JOIN THE FUN!

- Check the Chat window in Zoom (More options)
- Register to our HPE DEV Slack
  - https://slack.hpedev.io/
- Join the #munch-and-learn channel and start chatting
  - https://hpedev.slack.com/archives/C01GVQUPM3P
  - Upload pictures of your munchies
- Join #ezmeral-data-fabric channel and ask technical question to the community (including Ted & Ellen)
  - https://hpedev.slack.com/archives/CU3JRBTB7

HPE DISCOVER
VIRTUAL EXPERIENCE

WHAT'S A DATA FABRIC, ANYWAY?

Ted Dunning, Data Fabric CTO, HPE
@ted_dunning
ted.dunning@hpe.com

#HPEDISCOVER

A DATA FABRIC IS

A DATA FABRIC DOES

RUN THE RIGHT APPLICATIONS AT THE RIGHT TIME IN THE RIGHT PLACE AGAINST THE RIGHT DATA

Your business doesn’t actually happen in a data center

You can’t ignore geography

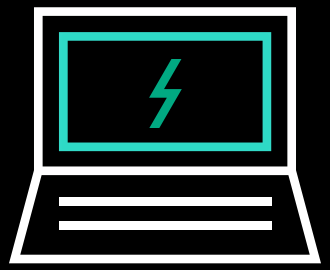
The value in your business is where your customers are

Let’s talk about some specific examples

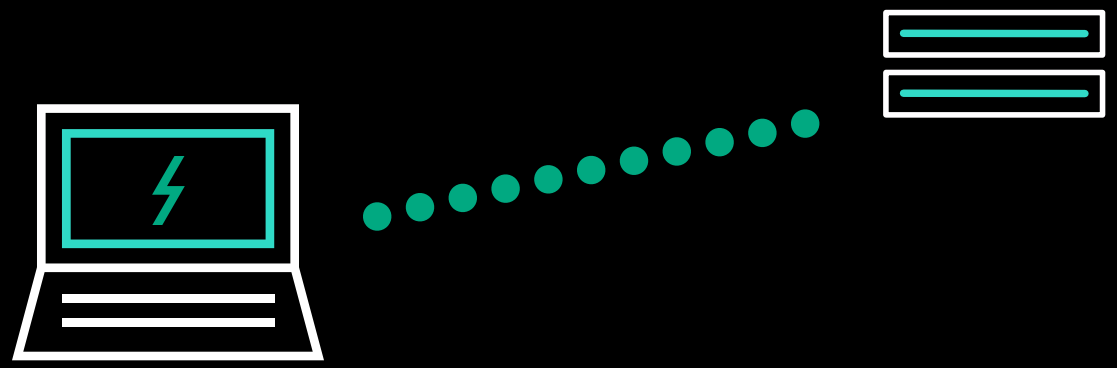
1

How does streaming video **actually work**?

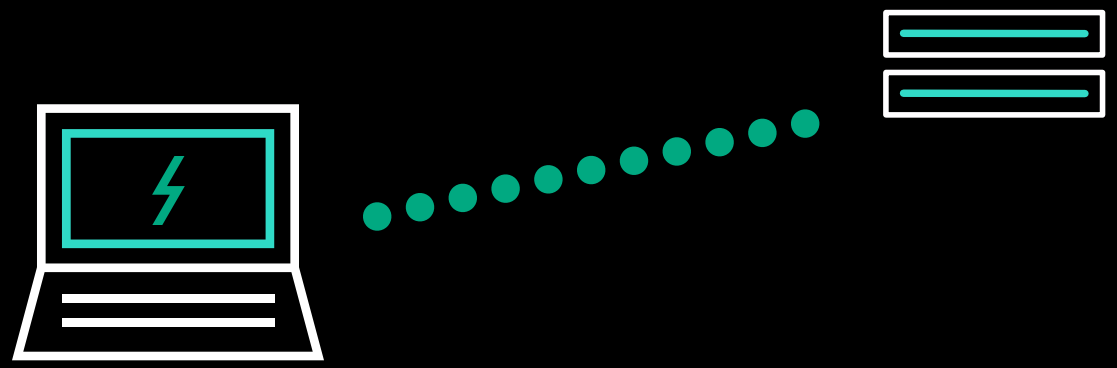
How does streaming video **actually work**?



Regional content server responds to content requests

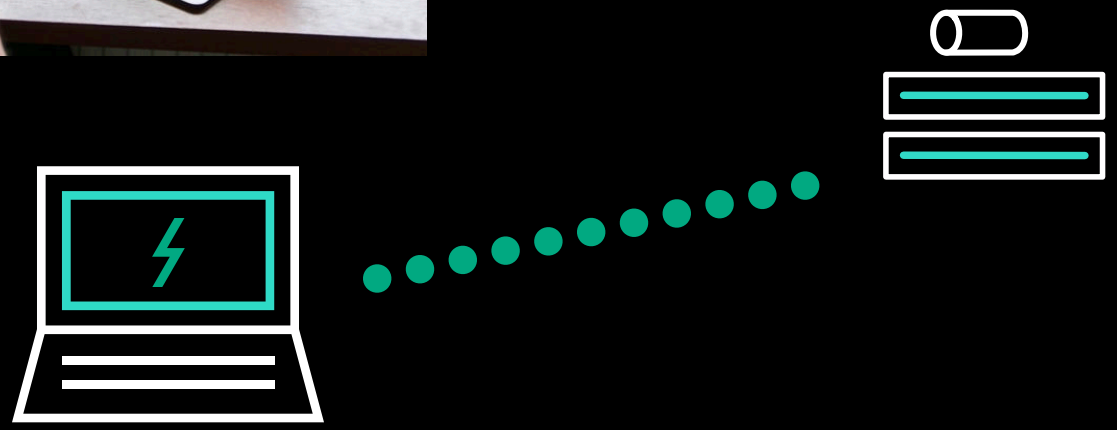


Regional content server responds to content requests

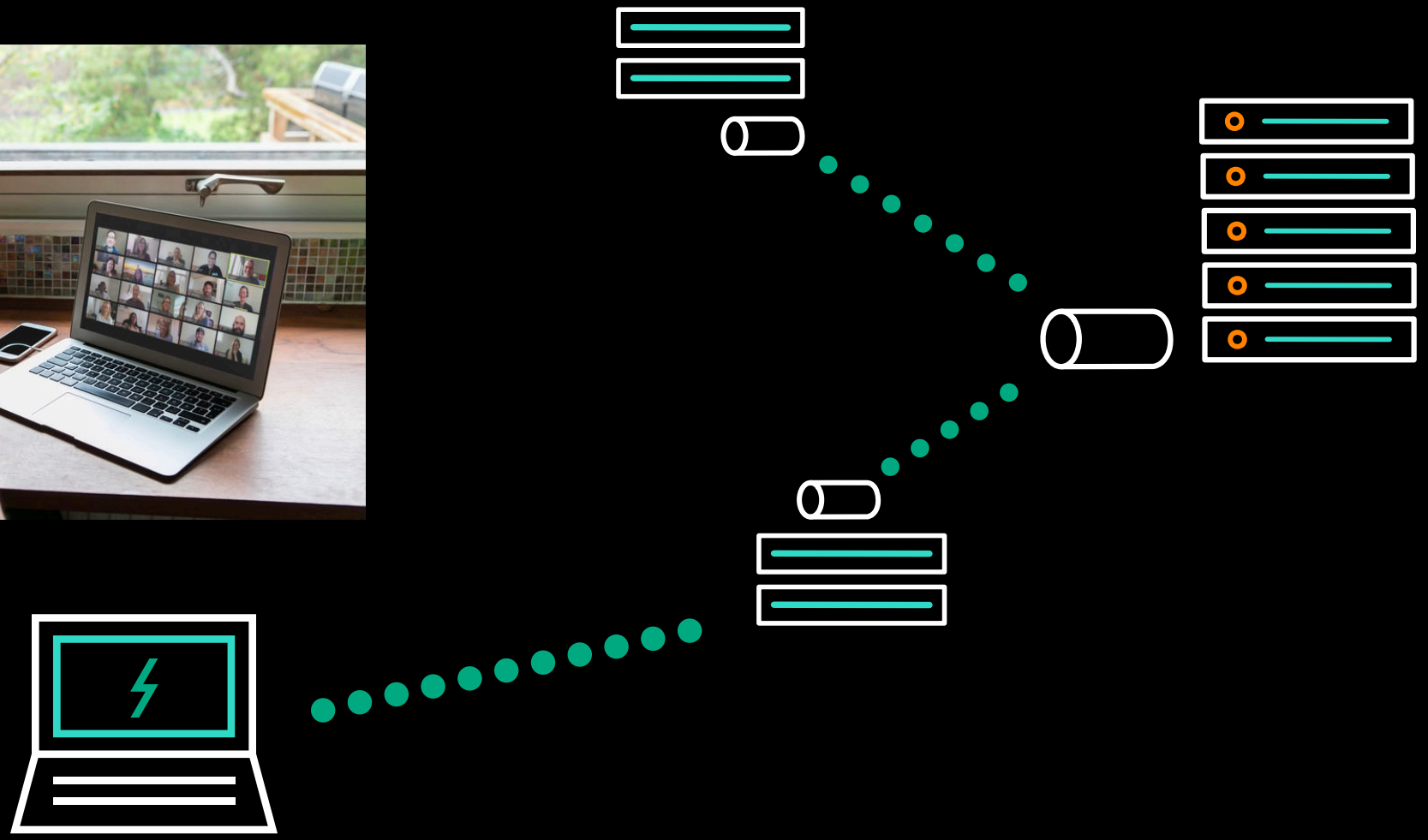


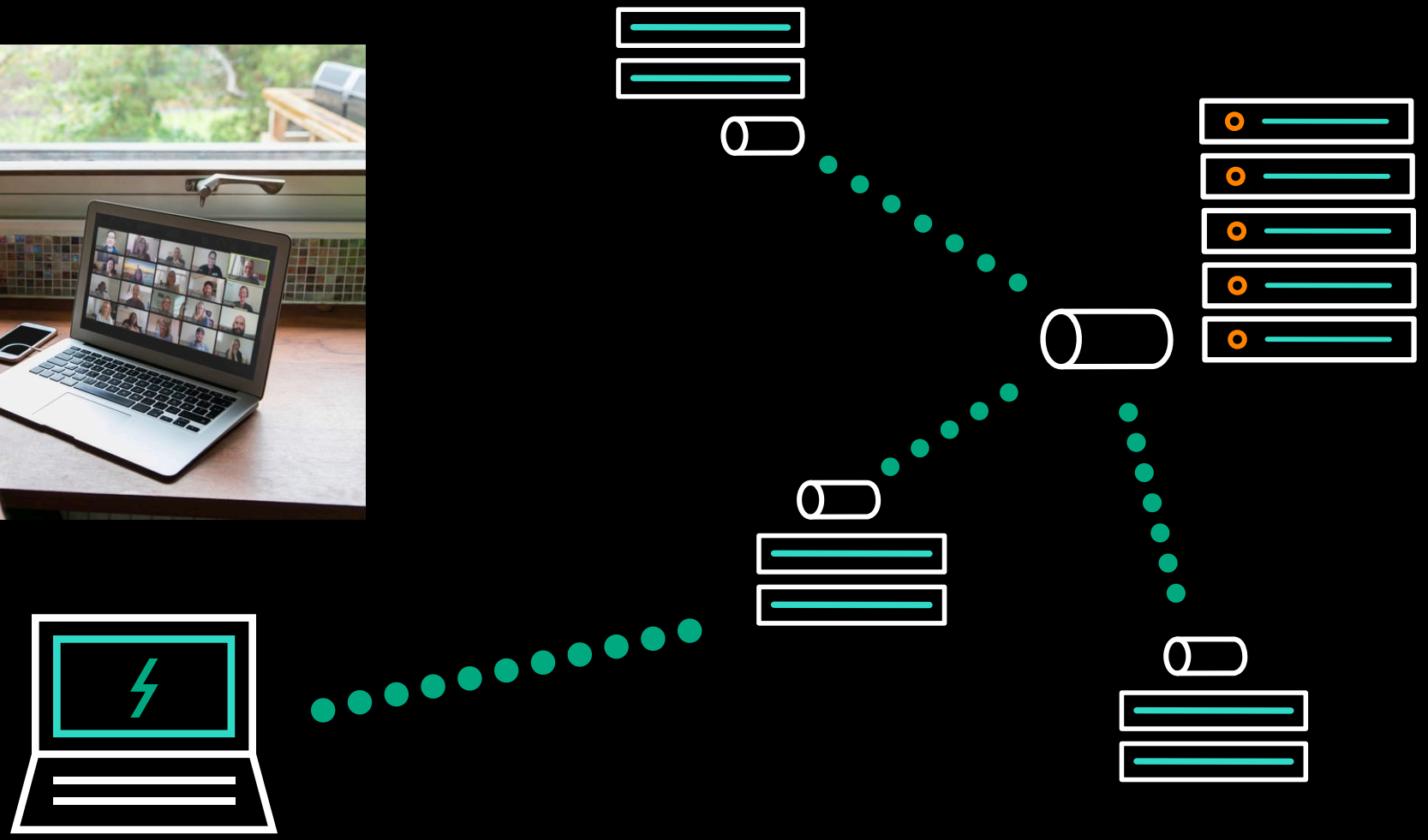
How do we know if it is working?

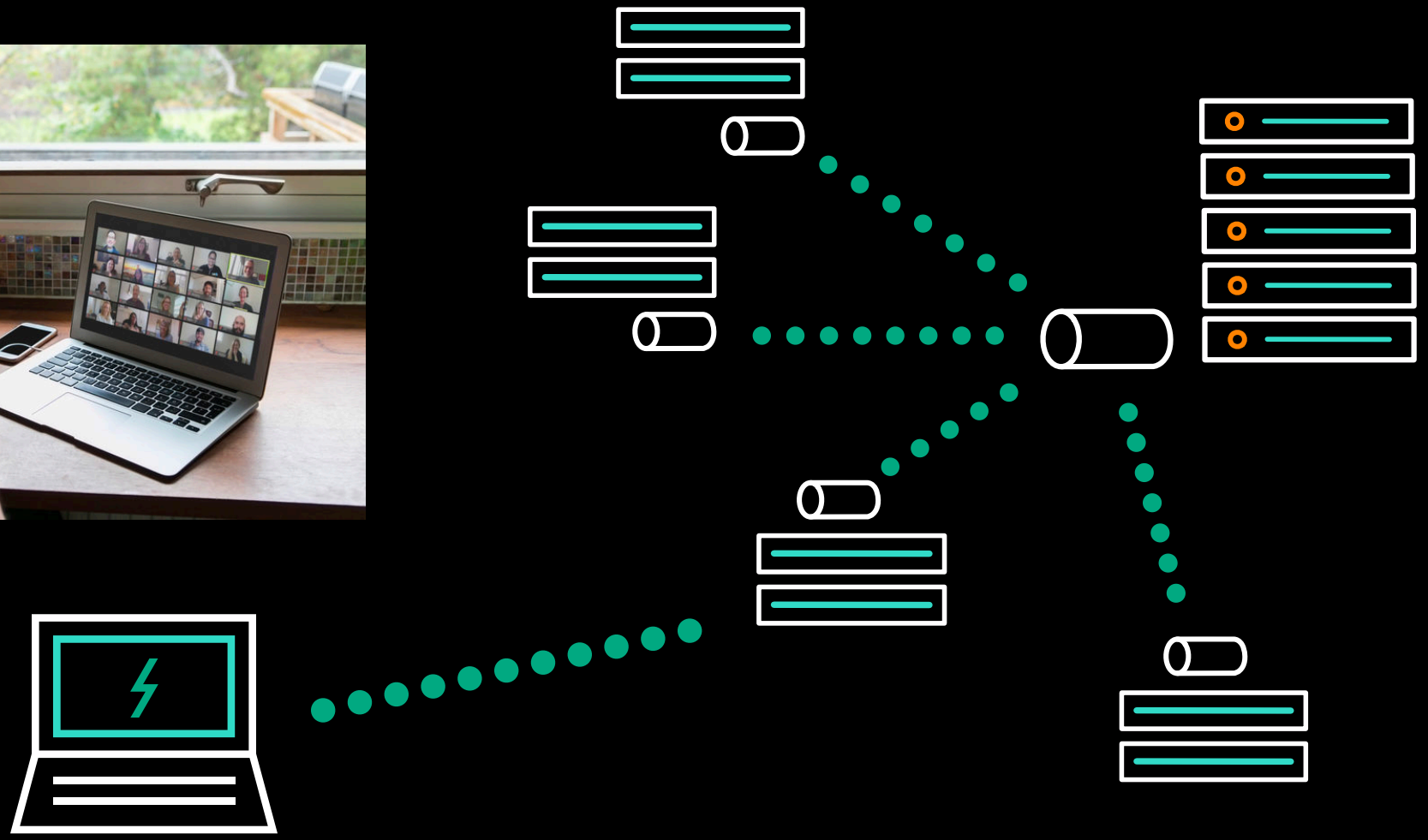
Collect telemetry in a message stream

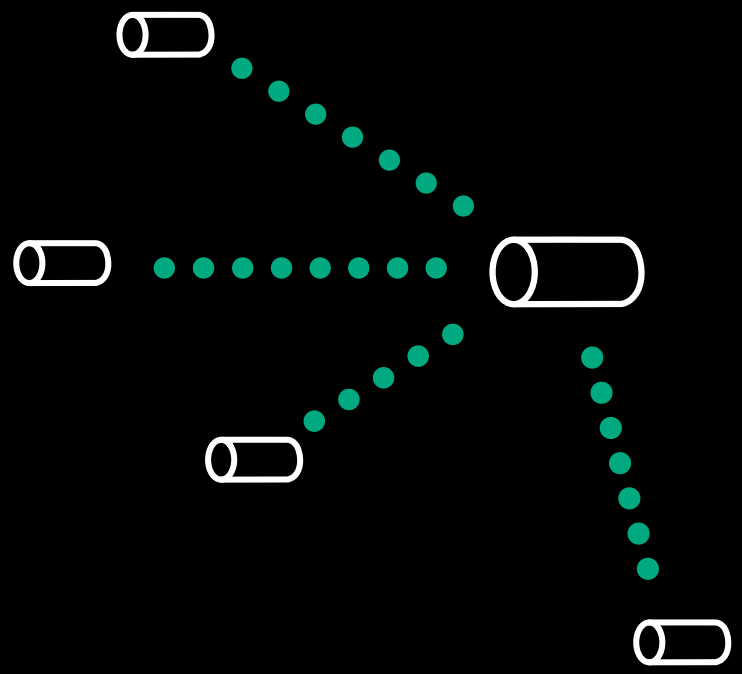


...and replicate to galactic HQ

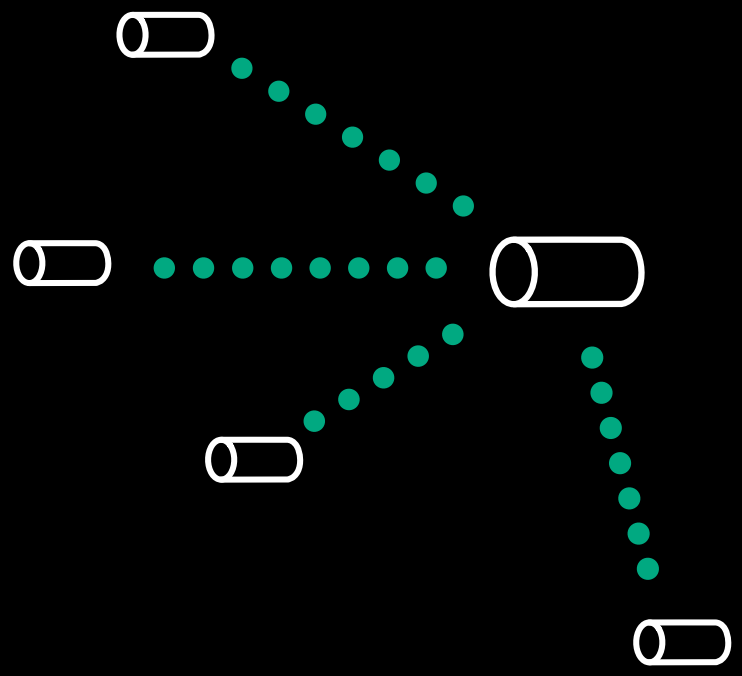








The important point is that **one data structure** can span the enterprise



The important point is that **one data structure** can span the enterprise

The replication is at a fabric level

The important point is that **one data structure** can span the enterprise

The replication is at a fabric level

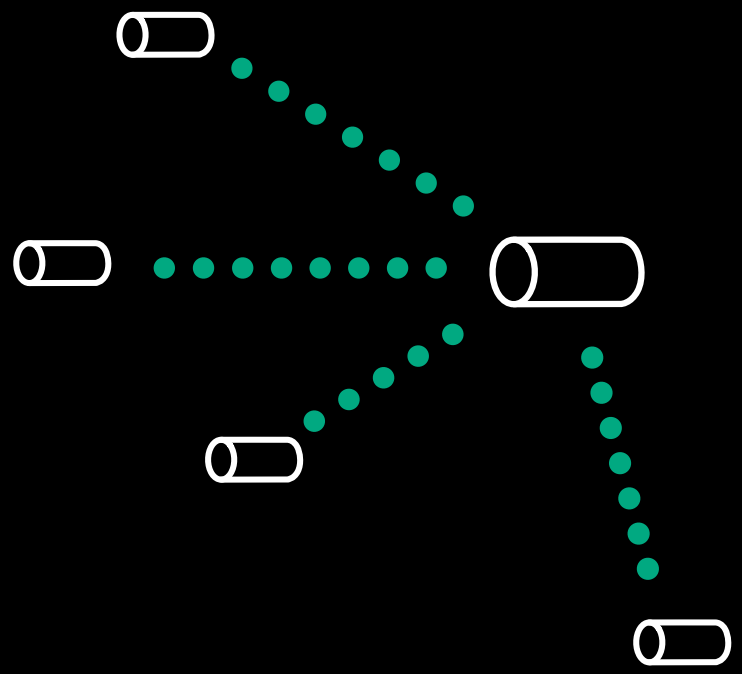
...the source doesn't know

The important point is that **one data structure** can span the enterprise

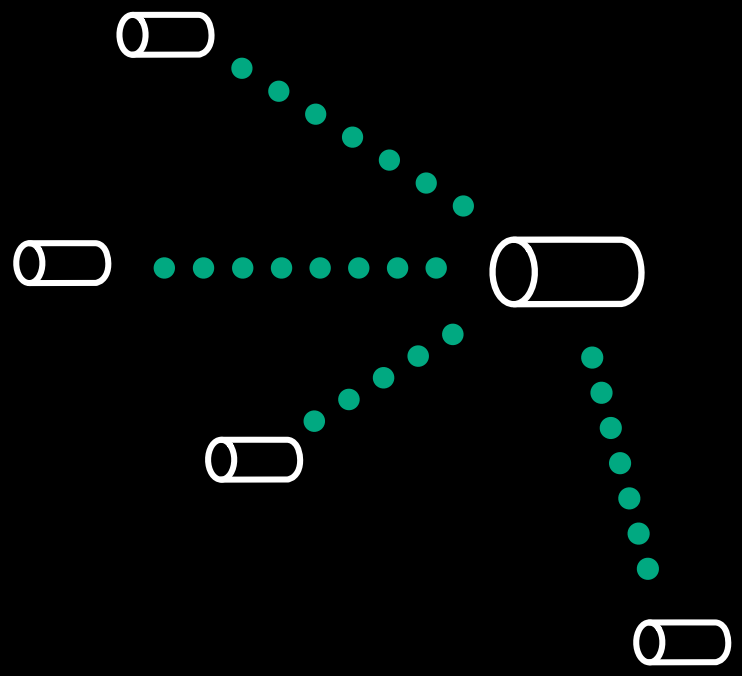
...and analytical apps get a **global view**

The replication is at a fabric level

...the source doesn't know

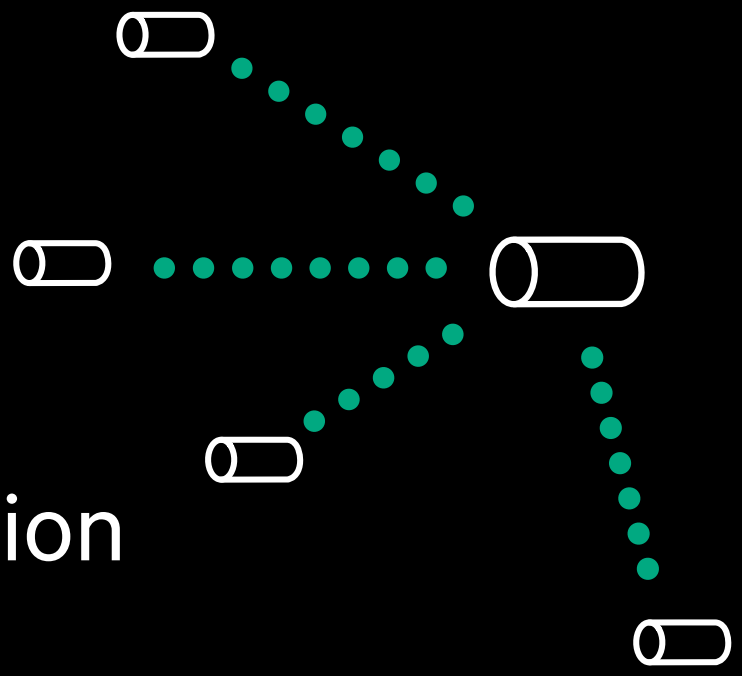


This idea isn’t limited to just media streaming

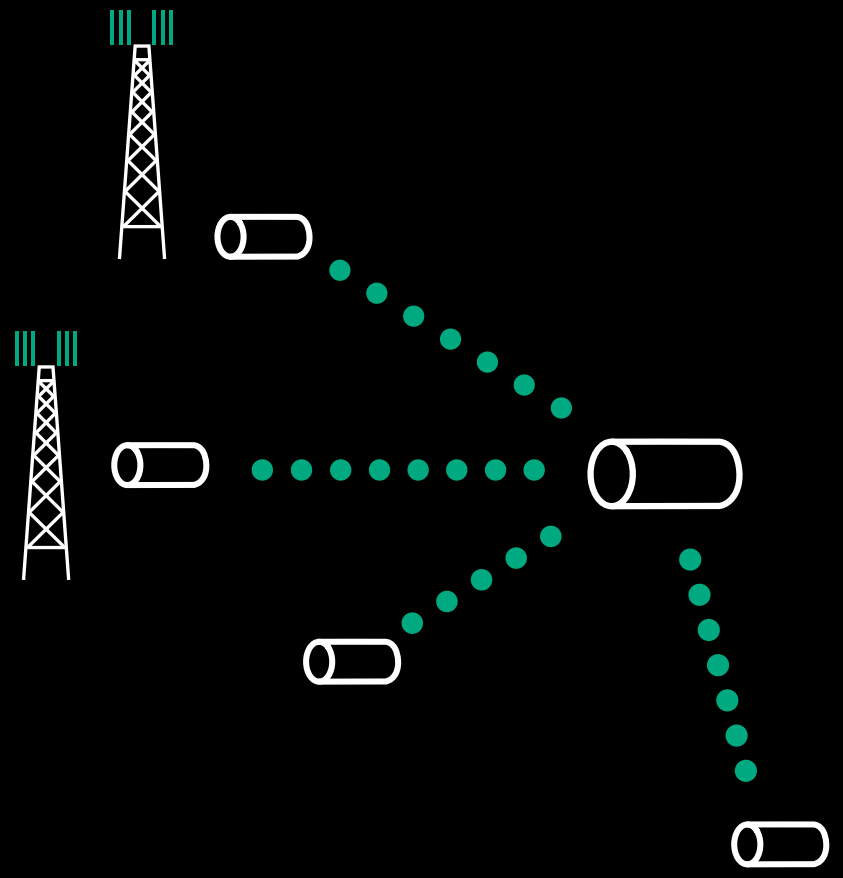


This idea isn't limited to just media streaming

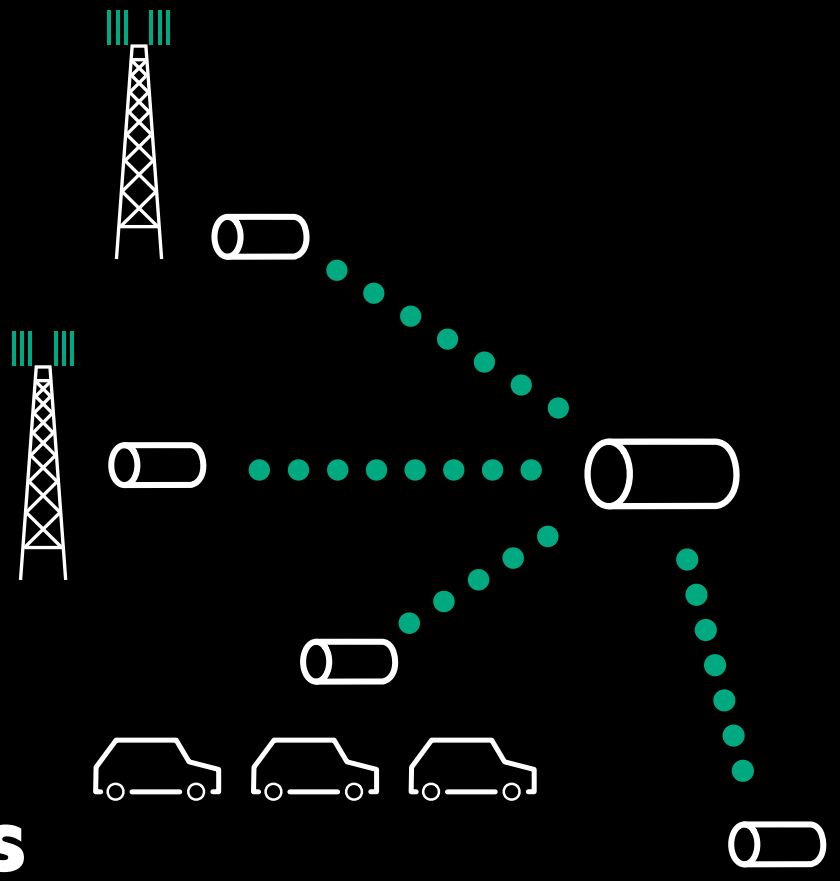
The telemetry can come from **any kind** of observation



