Failing my way to success with Azure

A story of the myriad ways I tried to cope with change, scale, and Azure

Microsoft MVP

itsalocke.com

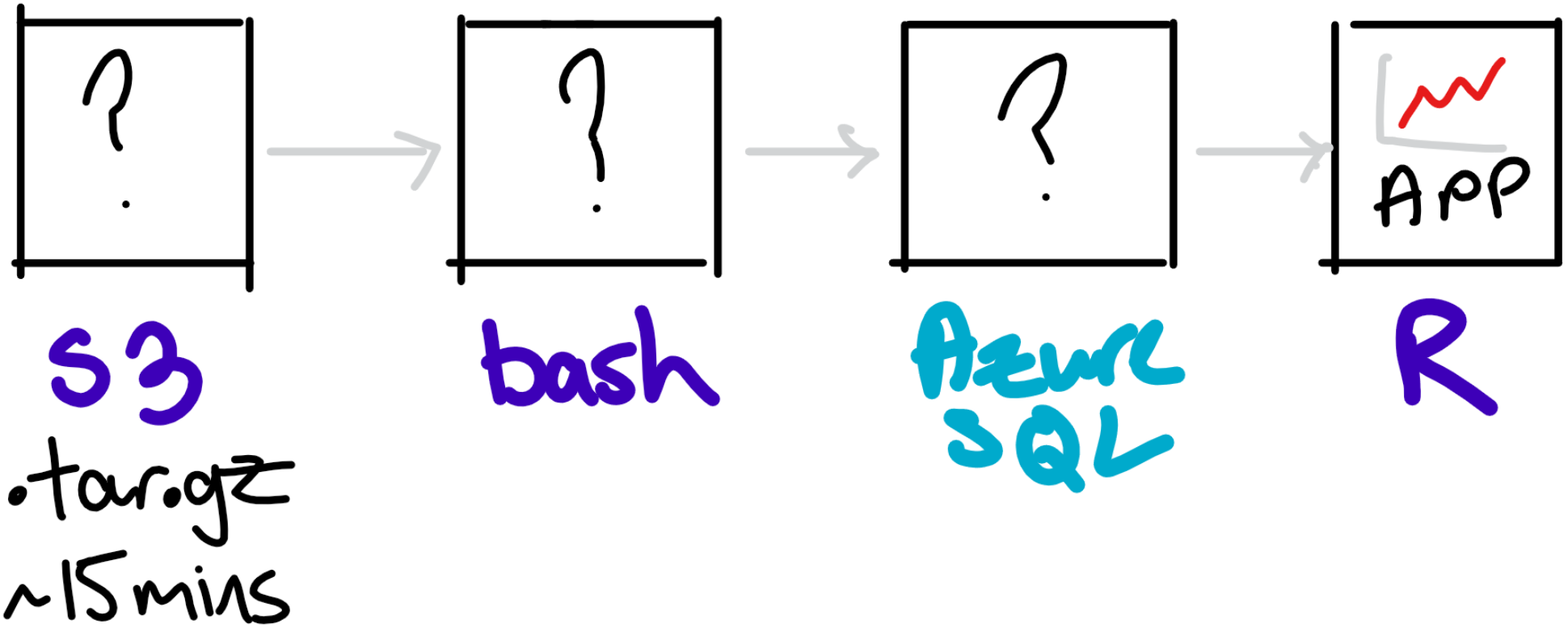
Steph Locke

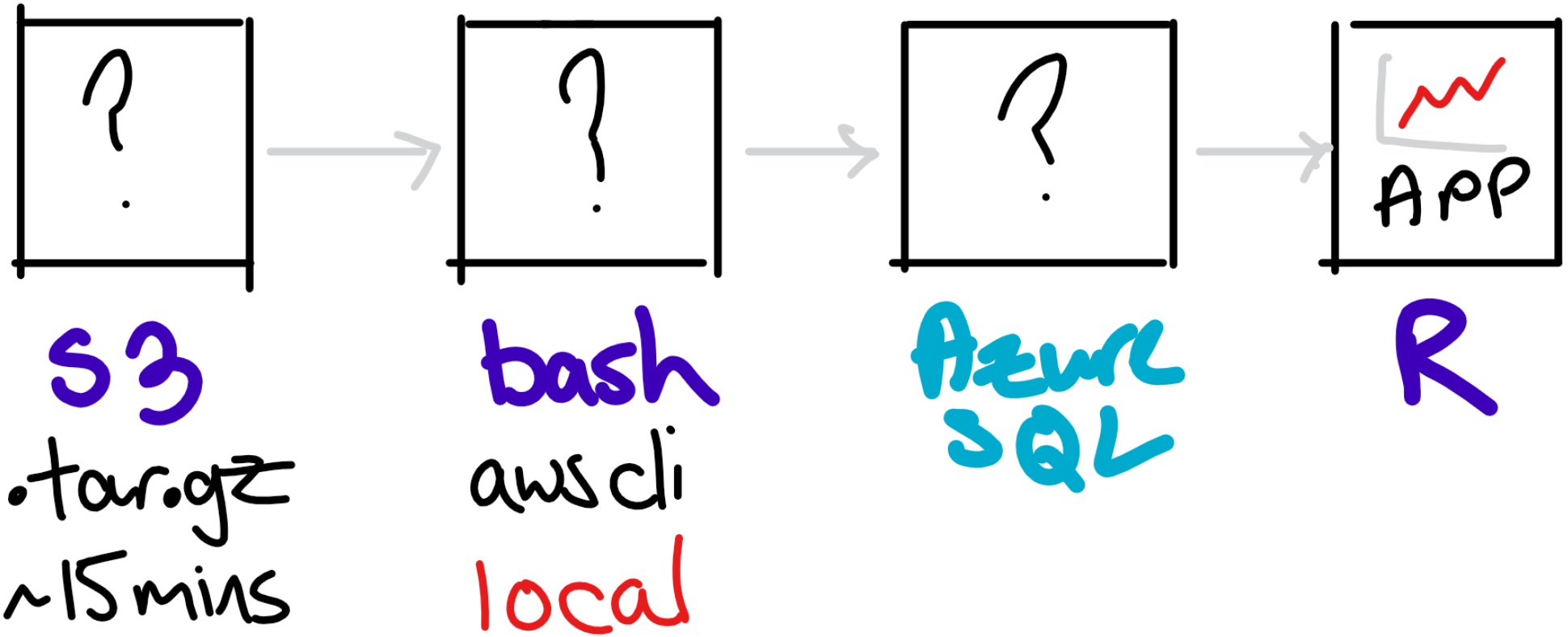
It's my birthday!

