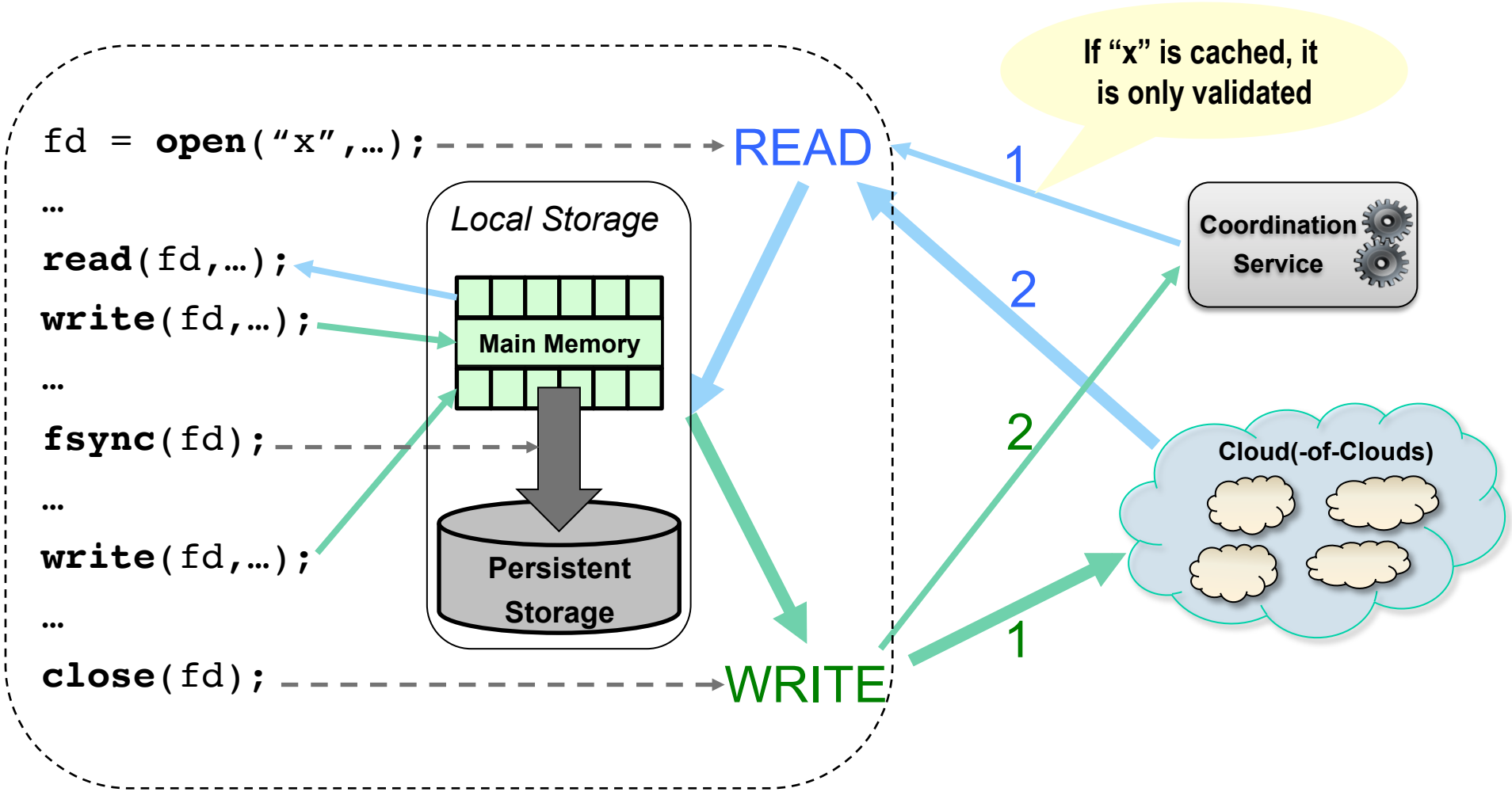
    http://github.com/cloud-of-clouds/mwmr-registers
- **How to build a complete system around it?**

SCFS: Shared Cloud-backed File System

[*USENIX ATC 2014*]