The rollout of **5G telecom networks** benefits the same way

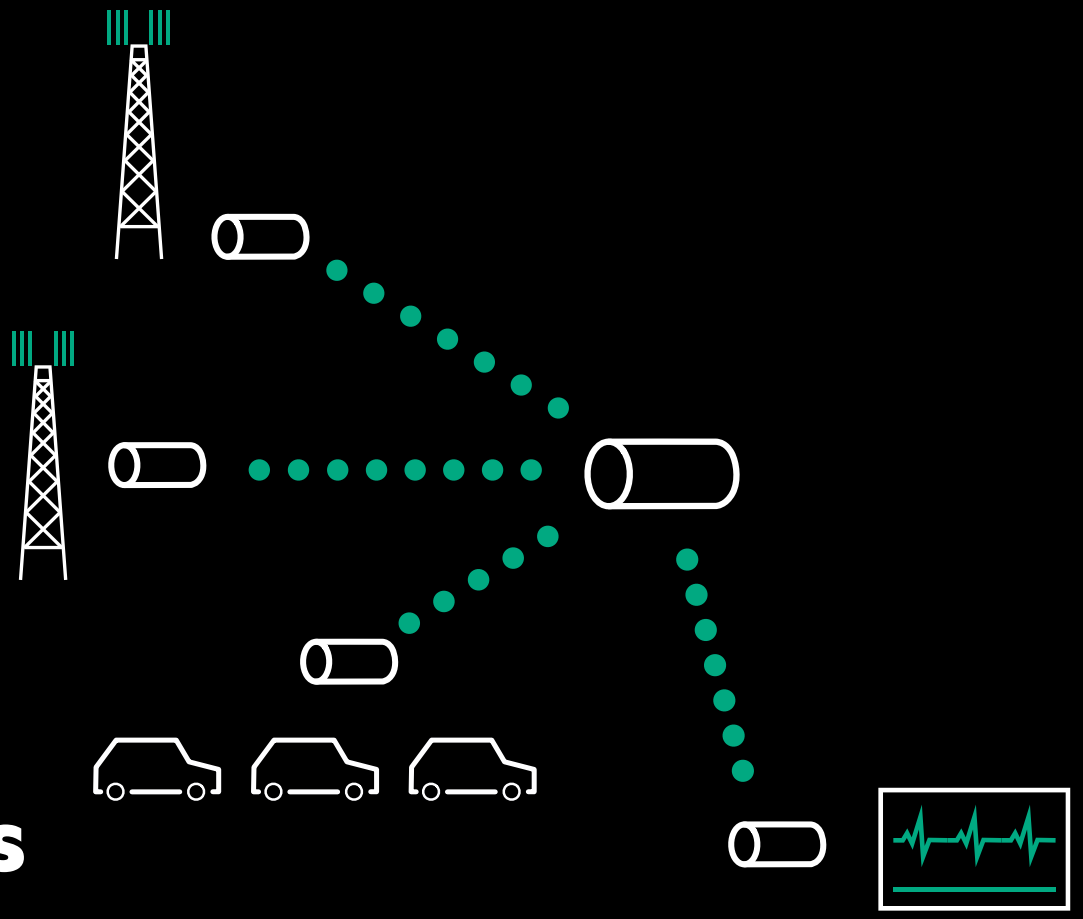


The rollout of **5G telecom networks** benefits the same way



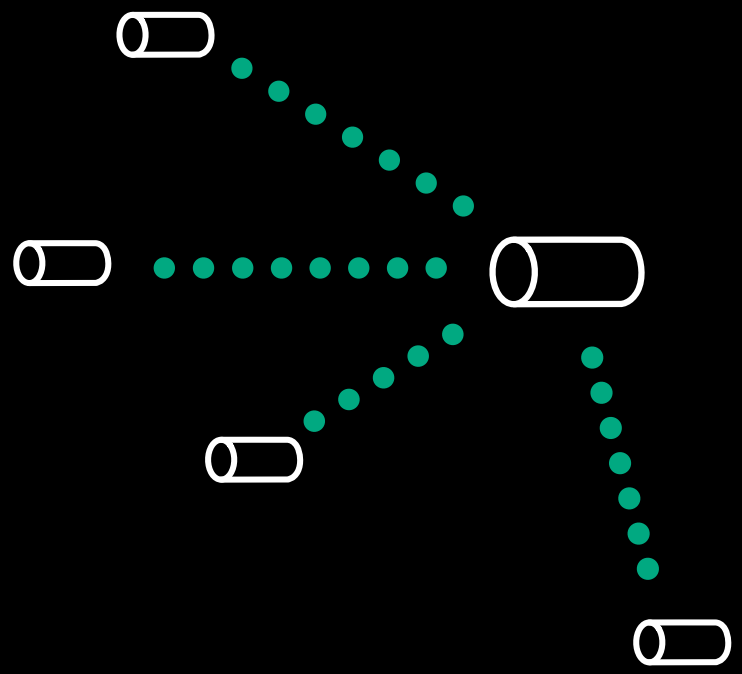
Manufacturing systems can do the same

The rollout of **5G telecom networks** benefits the same way

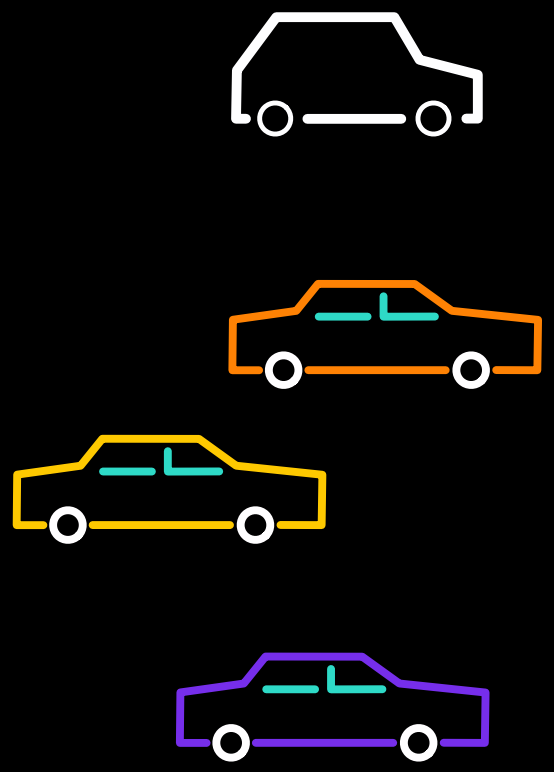


Manufacturing systems can do the same

...as can **medical devices**

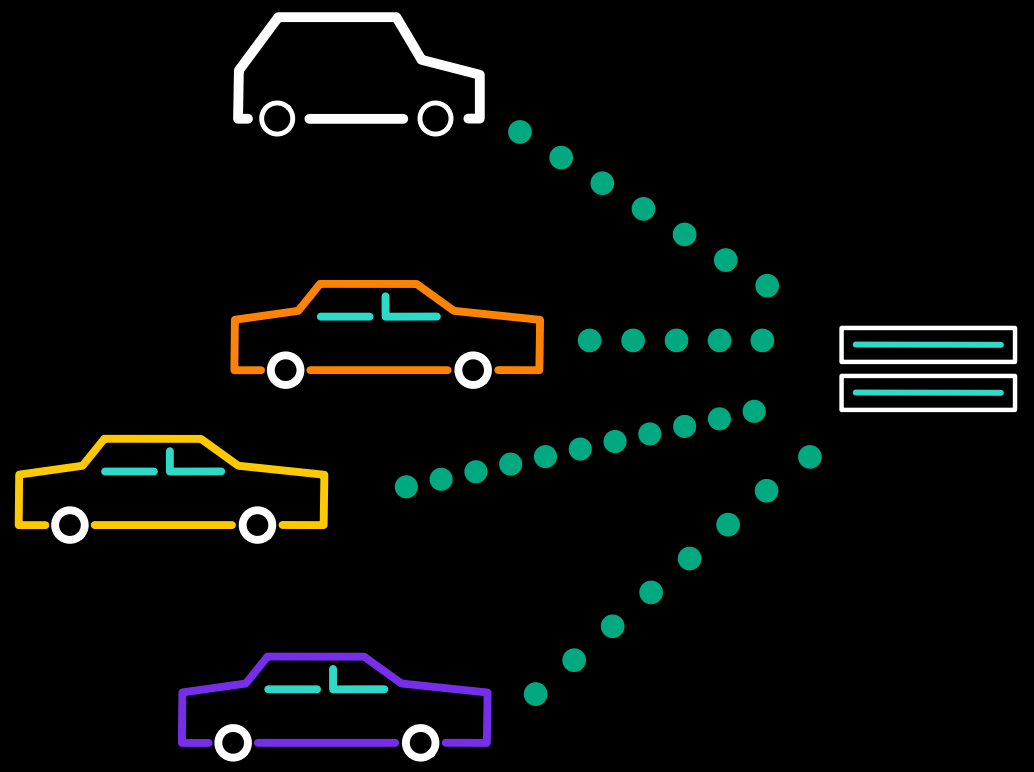


2

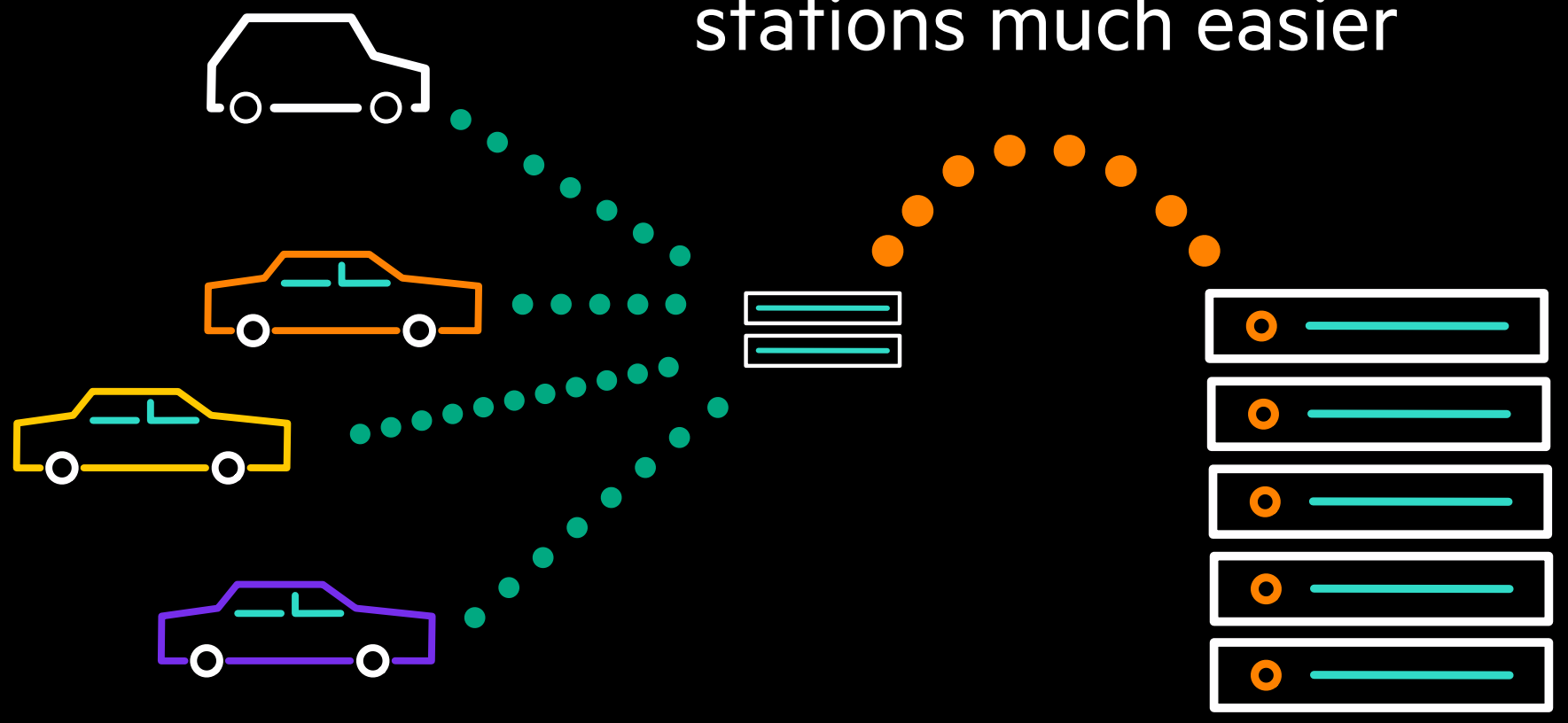


Developing autonomous vehicles requires vast amounts of real-world data

Data that can only be collected by **real cars** on **real roads**



A **data fabric** makes collecting this kind of data from field stations much easier



But the volumes are prodigious:

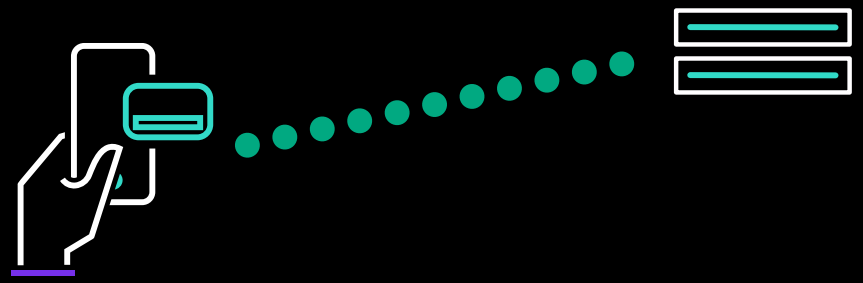
- Gigabytes/s per car
- Petabytes of new data per day
- Hundreds of petabytes retained

Having a **single fabric** that can span from edge to core makes these projects **practical**

3

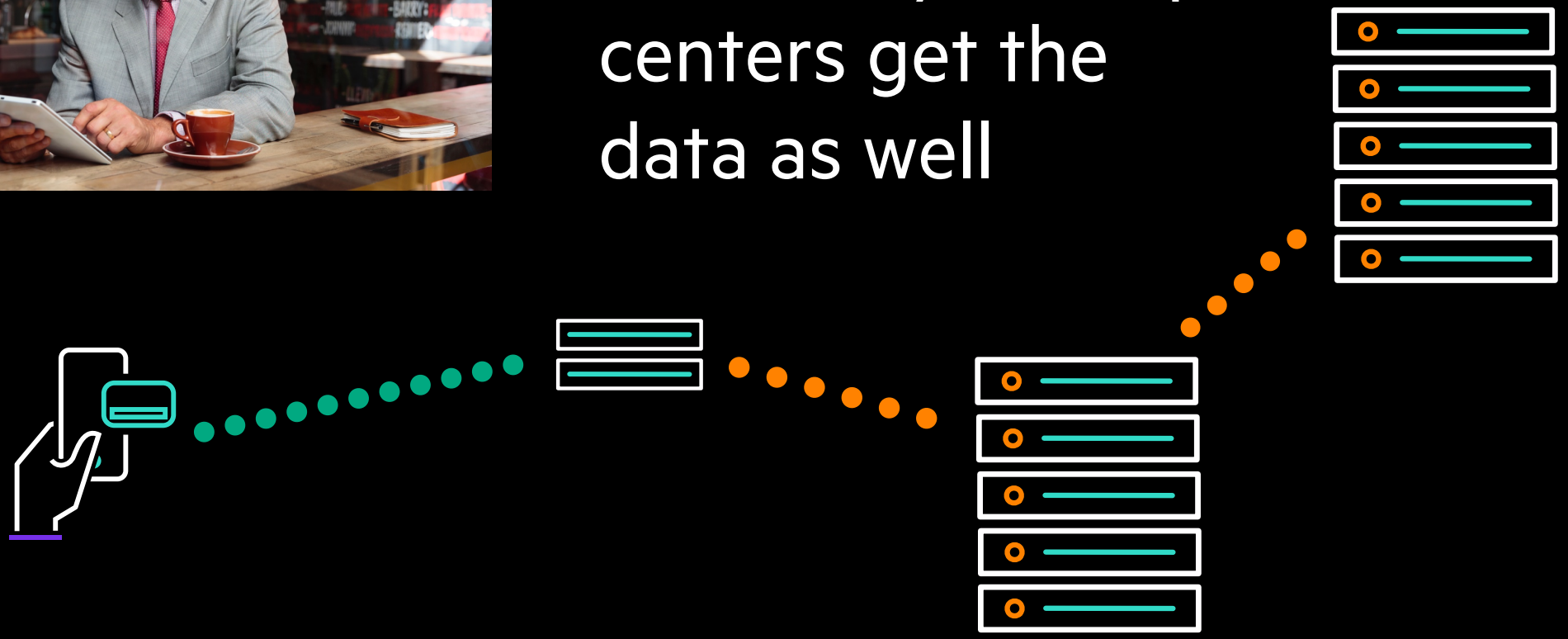
What happens behind the scenes when you pay for coffee in a café?

The point of sale system talks to a regional data center

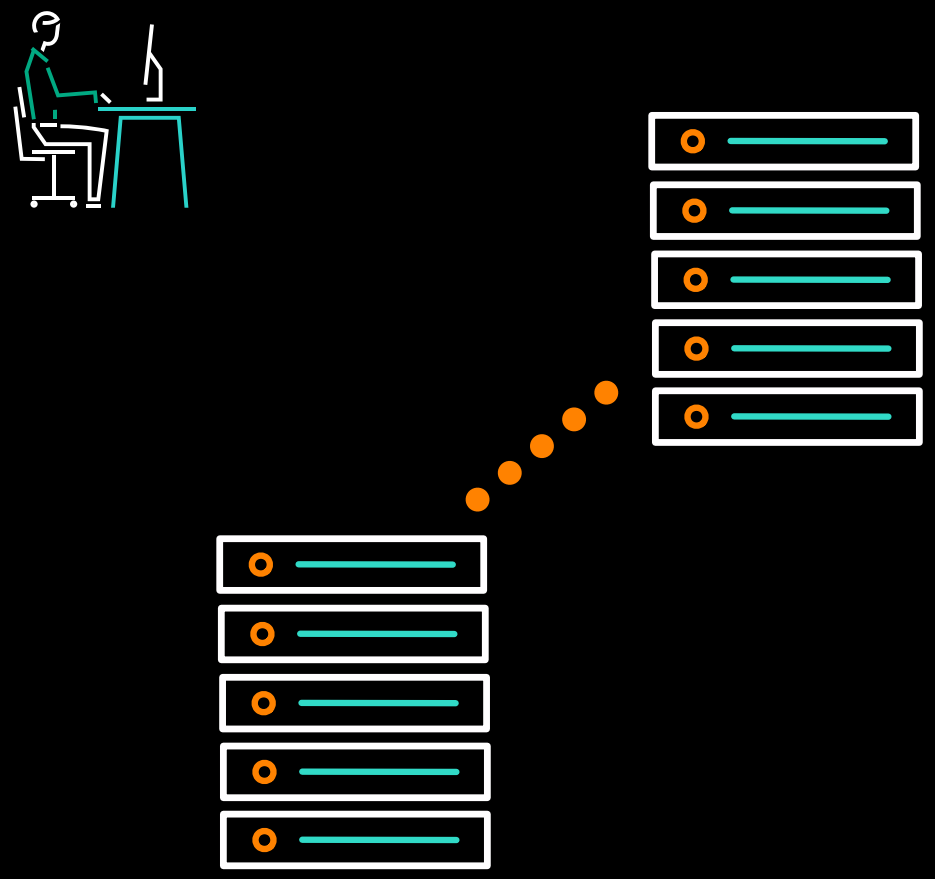


The decision to honor the transaction is made and data is forwarded

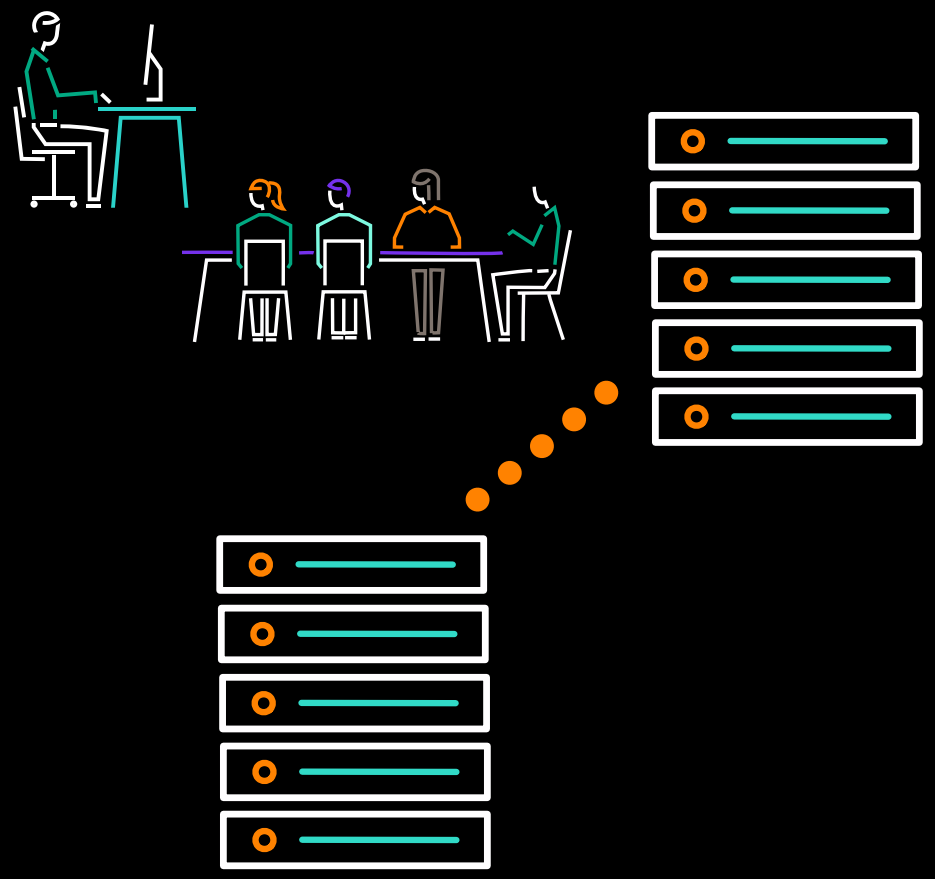
Ultimately backup centers get the data as well

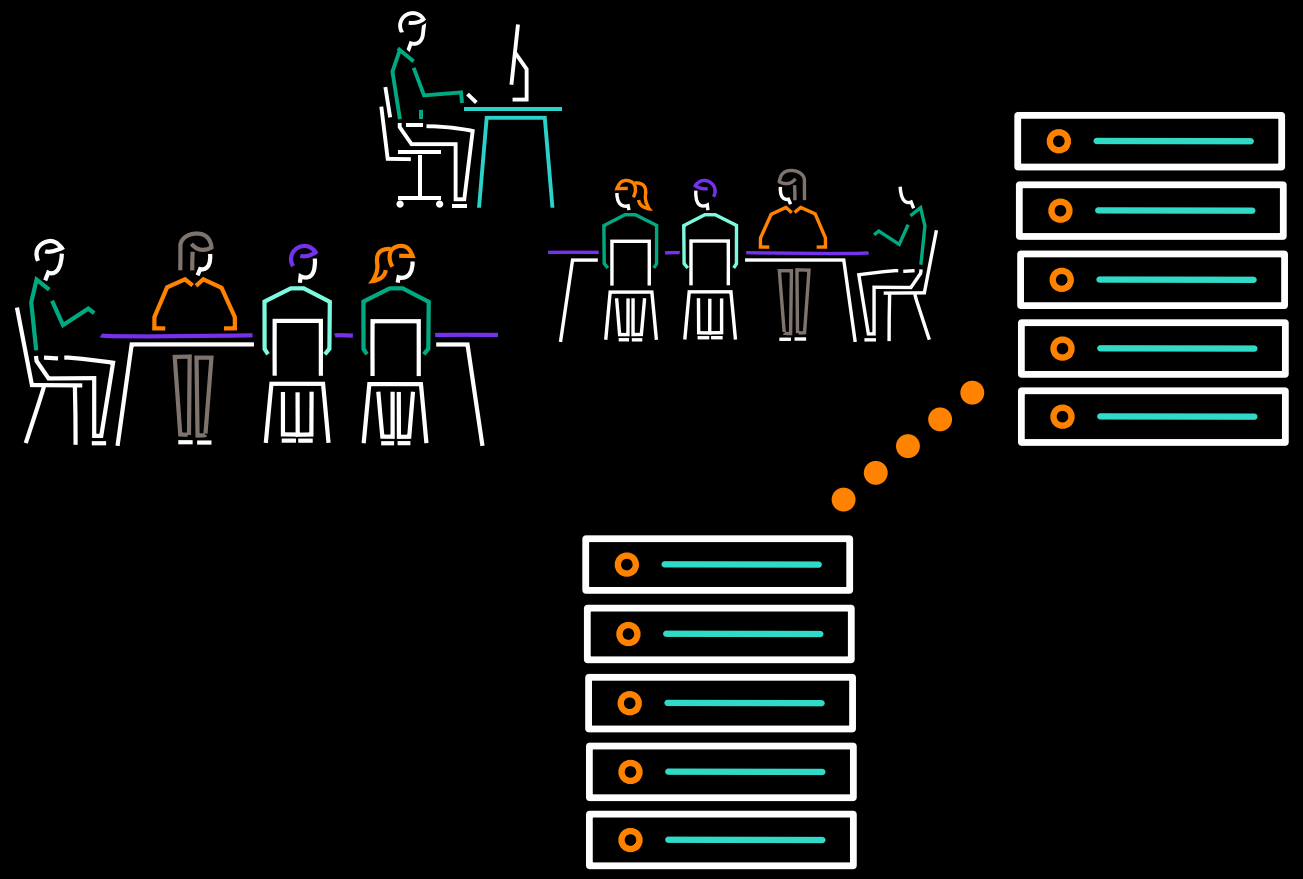


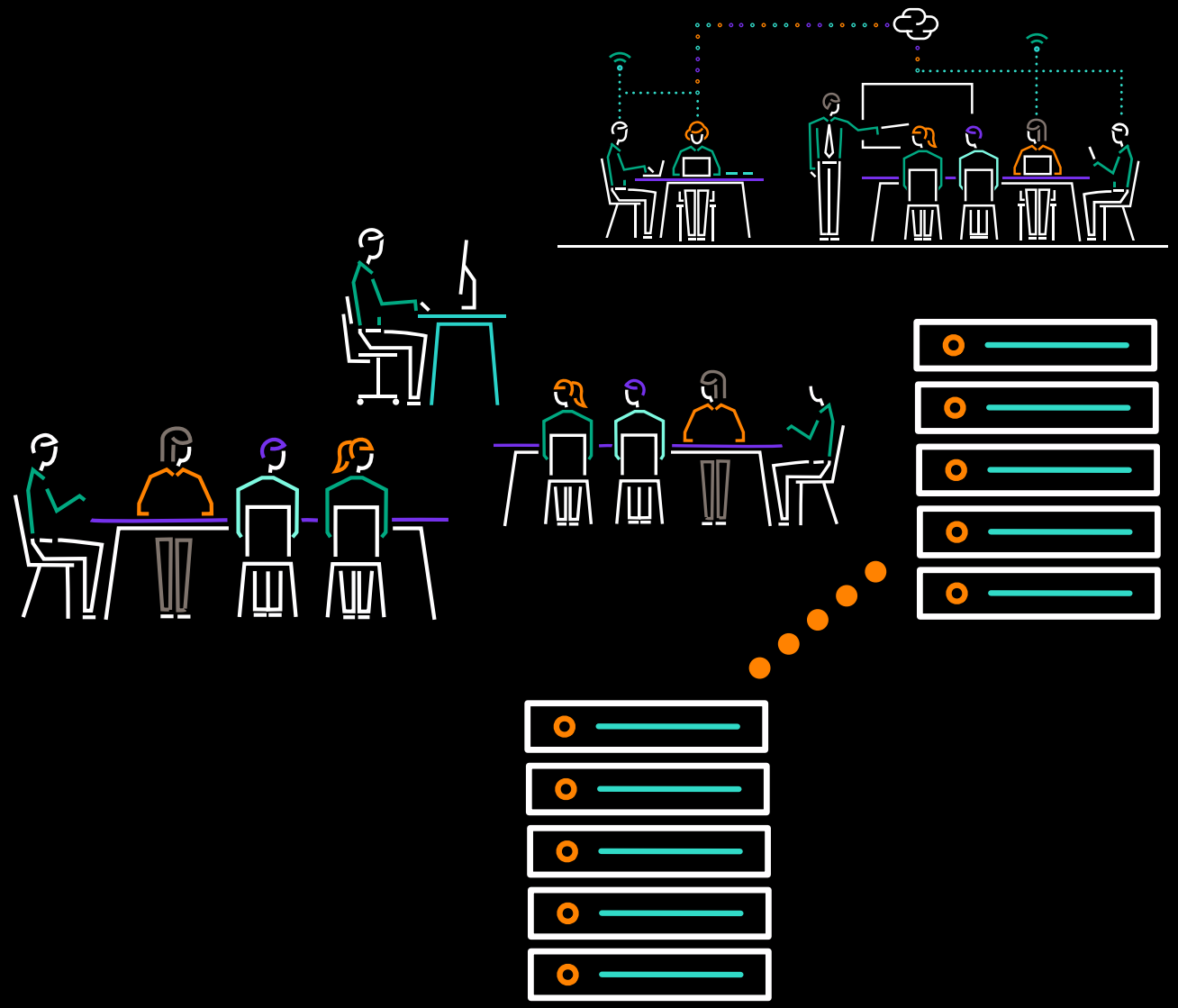
A **data scientist** builds models that will make future decisions