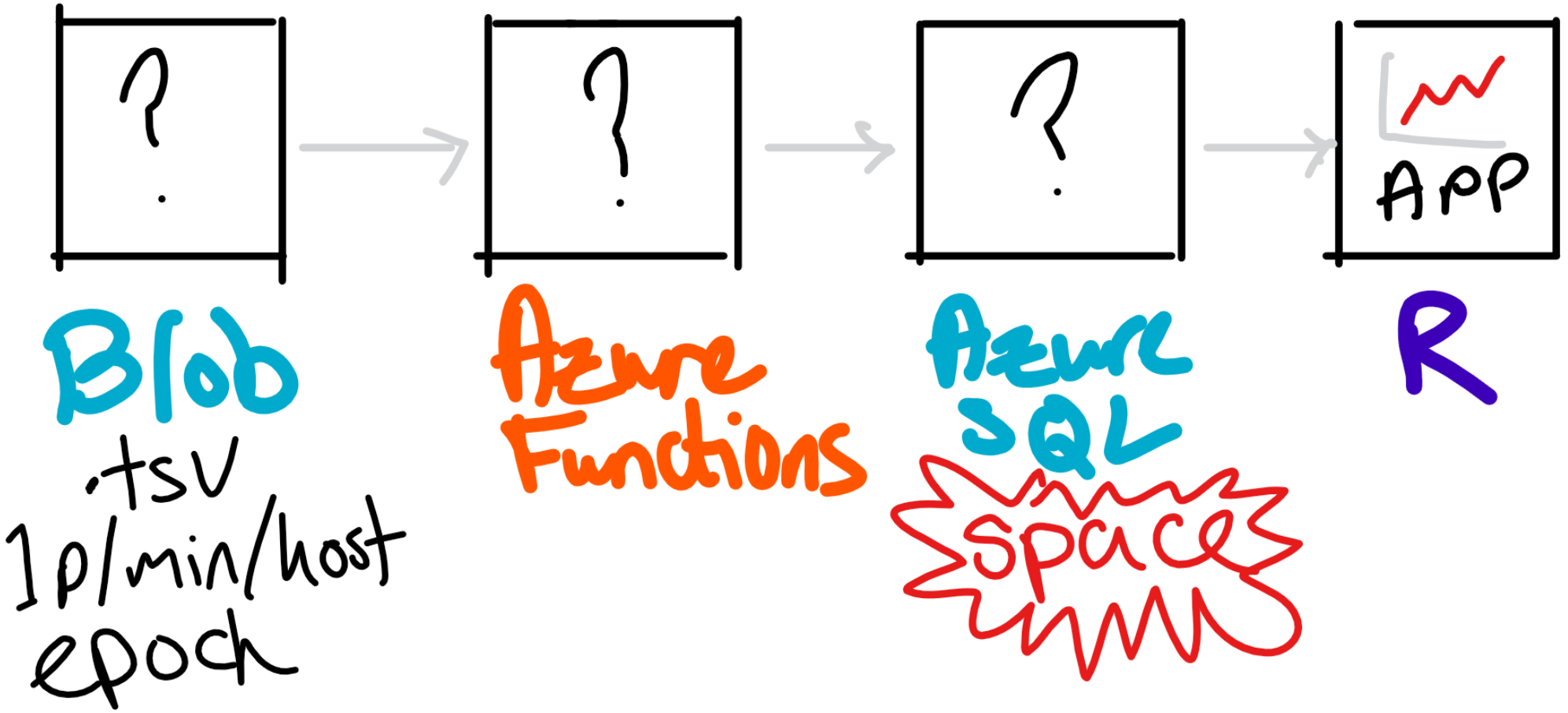
@stefflocke

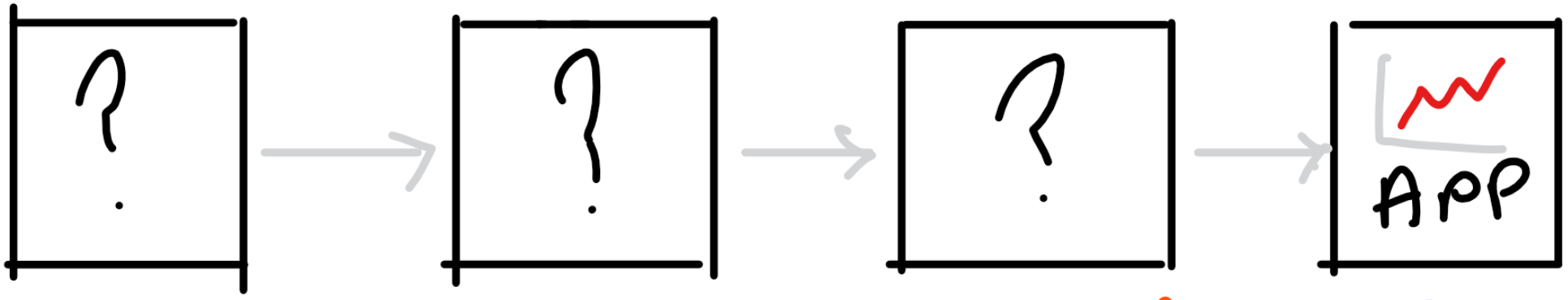
Data Science & Data Ops

Locke Data


?
?
?
APP
S3
bash
Azure SQL
R
.tar.gz
~15 mins


?
?
?
APP
S3
bash
Azure SQL
R
.tar.gz
~15mins
aws cli
local


?
?
?
APP
Blob
.tsv
1p/min/host
epoch
Azure Functions
Azure SQL
space
R



Blob

.tsv

1p/min/host

epoch

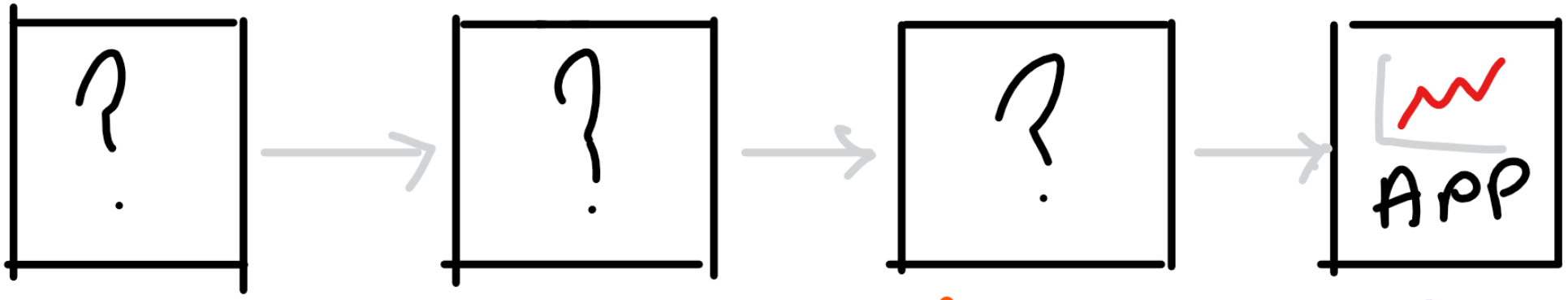
HDInsight

hive: slow

spark: slow