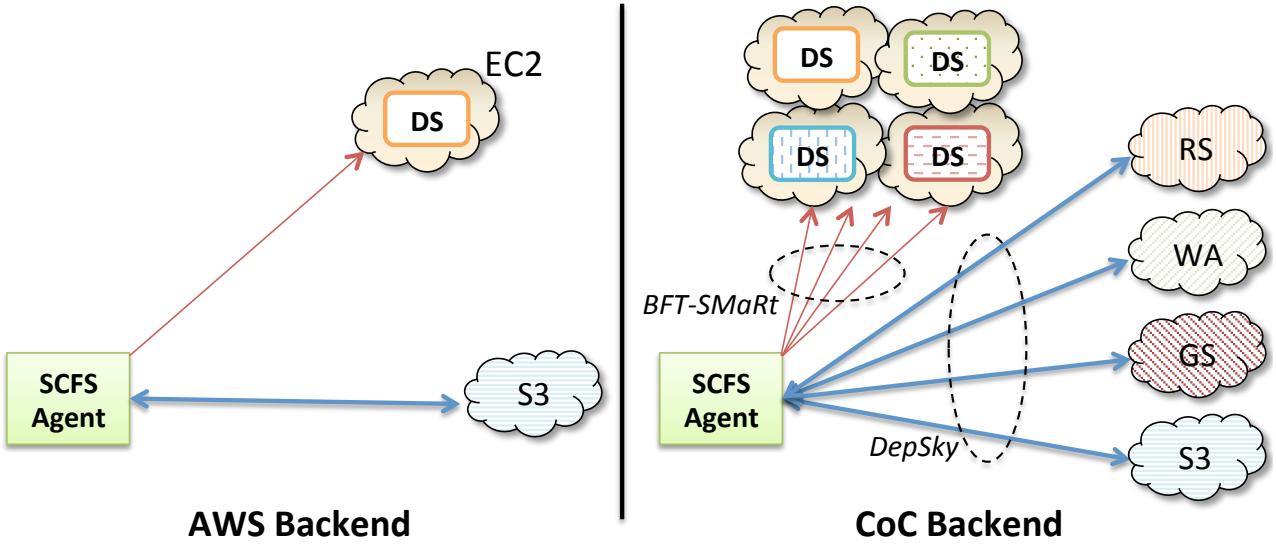


SCFS Consistency on Close



SCFS Backends

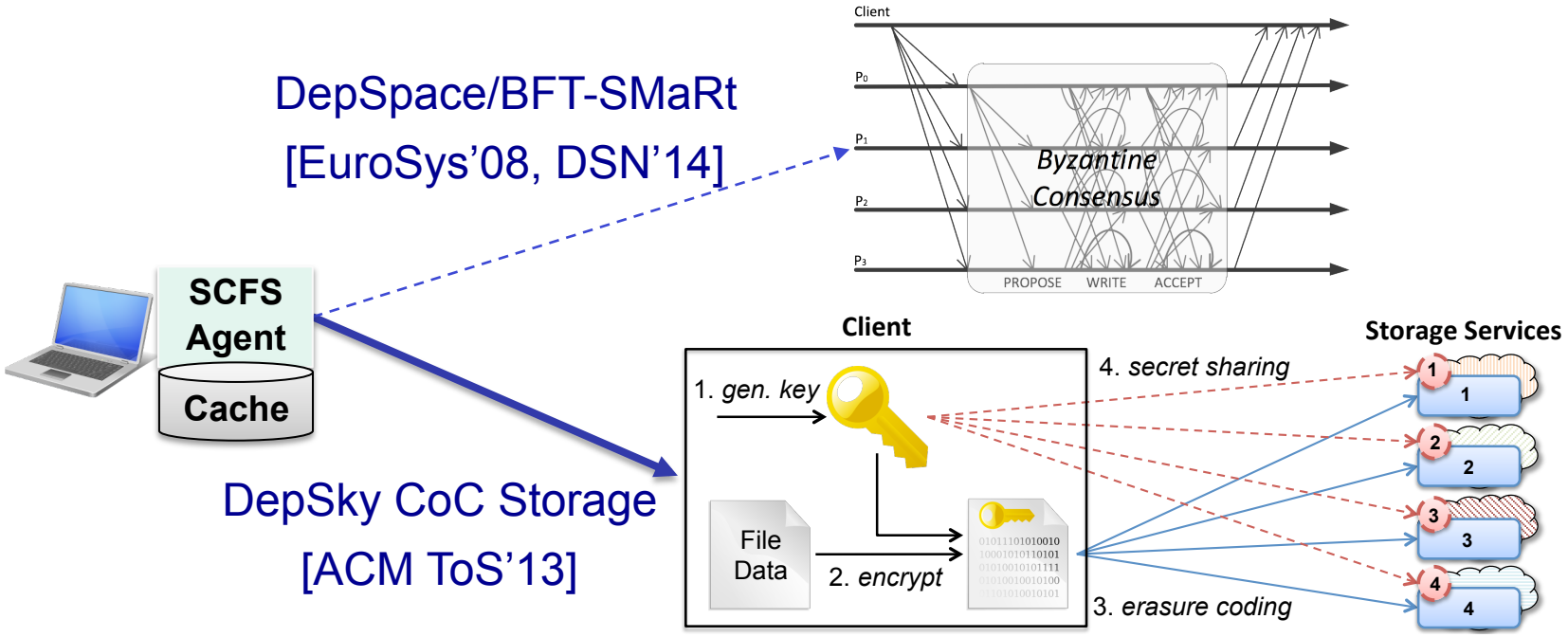
- SCFS can use different backends
  - i.e., different cloud storage and a coordination service plugin



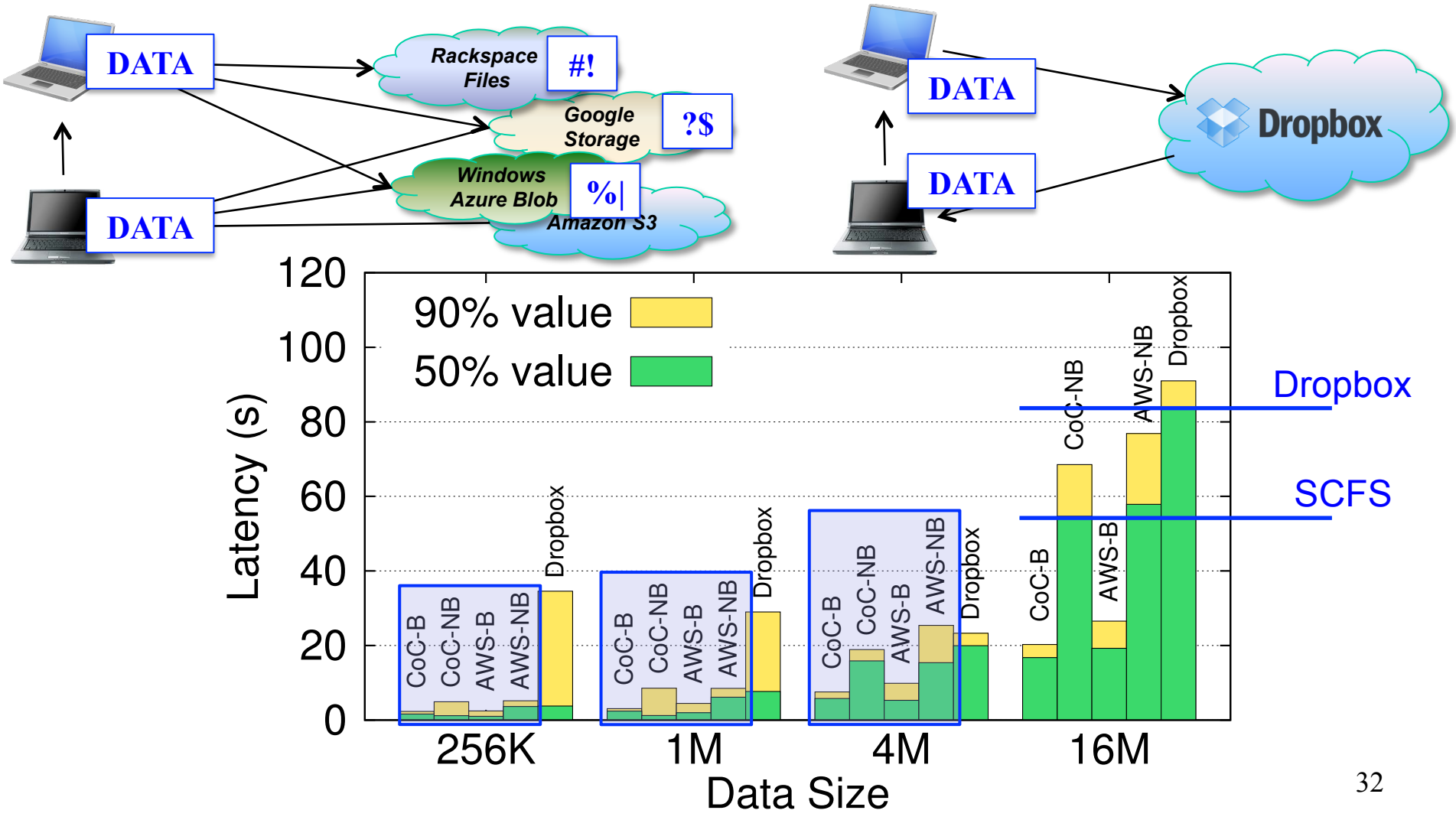
- Operation: **blocking**, non-blocking and non-sharing

The Cloud-of-Clouds Backend

- Does not require trust on any single cloud provider
  - SCFS works correctly as long as less than a third of the providers misbehave