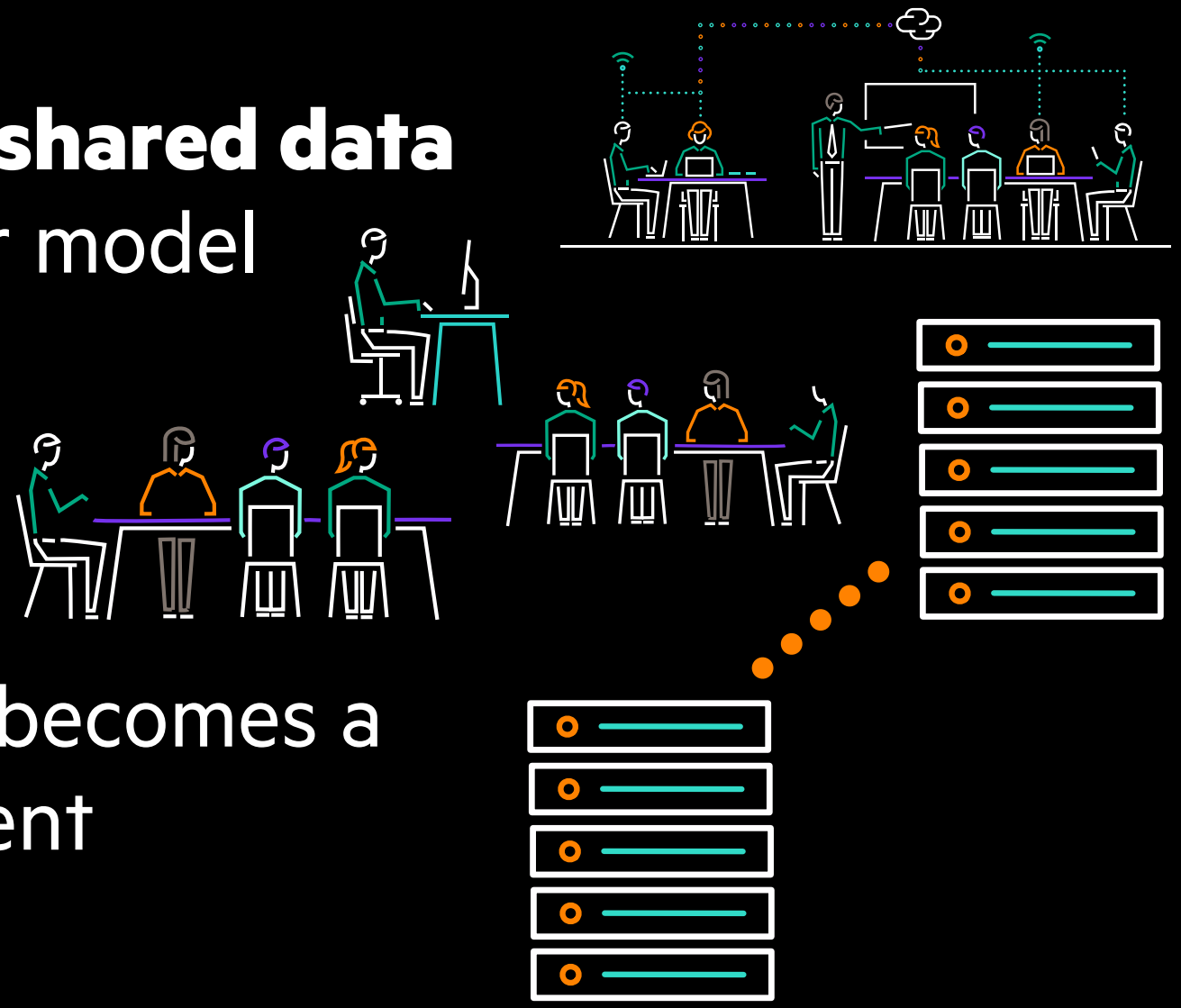
In fact, there may be thousands of data scientists all building models



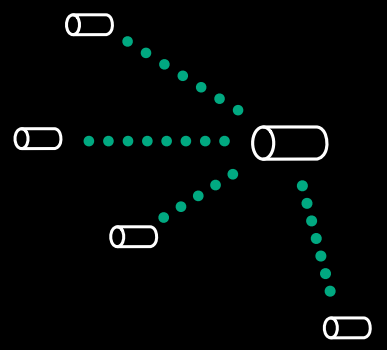




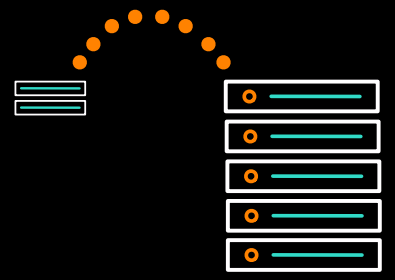
Working on **shared data** allows faster model learning



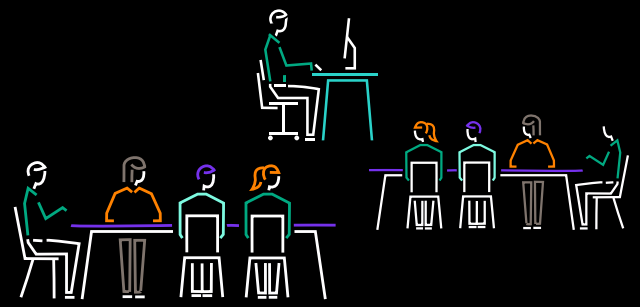
Multitenancy becomes a key requirement



Transparent data motion



Scale and speed

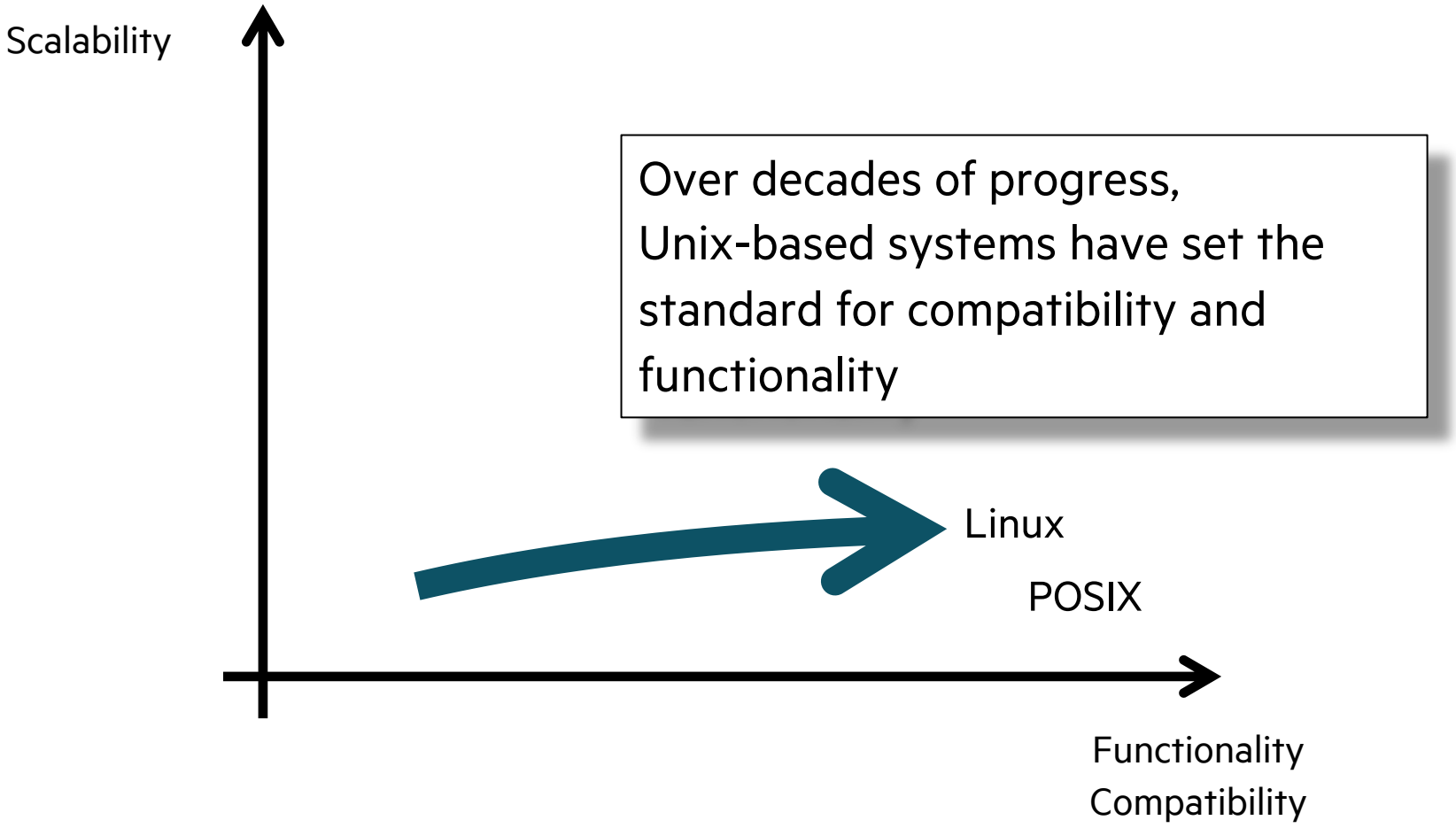


Sharing safely

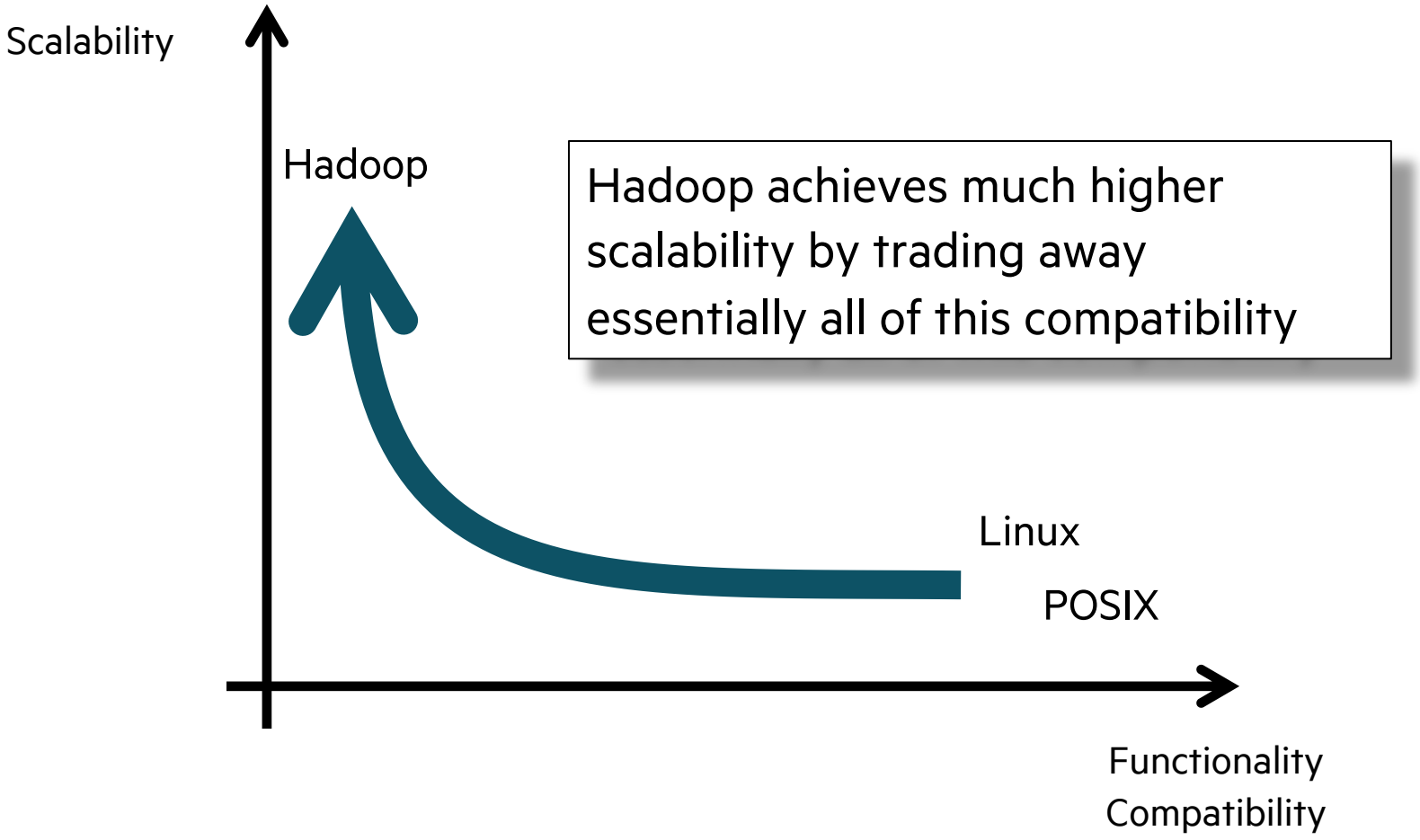
RUN THE RIGHT APPLICATIONS AT THE RIGHT TIME IN THE RIGHT PLACE AGAINST THE RIGHT DATA

Let’s dig in
and see some details

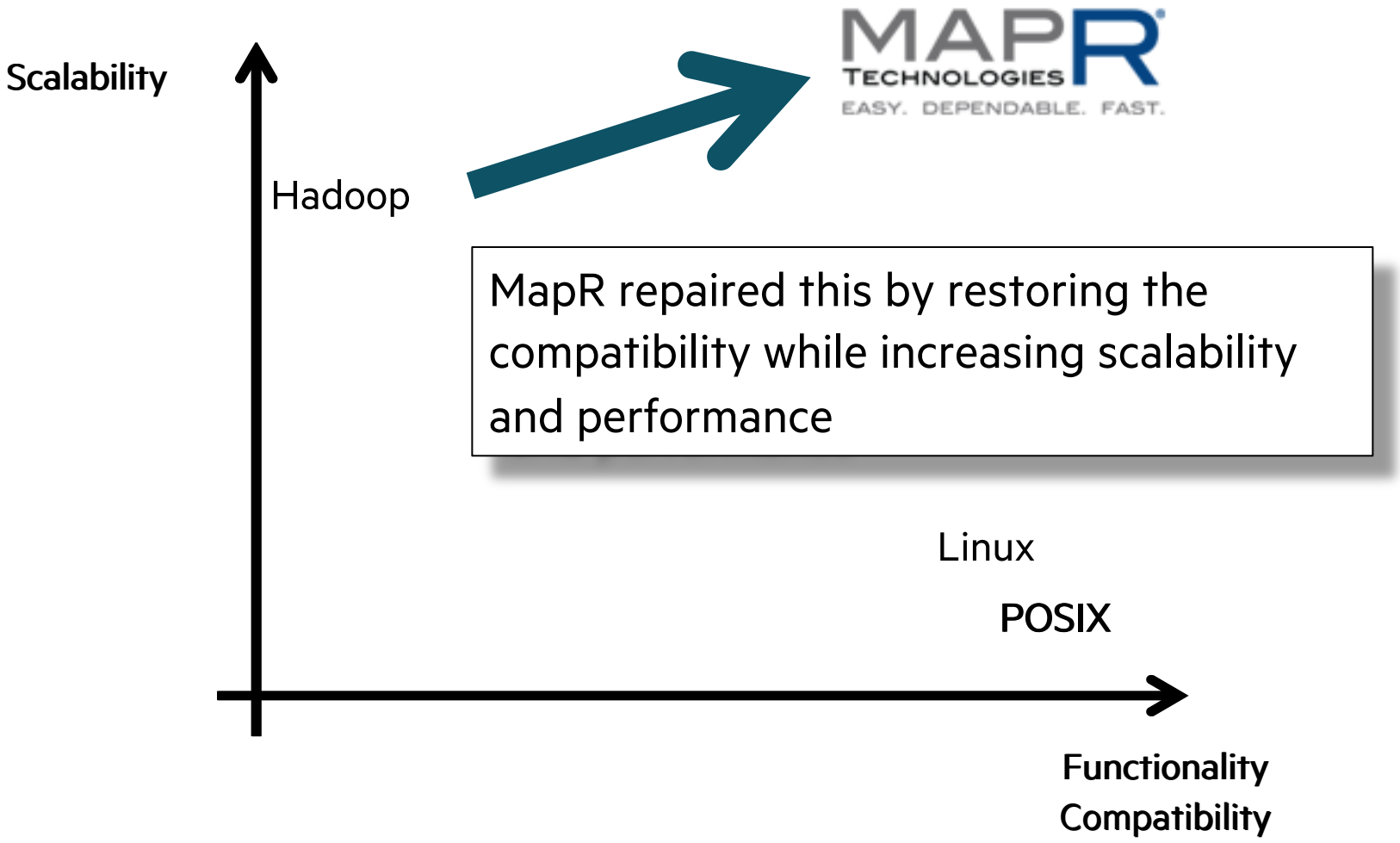
EVOLUTION OF DATA STORAGE



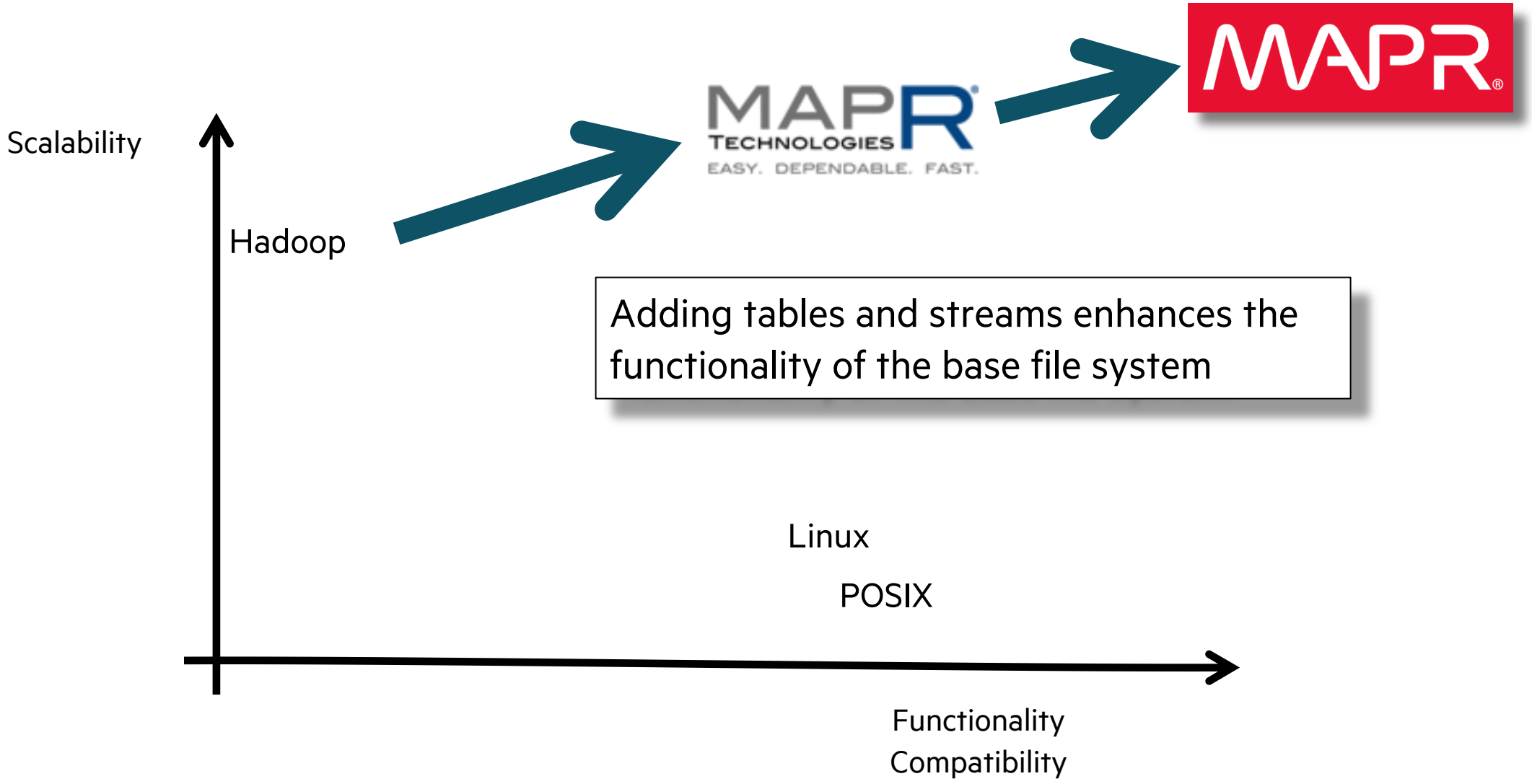
EVOLUTION OF DATA STORAGE



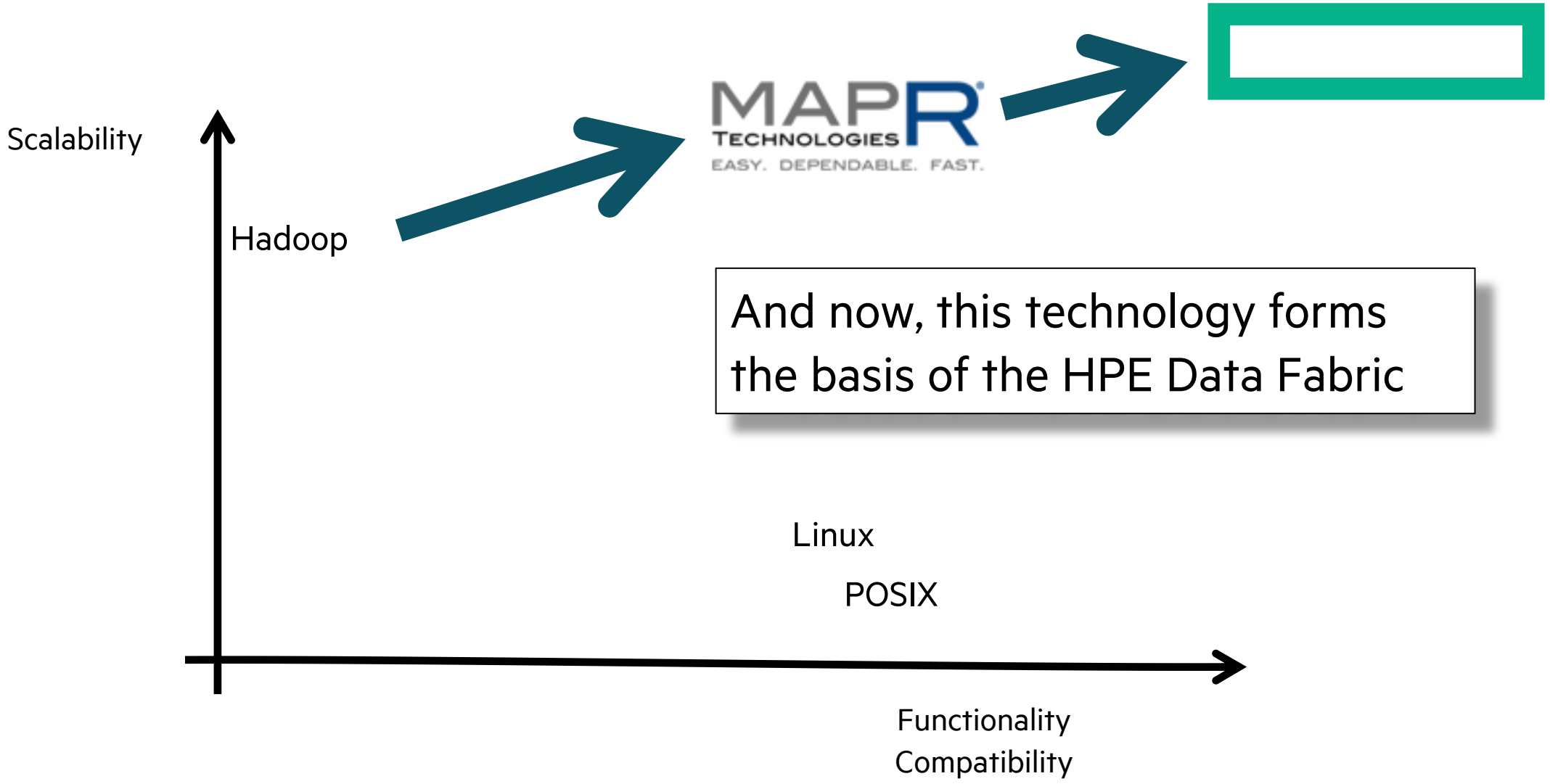
EVOLUTION OF DATA STORAGE



EVOLUTION OF DATA STORAGE



EVOLUTION OF DATA STORAGE



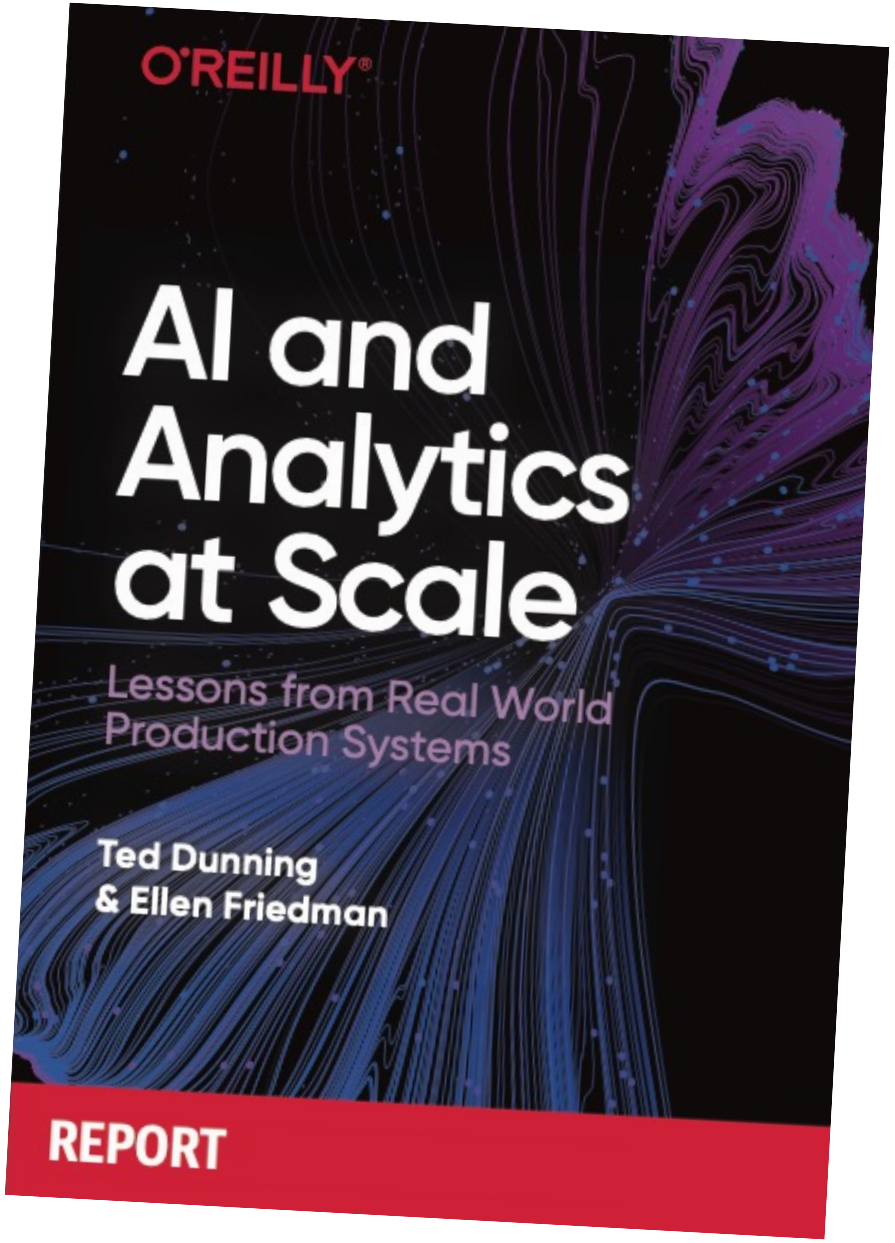
And now, this technology forms the basis of the HPE Data Fabric

New tools for a new world of data

NEW BOOK COMING

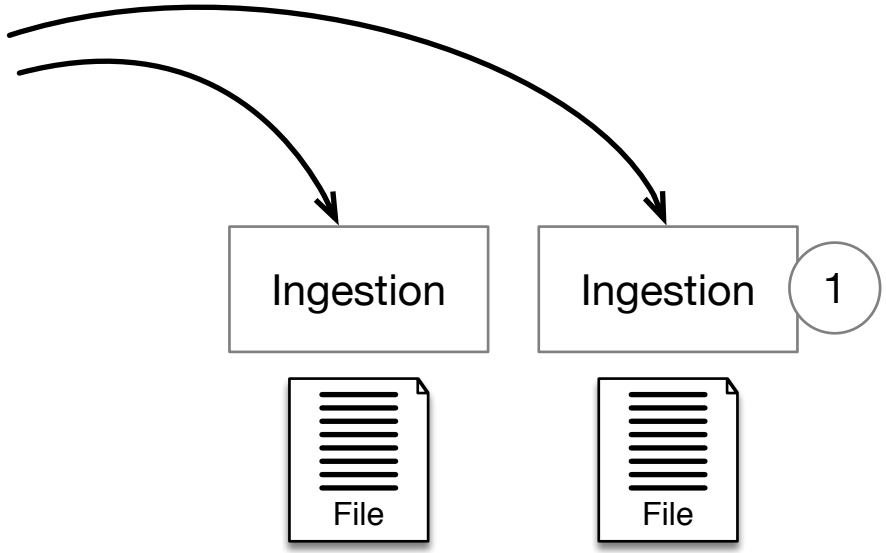
The tradeoffs people assume are unavoidable, in fact, are not.