overall: EXPENSIVE

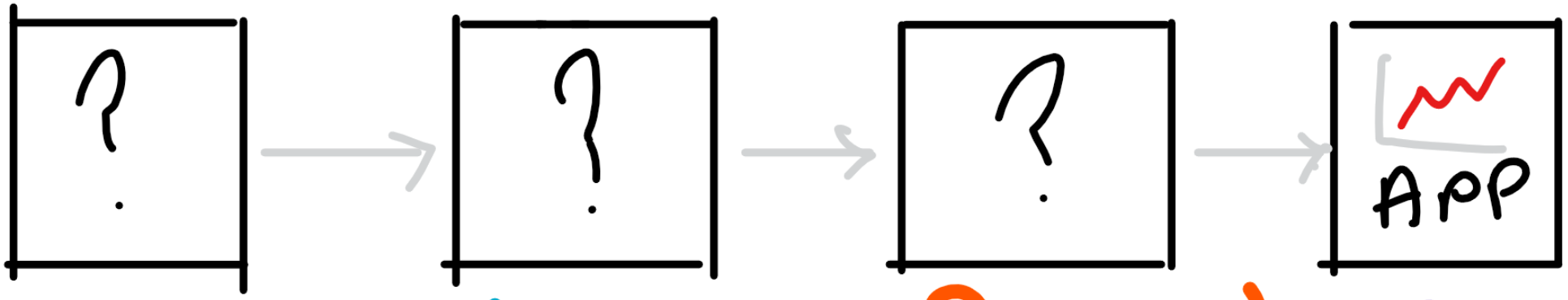
R



Blob
.tsv
1p/min/host
epoch

Azure Data Warehouse
Expensive
Identified encoding errors
Quicker than HDInsight

R

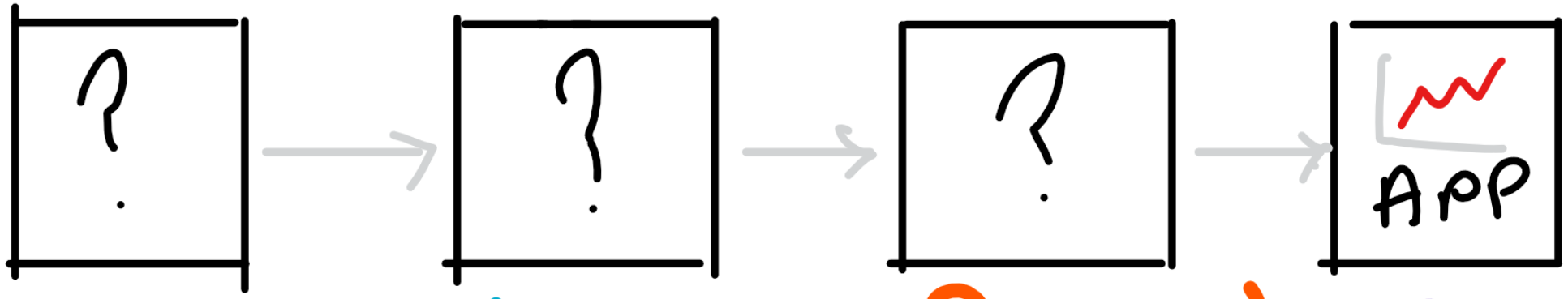


Blob
.tsv
1p/min/host
epoch

Azure Functions

Document DB
4x expected cost
Summarisation is hard

R



Blob

.tsv

1p/min/host

epoch

Azure Functions

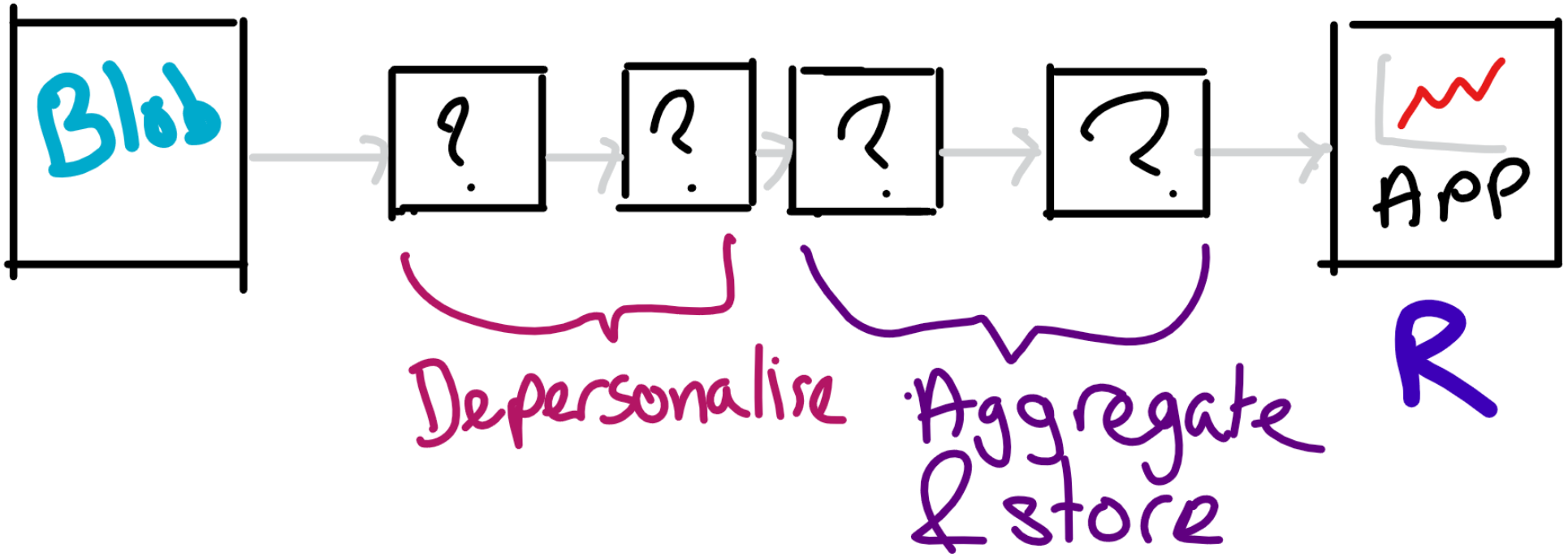