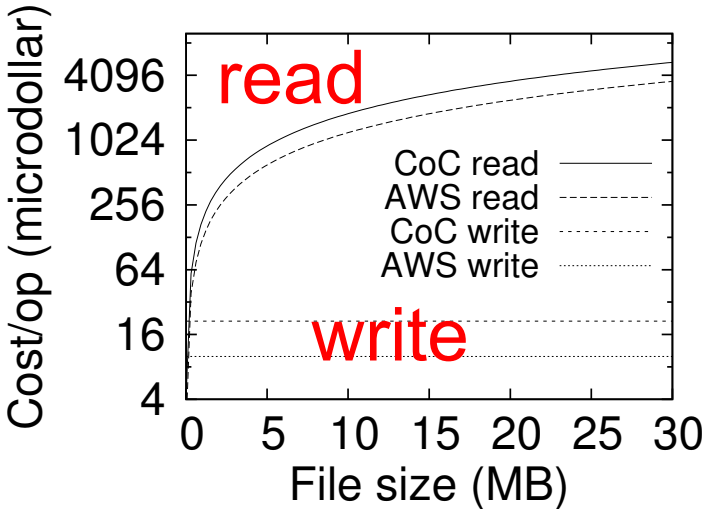
Sharing Latency: SCFS vs DropBox



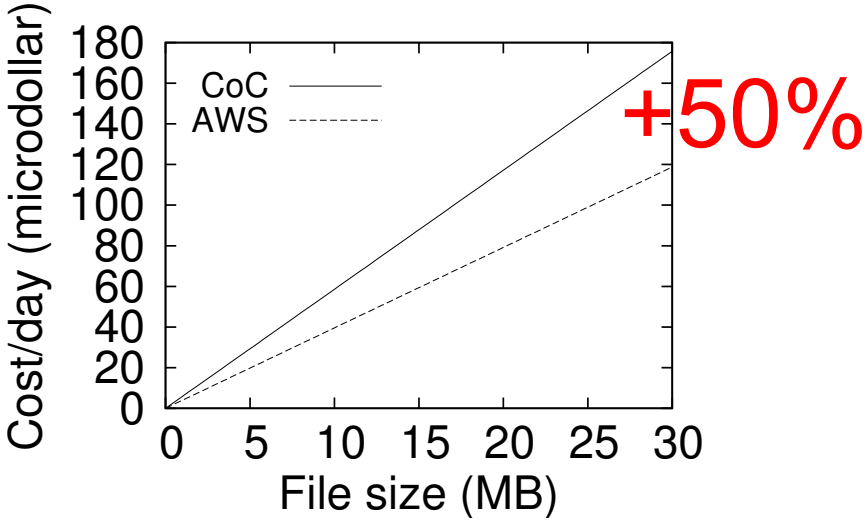
Financial Evaluation

VM Instance	EC2	EC2×4	CoC	Capacity
Large	$6.24	$24.96	$39.60	7M files
Extra Large	$12.96	$51.84	$77.04	15M files

(a) Operation costs/day and expected coordination service capacity.



(b) Cost per operation (log scale).



(c) Cost per file per day.

Ciências ULisboa

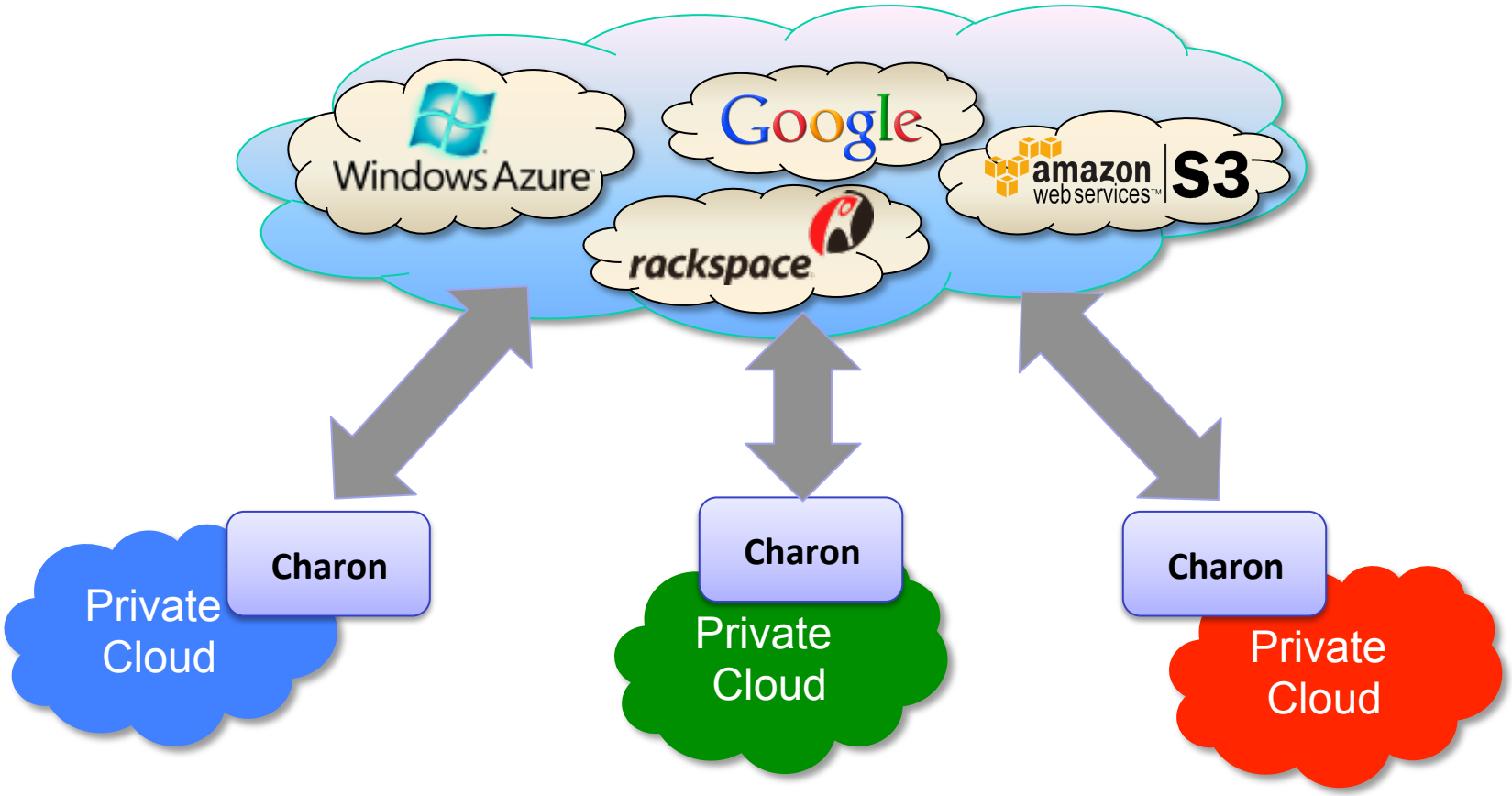
Navigators

Practical Considerations

- SCFS improves the transparency of cloud-of-clouds(-backed) storage
- It is implemented as a Linux FUSE FS:

  http://cloud-of-clouds.github.io/SCFS
- Limitations:
  - Does not work well with big files
  - Require computing instances to run the coordination service (e.g., Zookeeper replicas)
- **Can we do better?**