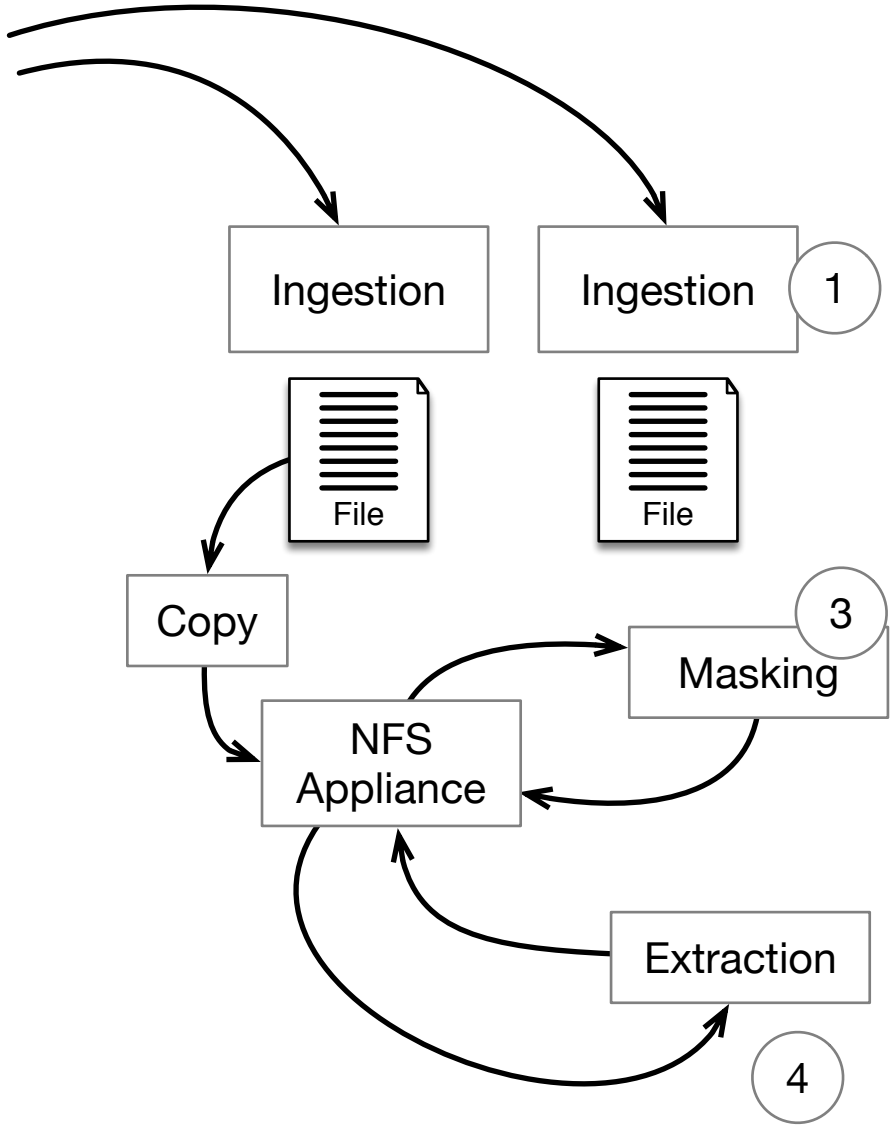
A unified data infrastructure can free you from these tradeoffs.

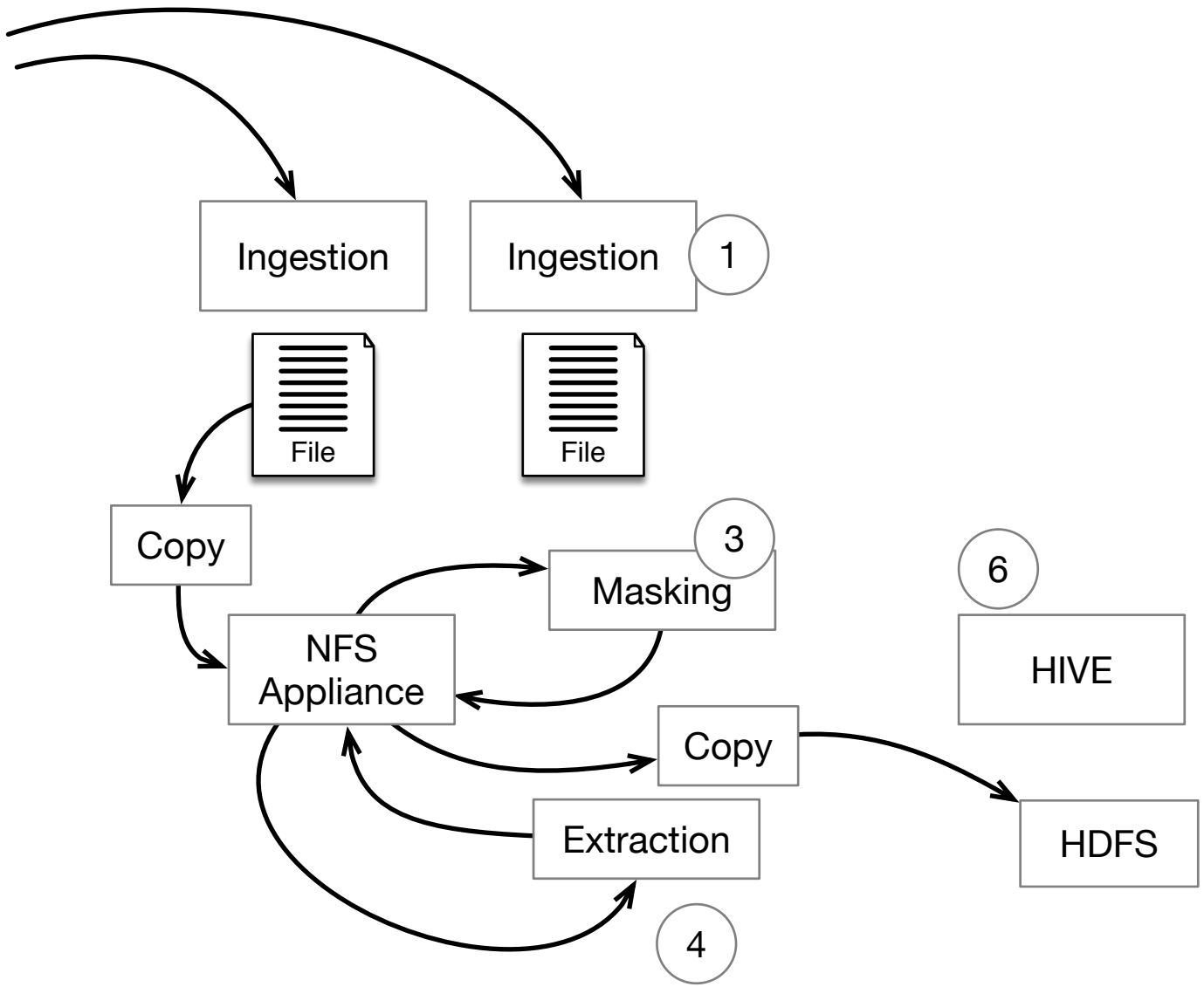


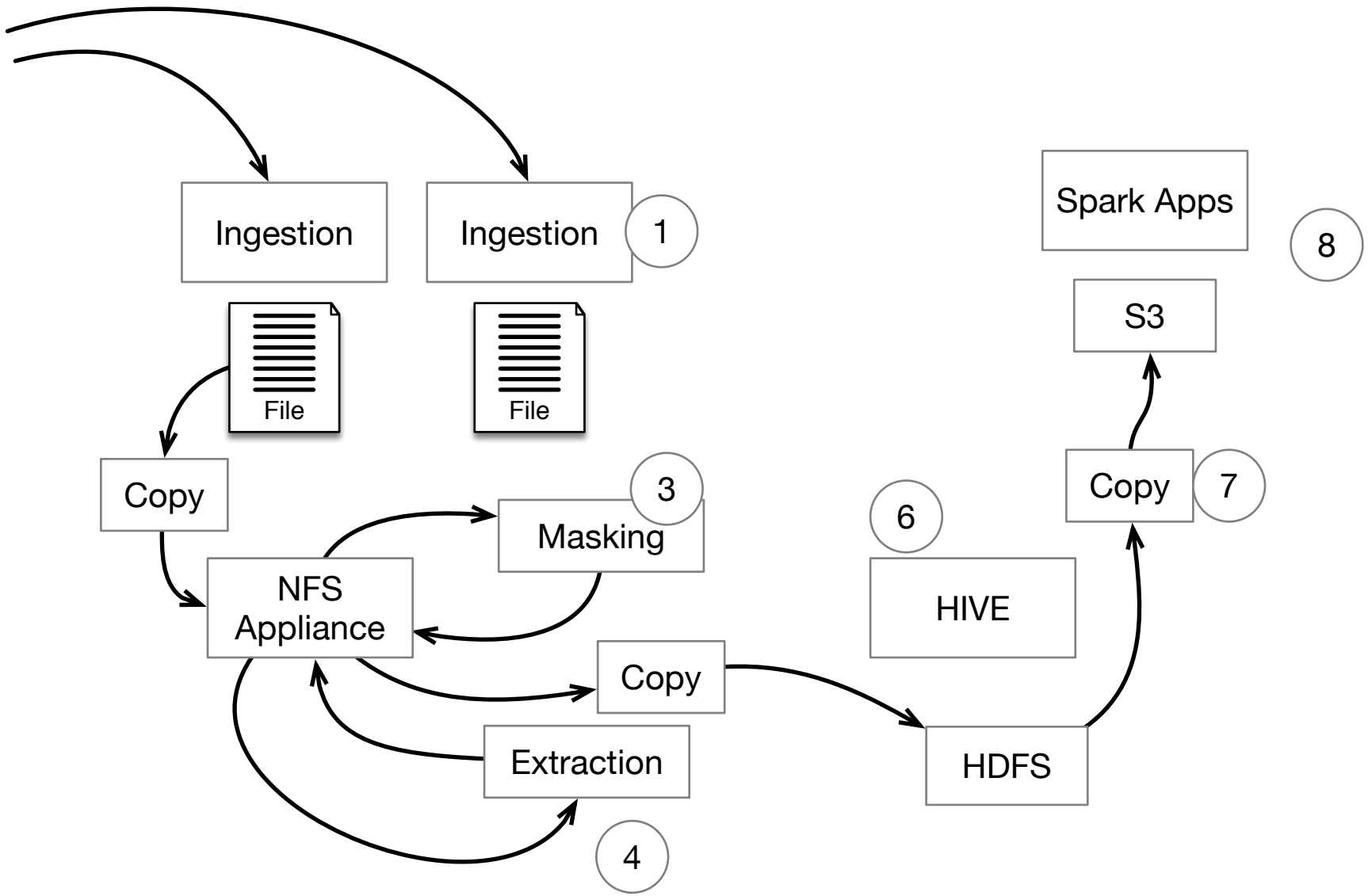
A RECENT EXAMPLE

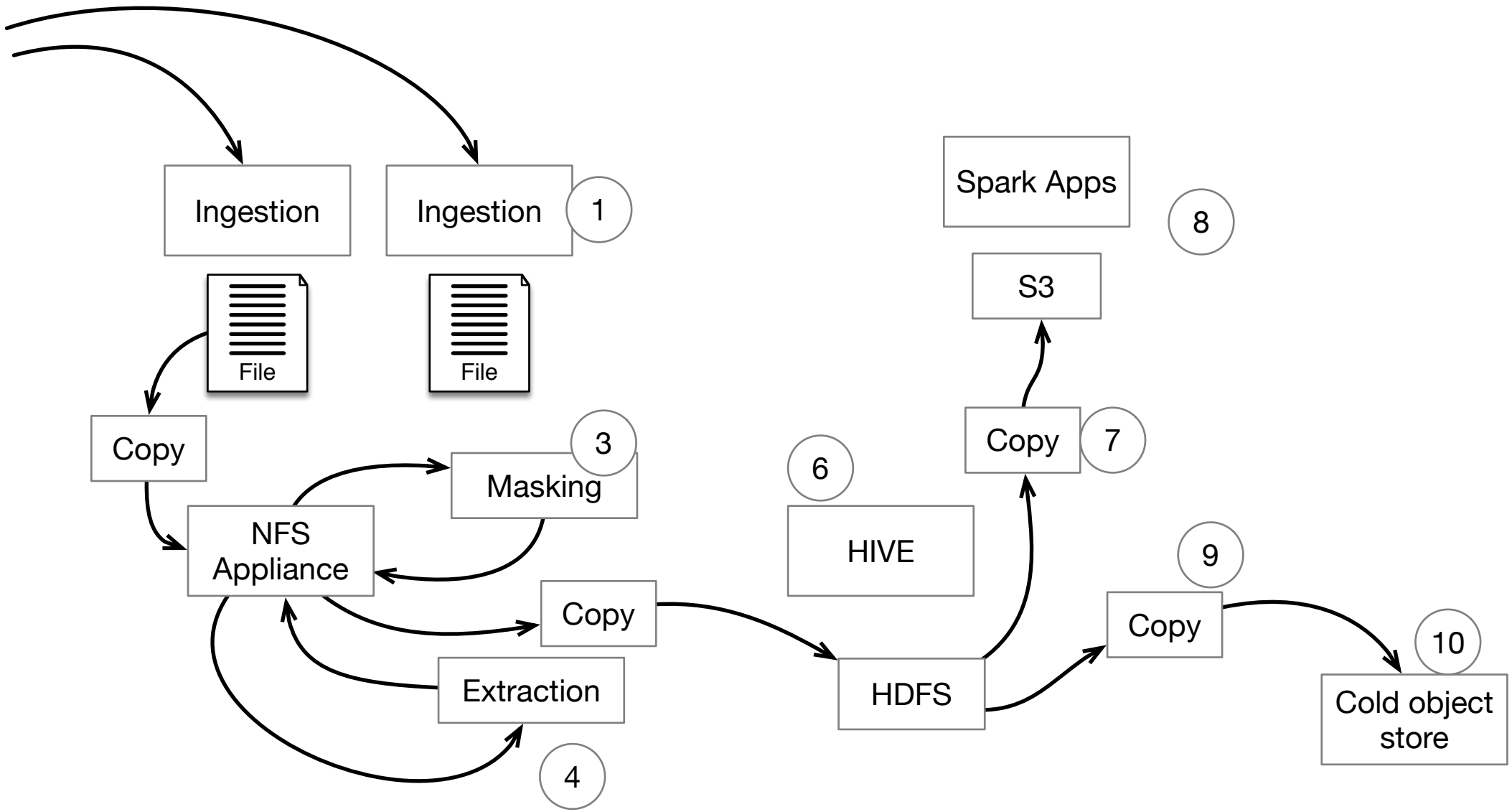
How ***you*** can build these systems

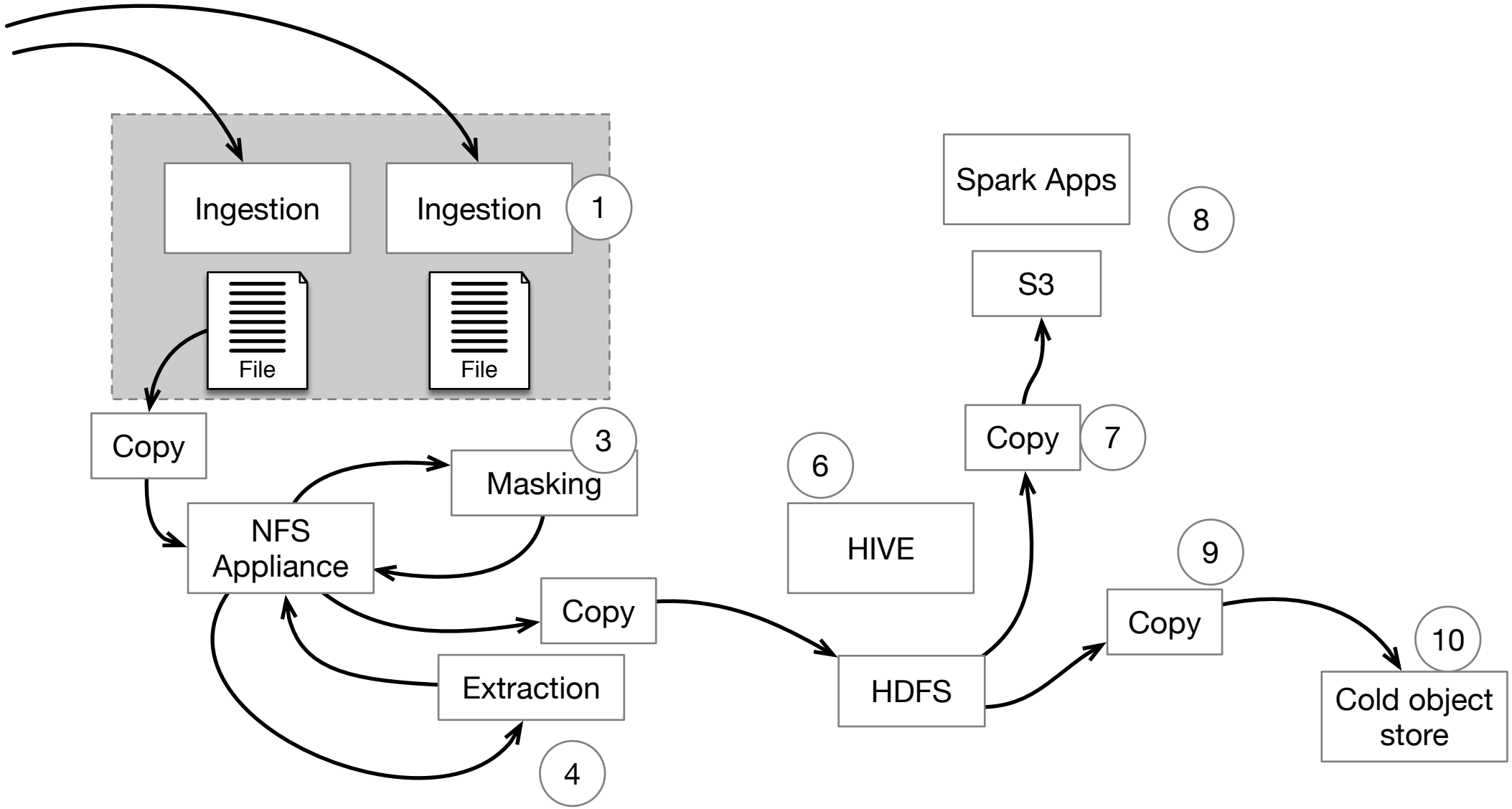

Ingestion
Ingestion
1
File
File

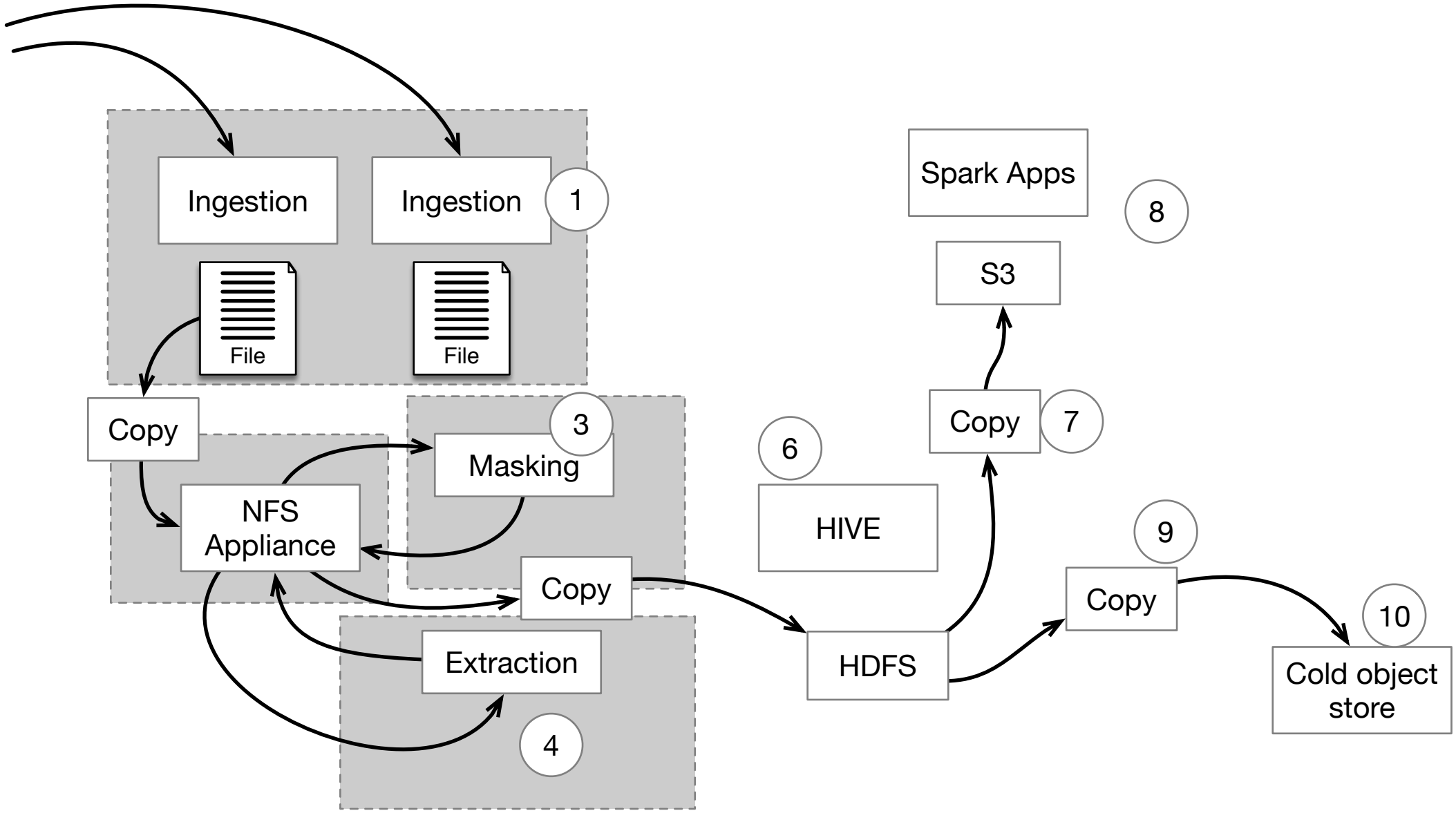

Ingestion
Ingestion
1
File
File
Copy
3
Masking
NFS
Appliance
Extraction
4

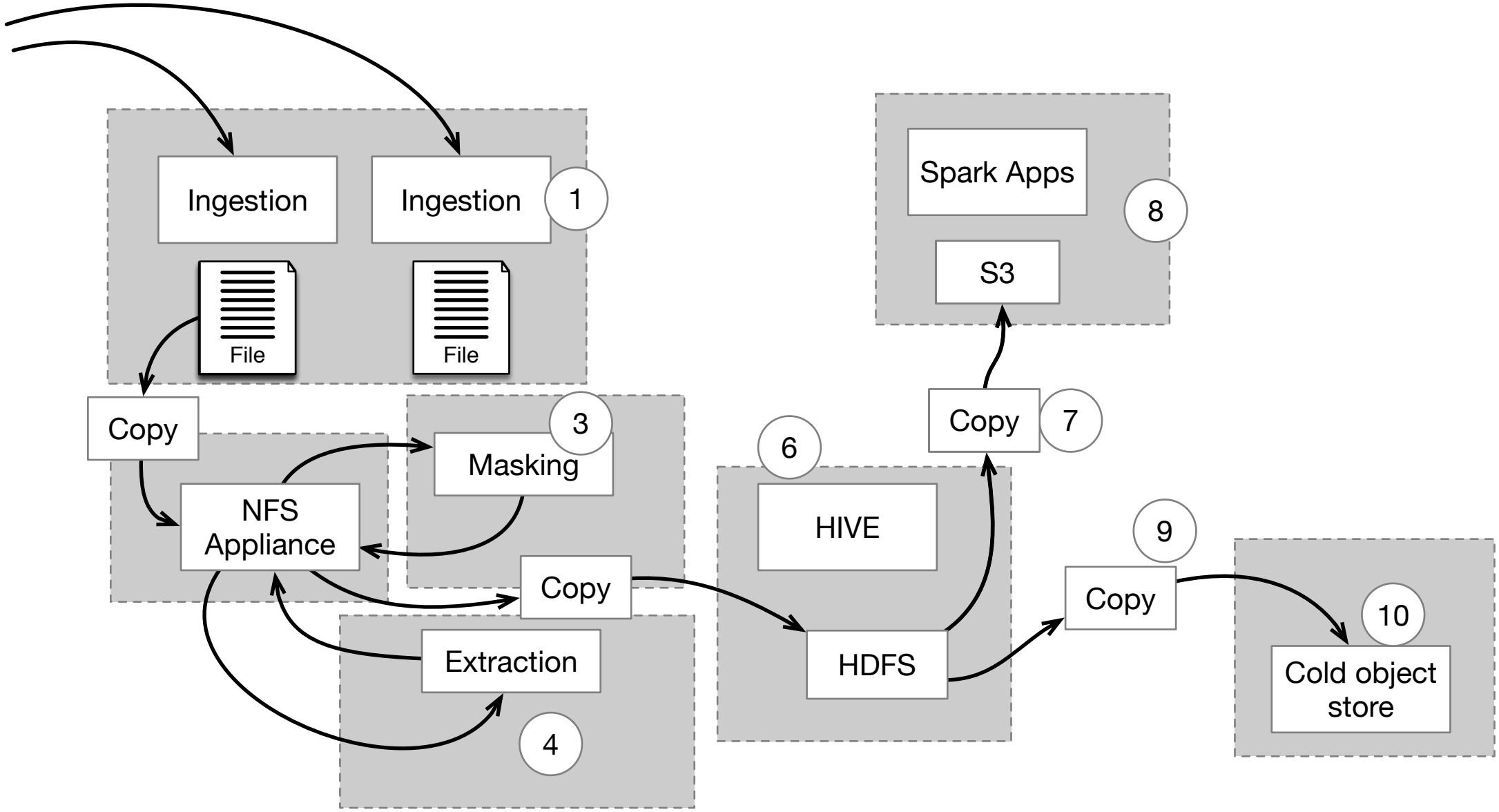

Ingestion
Ingestion
1
File
File
Copy
3
Masking
6
NFS
Appliance
HIVE
Copy
Extraction
HDFS
4

