5 Levels of High Availability

From Multi-instance to Hybrid Cloud

Rafał Leszko
@RafalLeszko
rafalleszko.com
Hazelcast

About me

- Cloud Software Engineer at Hazelcast
- Worked at Google and CERN
- Author of the book "Continuous Delivery with Docker and Jenkins"
- Trainer and conference speaker
- Live in Kraków, Poland

About Hazelcast

- Distributed Company
- Open Source Software
- 140+ Employees
- Products:
  - Hazelcast IMDG
  - Hazelcast Jet
  - Hazelcast Cloud

@Hazelcast

Agenda

- Introduction
- High Availability Levels
  - Level 0: Single Instance
  - Level 1: Multi Instance
  - Level 2: Multi Zone
  - Level 3: Multi Region
  - Level 4: Multi Cloud
  - Level 5: Hybrid Cloud
- Summary

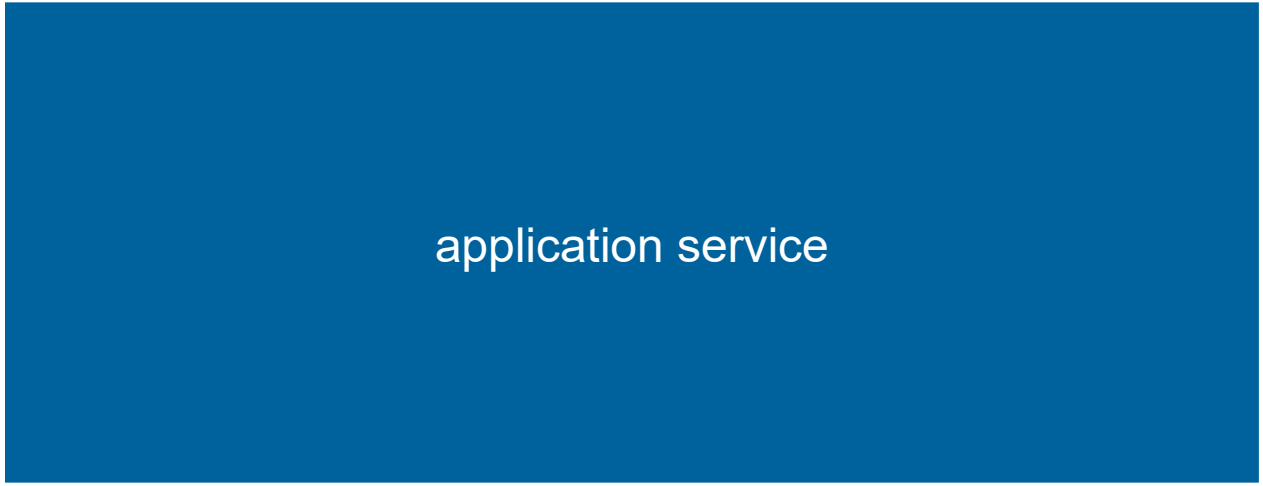
Introduction

hazelcast®

micro-service

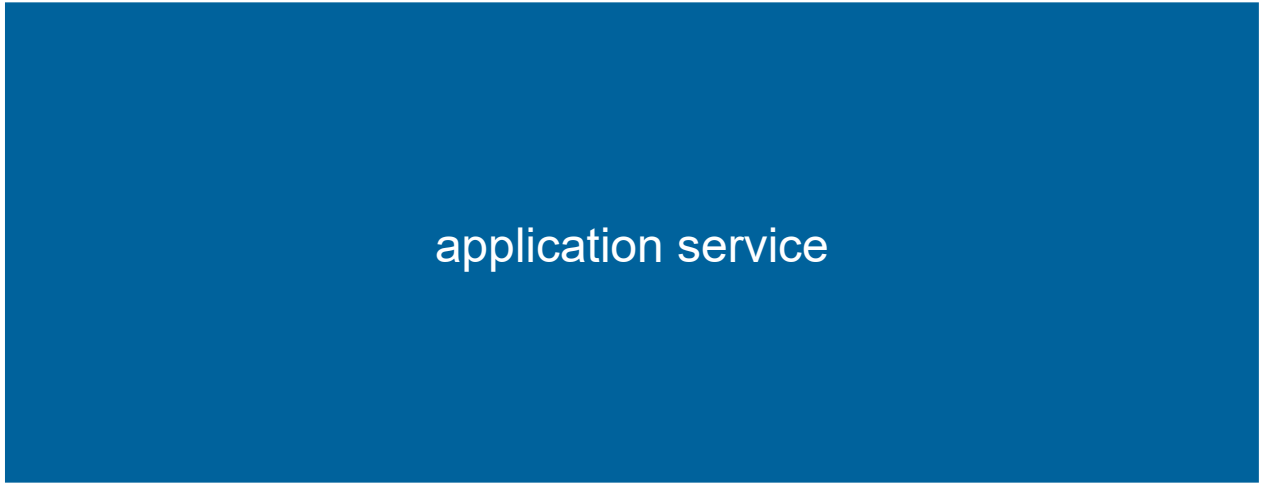
application service

micro?