Document DB

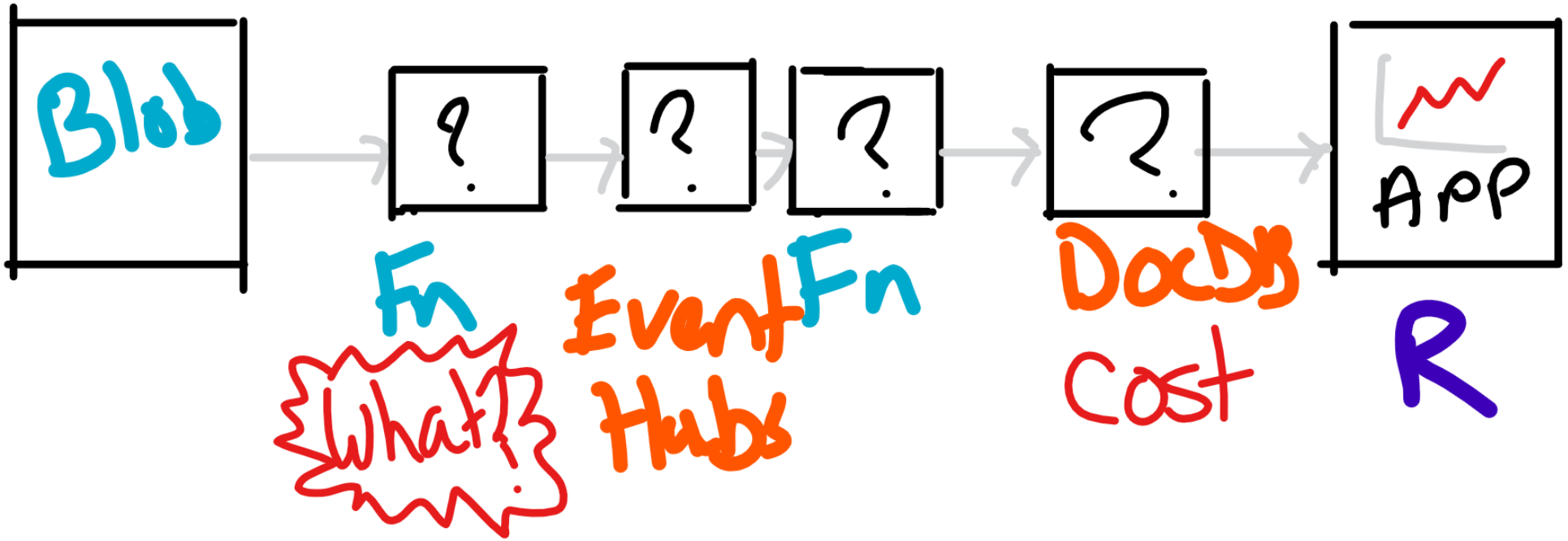
4x expected cost

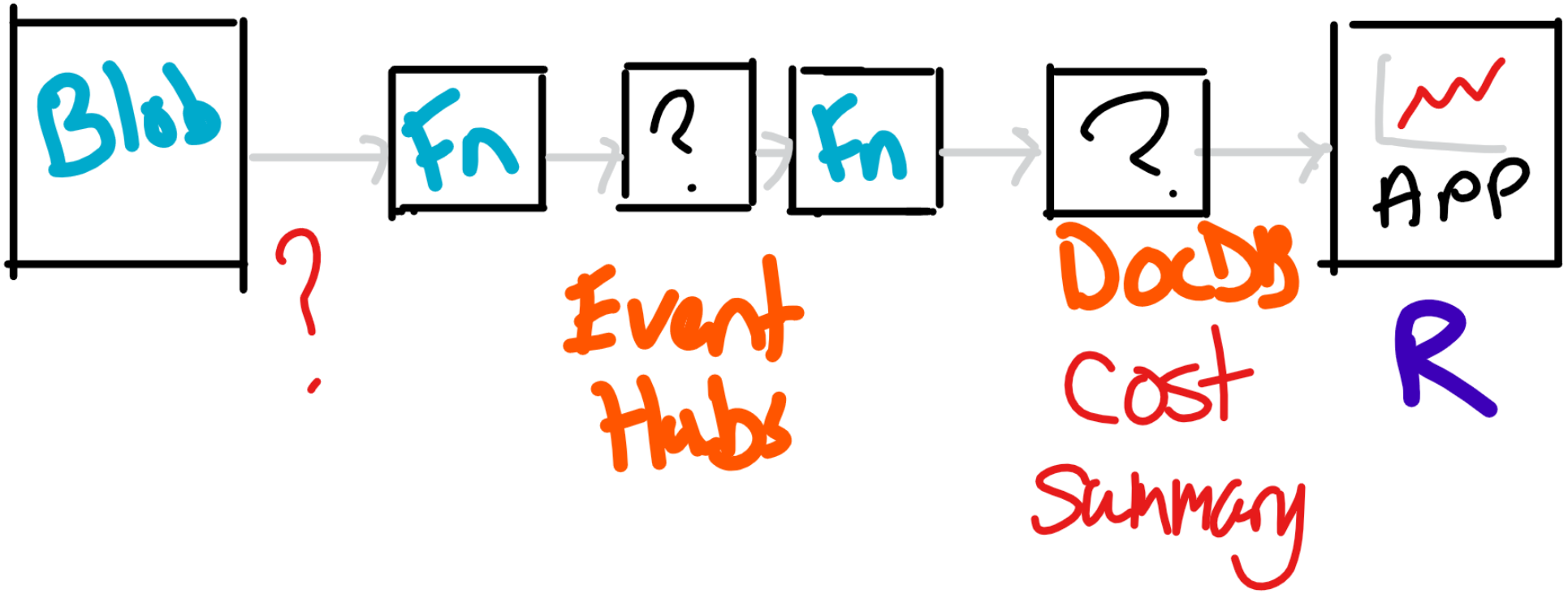
Summarisation is hard

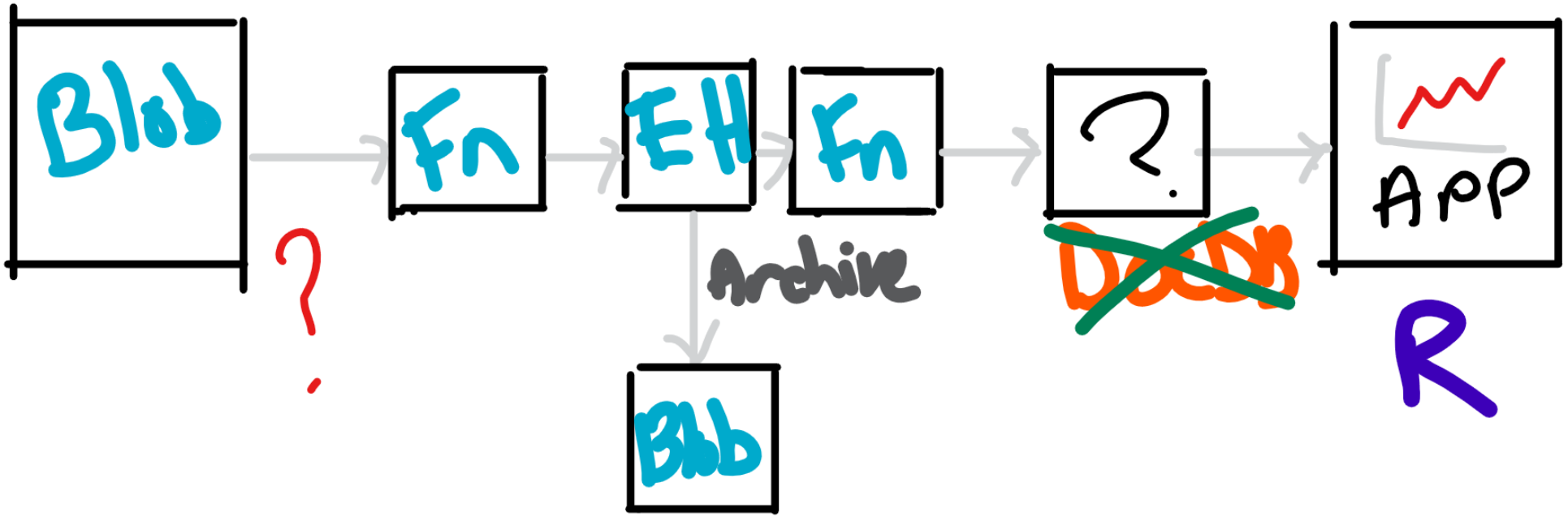
R

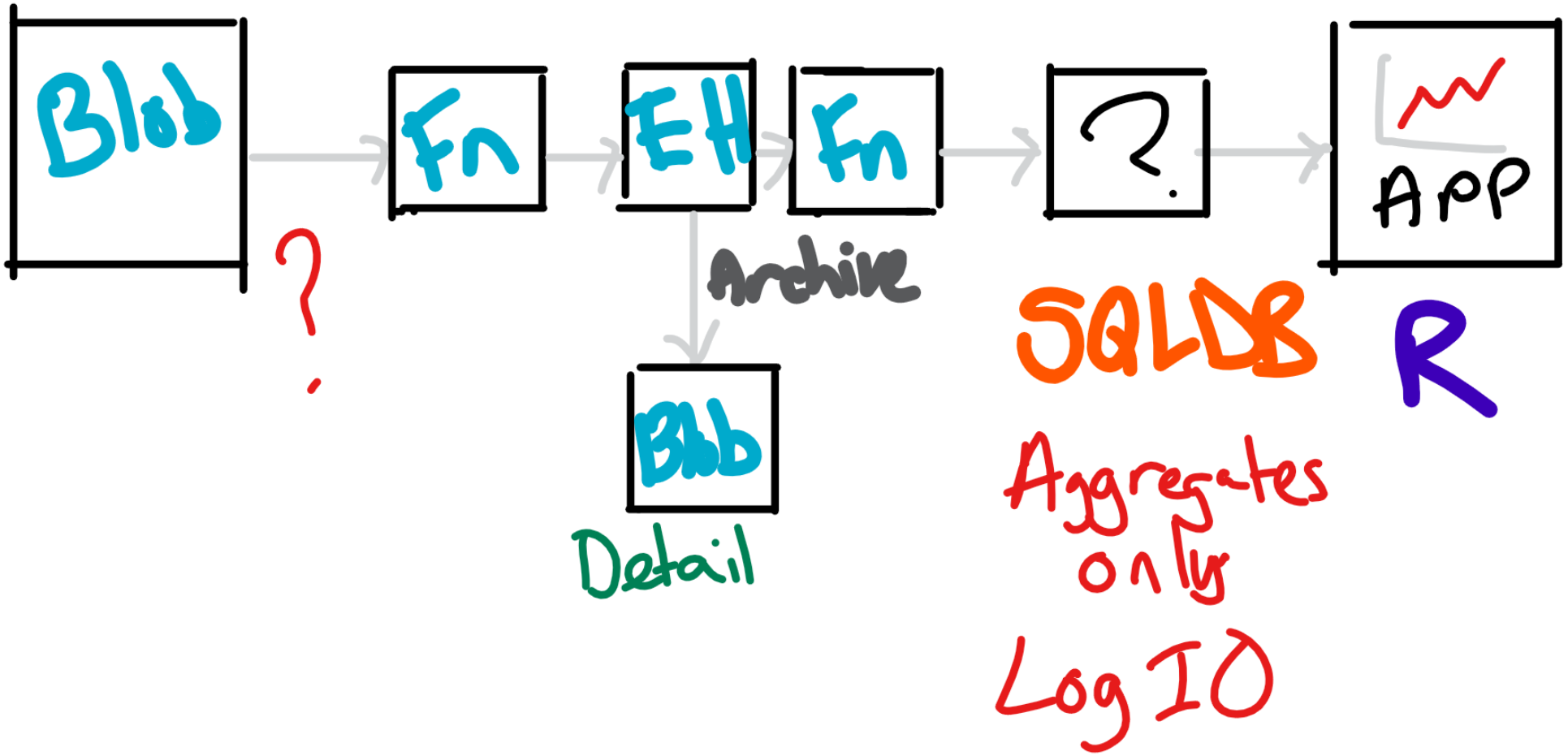
New Regions

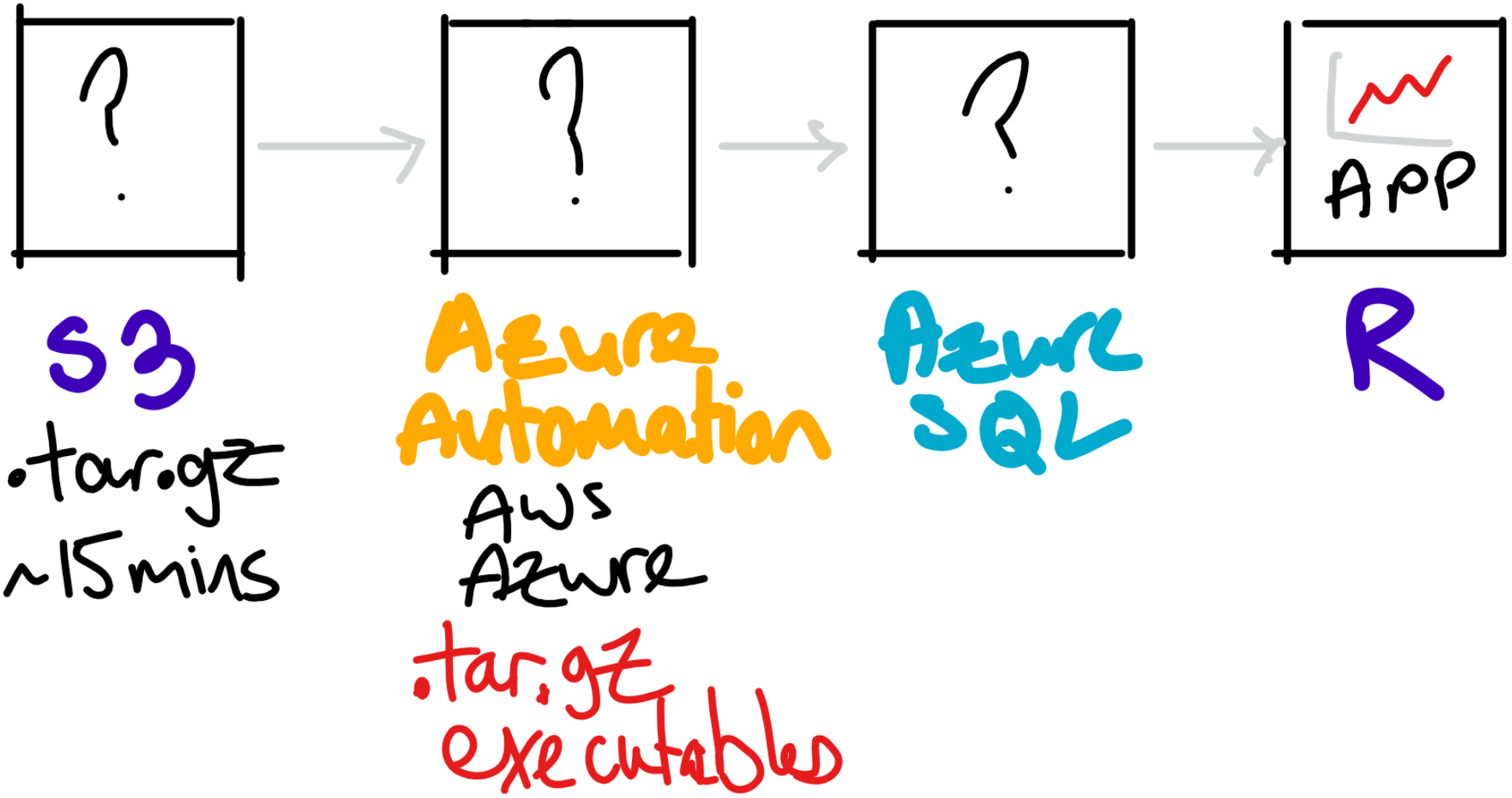

Blob
?
?
?
?
App
Depersonalise
Aggregate & store
R

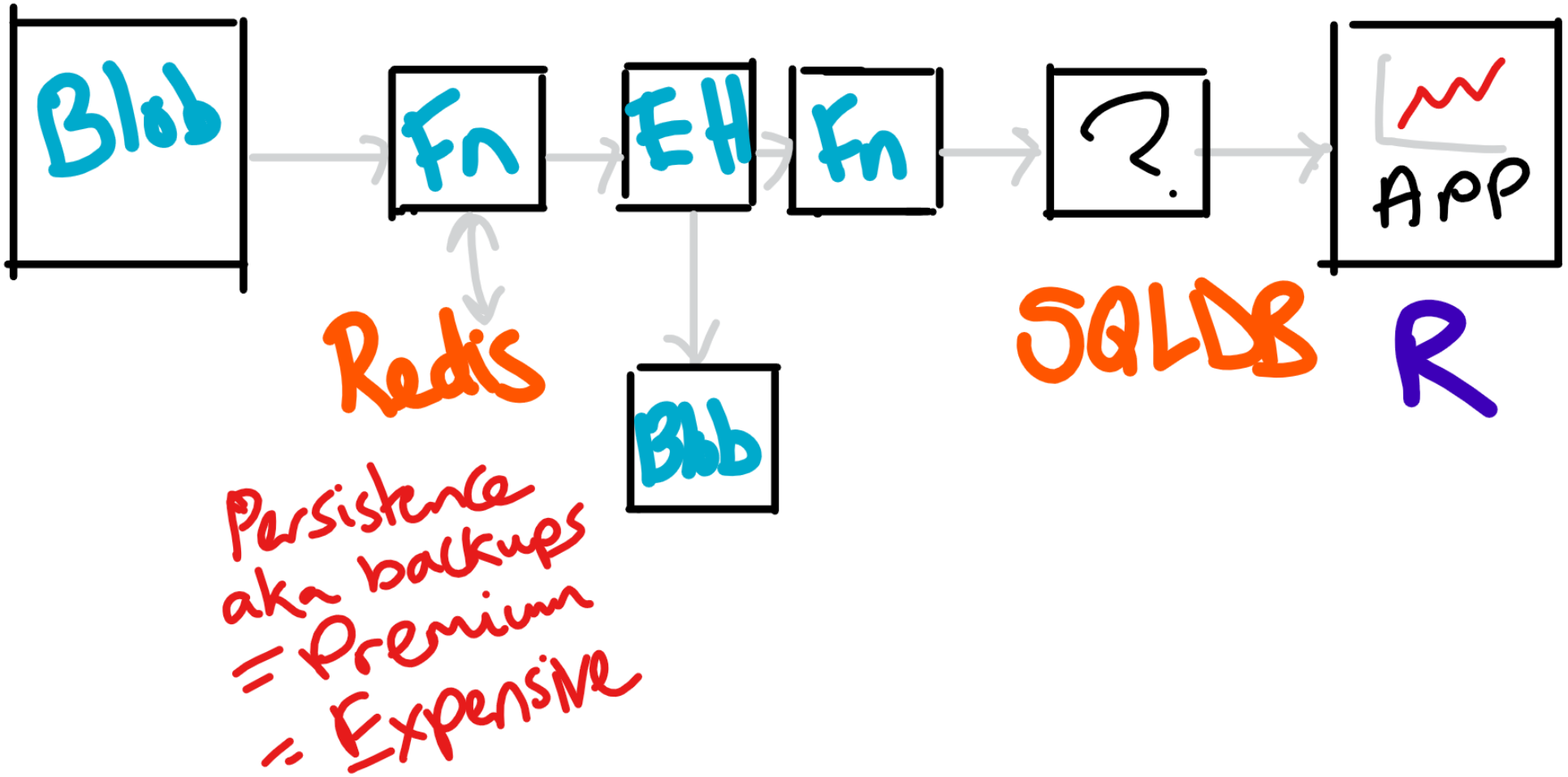

Blob
?
?
?
?
App
Fn
What?
EventFn
Hubs
DocDB
Cost
R