Charon Cloud-backed File System



+ big data storage and sharing

+multiple storage locations

+ serverless design

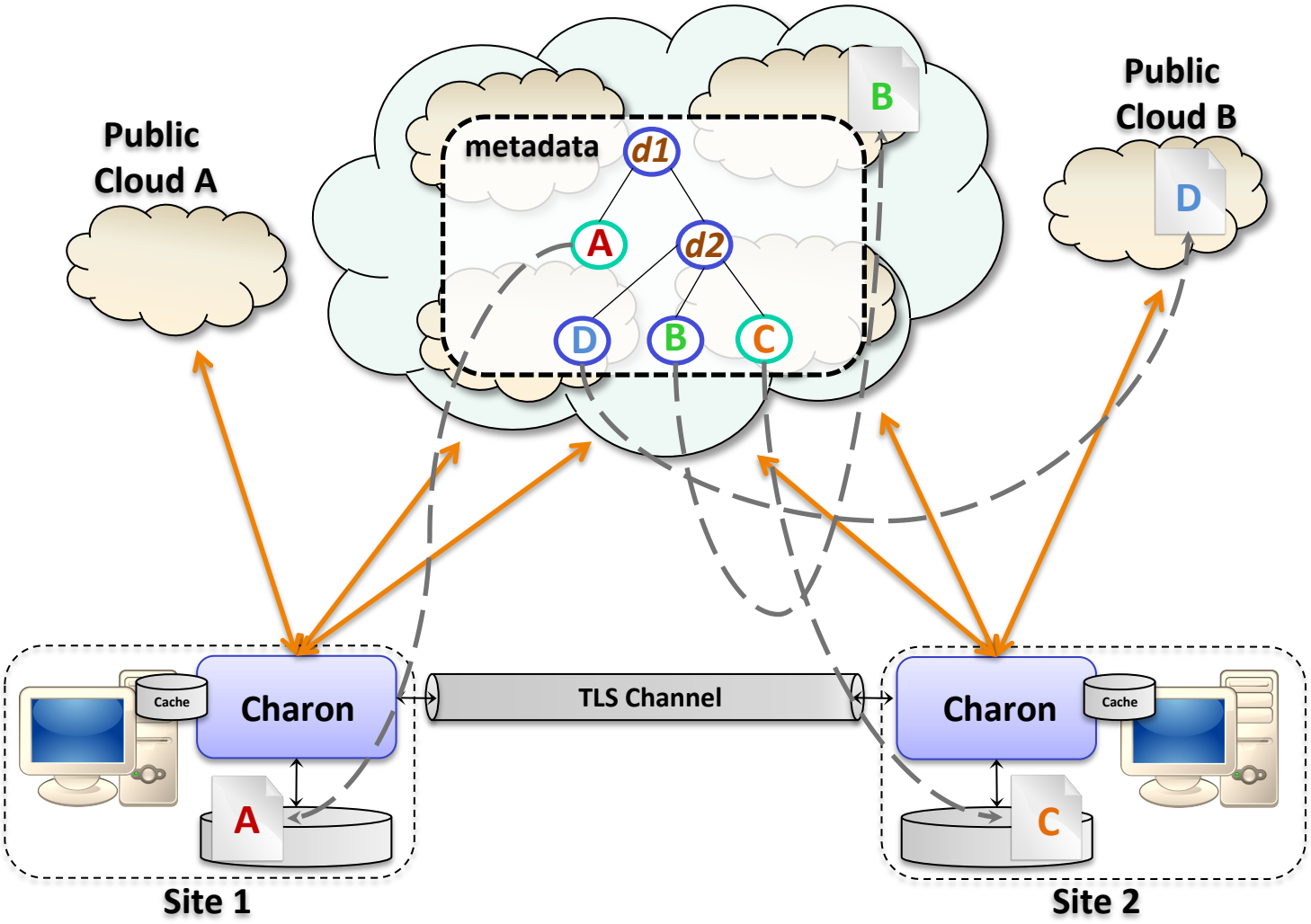
Ciências ULisboa

Navigators

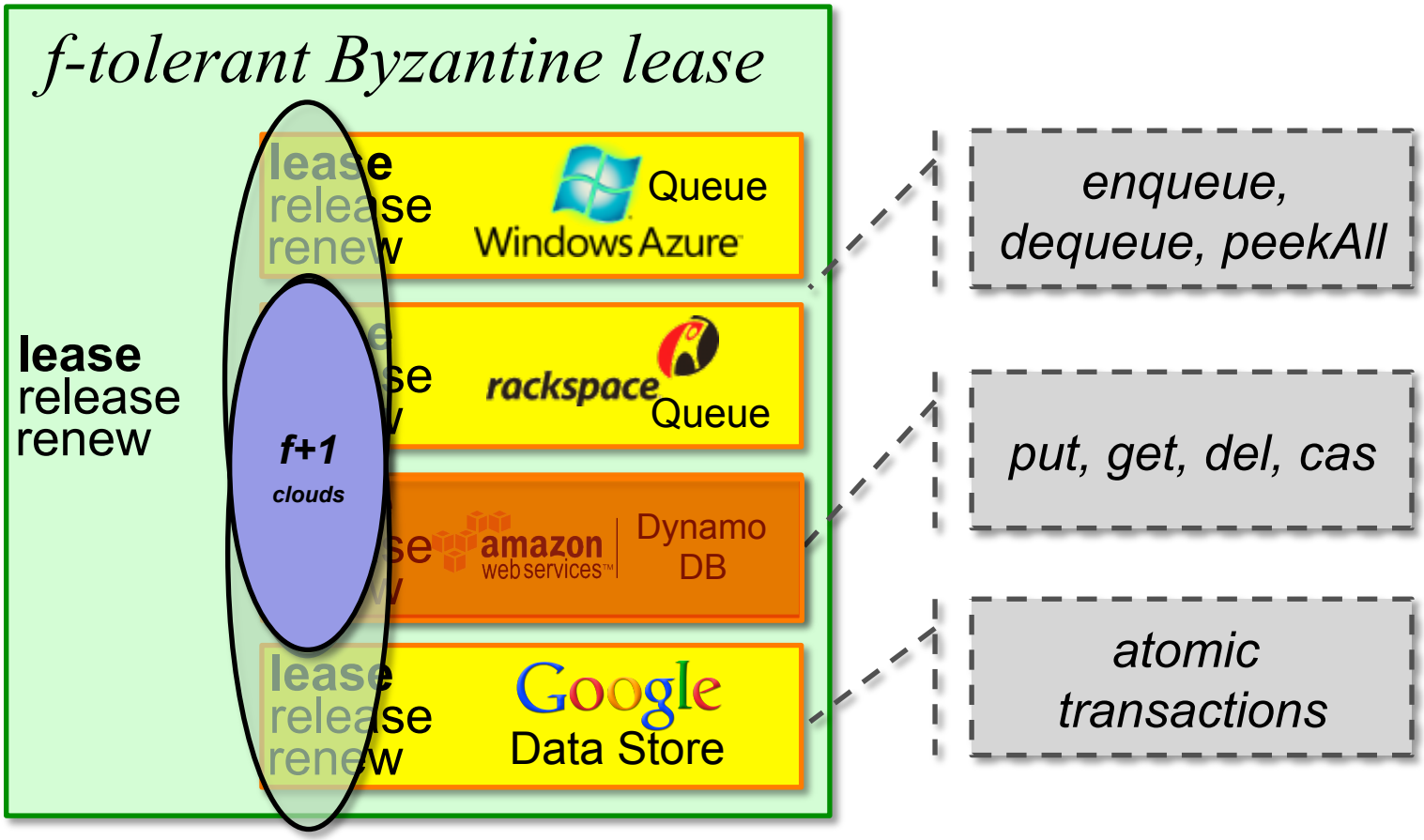
Charon Building Blocks

- **One-Shot Register**
  - Can be written only once
  - Used for storing file contents
- **Single-Writer Multi-Reader (SWMR) Register**
  - Can be updated, but not concurrently
  - Used for storing file system metadata
- **Lease Object**
  - Provide lock/unlock operations
  - Used for avoiding write-write conflicts in the system