hazelcast®


application service



hazelcast®

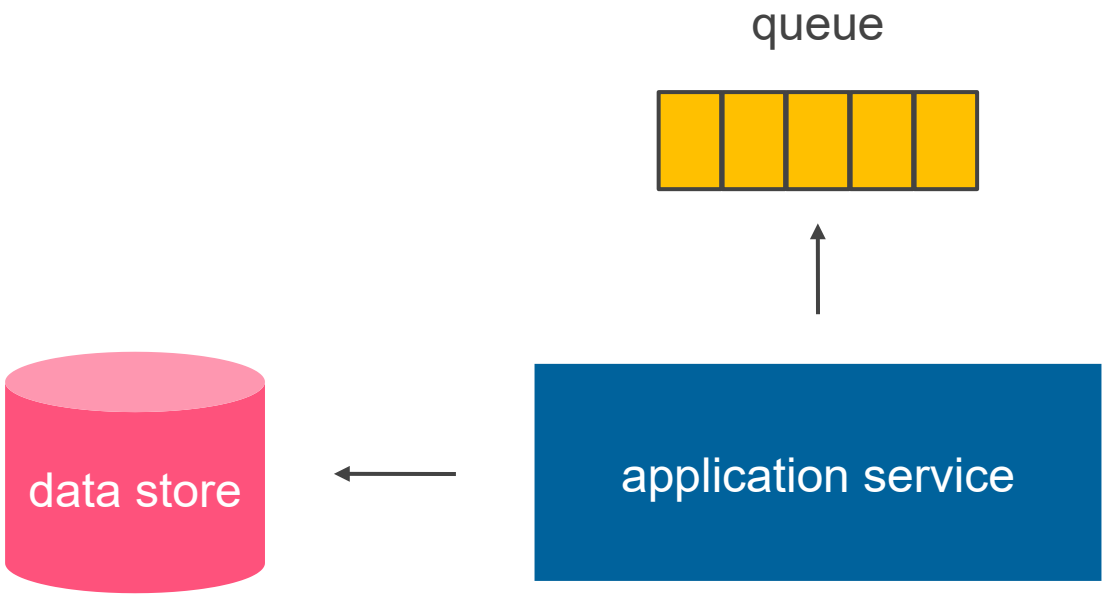
hazelcast®

hazelcast®

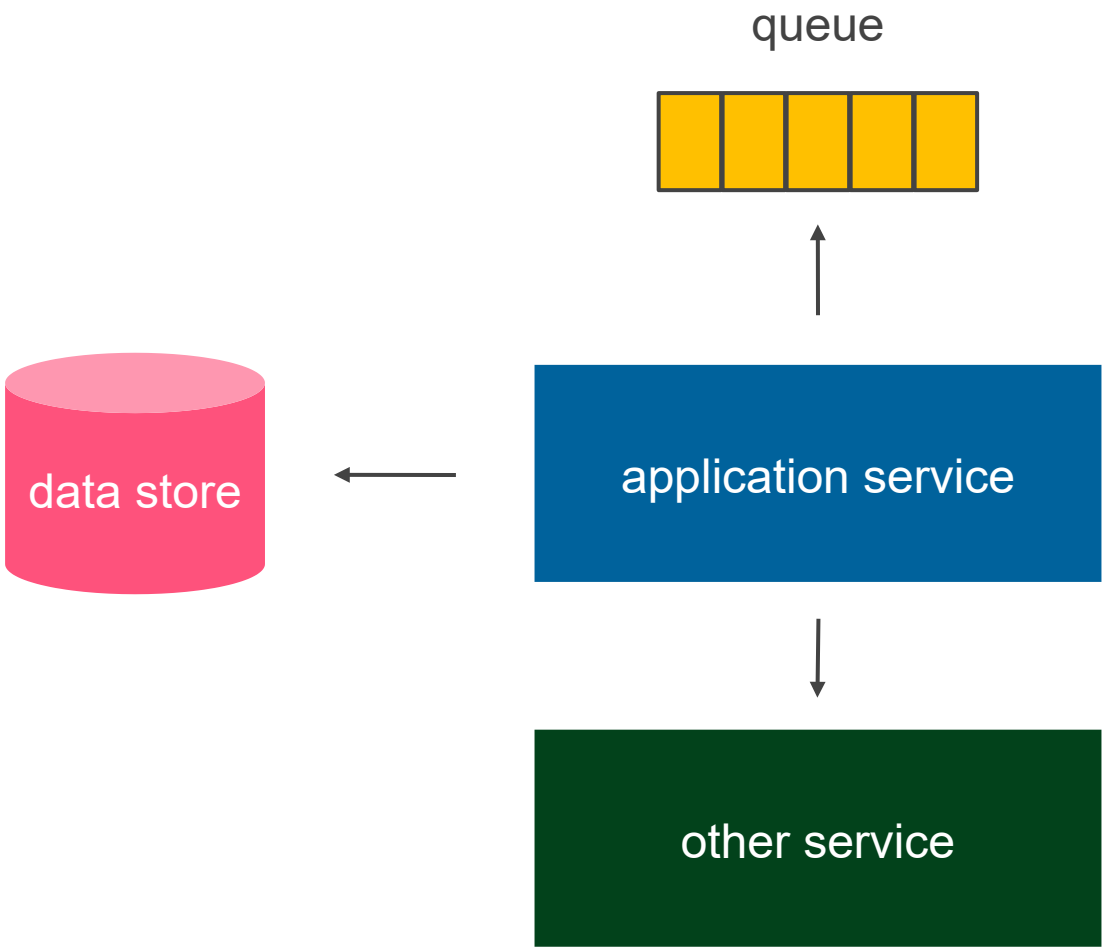
Stateless


GIVE PEACE
A CHANCE

hazelcast®



hazelcast®



hazelcast®

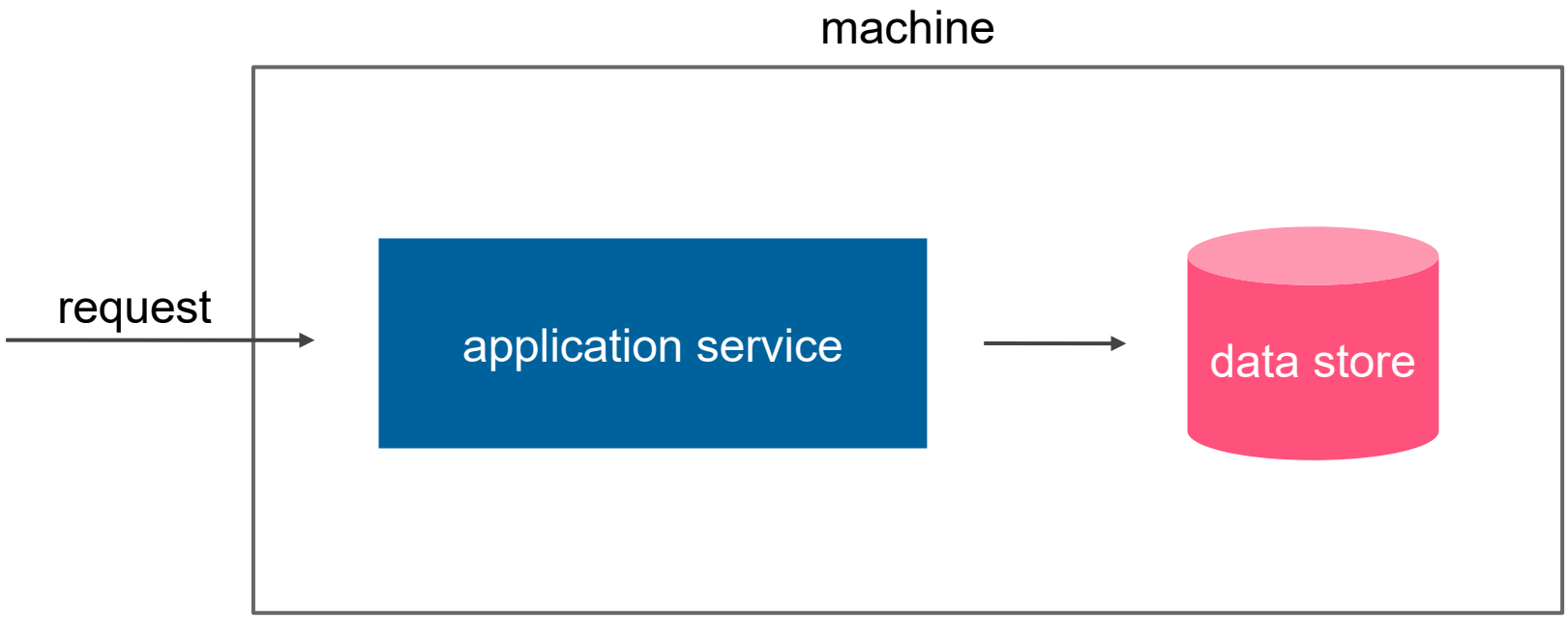
Data is the problem!

Agenda

- Introduction ✔
- High Availability Levels
  - Level 0: Single Instance
  - Level 1: Multi Instance
  - Level 2: Multi Zone
  - Level 3: Multi Region
  - Level 4: Multi Cloud
  - Level 5: Hybrid Cloud
- Summary