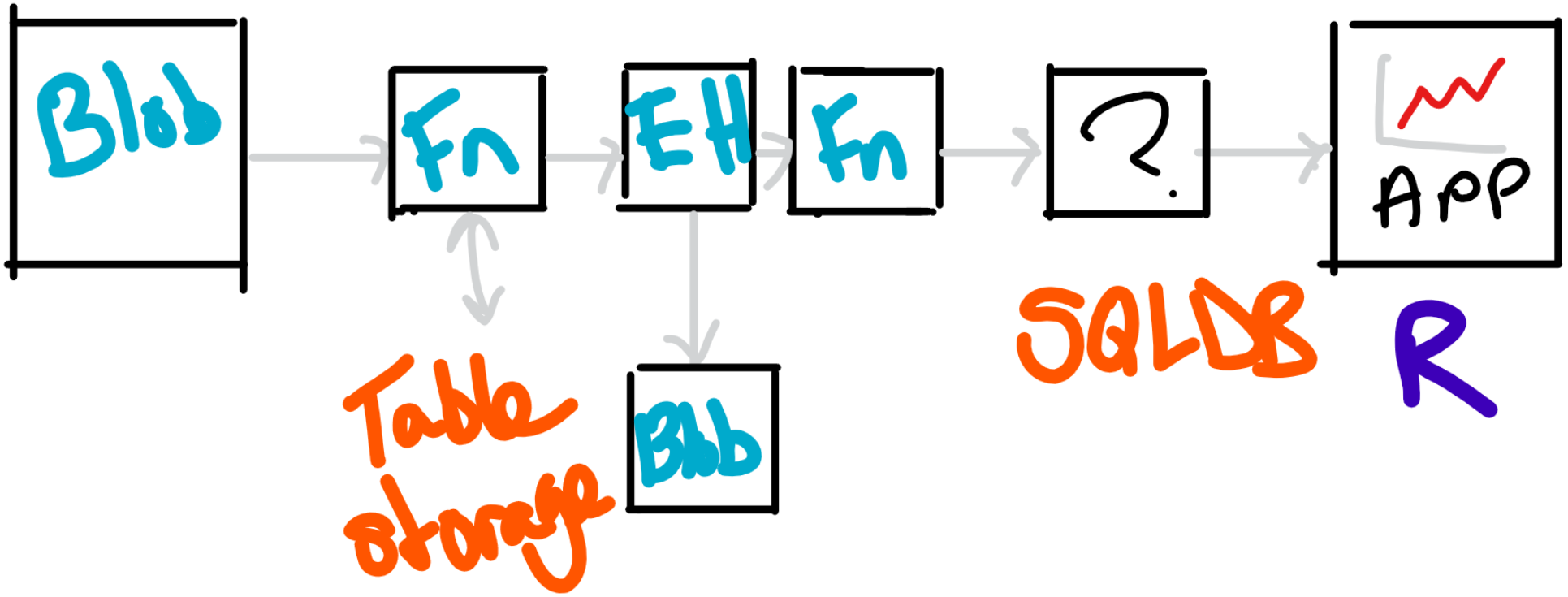

Blob
?
Fn
?
Event Hubs
Fn
?
DocDB
Cost
Summary
App
R

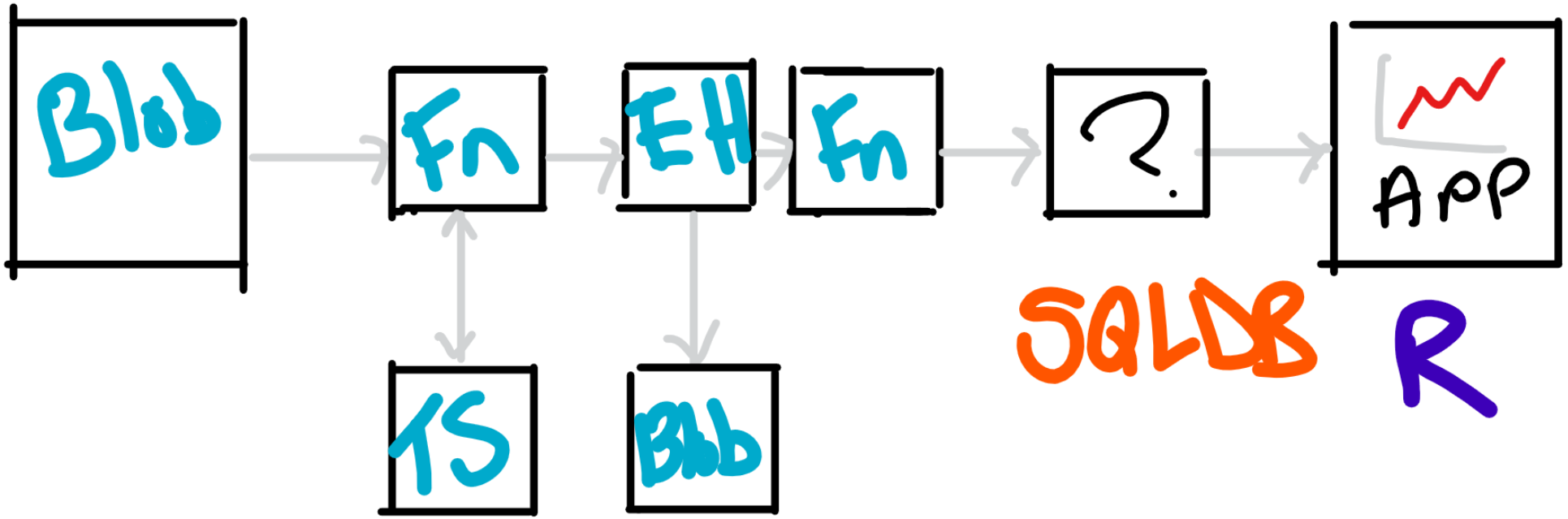

Blob
?
Fn
EH
Fn
Archive
Blob
?
App
R

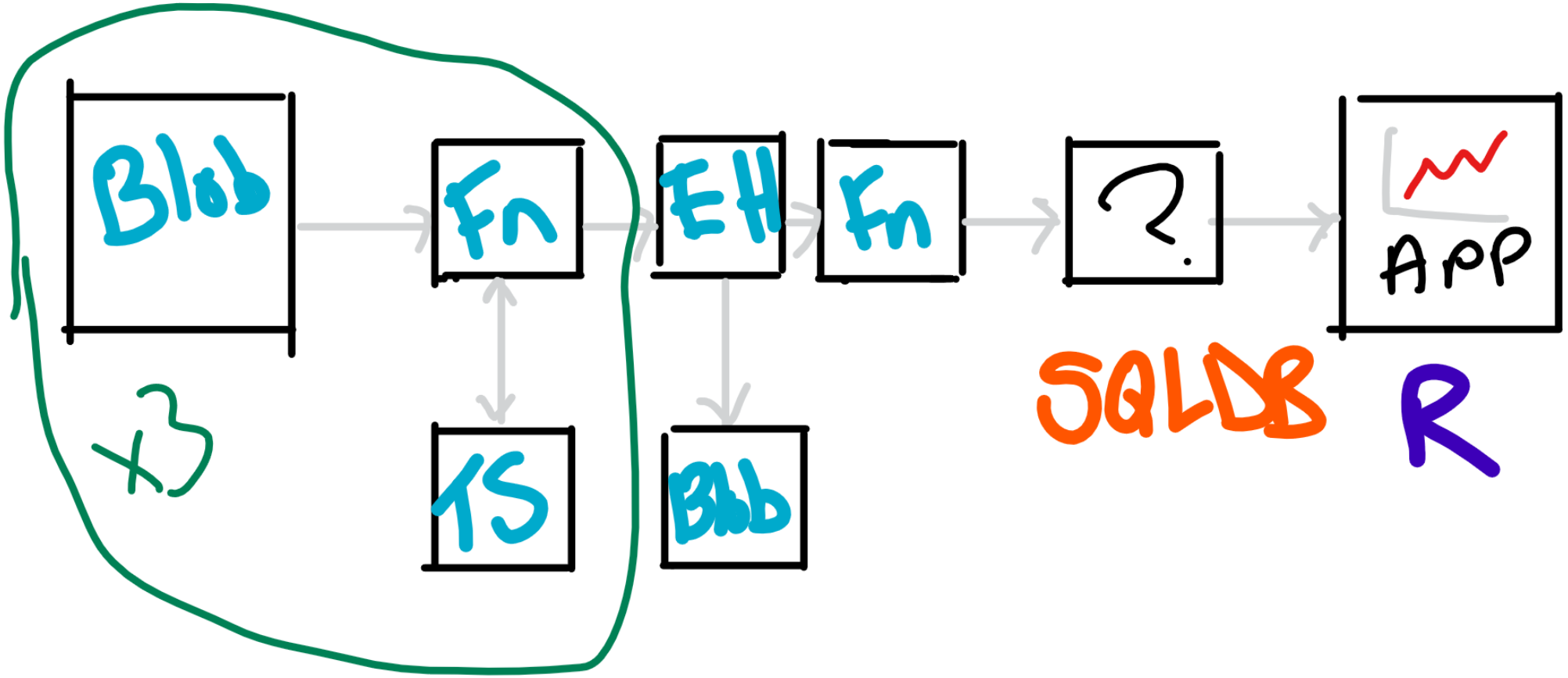

Blob
?
Fn
EH
Fn
Archive
Blob
Detail
?
SQLDB
Aggregates only
Log IO
App
R

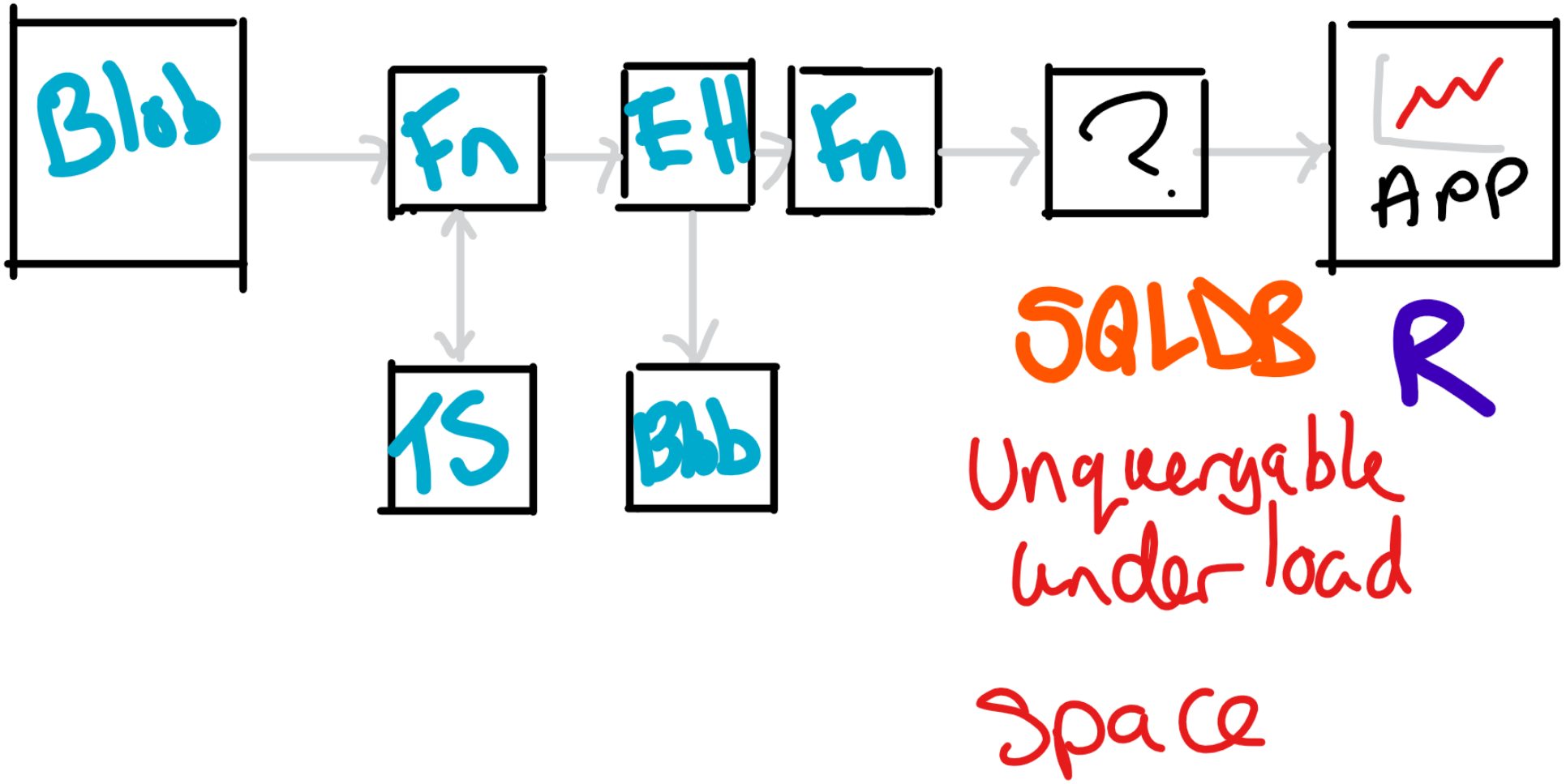

?
?
?
APP
S3
.tar.gz
~15mins
Azure Automation
AWS
Azure
.tar.gz
executables
Azure SQL
R