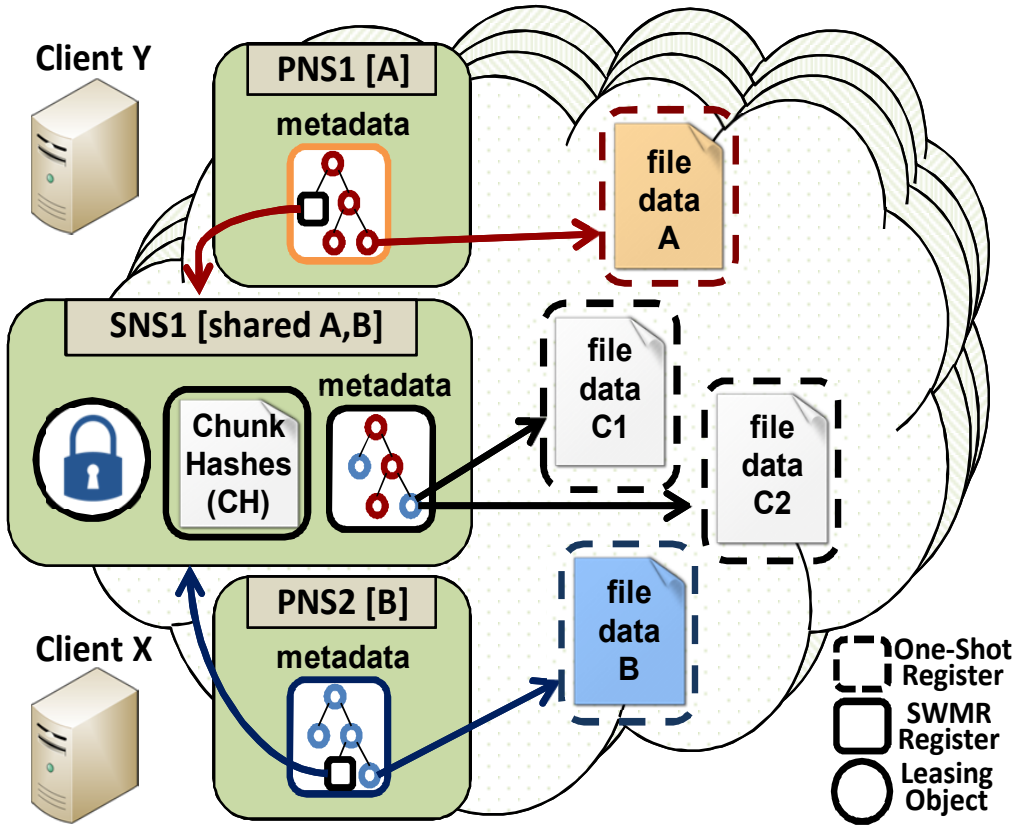
Charon Architecture



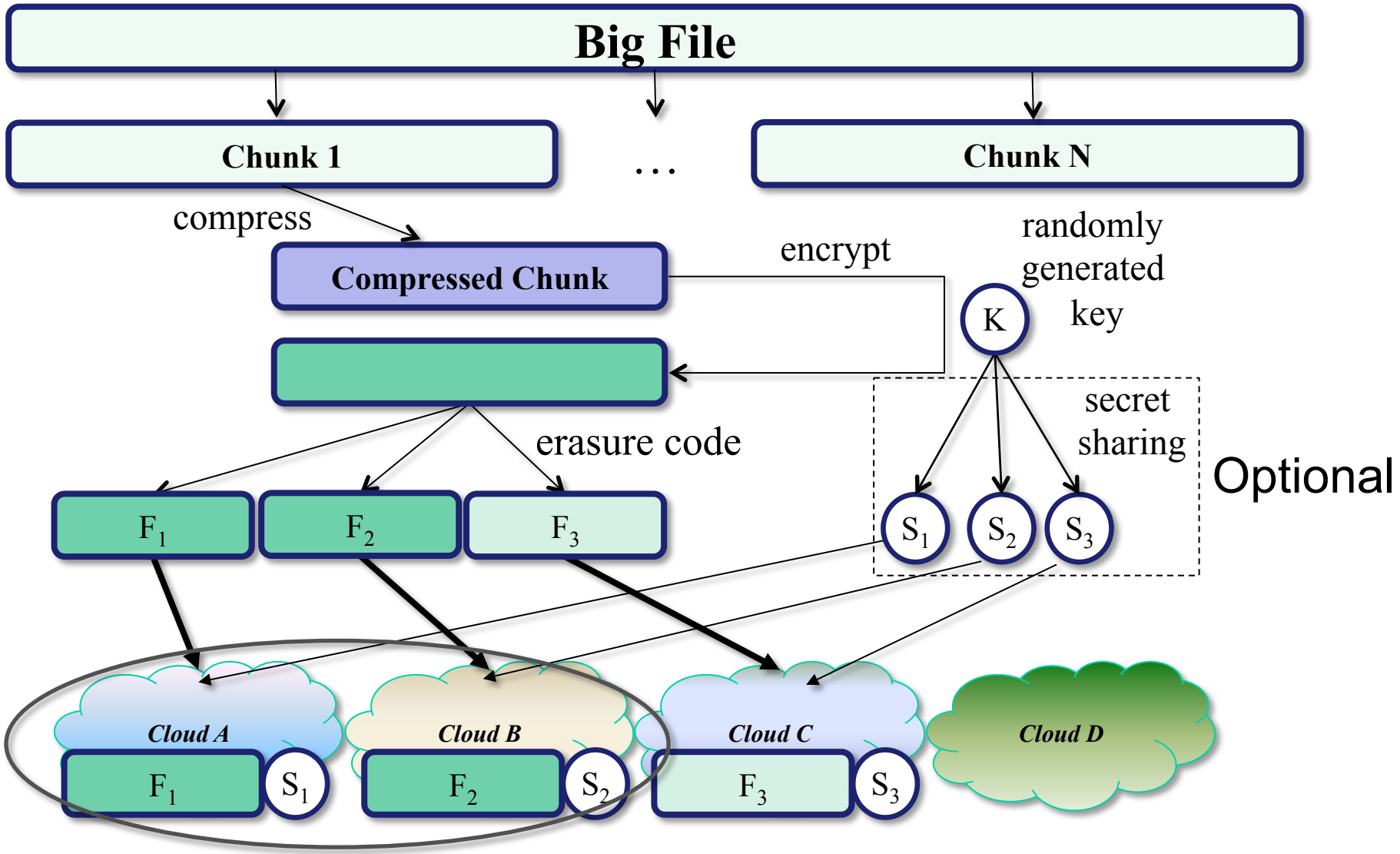
Avoiding write-write Conflicts without External Coordination