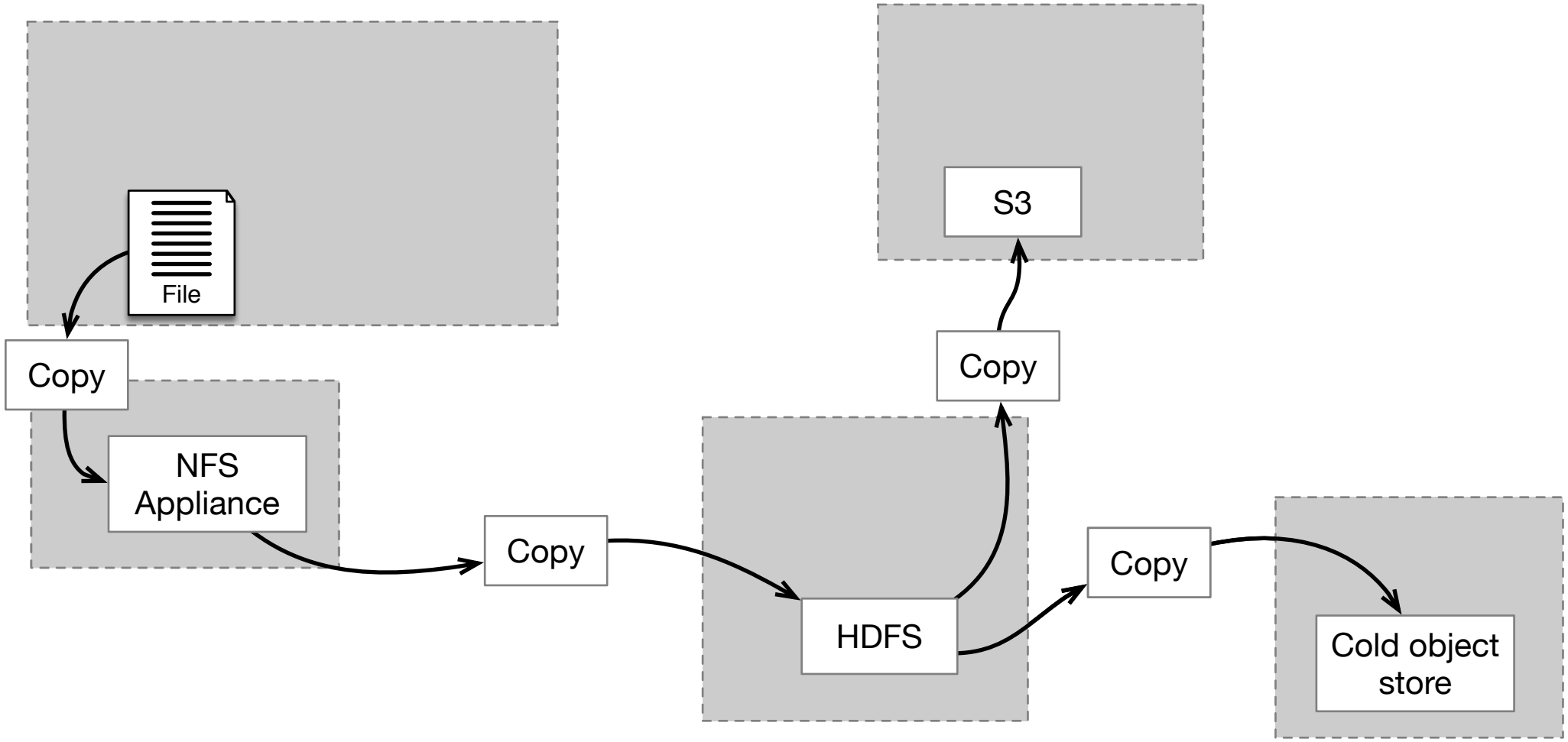

Ingestion
Ingestion
1
File
File
Copy
NFS
Appliance
Masking
3
Copy
Extraction
4
HDFS
HIVE
6
Copy
7
S3
Spark Apps
8

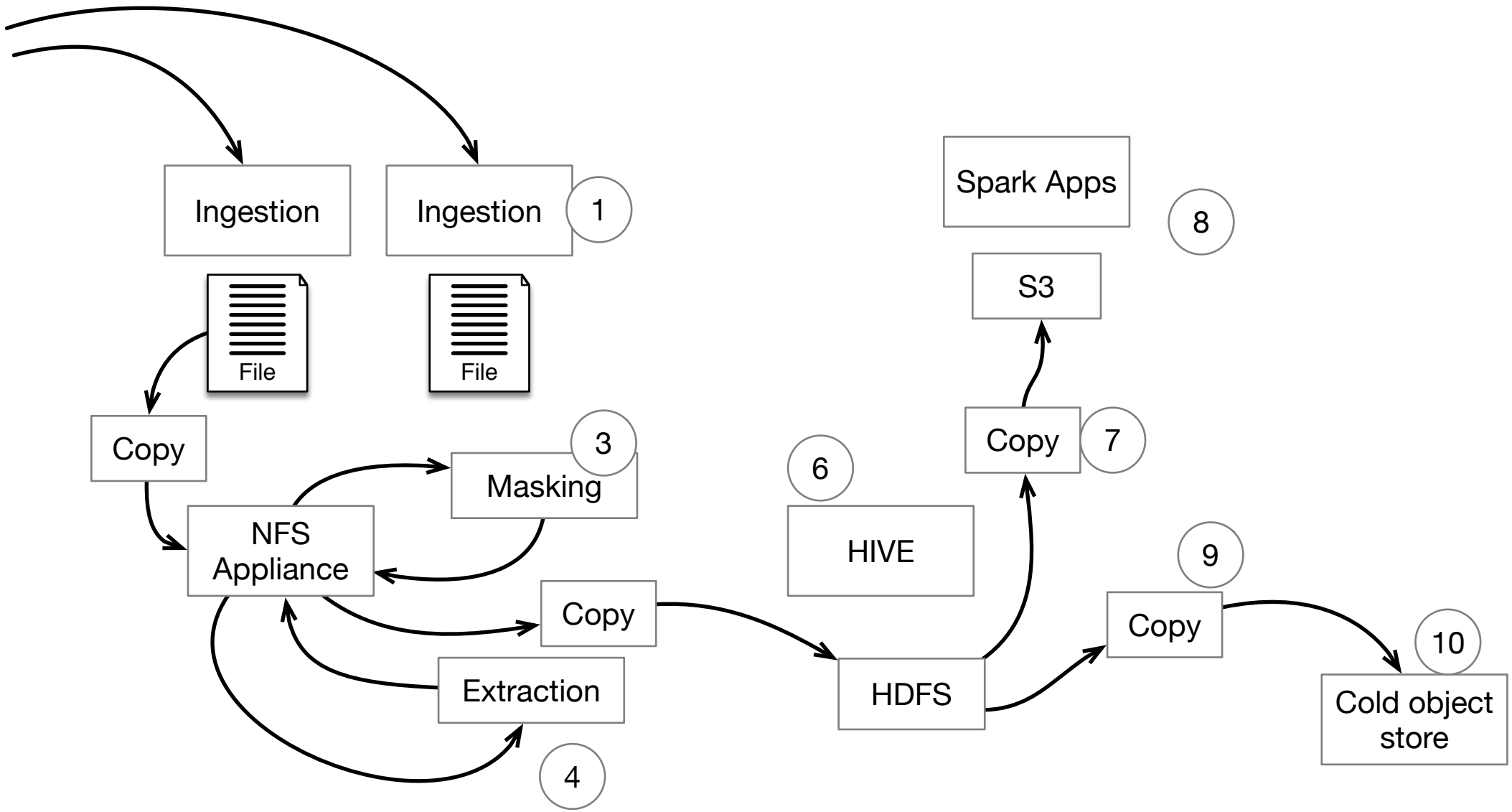

Ingestion
Ingestion
1
File
File
Copy
NFS Appliance
Masking
3
Copy
Extraction
4
HDFS
6
HIVE
Copy
7
S3
Spark Apps
8
Copy
9
Cold object store
10

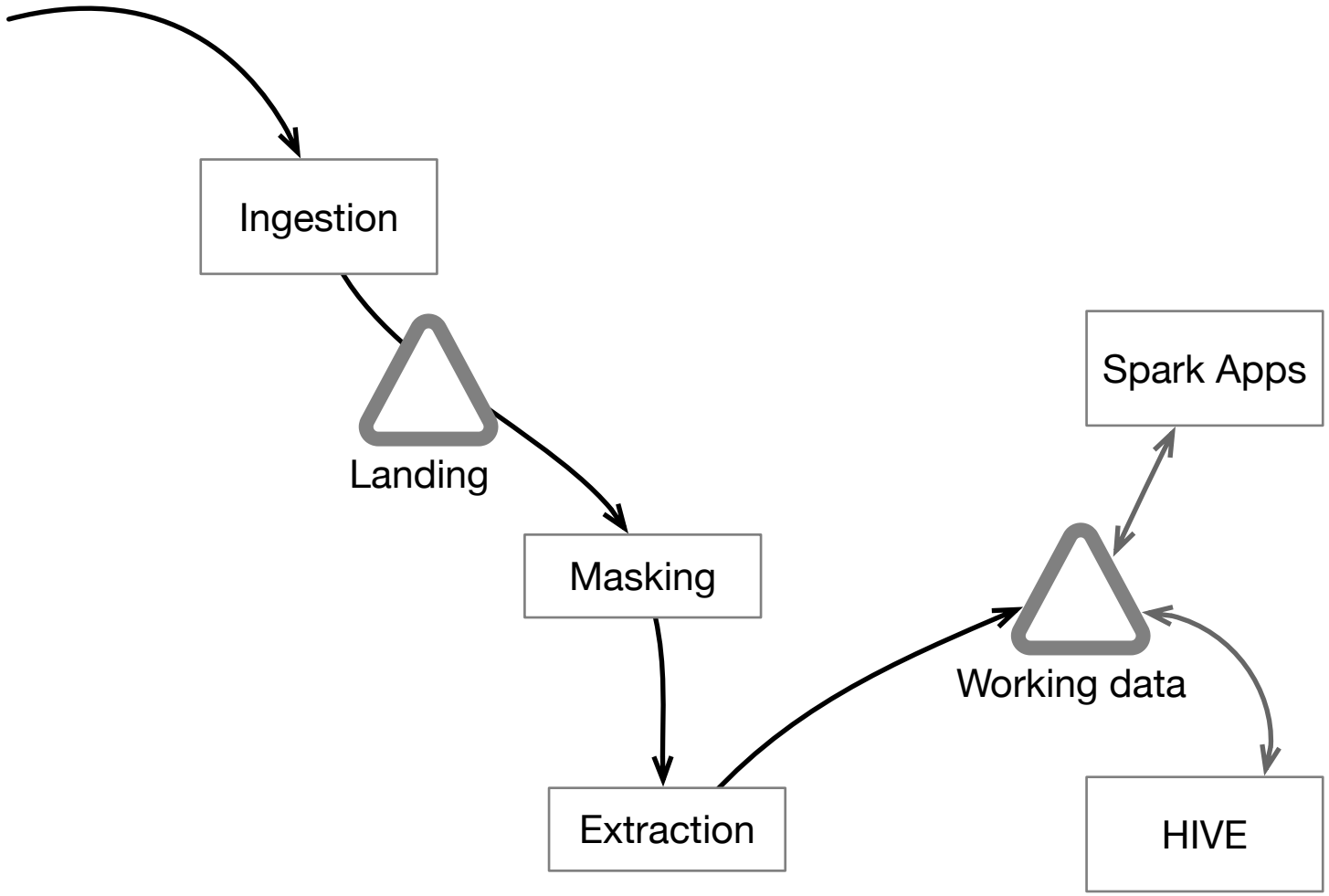

Ingestion
Ingestion
1
File
File
Copy
3
Masking
NFS Appliance
Copy
Extraction
4
6
HIVE
HDFS
Spark Apps
8
S3
Copy
7
Copy
9
10
Cold object store

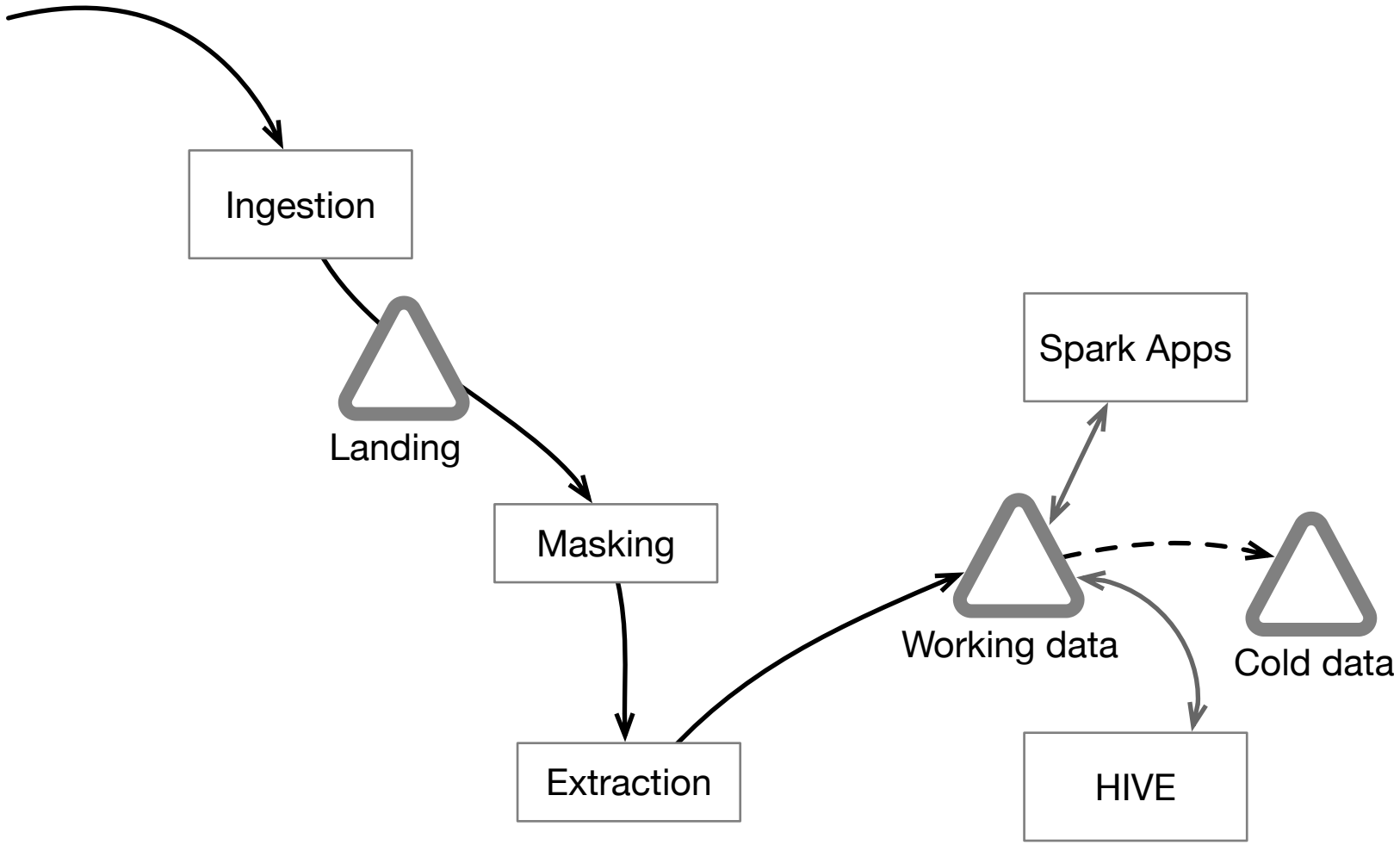

Ingestion
Ingestion
1
File
File
Copy
NFS Appliance
3
Masking
Copy
Extraction
4
6
HIVE
HDFS
Spark Apps
8
S3
Copy
7
Copy
9
10
Cold object store

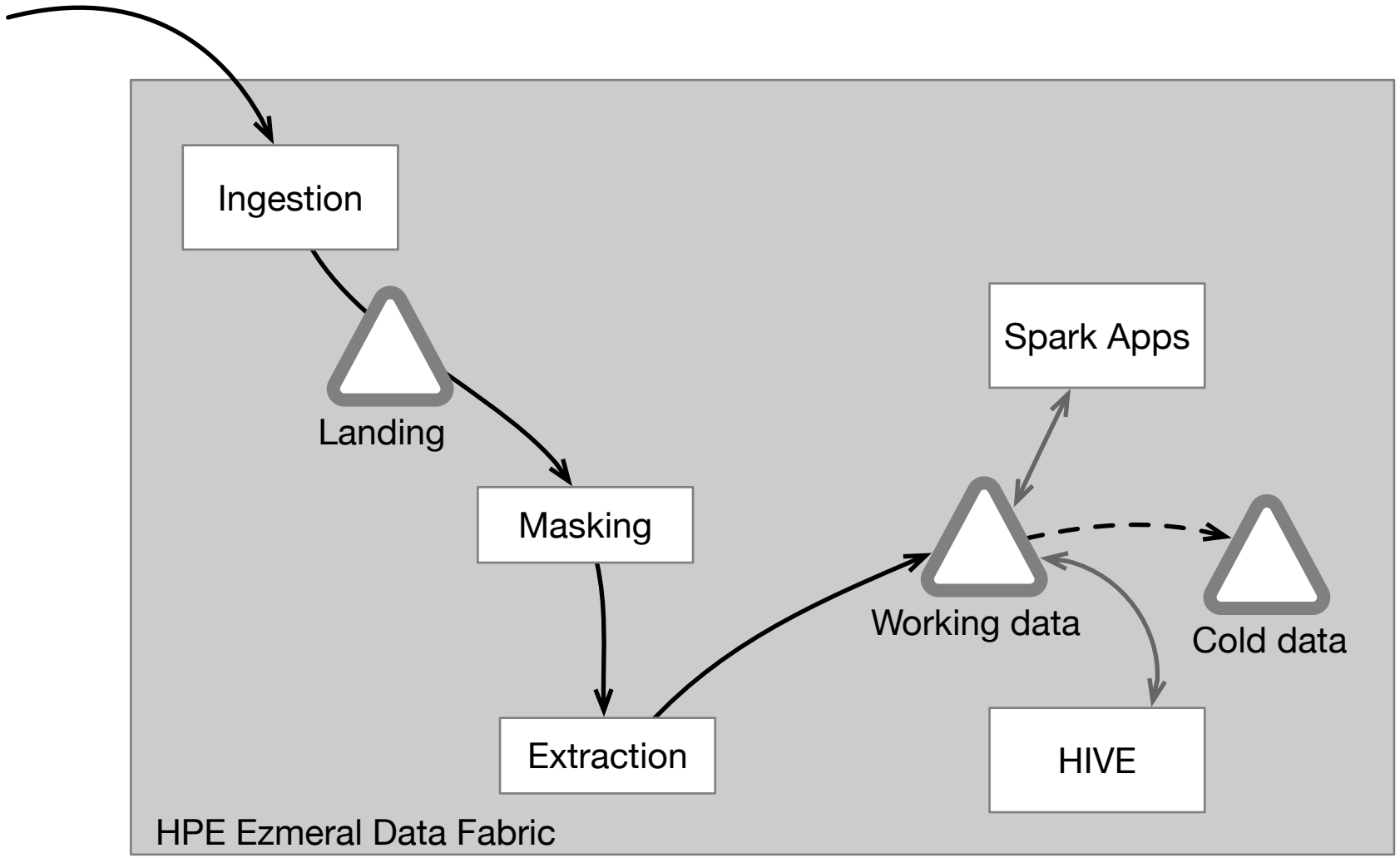

Ingestion
Ingestion
1
File
File
Copy
NFS Appliance
3
Masking
Copy
Extraction
4
6
HIVE
HDFS
Copy
7
Spark Apps
8
S3
9
Copy
10
Cold object store


File
Copy
NFS
Appliance
Copy
HDFS
Copy
S3
Copy
Cold object
store

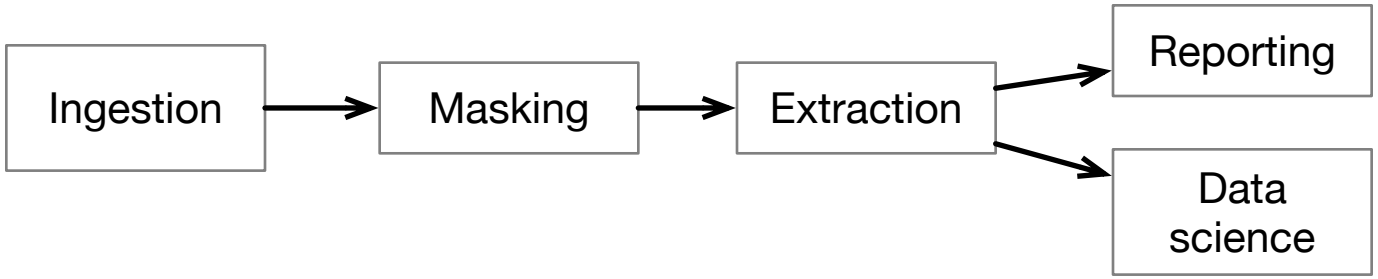

Ingestion
Ingestion
1
File
File
Copy
NFS Appliance
Masking
3
Copy
Extraction
4
HDFS
6
HIVE
Copy
7
S3
Spark Apps
8
Copy
9
Cold object store
10


Ingestion
Landing
Masking
Extraction
Working data
Spark Apps
HIVE

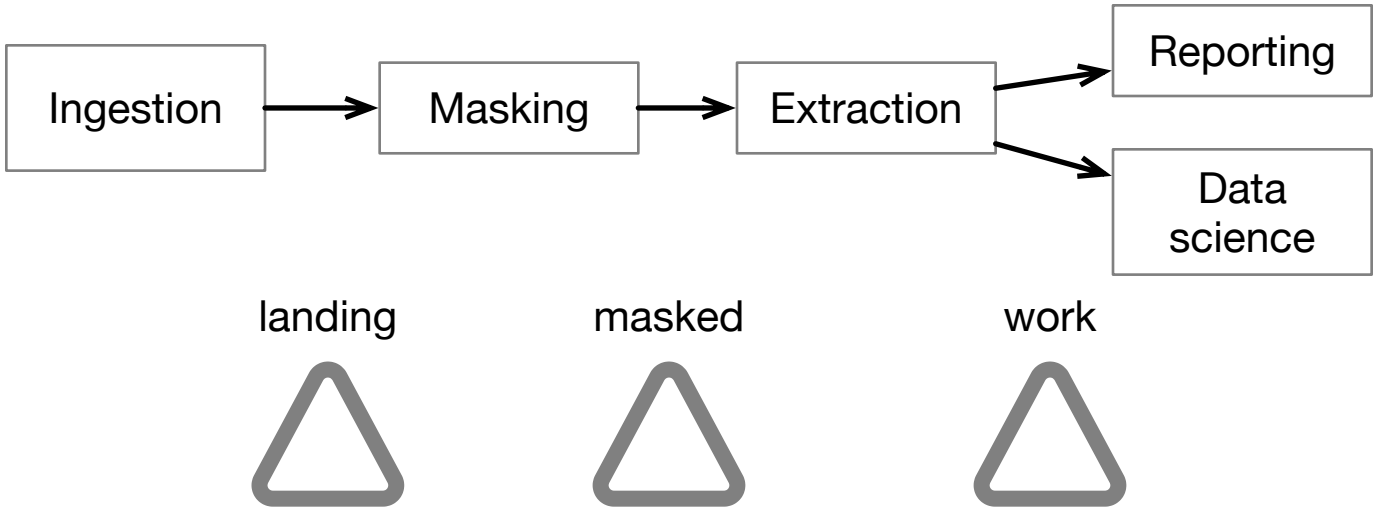

Ingestion
Landing
Masking
Extraction
Spark Apps
Working data
Cold data
HIVE


Ingestion
Landing
Masking
Extraction
Spark Apps
Working data
Cold data
HIVE
HPE Ezmeral Data Fabric

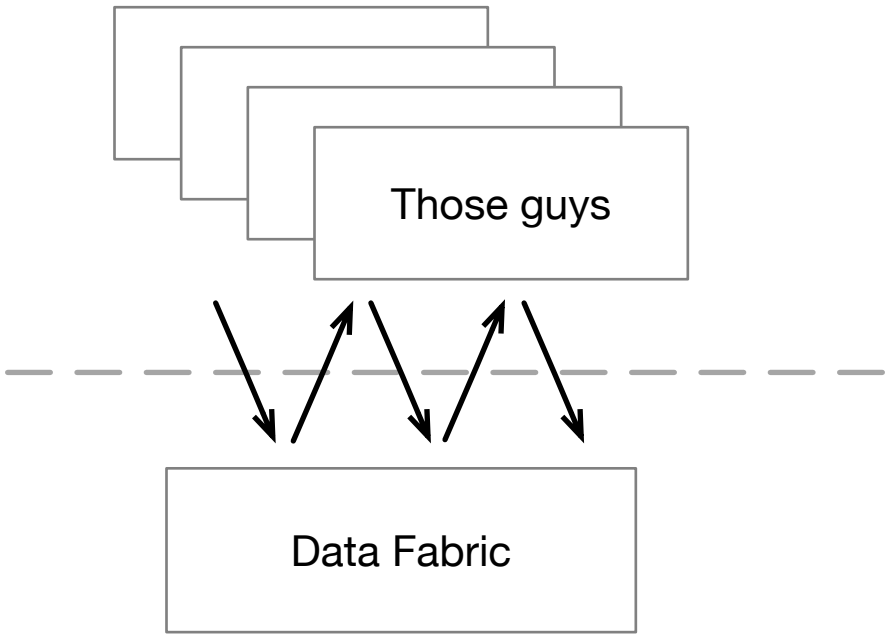
DEVOPS VIEW



DEVOP VIEW

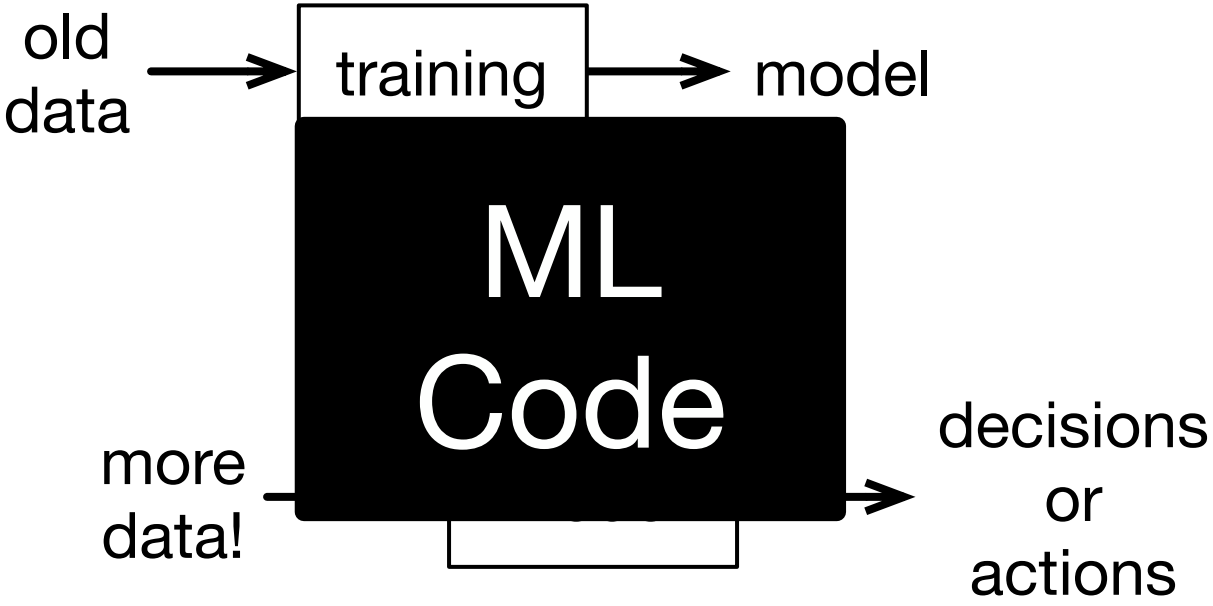

Ingestion
Masking
Extraction
Reporting
Data science
landing
masked
work

INFRASTRUCTURE TEAM VIEW

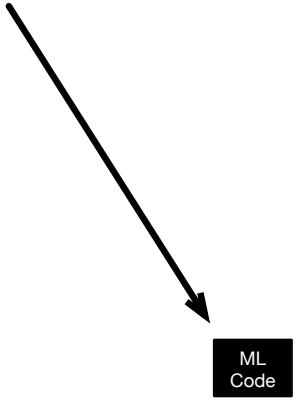


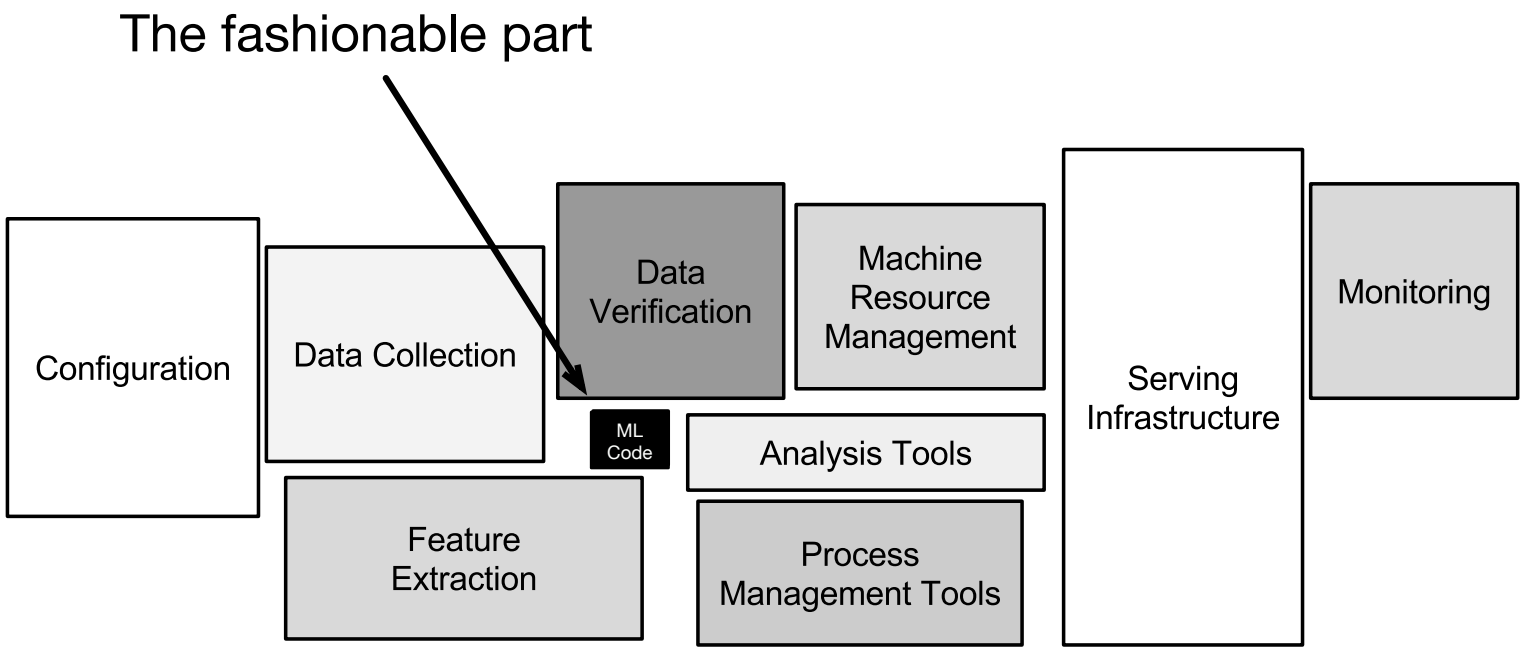
90% of the effort in successful machine learning isn’t the learning ...