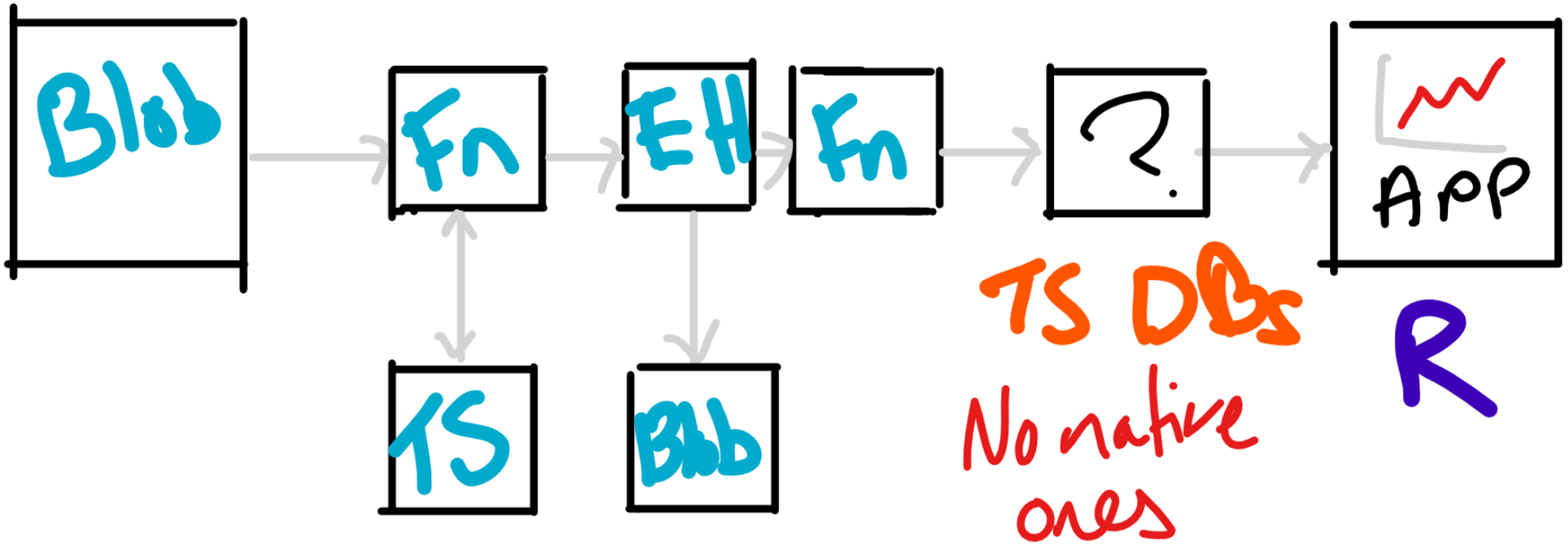

Blob
Fn
EH
Fn
?
APP
Redis
Blob
SQLDB
R
Persistence
aka backups
= Premium
= Expensive