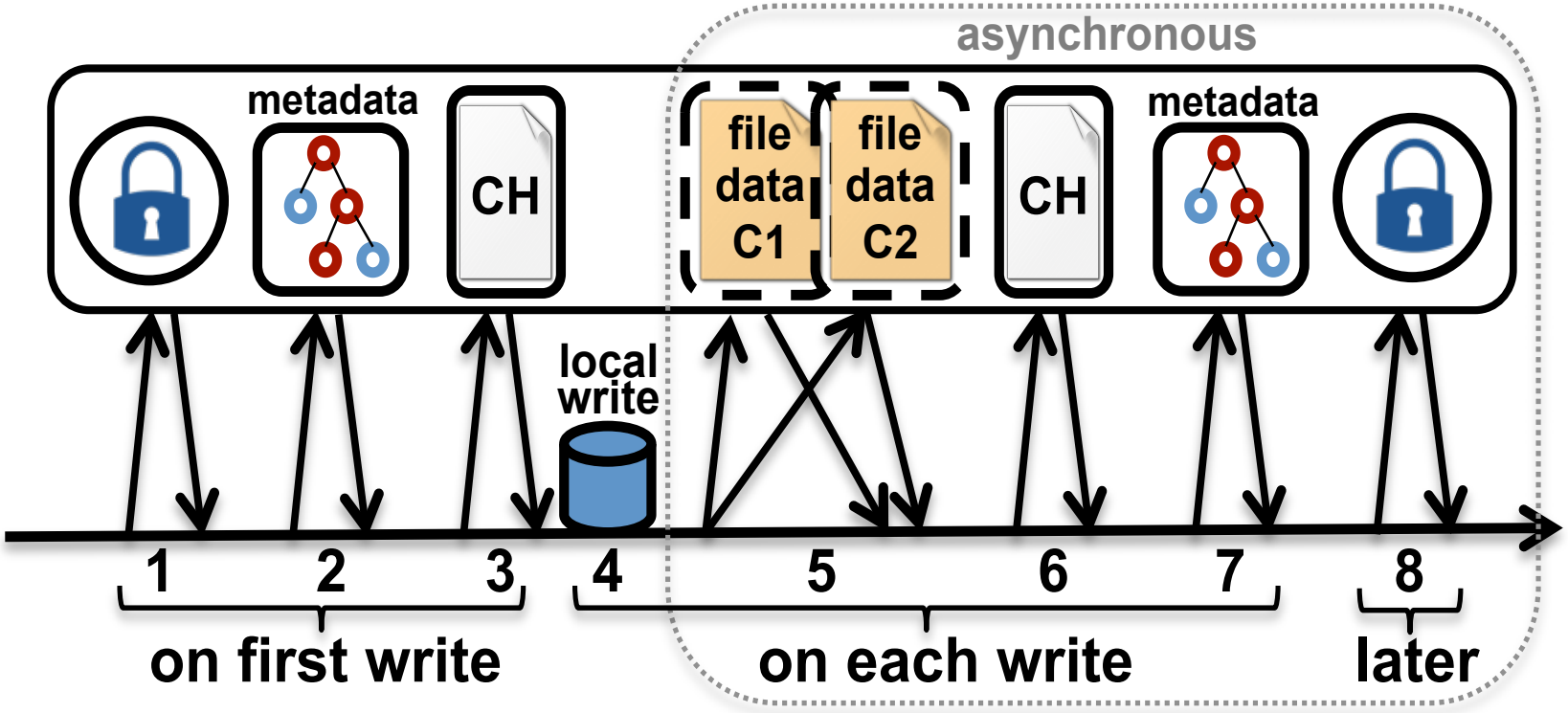
Personal and Shared Namespaces



Confidentiality & Storage-Efficiency



Processing Writes



FS Microbenchmarks

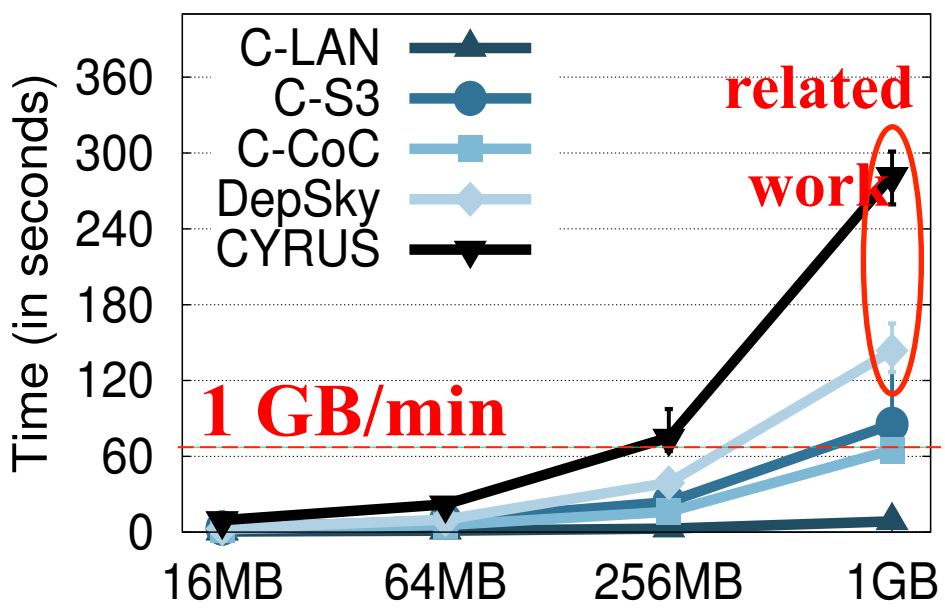
Metadata-intensive operations (ops/s)

Operation	ext4	NFS	S3QL	SCFS	CHARON
Create	2618	192	105	2	485
Delete	1895	2518	486	4	1258
Stat	15299	20881	5995	9	12925
MakeDir	14998	16664	4242	14	13665
DeleteDir	11998	6785	950	5	8665
ListDir	18759	17426	604	6	9894

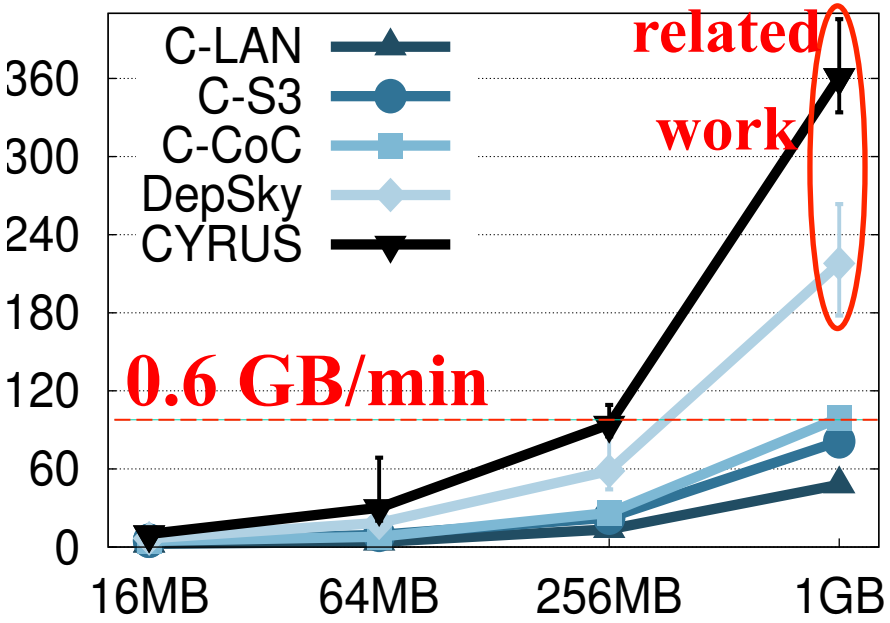
Data-intensive operations (MB/s)

Operation	ext4	NFS	S3QL	SCFS	CHARON
seqRead	215	211	214	200	194
randRead	210	205	207	191	186
seqWrite	125	125	10	17	36