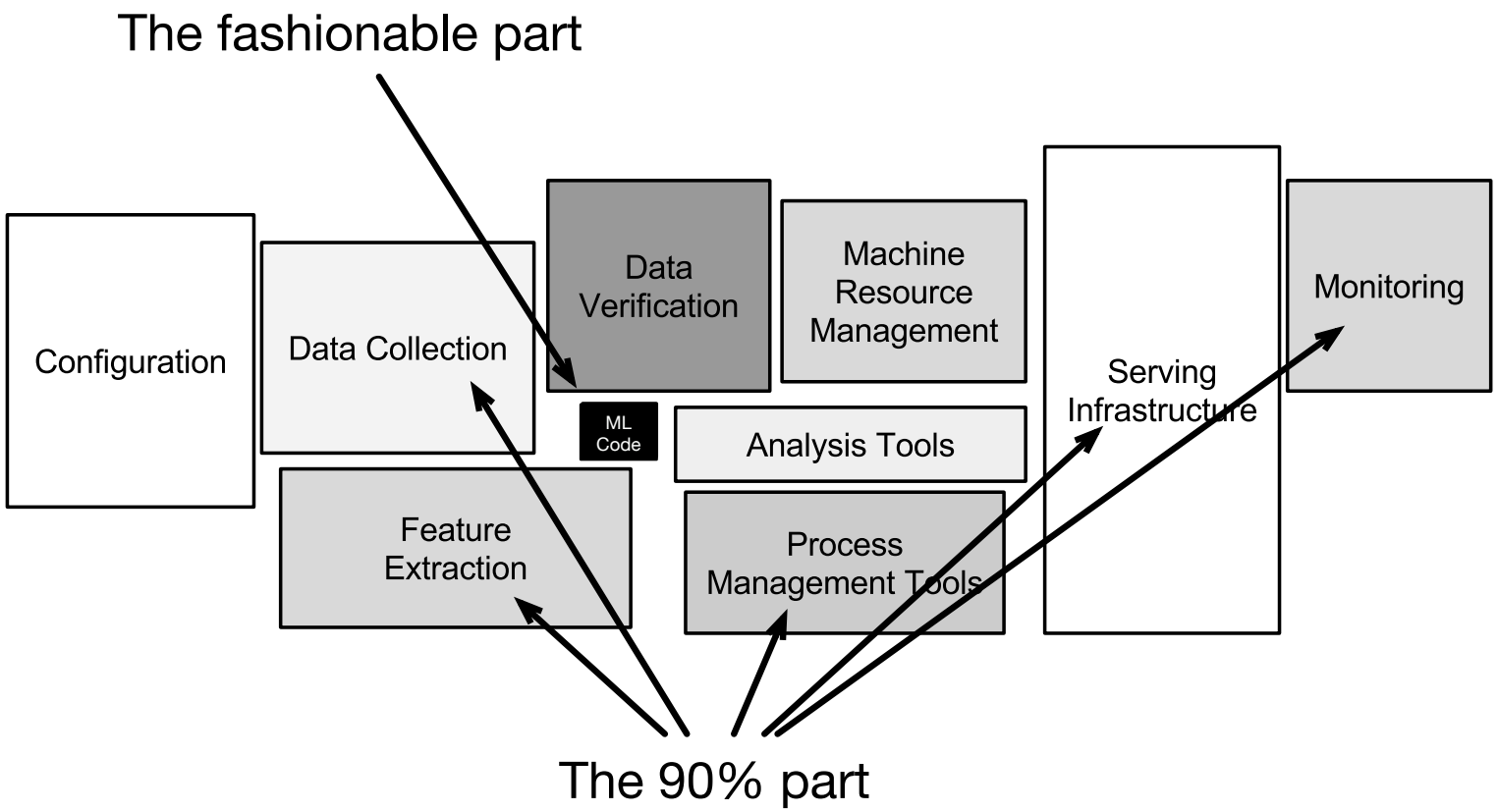
It’s the logistics

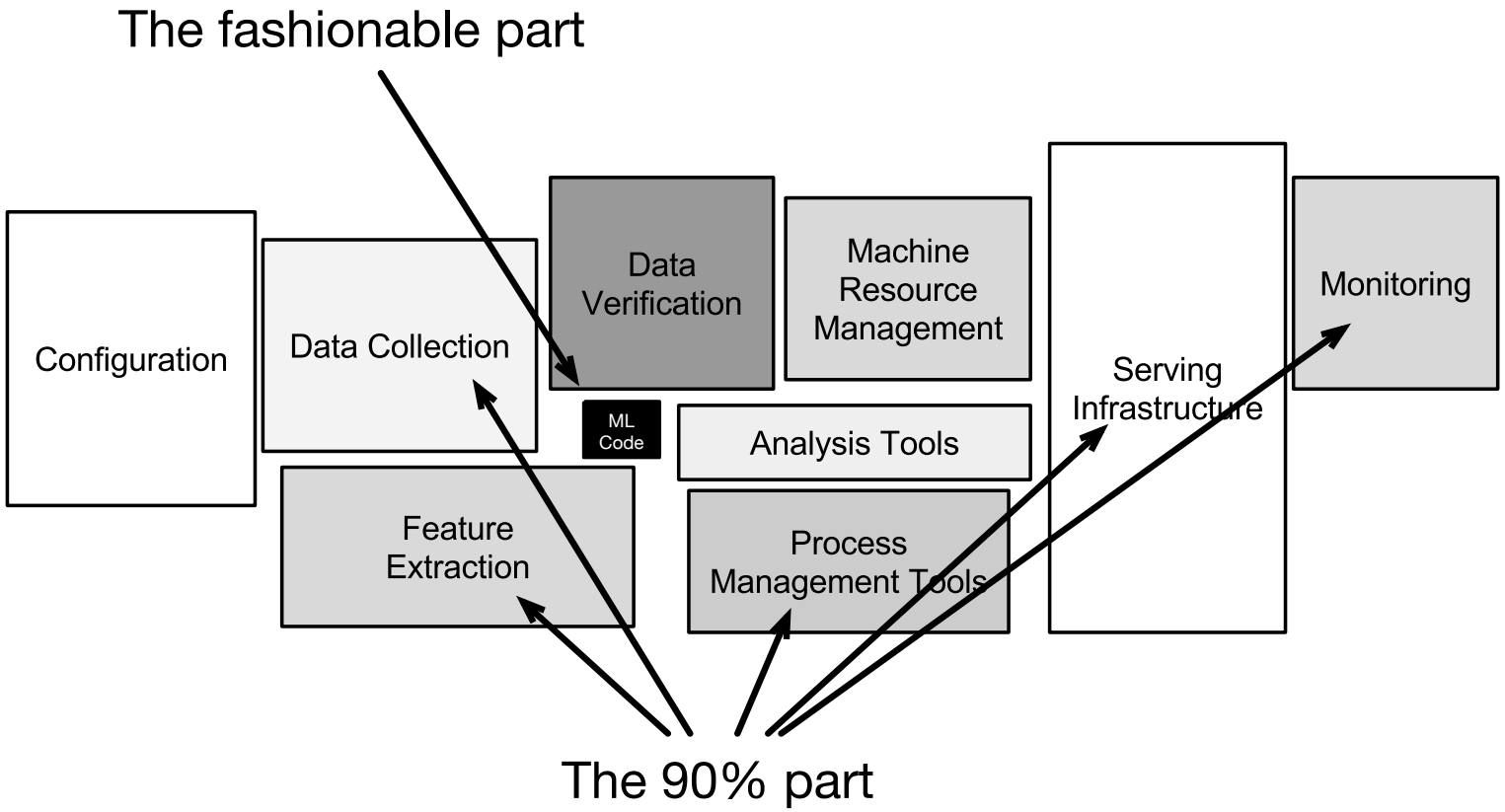

old
data
training
model
ML
Code
more
data!
decisions
or
actions

The fashionable part




The fashionable part
Configuration
Data Collection
Data Verification
Machine Resource Management
Serving Infrastructure
Monitoring
ML Code
Analysis Tools
Feature Extraction
Process Management Tools

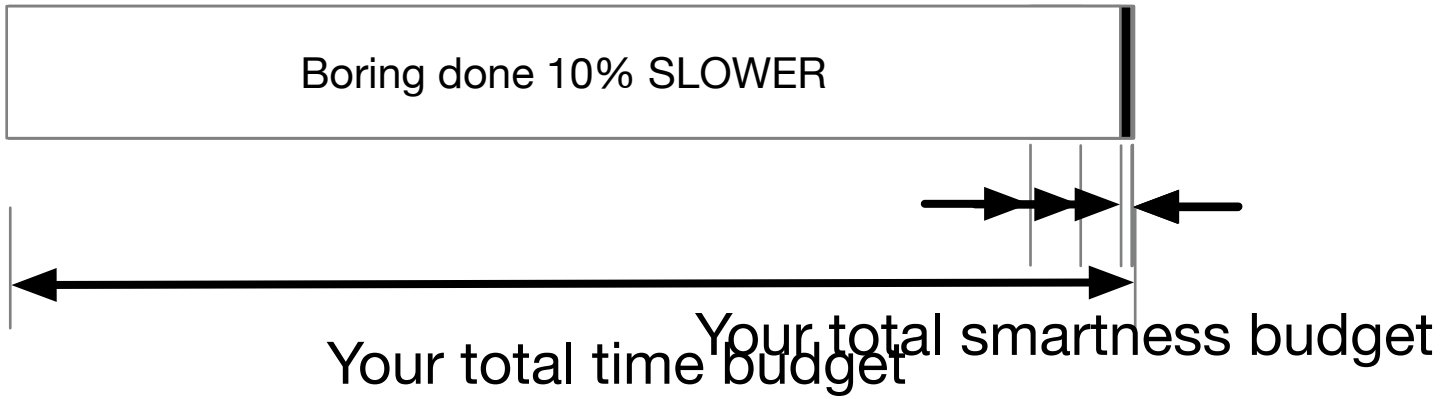

The fashionable part
Configuration
Data Collection
Data Verification
Machine Resource Management
Serving Infrastructure
Monitoring
ML Code
Analysis Tools
Feature Extraction
Process Management Tools
The 90% part



Hidden Technical Debt in Machine Learning Systems
https://papers.nips.cc/paper/5656-hidden-technical-debt-in-machine-learning-systems.pdf

So why focus on the boring bits?

So why focus on the boring bits?



So why focus on the boring bits?

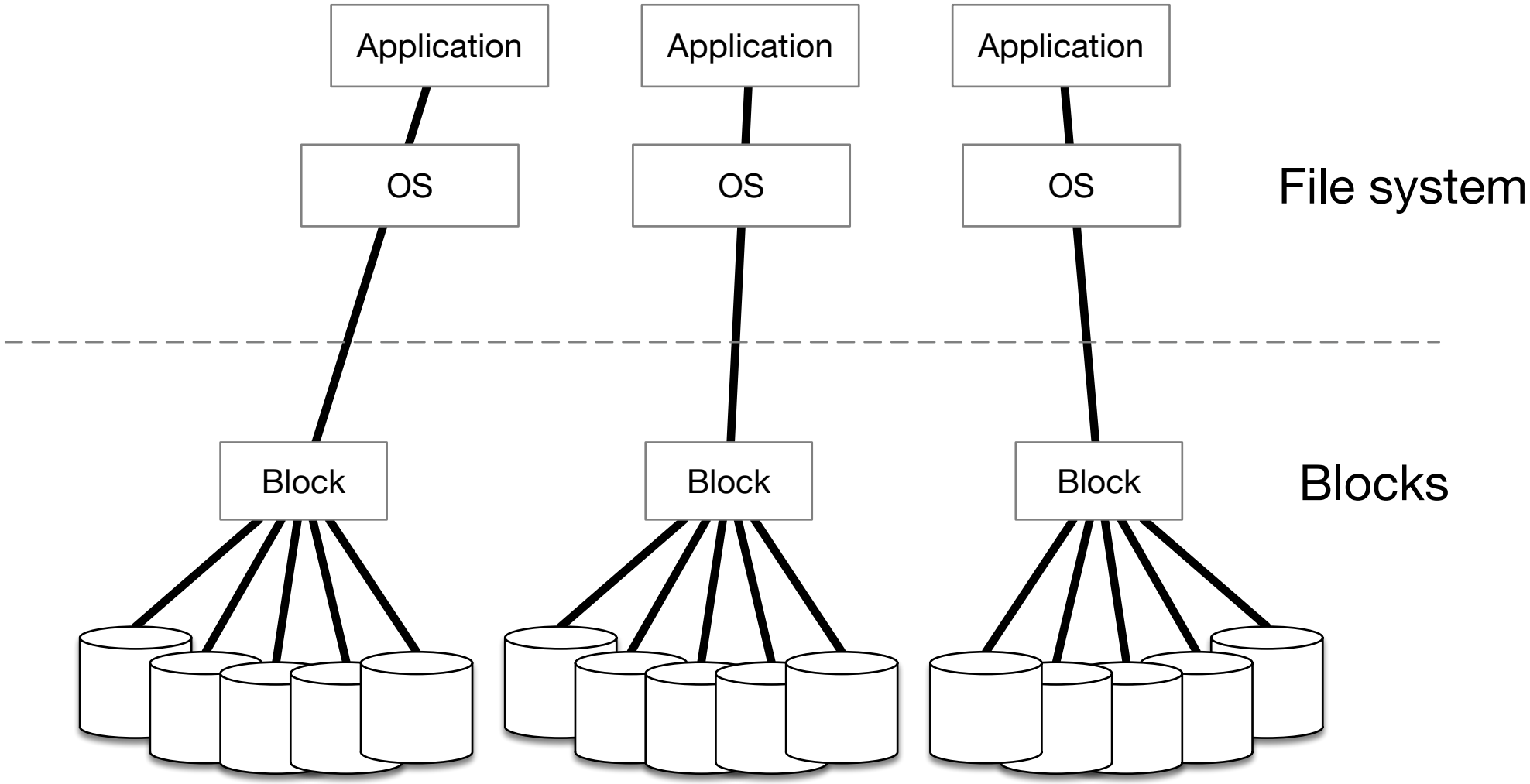
So why focus on the boring bits?

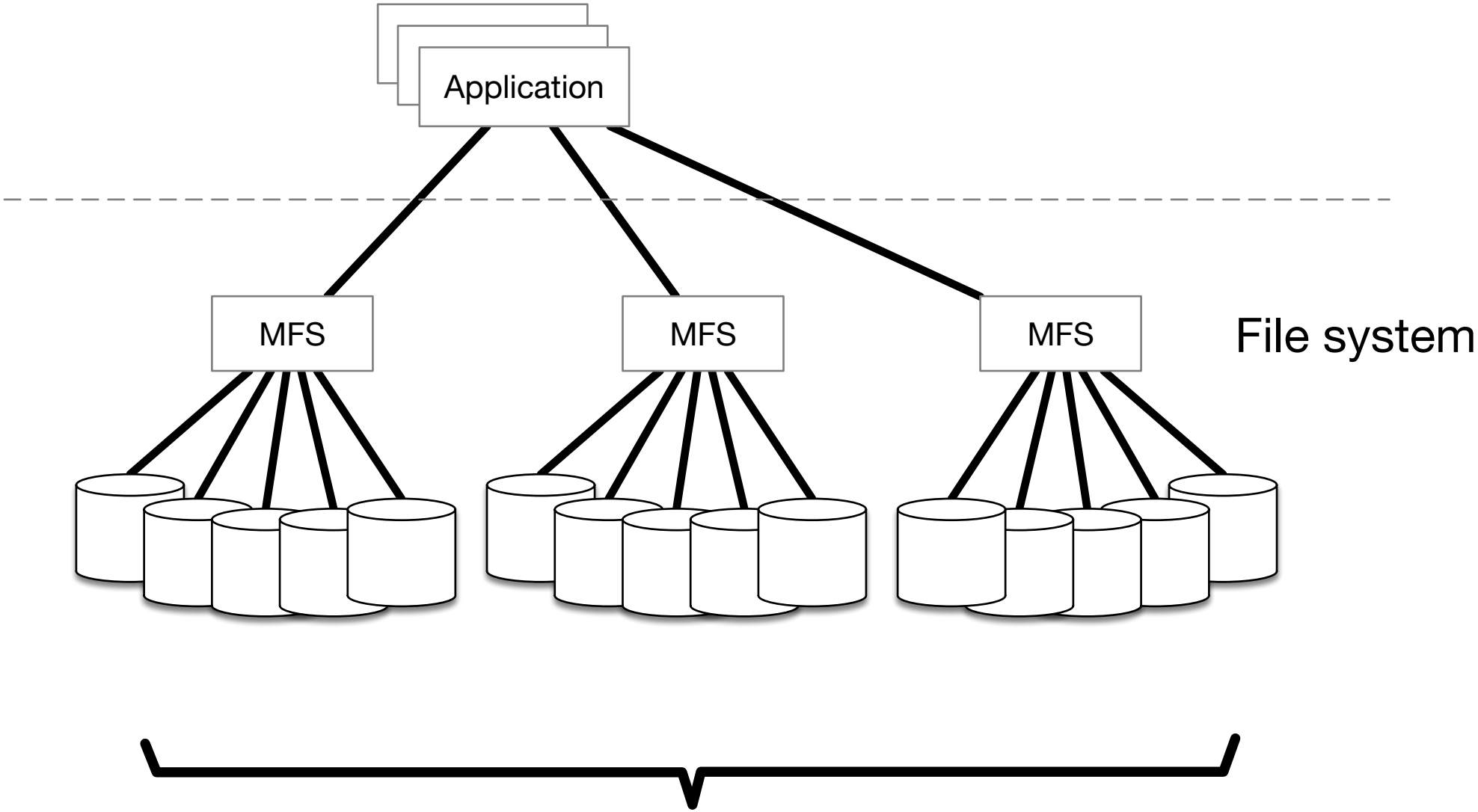
Because they will eat the clever bits!

So why focus on the boring bits?

Because they will eat the clever bits!

Also, systems aren’t ready for business until they are properly boring

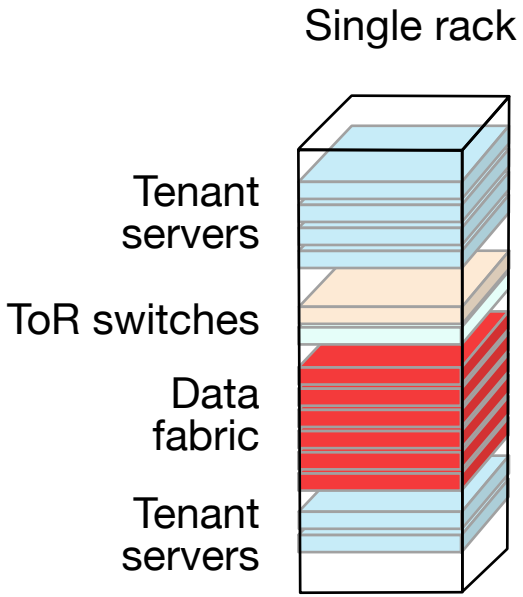

Application
Application
Application
OS
OS
OS
File system
Block
Block
Block
Blocks



Objects span multiple servers

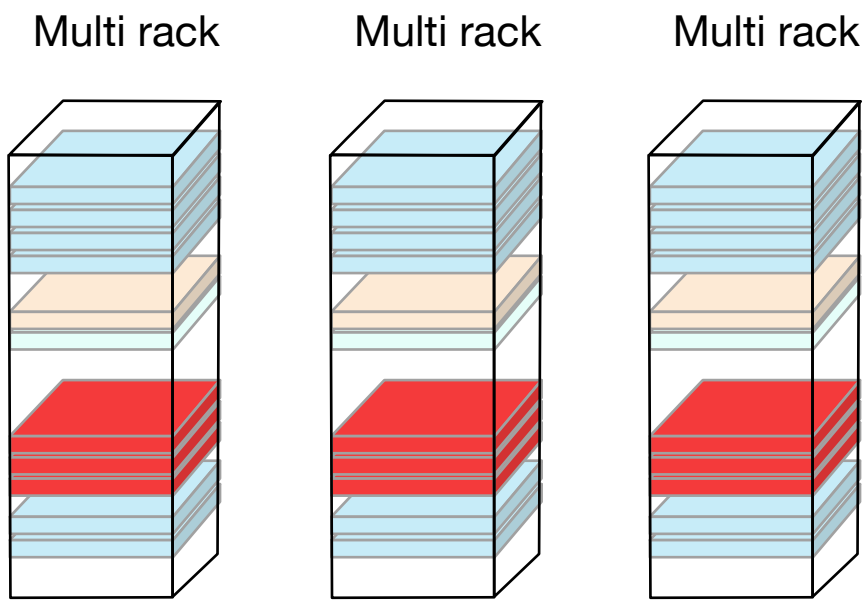
DESIGN ELEMENT - SCALE

SCALING DOWN



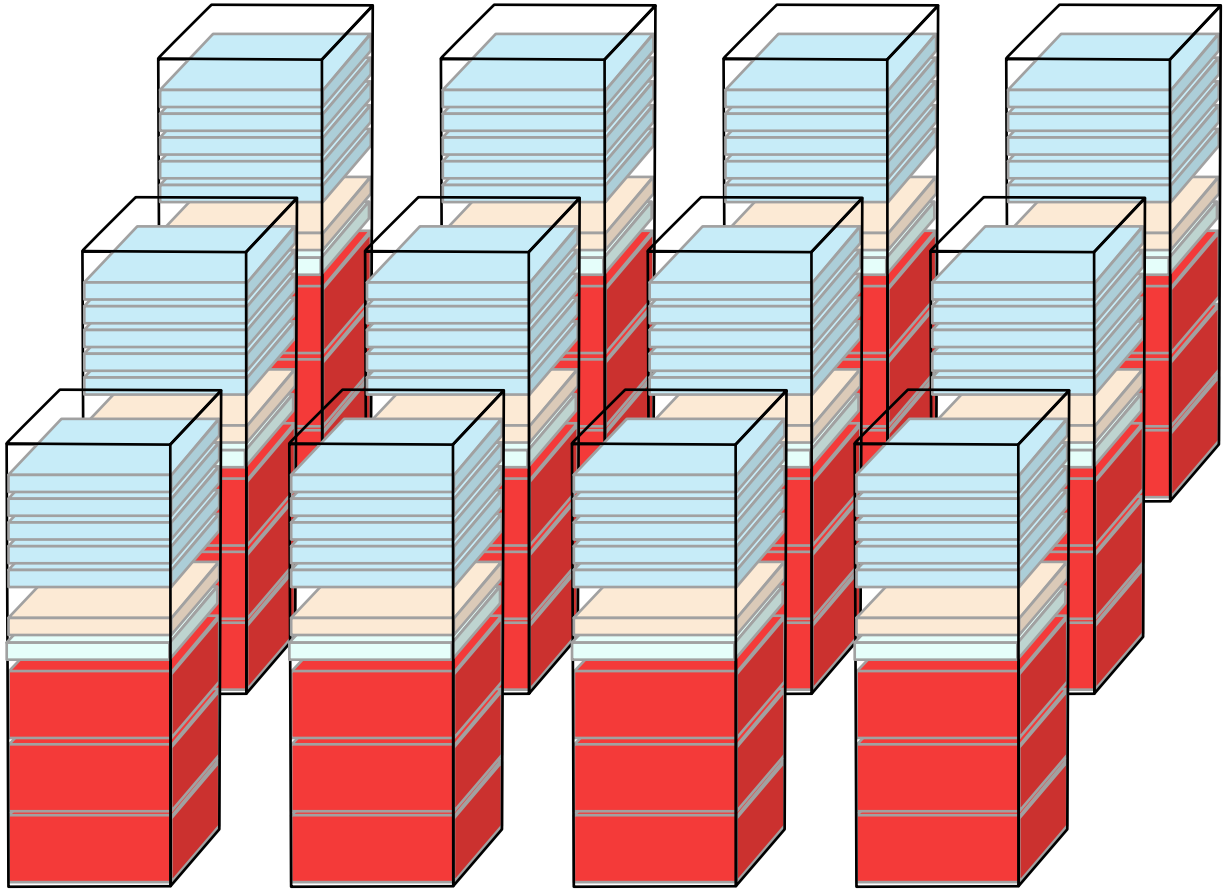
Single rack configuration is self-contained insofar as data fabric is concerned