Level 0: Single Instance

Level 0: Single Instance



hazelcast®

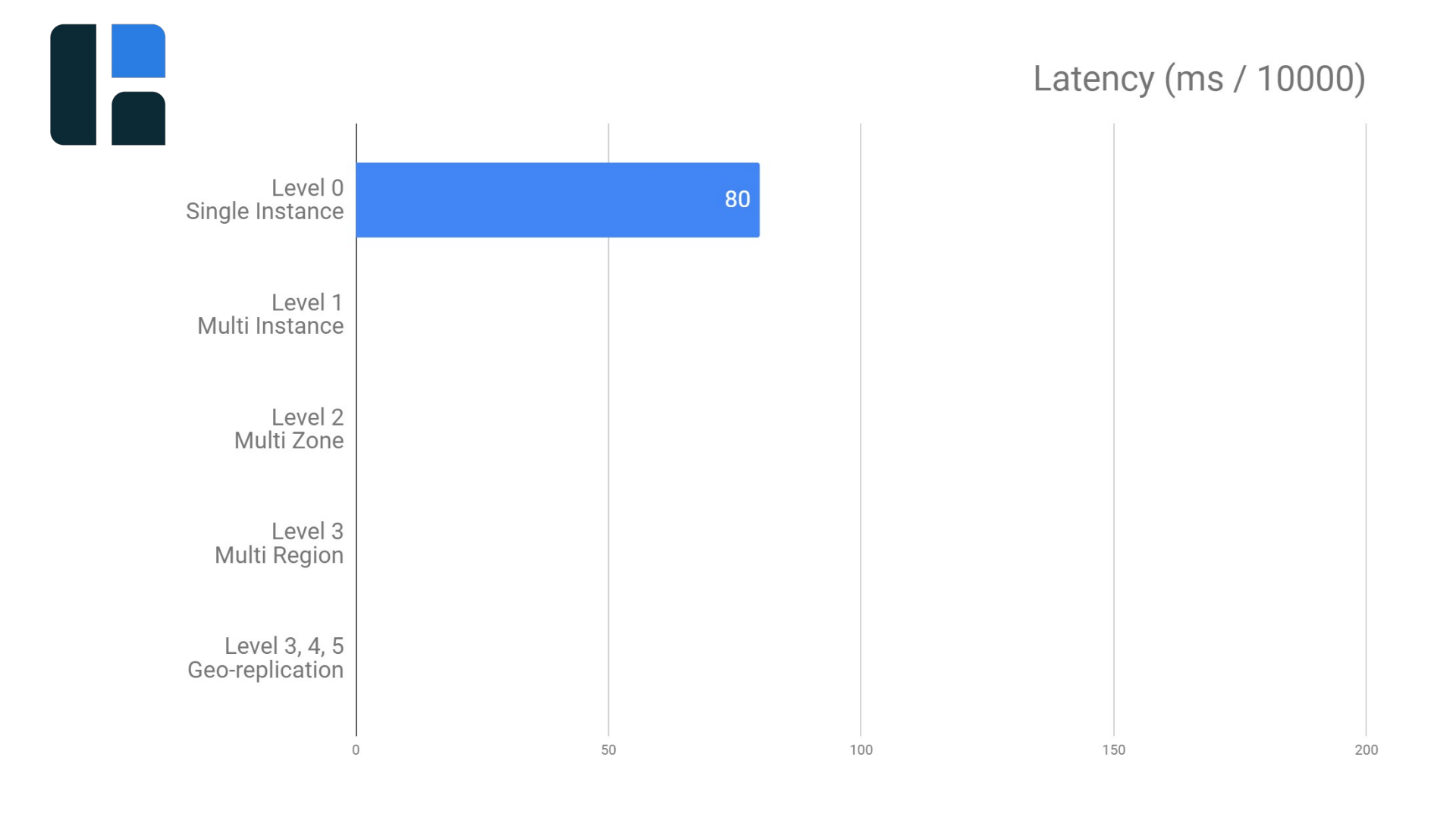
Level 0: Single Instance

YOLO!

hazelcast®

LATENCY EXPERIMENT

hazelcast®


Latency (ms / 10000)
Level 0
Single Instance
80
Level 1
Multi Instance
Level 2
Multi Zone
Level 3
Multi Region
Level 3, 4, 5
Geo-replication
0
50
100
150
200

What does "Level 0: Single Instance" mean to You?

No high availability!

No scalability!

Super low latency:

- in-process memory
- no network
- local file system

Data consistency

hazelcast®

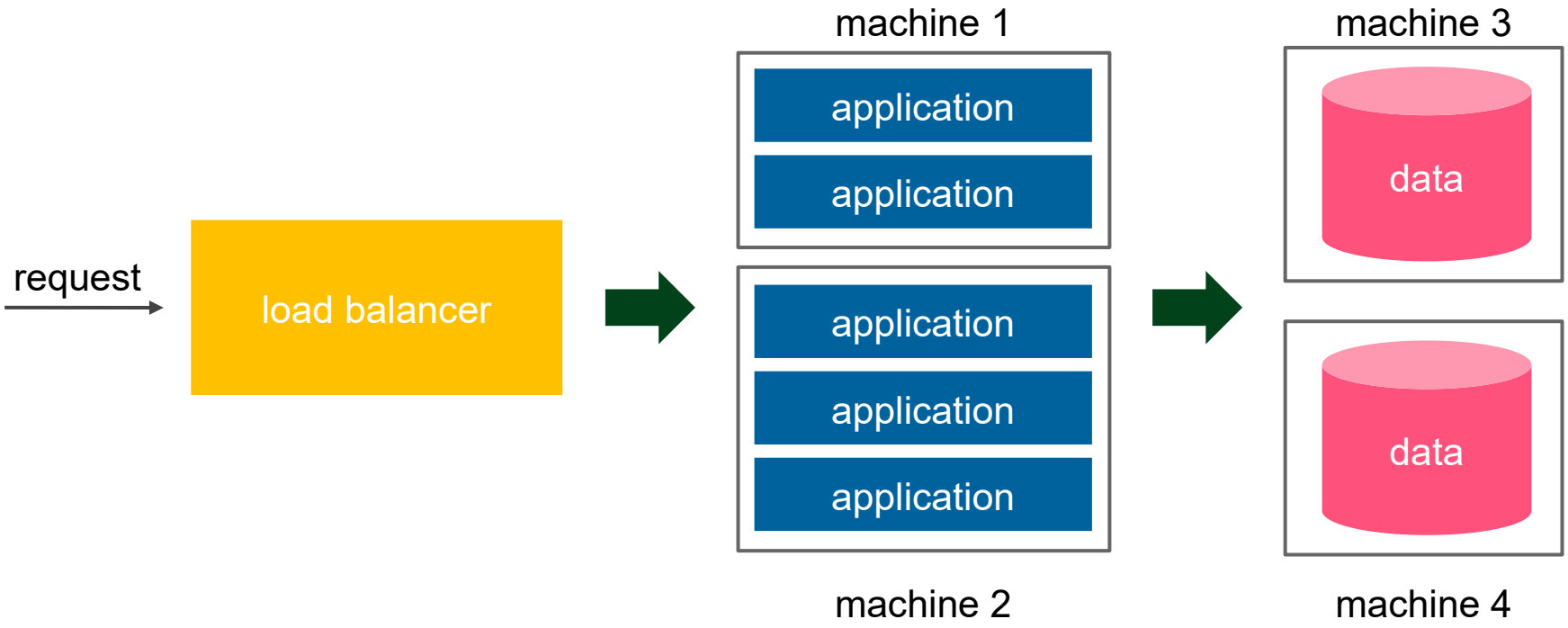
Agenda

- Introduction ✔
- High Availability Levels
  - Level 0: Single Instance ✔
  - Level 1: Multi Instance
  - Level 2: Multi Zone
  - Level 3: Multi Region
  - Level 4: Multi Cloud
  - Level 5: Hybrid Cloud
- Summary