Upload/Download Latency



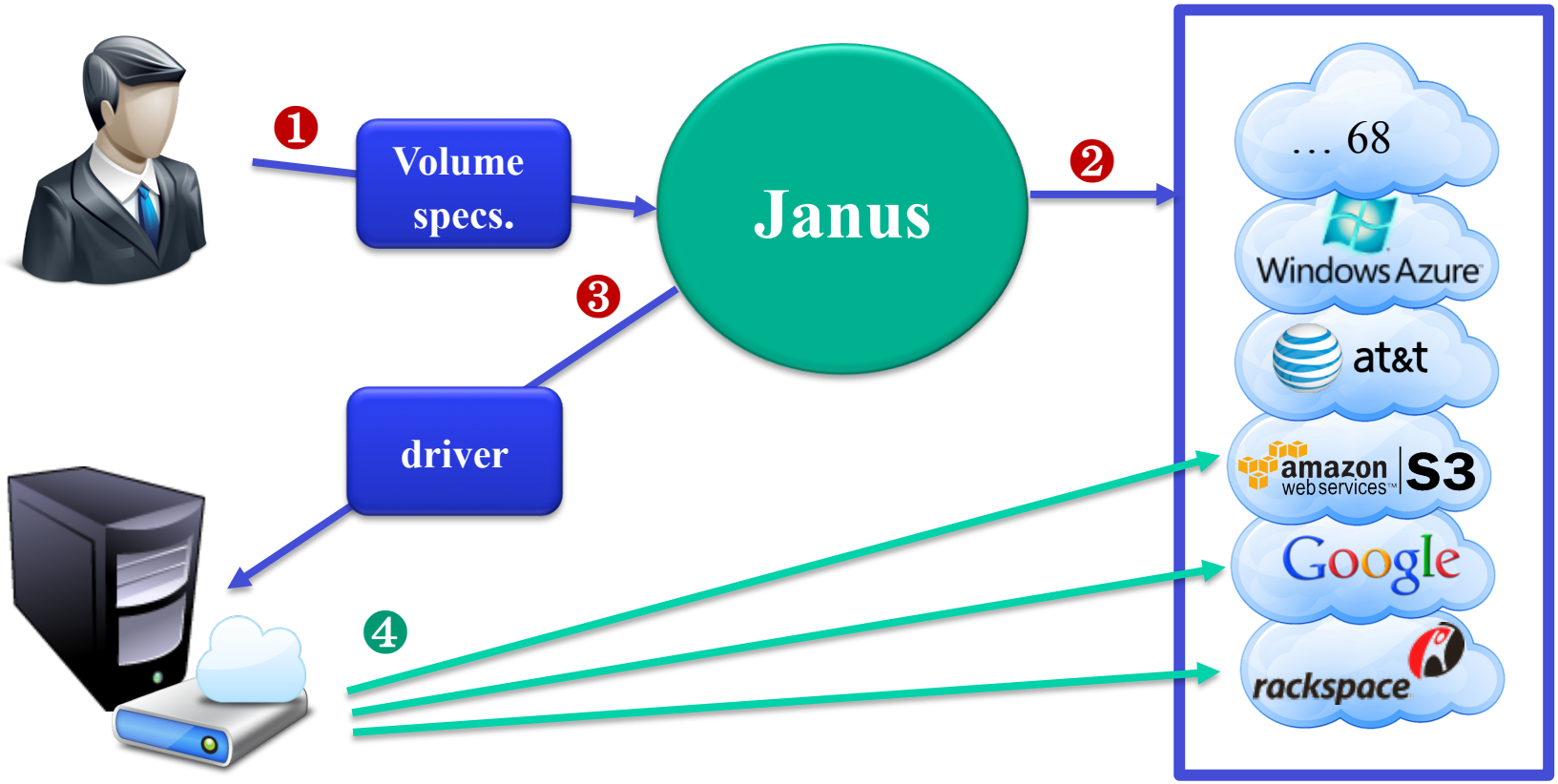
(a) Non-cached read.



(b) Write and upload.

Cloud-of-clouds as a Service

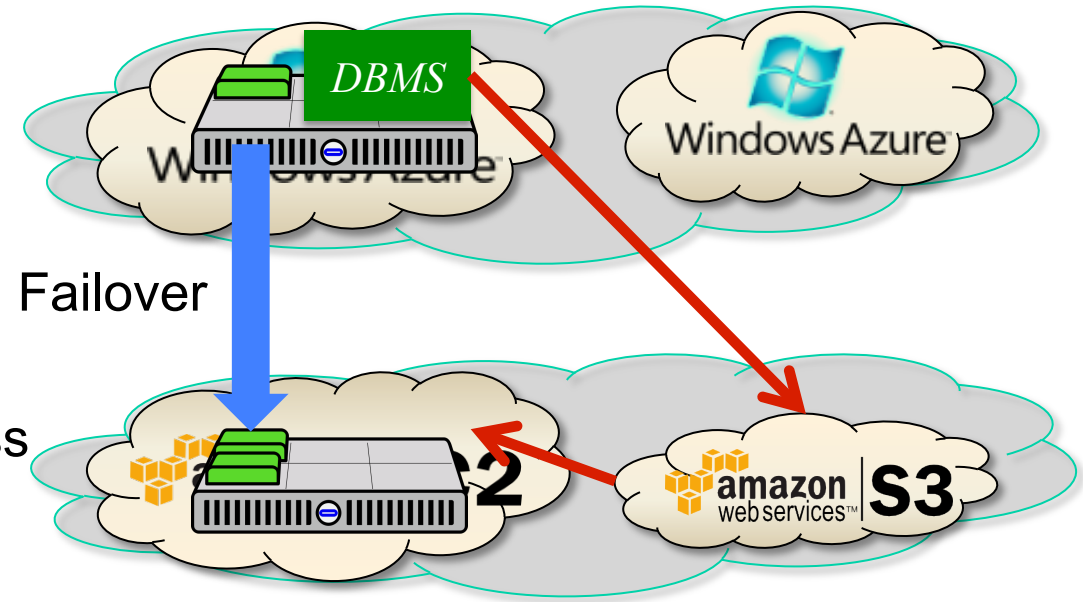
Janus: User-defined cloud-backed storage volumes



Cloud-of-clouds Computing
(very briefly)

Disaster Recovery/Tolerance

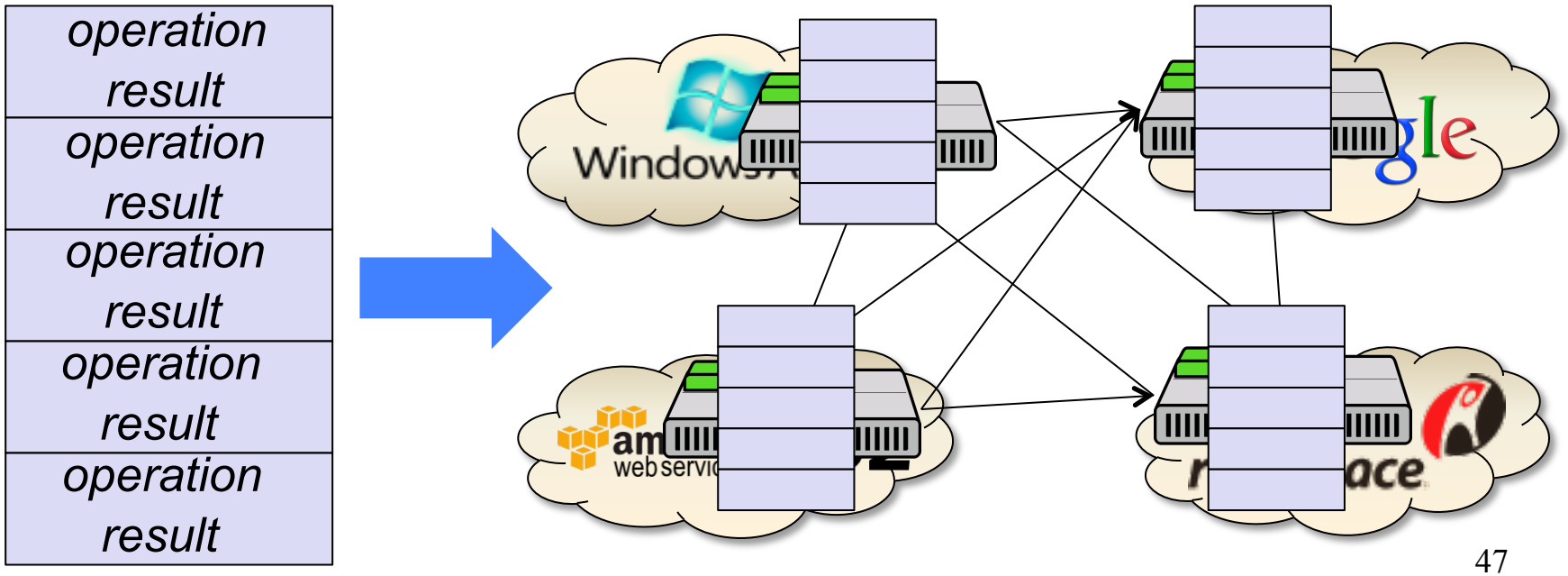
- Services in cloud A can have a backup in cloud B
- For less than **€1/month** it is possible to keep up to 30GB of data in another cloud w/ a RPO of 3 min