SCALING DOWN



Multi rack configuration can amortize minimum # nodes

SCALING UP

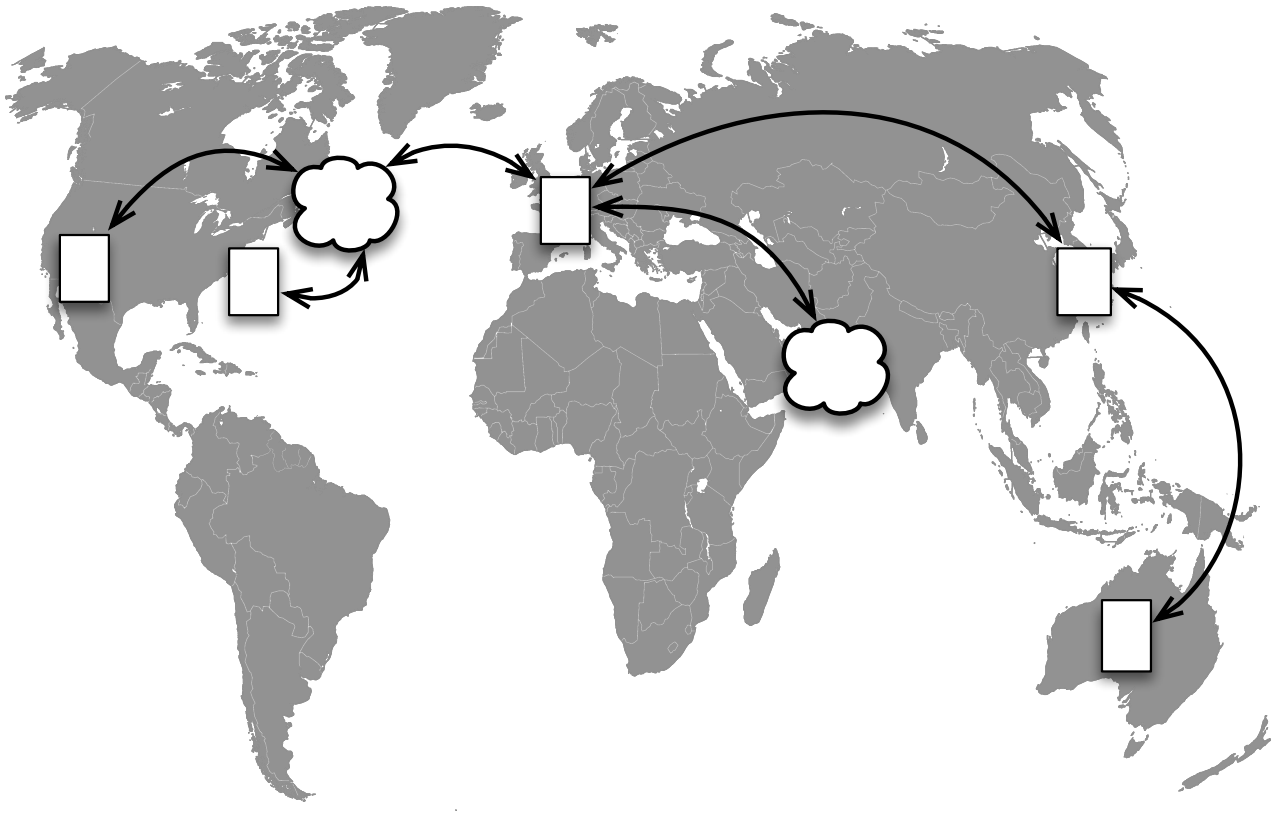


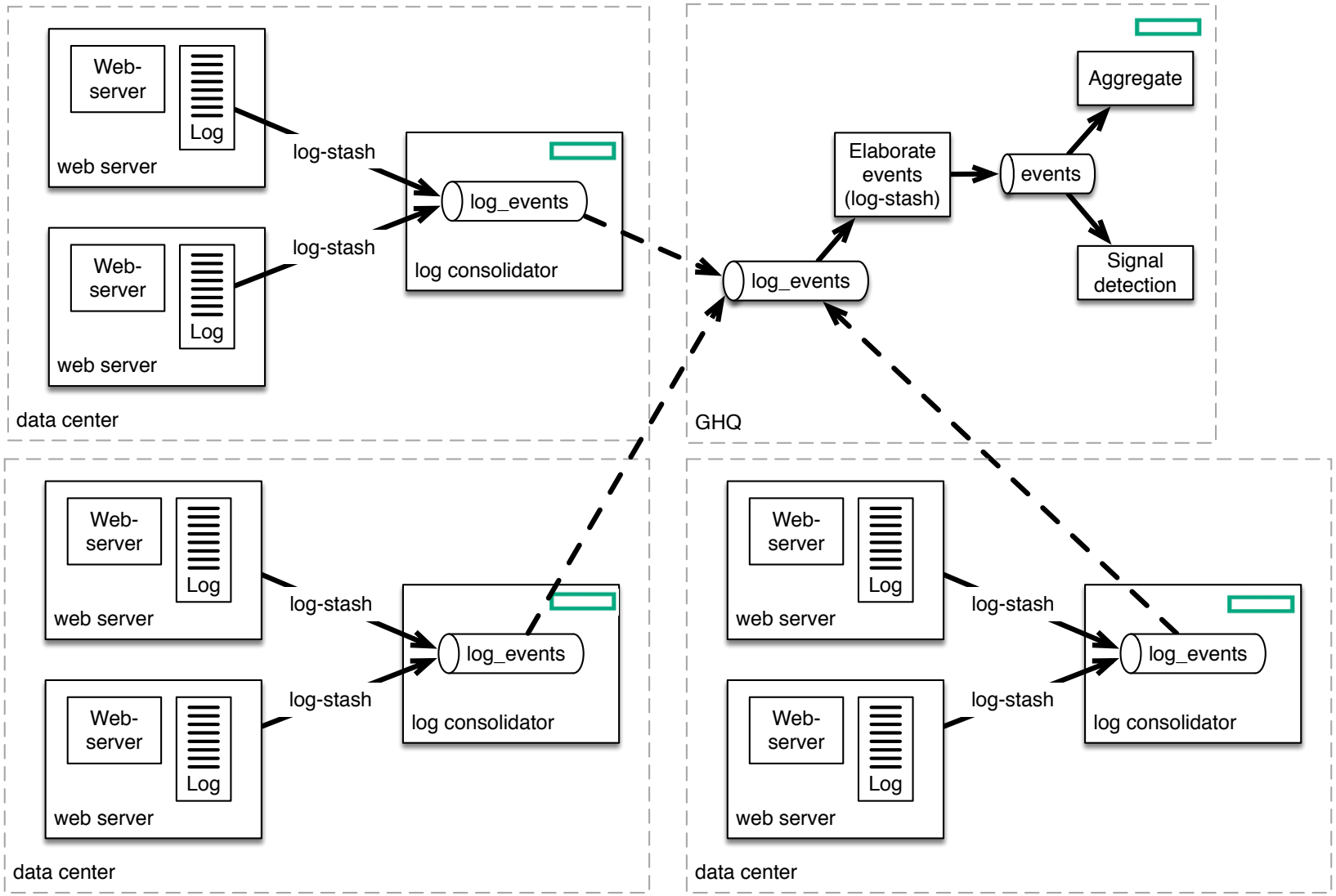
At very large scale, dense units become preferred

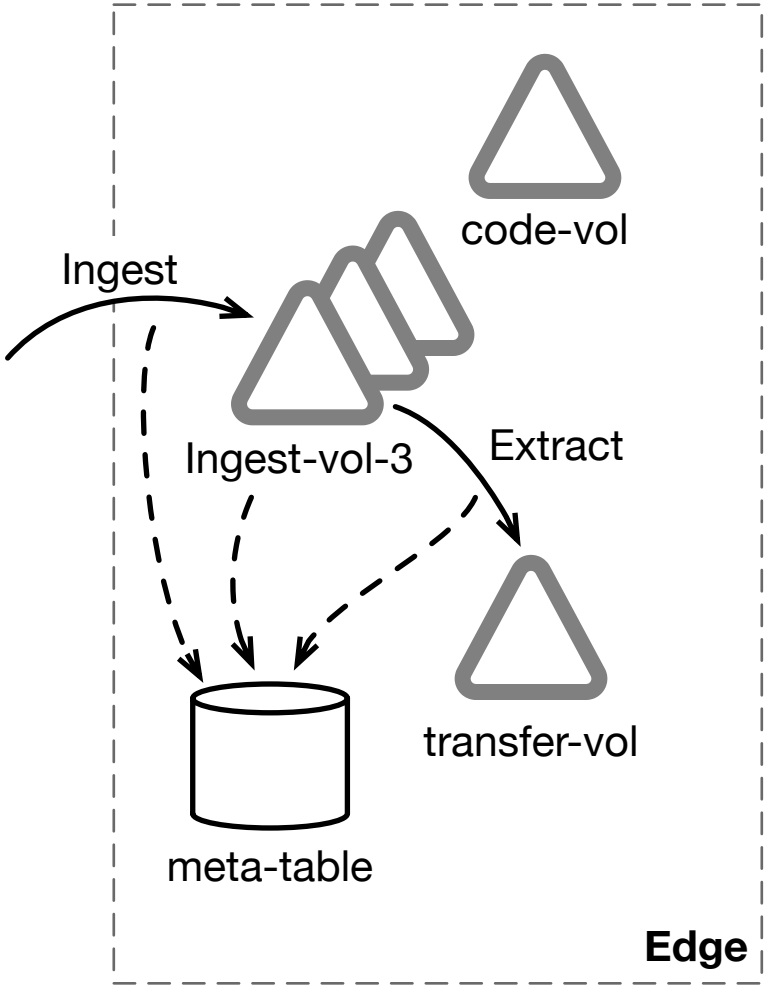
This is a rough rendition of a 24 PB system

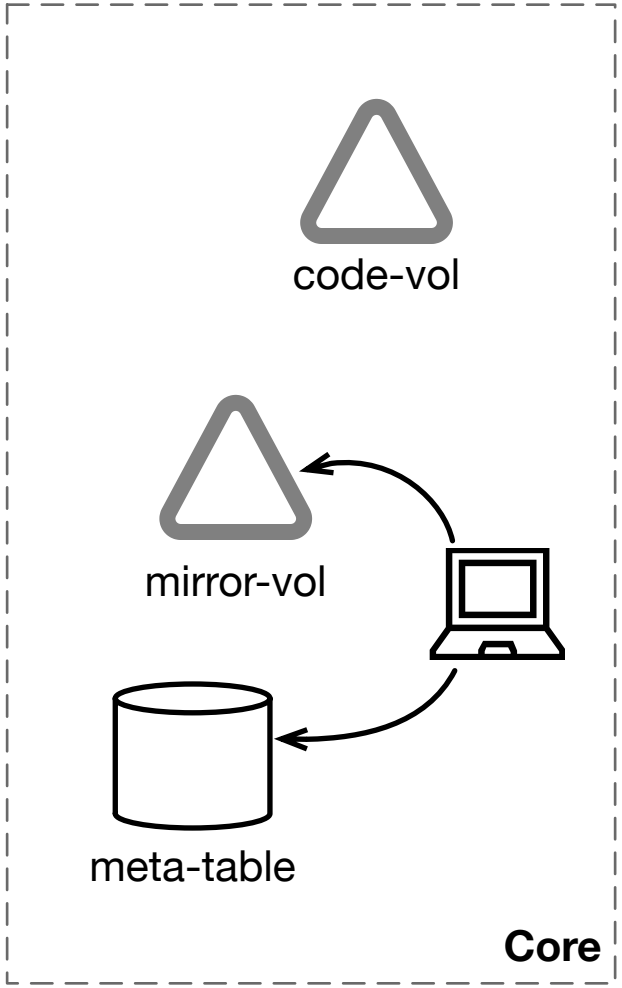
Some customers have >500PB systems

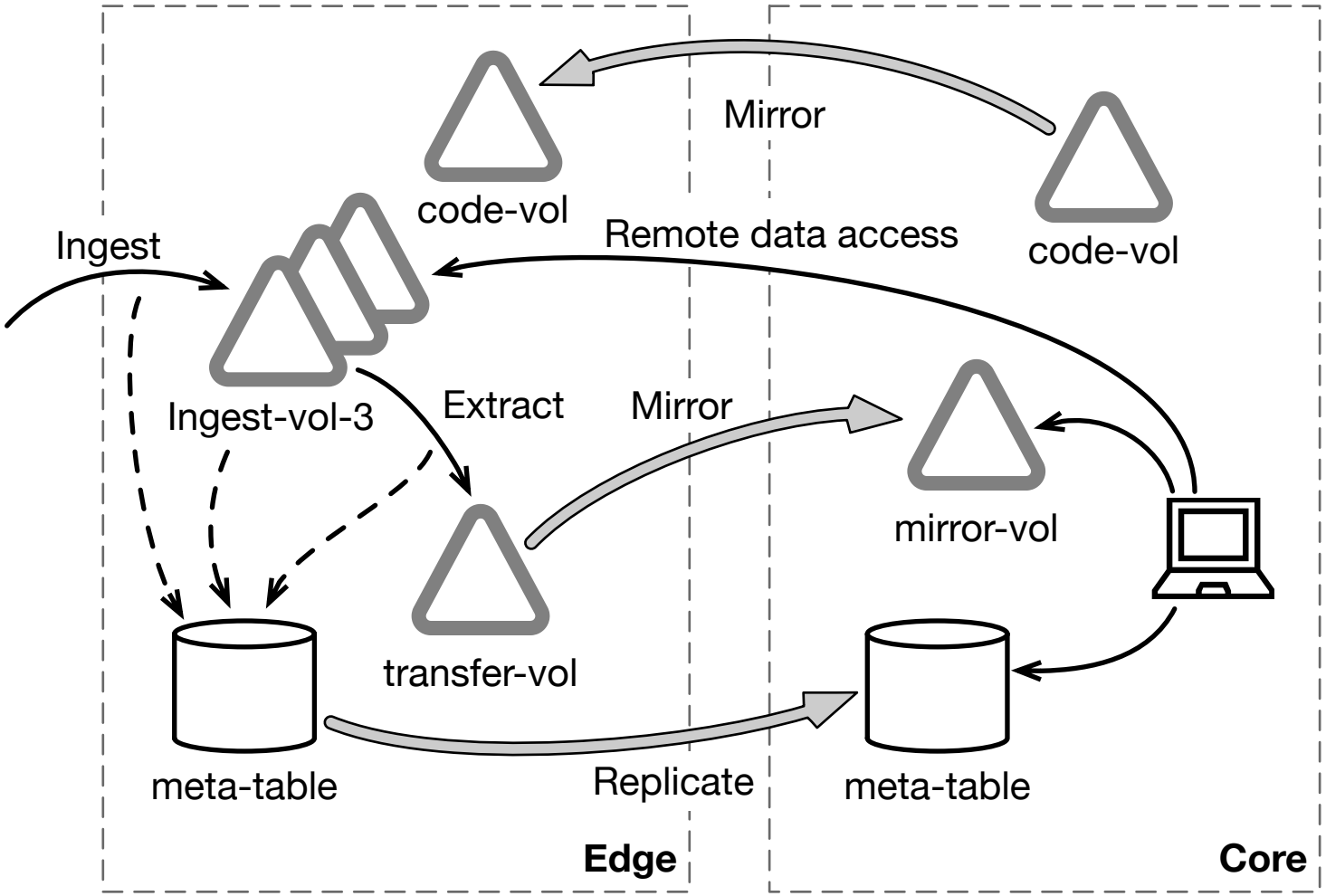
DESIGN ELEMENT – GEO-DISTRIBUTION




Web-
server
Log
web server
log-stash
log_events
log consolidator
data center
Aggregate
Elaborate
events
(log-stash)
events
Signal
detection
log_events
GHQ
data center
data center

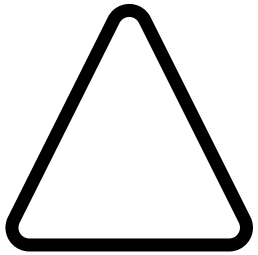

code-vol
Ingest
Ingest-vol-3
Extract
transfer-vol
meta-table
Edge


code-vol
mirror-vol
meta-table
Core

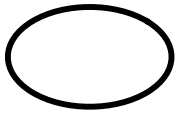

Mirror
code-vol
code-vol
Ingest
Remote data access
Ingest-vol-3
Extract
Mirror
mirror-vol
transfer-vol
meta-table
Replicate
meta-table
Edge
Core

DESIGN ELEMENT - VOLUMES

Volumes look like directories to programs and most users



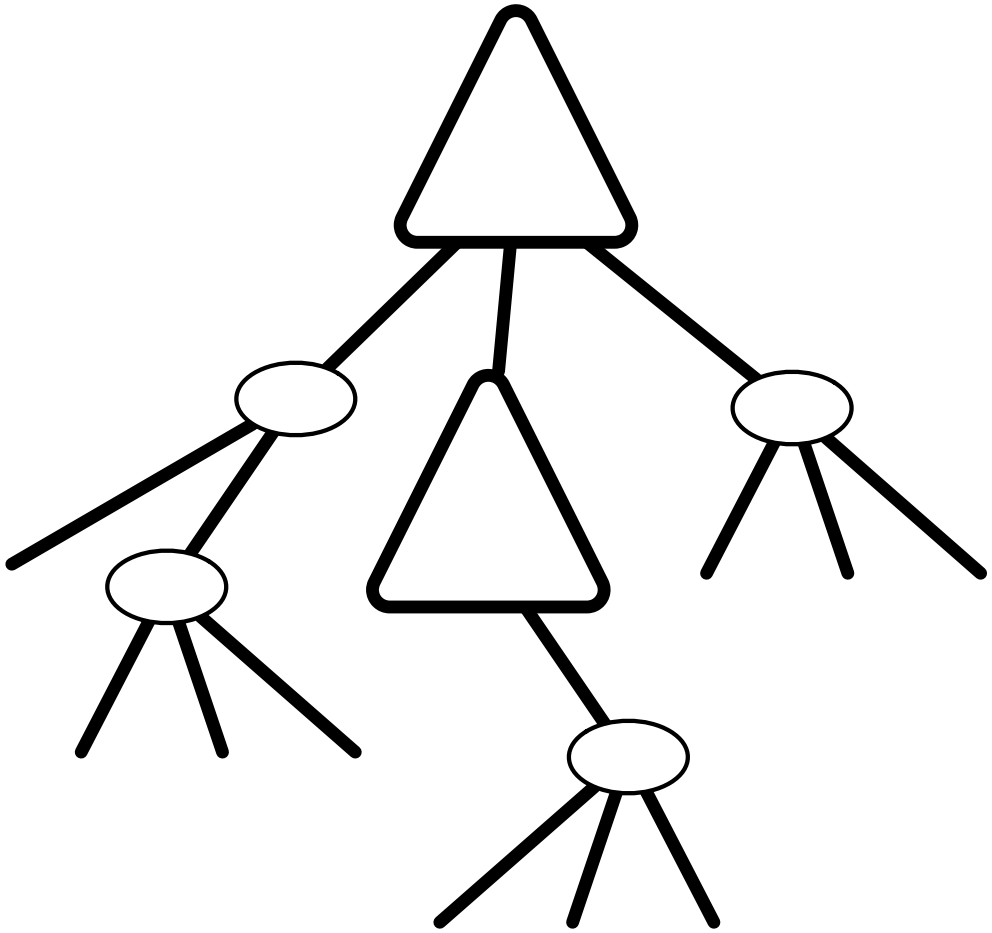
Volume

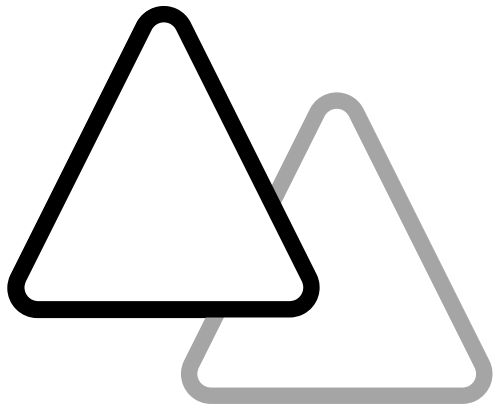


Directory

But they give us management powers of how data is stored

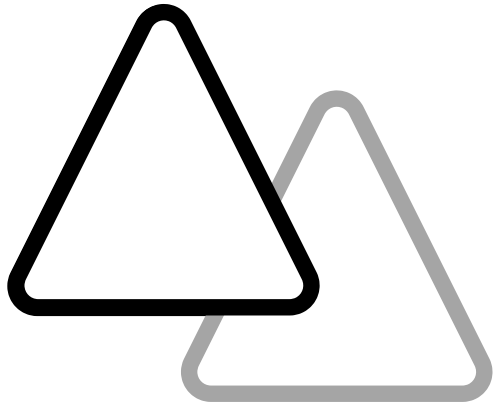
Volumes form trees just like directories.



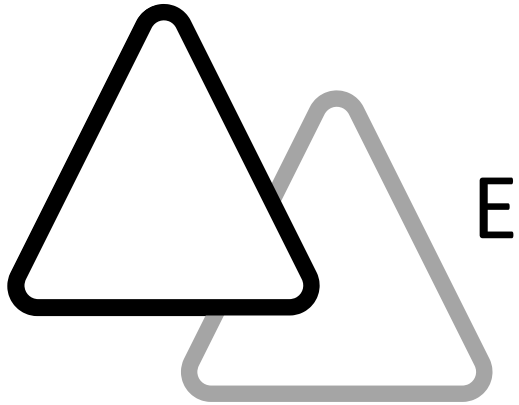


And we can have shadows of volumes that have different properties

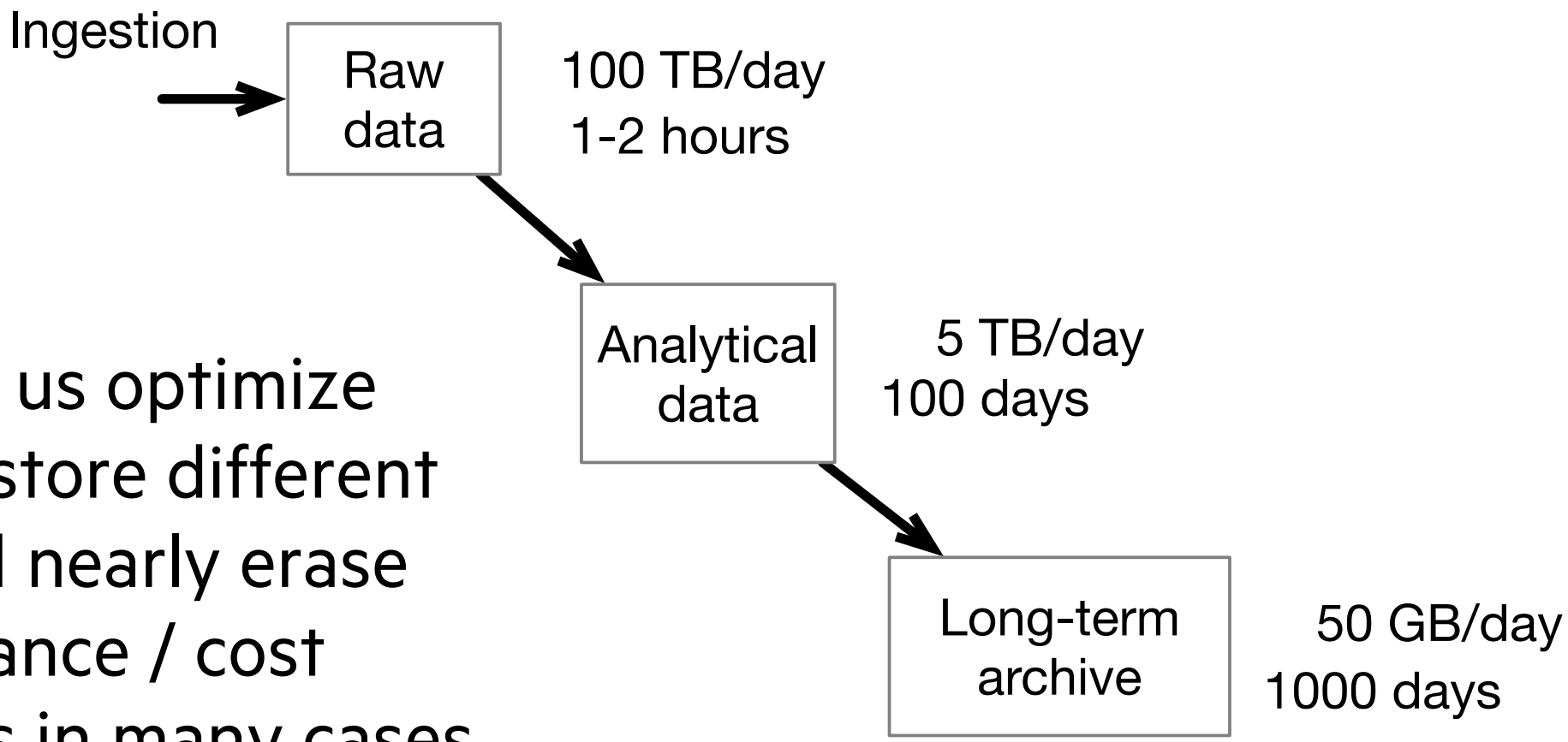
Replicated



Replicated



Erasure coded



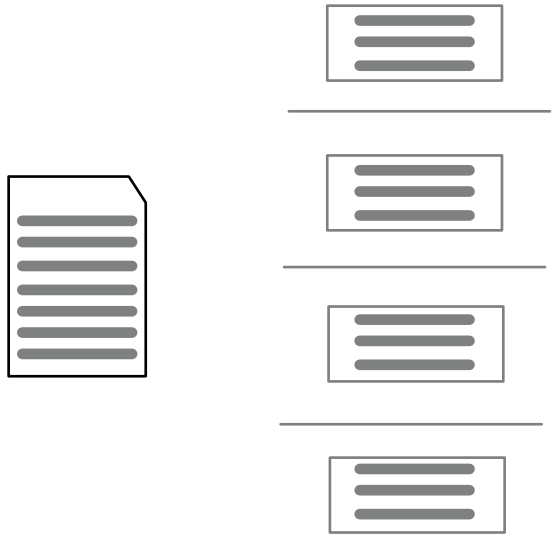
This lets us optimize how we store different data and nearly erase performance / cost tradeoffs in many cases