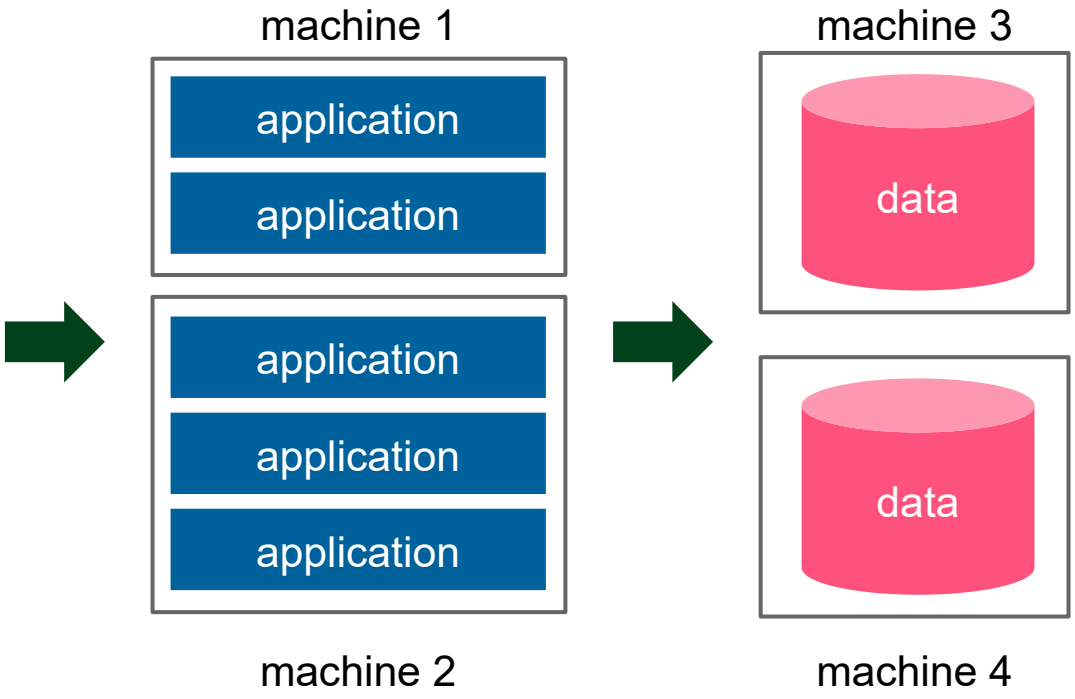
Level 1: Multi Instance

If one machine is down,
the system is still available

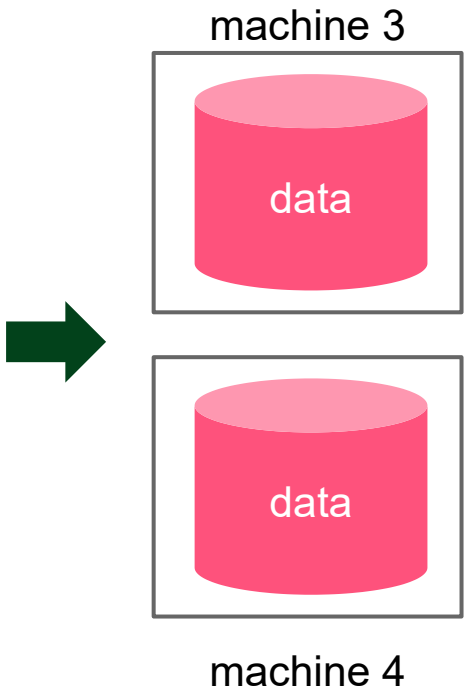
Level 1: Multi Instance



Level 1: Multi Instance