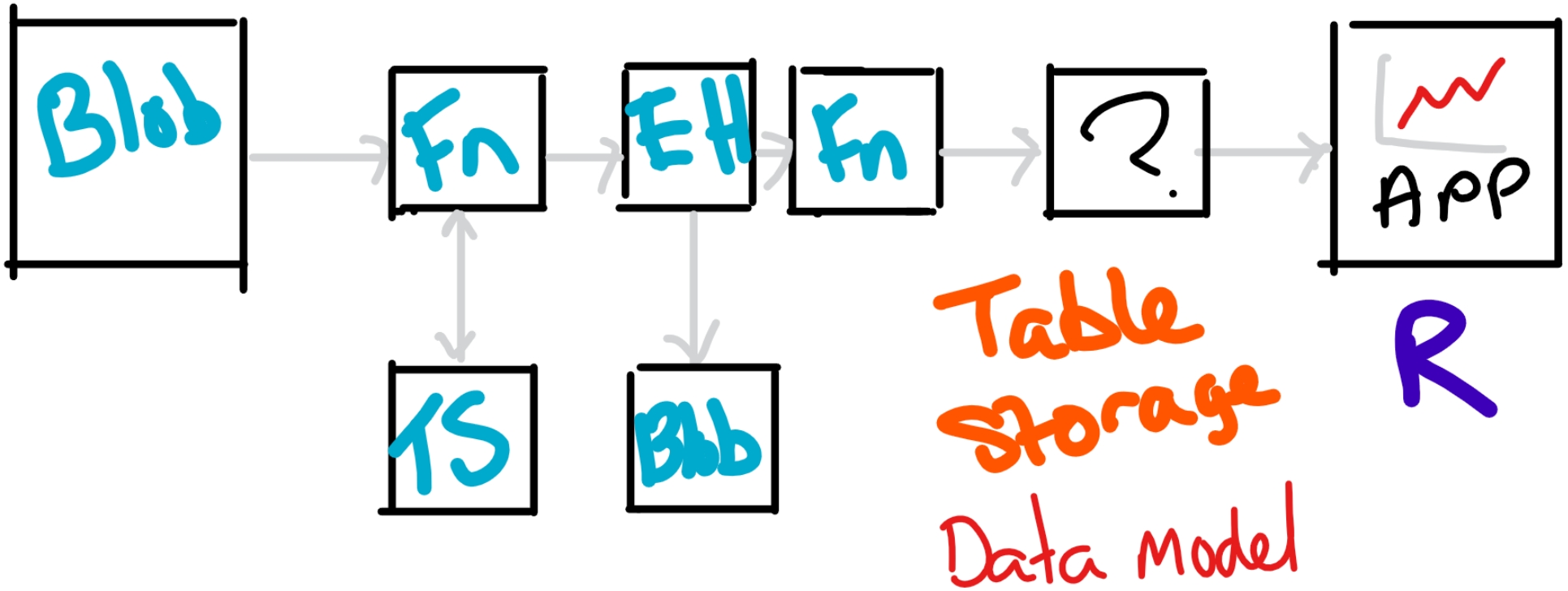

Blob
Fn
EH
Fn
?
APP
Table storage
Blob
SQLDB
R

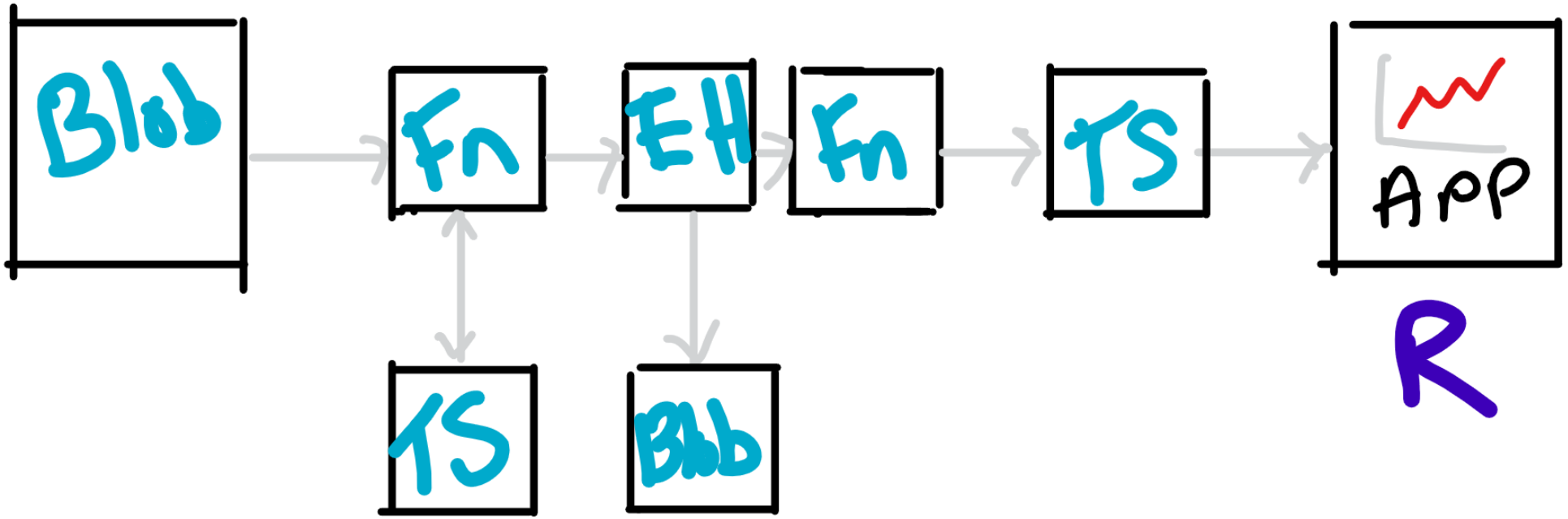

Blob
Fn
EH
Fn
?
App
IS
Blob
SQLDB
R

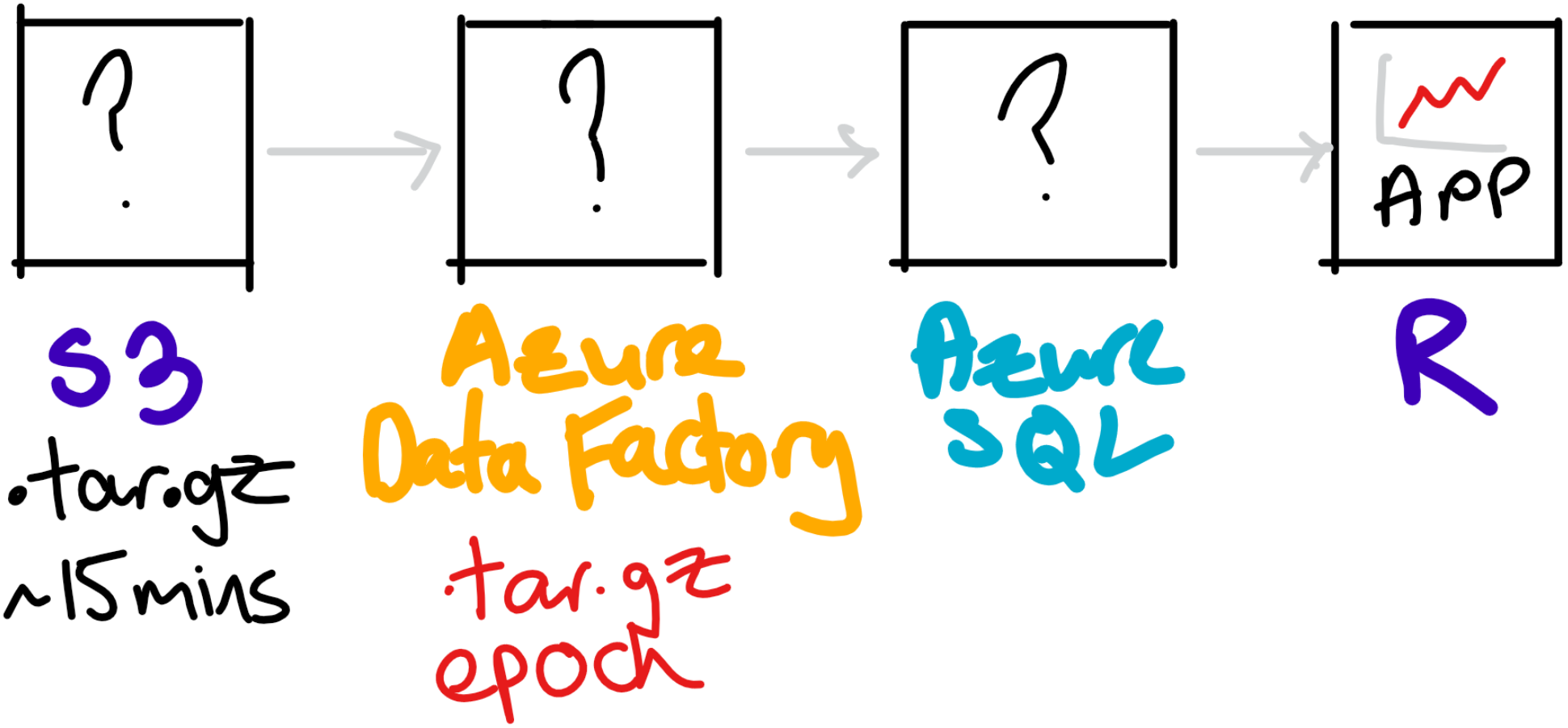

Blob
Fn
EH
Fn
?
APP
x3
IS
Blob
SQLDB
R

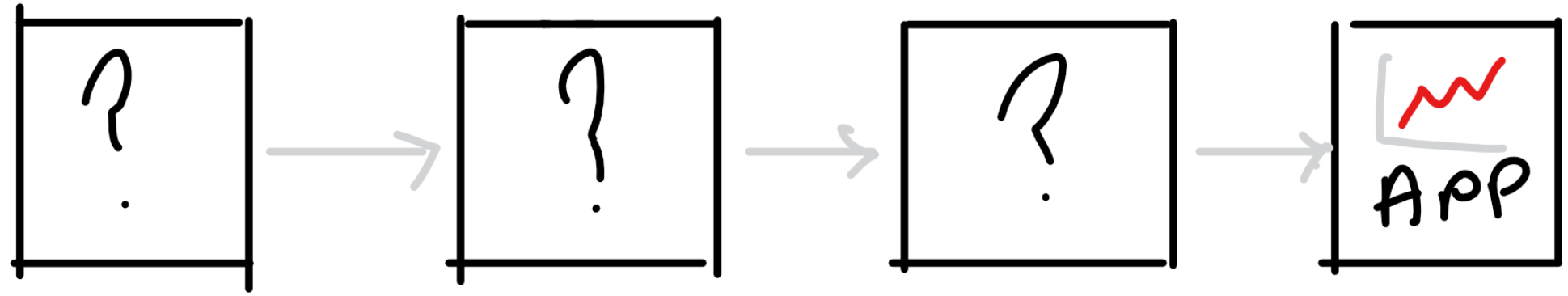

Blob
Fn
EH
Fn
?
APP
IS
Blob
SQLDB
R
Unqueryable
under-load
space


Blob
Fn
EH
Fn
?
App
TS
Blob
TS DBs
No native ones
R


Blob
Fn
EH
Fn
?
APP
IS
Blob
Table storage
R
Data model


Blob
Fn
EH
Fn
TS
App
TS
Blob
R


?
?
?
APP
S3
.tar.gz
~15mins
Azure
Data Factory
.tar.gz
epoch
Azure
SQL
R



S3

.tar.gz

~15mins

Azure Functions

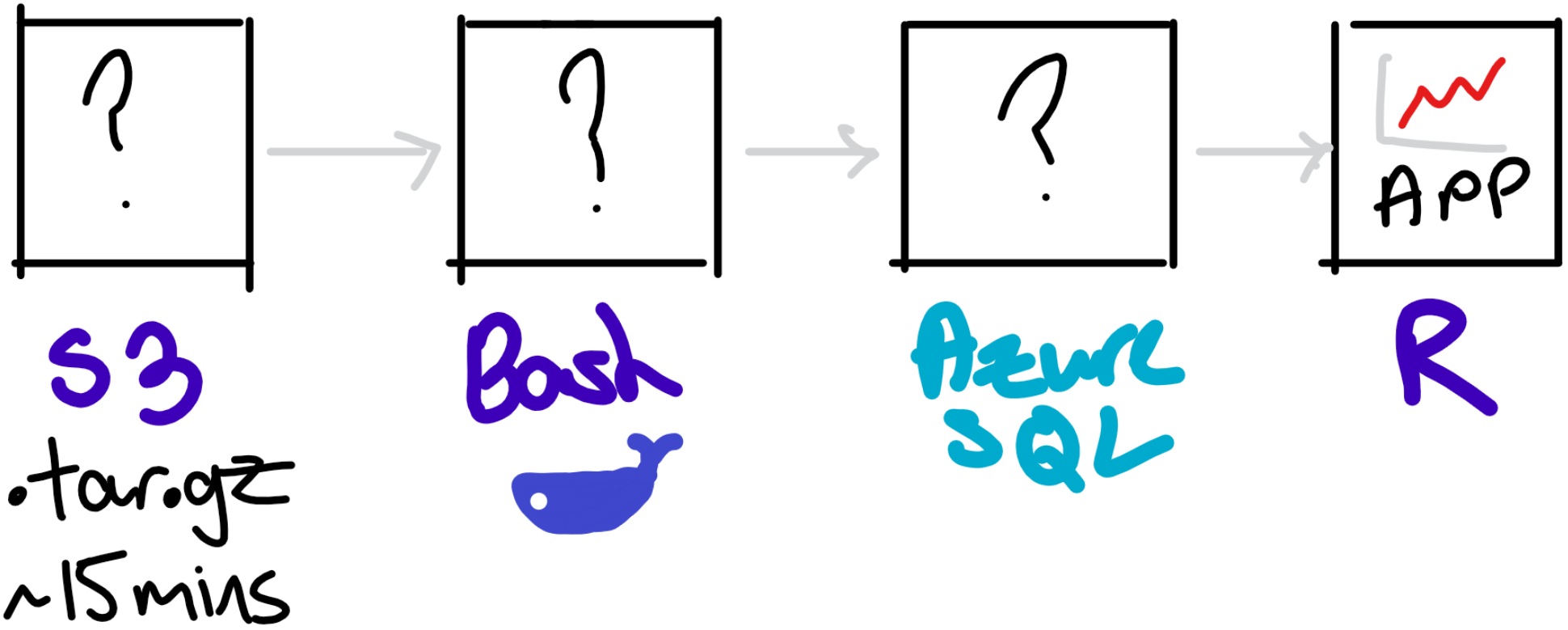