HOW FILES AND TABLES WORK

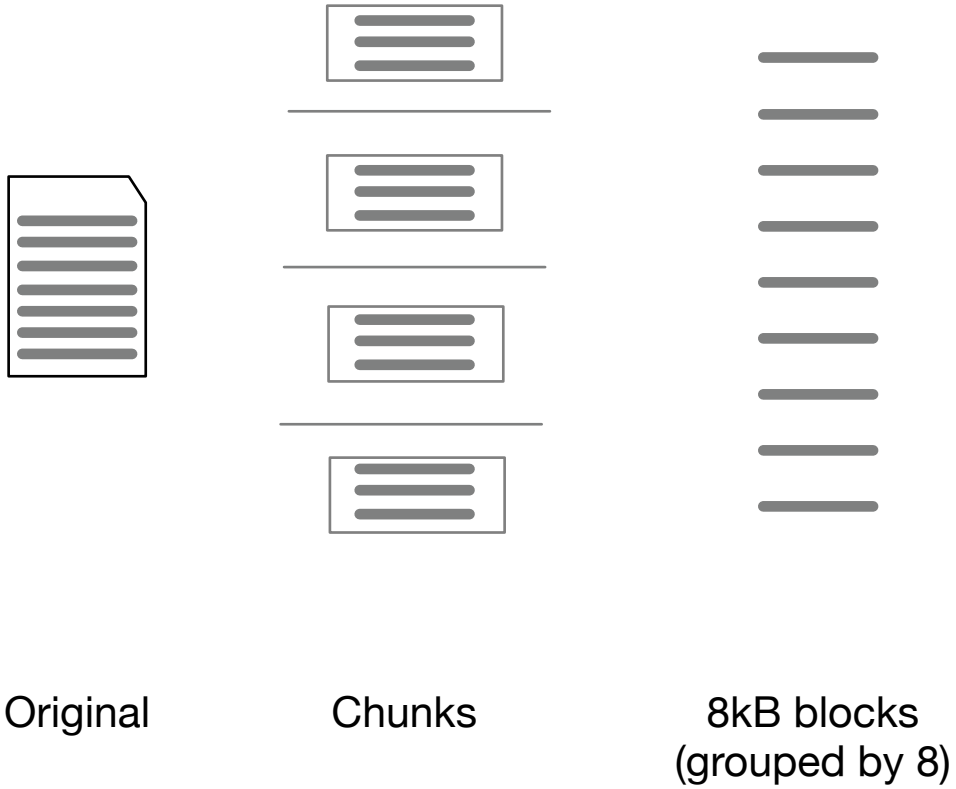
HOW FILES ARE STORED

Original

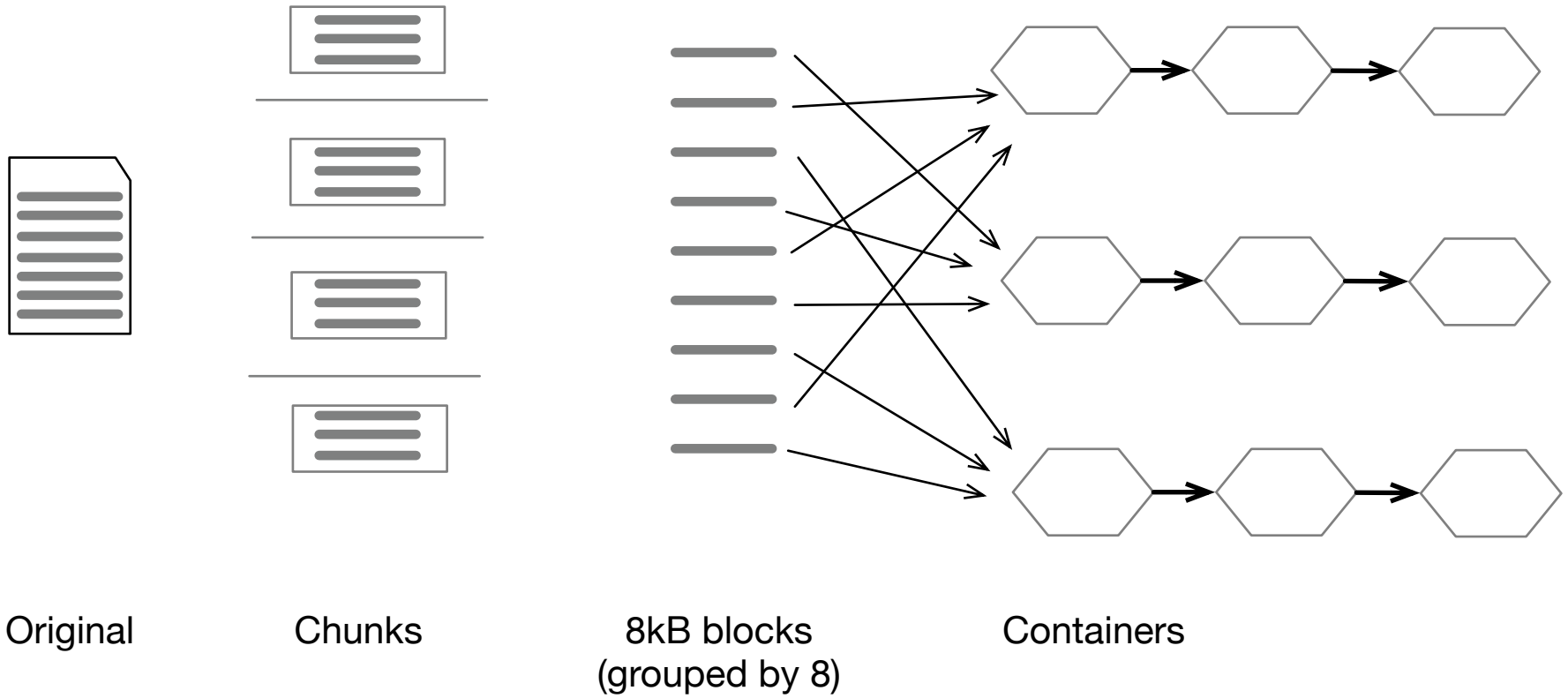
HOW FILES ARE STORED

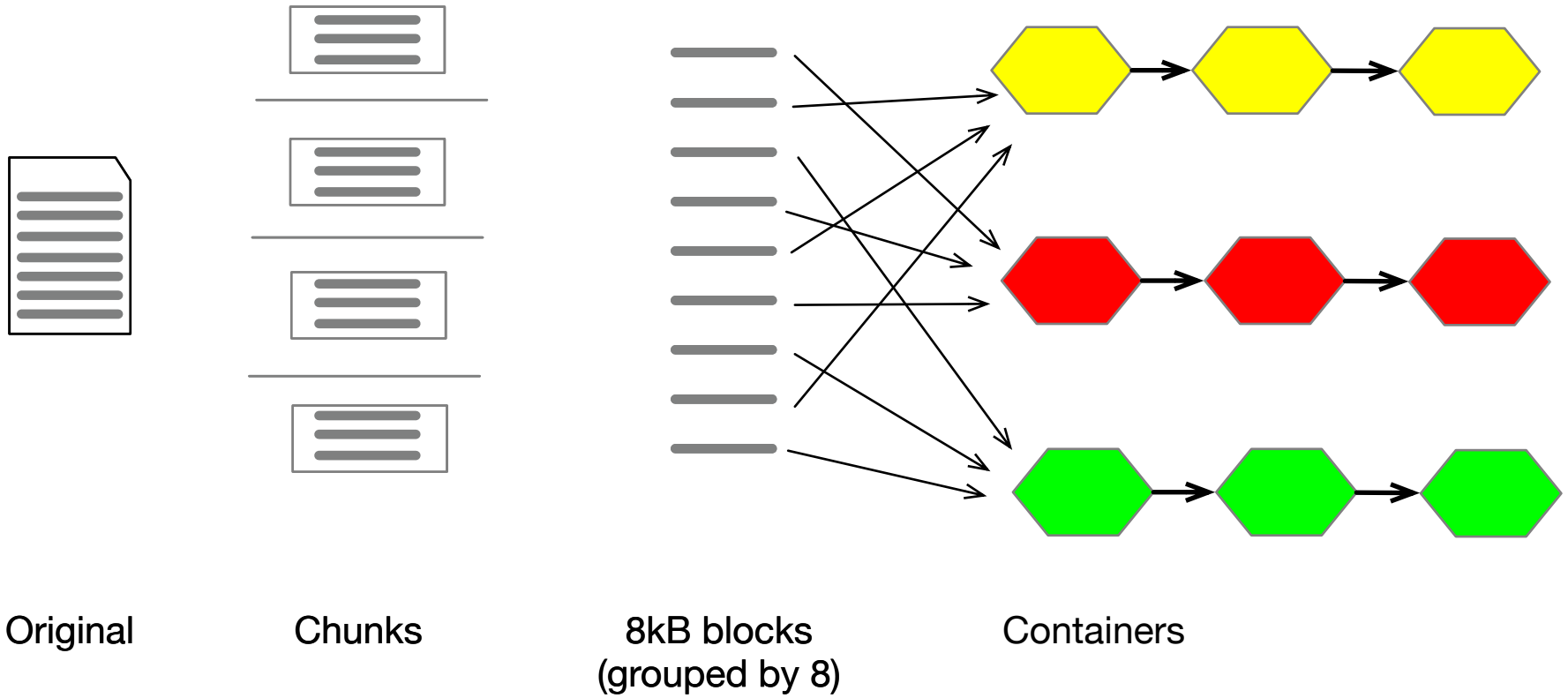


HOW FILES ARE STORED

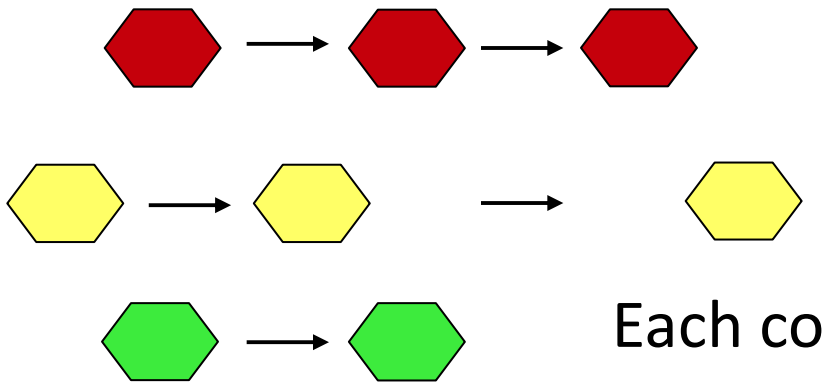


HOW FILES ARE STORED




Original
Chunks
8kB blocks
(grouped by 8)
Containers

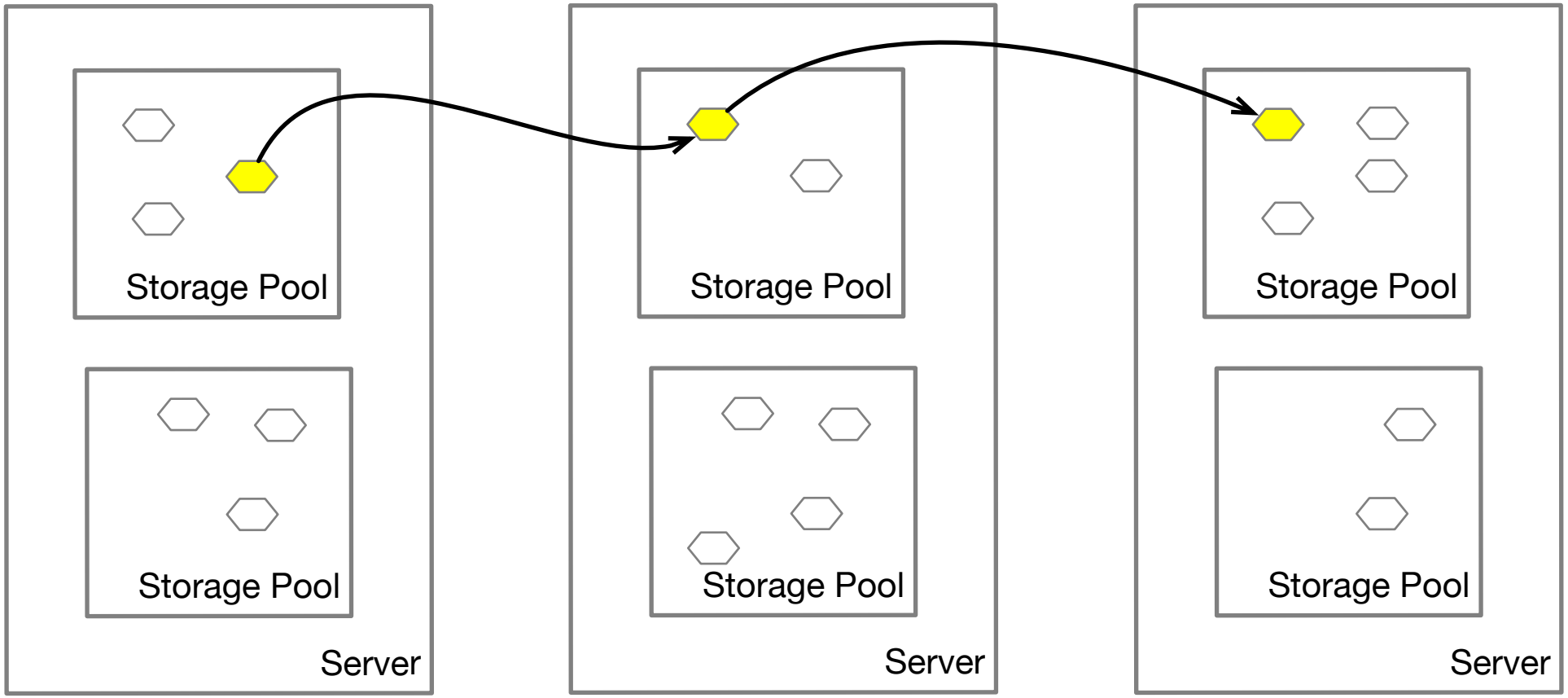
FABRIC CONTAINERS

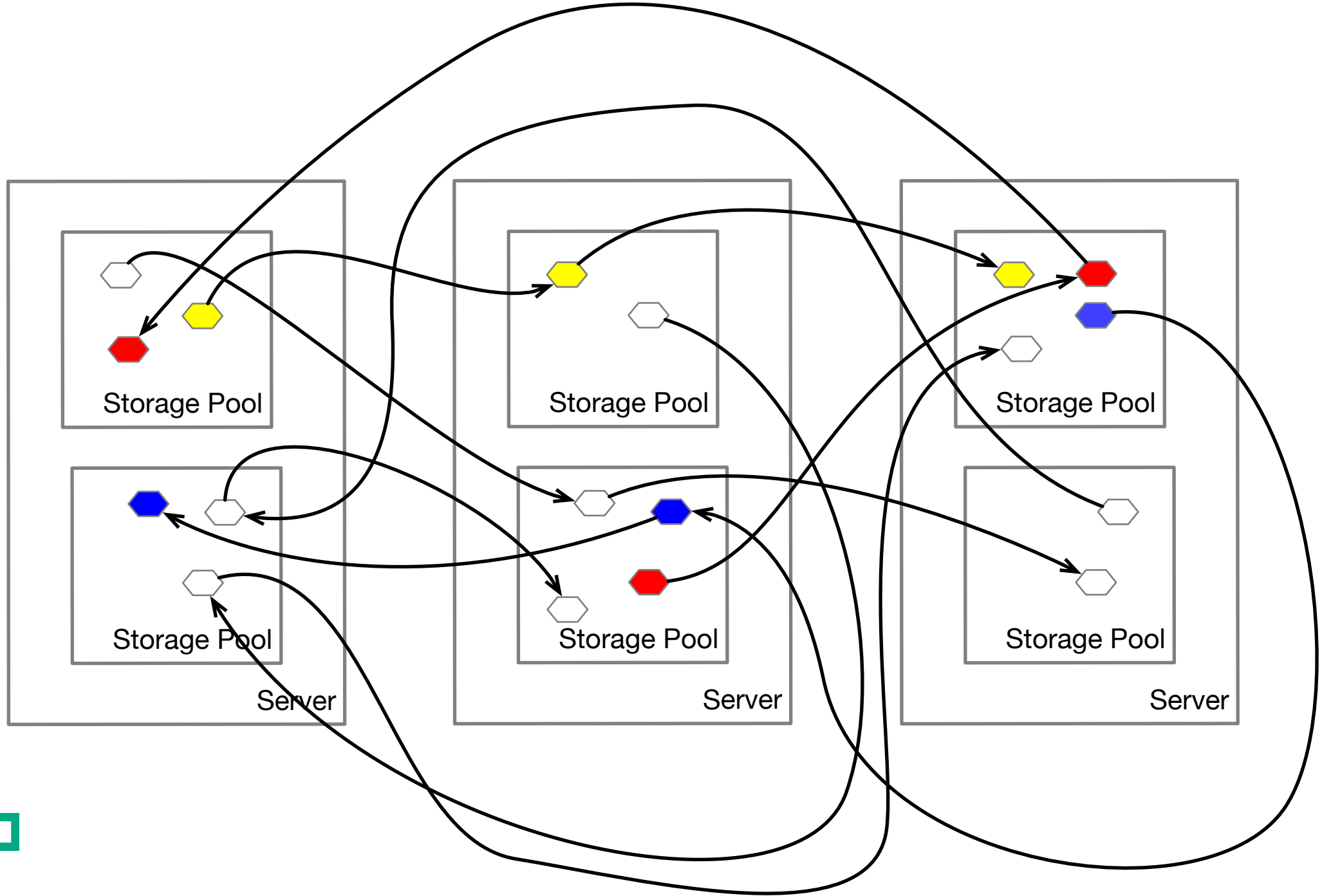


Each container has a replication chain

Updates are transactional

Failures are handled by rearranging replication


Storage Pool
Storage Pool
Server
Storage Pool
Storage Pool
Server
Storage Pool
Storage Pool
Server


Storage Pool
Storage Pool
Storage Pool
Server
Storage Pool
Storage Pool
Server
Storage Pool
Storage Pool
Server

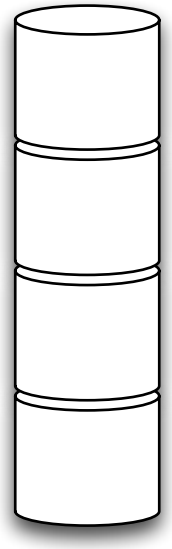
BUT WAIT, THERE'S MORE

- Directories and files are implemented in terms of B-trees
  - Key is offset, value is data blob
  - Internal transactional semantics guarantees safety and consistency
  - Layout algorithms give very high layout linearization

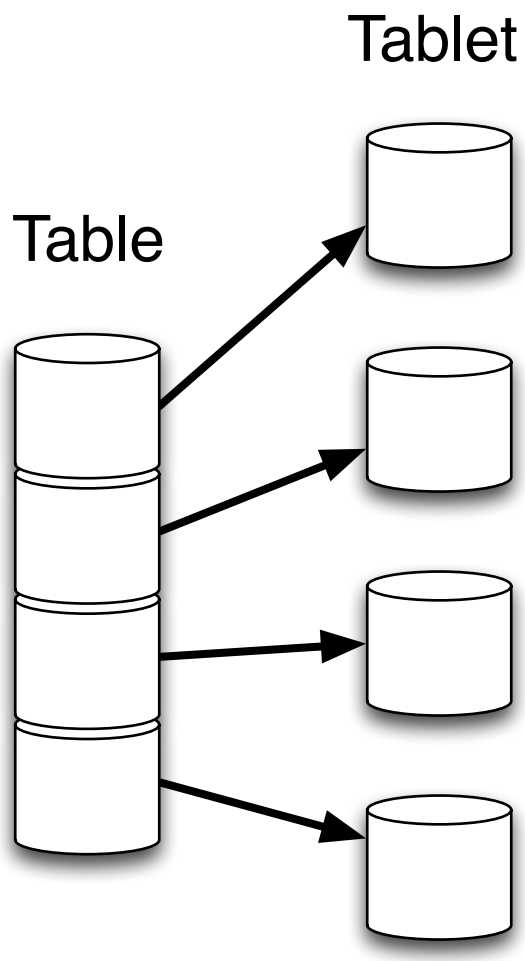
- Tables are implemented in terms of B-trees
  - Twisted B-tree implementation allows virtues of log-structured merge tree without the compaction delays
  - Tablet splitting without pausing, integration with file system transactions

- Common security and permissions scheme

Table

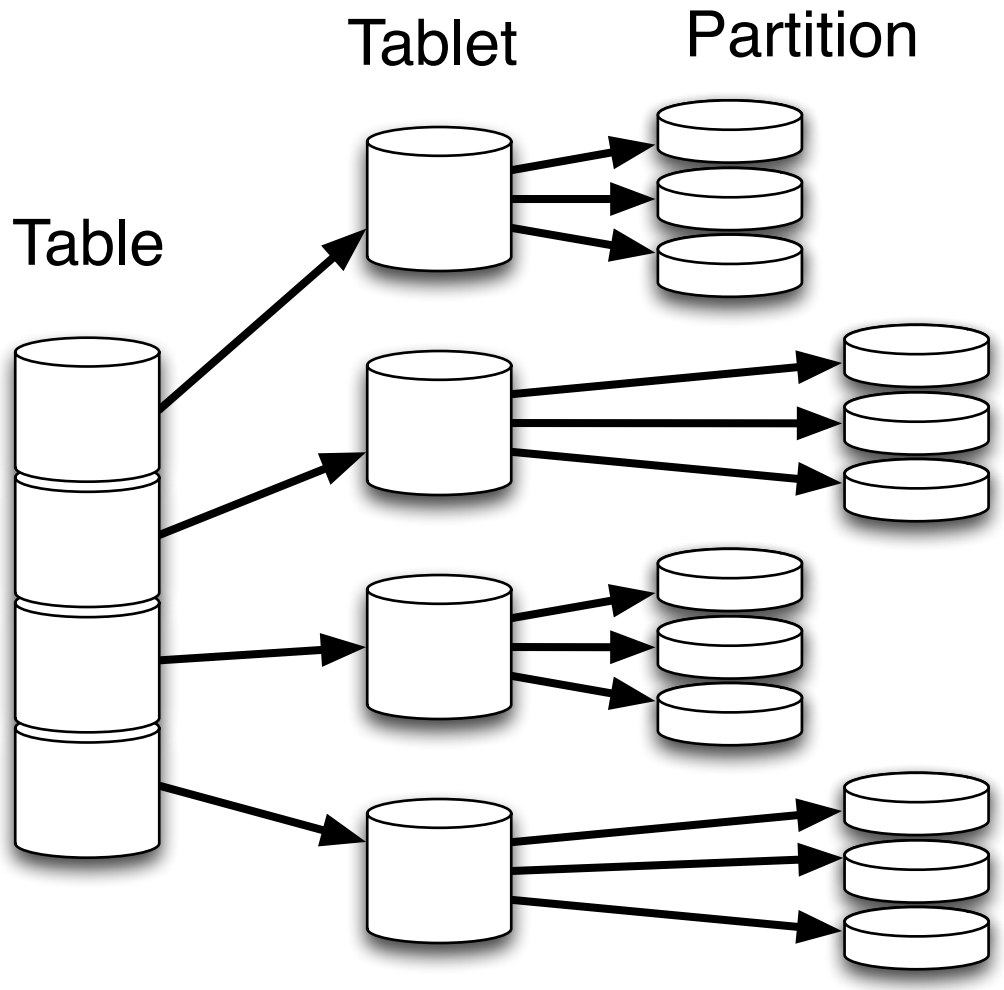


MAPR DB TABLES



Similar to LSM implementations, tables are decomposed by key ranges

MAPR DB TABLES



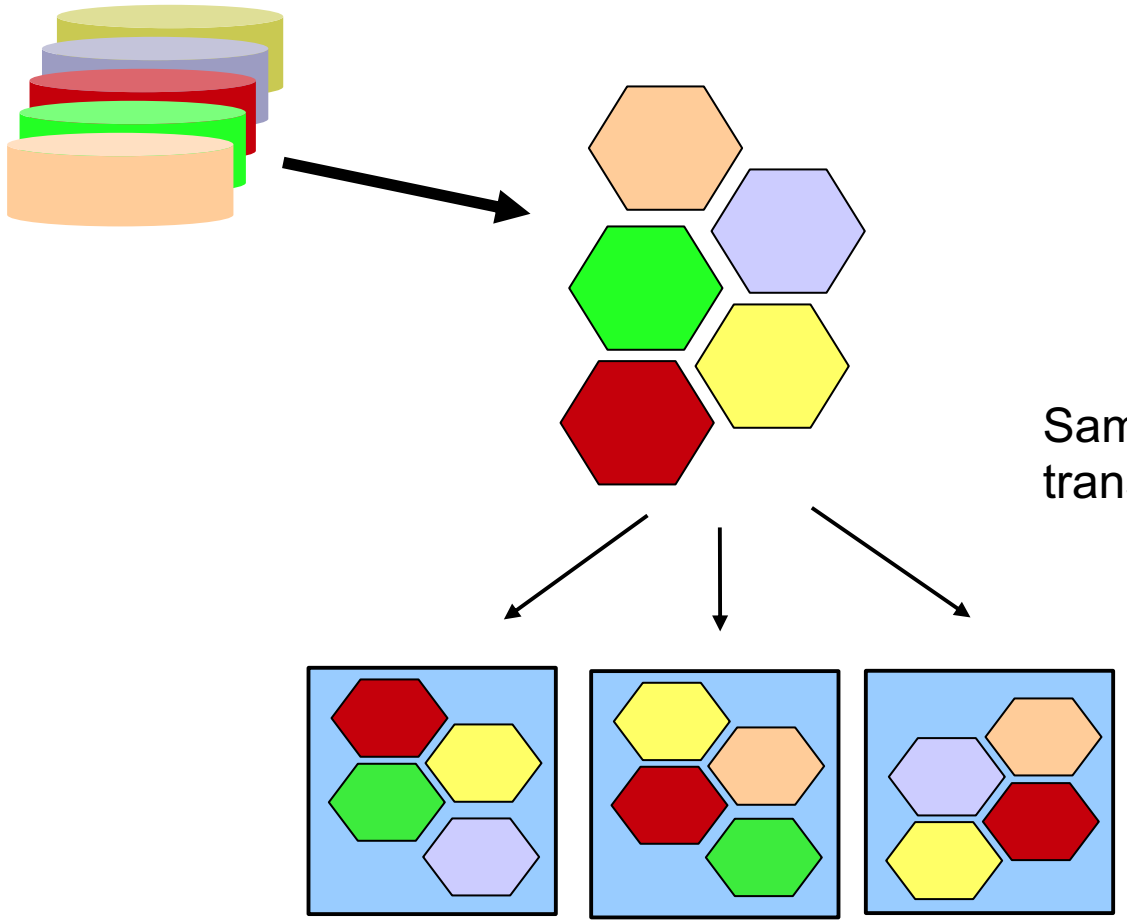
Similar to LSM implementations, tables are decomposed by key ranges

Distinct from HBase and Level DB, MapR tables used *fixed* number (greater than 1) of decompositions

Very unusually, relative to LSM and cousins, data structures at the leaf are mutable

RE-USE OF PROVEN TECHNOLOGY

Partitions are distributed just like file chunks



Same replication and transaction technology

HOW IT LOOKS

EXAMPLE

```
[tdunning@se-node10 ~]$ ls -F
apache-kylin-0.7.2-incubating-src-source-release.tar.gz
apache-kylin-0.7.2-incubating-src-source-releas d5
apache-kylin-0.7.2-incubating-src-source-releas ha1
bar/
build.xml
car-data.csv
cooc.parquet/
counts.parquet/
deep.json
drillbit.log
edges.ssv
edges.tsv.bak
r.csv/
README
Rplots.pdf
s1@
s2@
schema.json
sf-city-lots-json/
side-log
src/
t1@
tags.json
time_to_60.view.drill*
```

Streams

Files

Table

Directories

Cluster

```
[tdunning@se-node10 ~]$ pwd
/mapr/se1/user/tdunning
[tdunning@se-node10 ~]$
```

Volume mount point

Cluster

```
[tdunning@se-node10 ~]$ pwd
/mapr/se1/user/tdunning
[tdunning@se-node10 ~]$
```

Volume mount point

THANK YOU

ted.dunning@hpe.com

CONFIDENTIAL DISCLOSURE AGREEMENT

- The information in this presentation is authorized for general distribution but it should not be modified
- Do not remove any classification labels, warnings or disclaimers on any slide or modify this presentation to change the classification level
- Do not remove this slide from the presentation
- HPE does not warrant or represent that it will introduce any product to which the information relates
- The information contained herein is subject to change without notice
- HPE makes no warranties regarding the accuracy of this information
- The only warranties for HPE products and services are set forth in the express warranty statements accompanying such products and services
- Nothing herein should be construed as constituting an additional warranty
- HPE shall not be liable for technical or editorial errors or omissions contained herein
- Strict adherence to the HPE Standards of Business Conduct regarding this classification level is critical

THANK YOU FOR BEING PART OF THE HPE DEV COMMUNITY

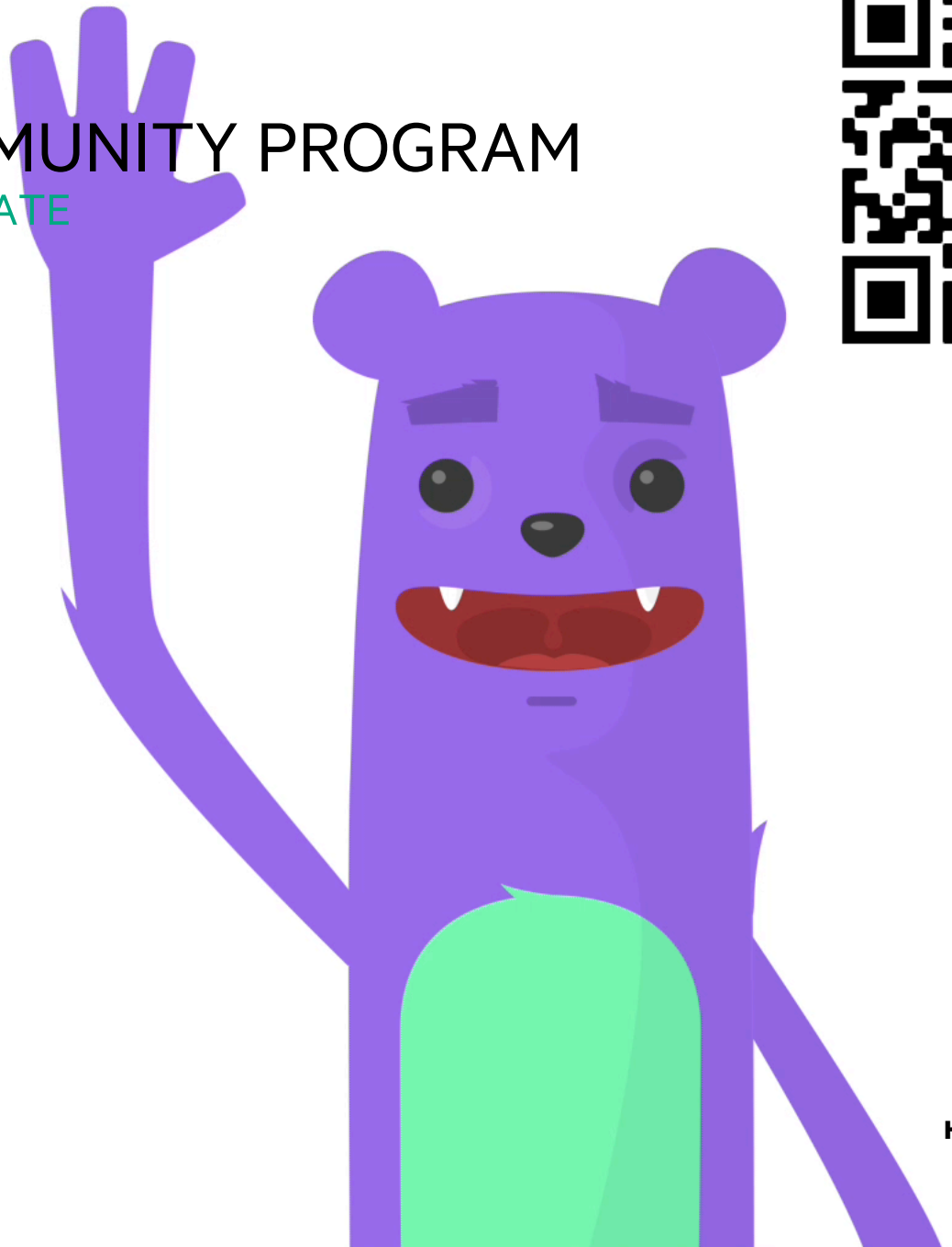
- Monthly Newsletter: https://developer.hpe.com/newsletter-signup

- Developer Survey: https://www.developereconomics.net/?member_id=hpe
  - Make your voice heard
  - Closes on February 8th

- Next Munch & Learn session
  - February 24th 5PM CET
  - The benefits of a Container Platform and Machine Learning (ML) Ops
  - Speaker: Tom Phelan, HPE Fellow/CPO Ezmeral Container Platform
  - Industry Host: Nigel Poulton
  - Registration link: https://hpe.zoom.us/meeting/register/tJYkdequqTwsE9LuPAgPDbV-mf1V7jq23Mxj

HPE DEVELOPER COMMUNITY PROGRAM

BUILD | COMMUNICATE | COLLABORATE



hpedev.io

hpedev.slack.com

Monthly Newsletter

mailto:hpedev@hpe.com

@HPE_DevCom

Yammer Group (Internal)

Yammer Group (Partner)

Workshops-on-Demand

HPE DEV