Recovers take less than 40s per GB

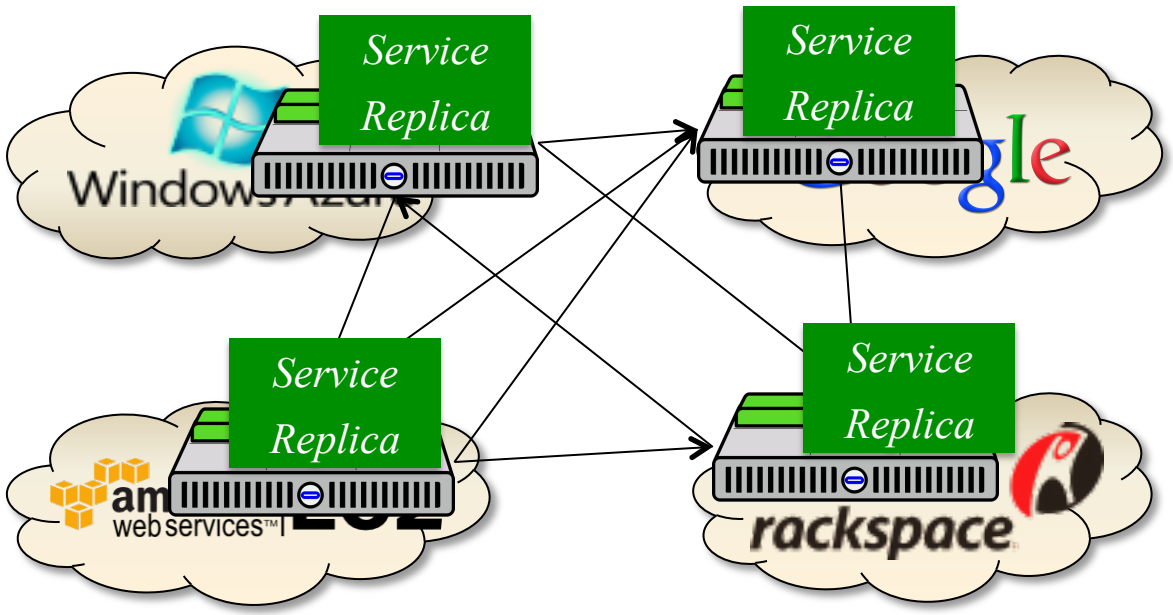
Permissioned Blockchains and **Smart Contracts**

- Permanent **decentralized ledger** used for recording transactions; requires a practical BFT consensus
- Ex.: Hyperledger, Symbiont Assembly, Ethereum, …



Geo-replicated Services

- Largely used in (single domain) internet-scale apps
- Wide-area Crash or Byzantine fault tolerant replication protocols are needed



Ciências ULisboa

Navigators

Final Remarks

- In 2010
  - Paxos and Zookeeper were not very popular
  - RAFT didn't exist
  - Bitcoin was a crazy new idea
  - Blockchain and smart contracts were mostly ignored
- The world is changing fast, and advanced distributed applications are here to stay...
- What will happen in 2020?

Ciências ULisboa

Navigators

Further reading…

Storage

- Bessani, Correia, Quaresma, André, Sousa. **DepSky: Dependable and Secure Storage in the Cloud of Clouds**. ACM Transactions on Storage. 2013. (preliminary version on ACM EuroSys'11).
- Oliveira, Mendes, Bessani. **Exploring Key-Value Stores in Multi-Writer Byzantine-Resilient Register Emulations.** OPODIS'16.
- Mendes, Oliveira, Cogo, Neves, Bessani. **Charon: A Dependable Cloud-Backed System for Storing and Sharing Big Data.** Under Submission. 2016.
- Bessani, Mendes, Oliveira, Neves, Correia, Pasin, Verissimo. **SCFS: A Shared Cloud-backed File System**. USENIX ATC'14.

Computing

- Sousa, Bessani. **Separating the WHEAT from the Chaff: An Empirical Design for Geo-Replicated State Machines.** IEEE SRDS'15.
- Bessani, Sousa, Alchieri. **State Machine Replication for the Masses with BFT-SMaRt**. IEEE/IFIP DSN'14.

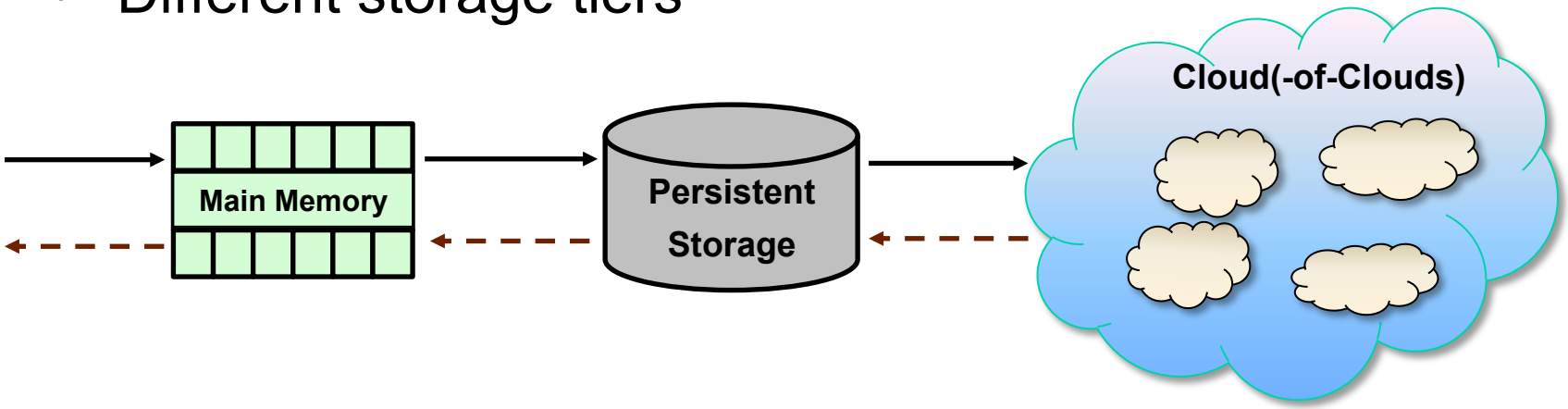
TClouds

BIOBANKCLOUD

SUPER CLOUD

Persistence and Durability

- Different storage tiers



- Different durability levels for different *syscalls*

Level	*Location*	*Latency*	*Fault tol.*	*Sys. call*
0	main memory	microsec	none	write
1	local disk	millisec	crash	fsync
2	cloud	seconds	local disk	close
3	cloud-of-clouds	seconds	f clouds	close