S3

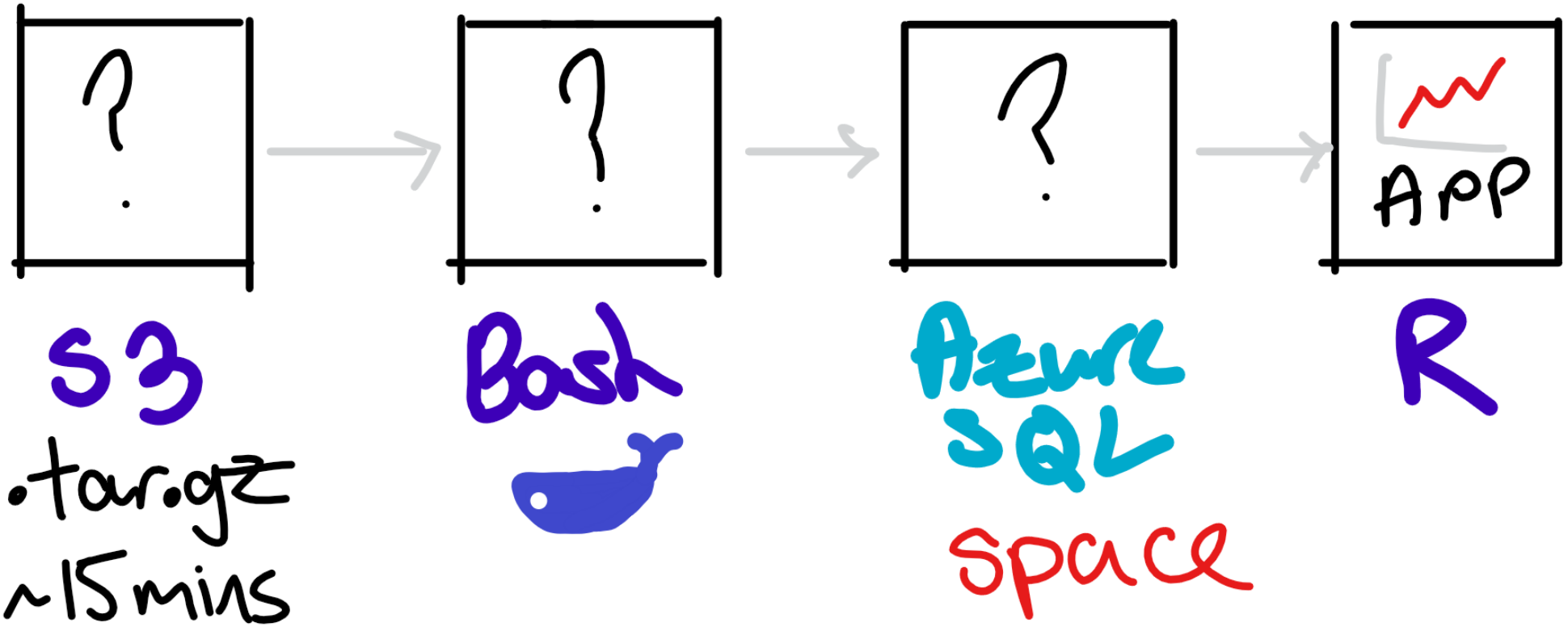
C#

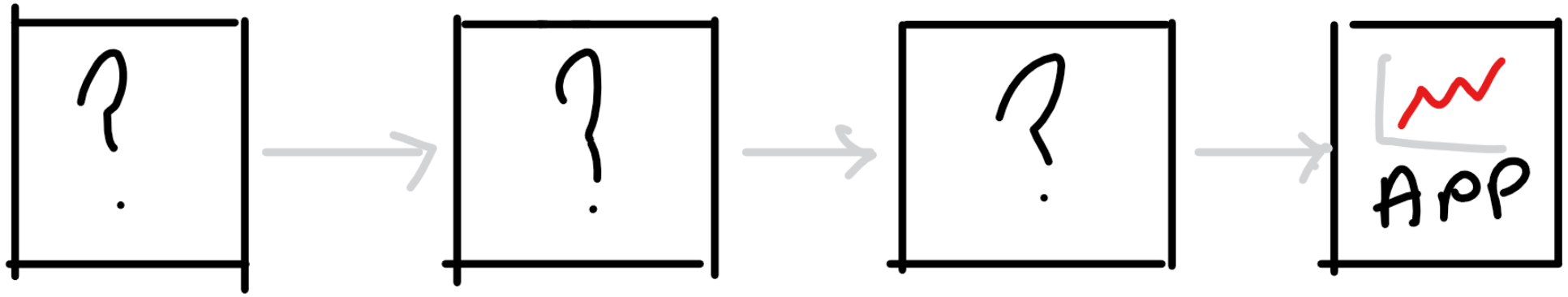
.tar.gz

Azure SQL

R


?
?
?
APP
S3
.tar.gz
~15mins
Bash
Azure SQL
R

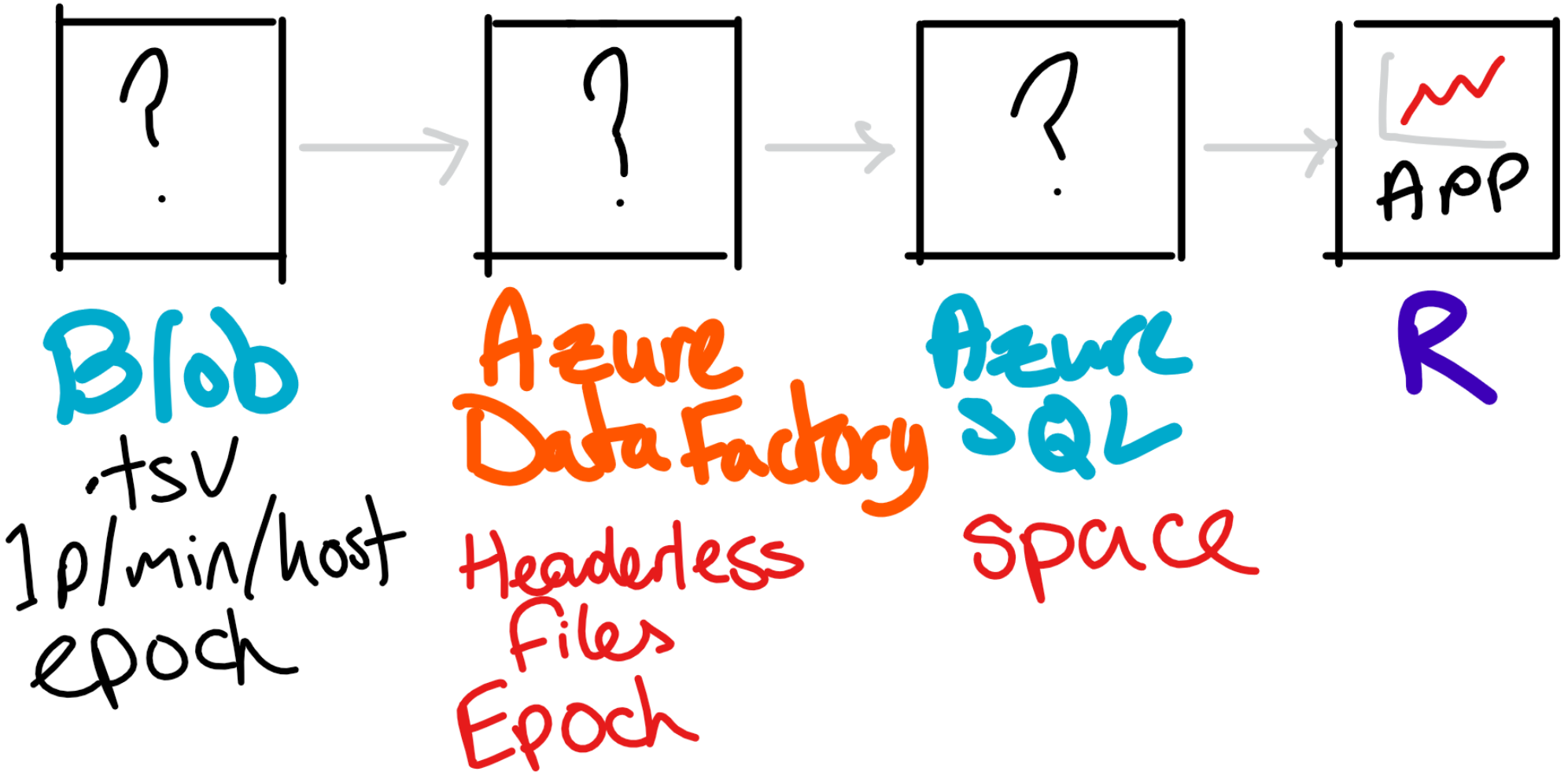

?
?
?
APP
S3
.tar.gz
~15mins
Bash
Azure SQL
space
R

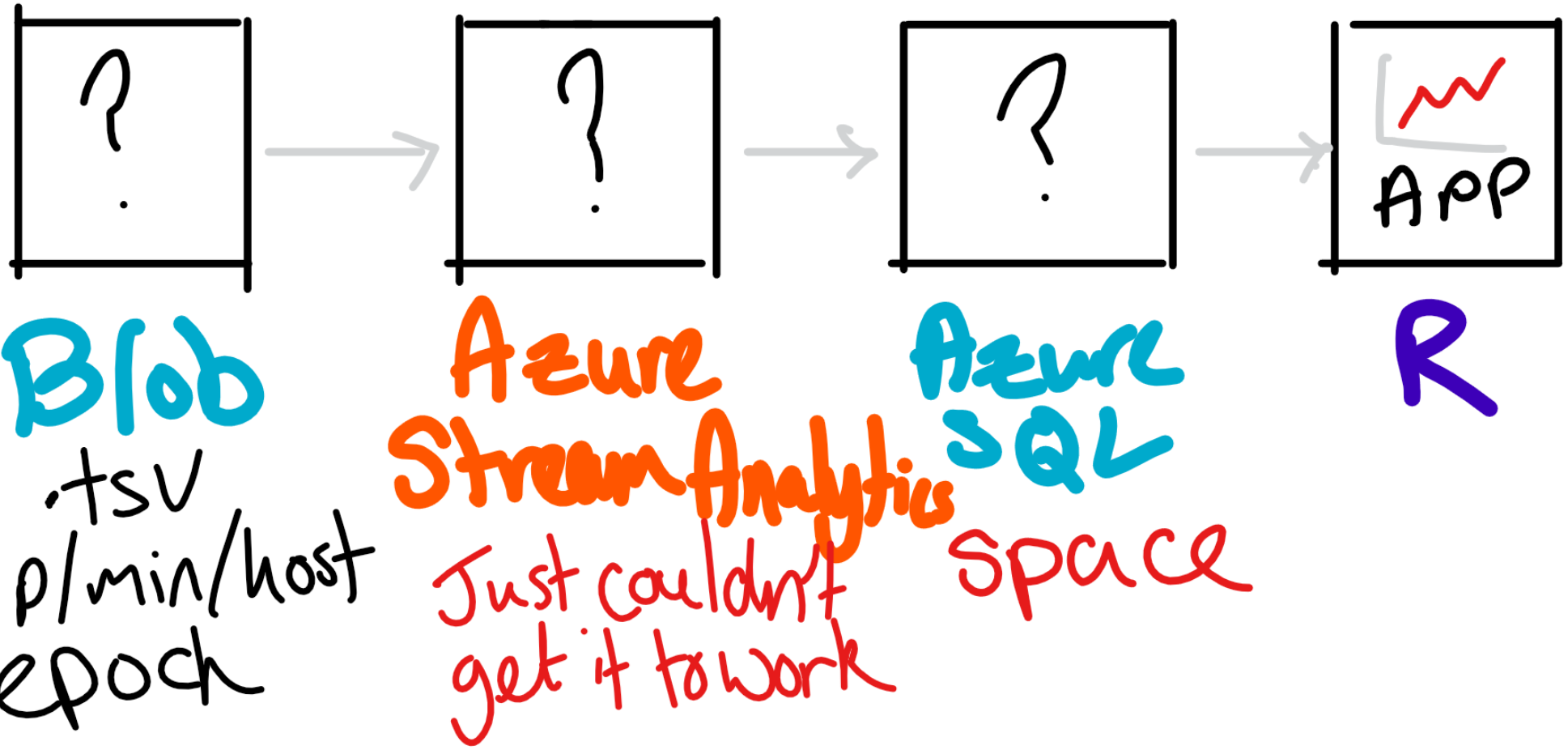


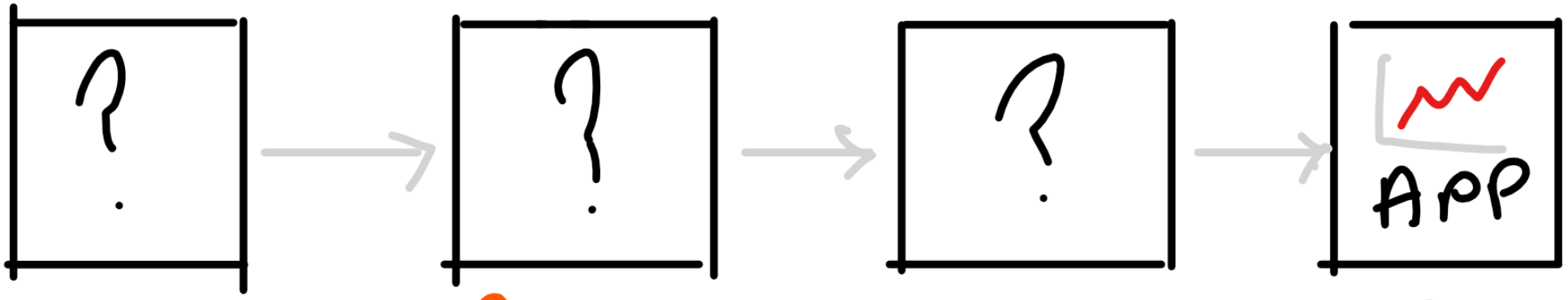
Blob
.tsv
1p/min/host
epoch

Azure SQL
space

R


?
?
?
APP
Blob
.tsv
1p/min/host
epoch
Azure
Data Factory
Headerless
files
Epoch
Azure
SQL
space
R


?
?
?
APP
Blob
.tsv
1p/min/host
epoch
Azure
Stream Analytics
Just couldn't
get it to work
Azure
SQL
space
R



Blob
.tsv
1p/min/host
epoch

Azure Functions

Azure SQL
space

R