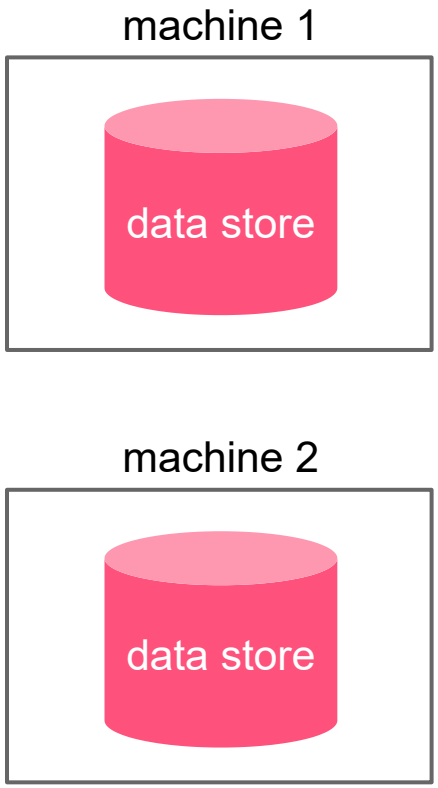
Level 1: Multi Instance



Level 1: Multi Instance

Assumptions:

- Local network
- Fast
- Reliable

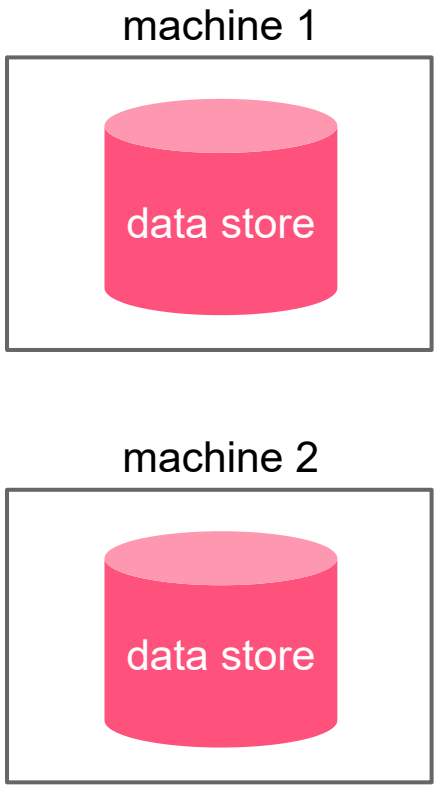


For example:

- EC2 Instances in the same availability zone
- GCP VM instances in the same zone
- Your on-premises server machines connected with LAN

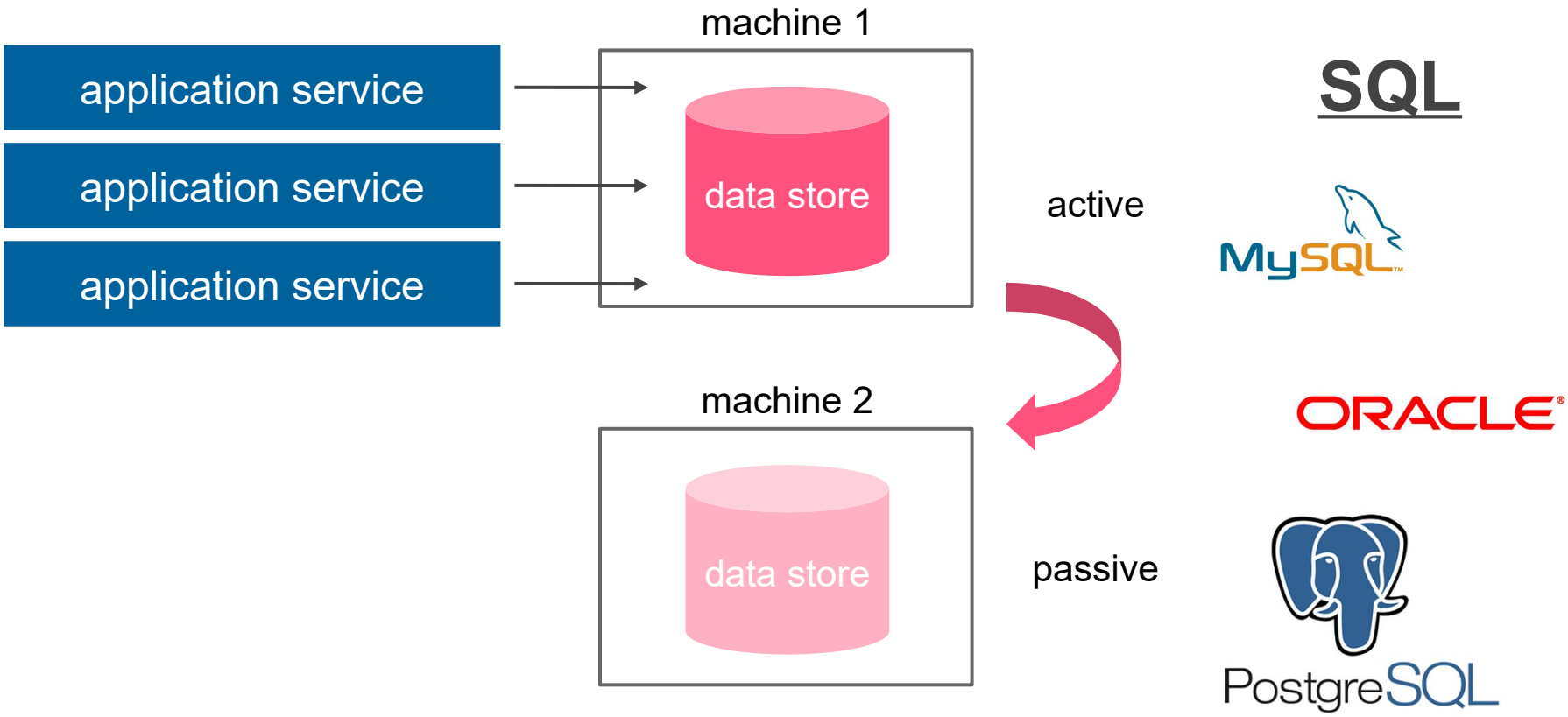

70
80
RI
RJ
RK
MAIL SERVERS
DATABASE SERVERS
PC DESKTOP
WORLD WIDE WEB SERVERS
PC ENGINEERING LICENSE SERVERS

Level 1: Multi Instance

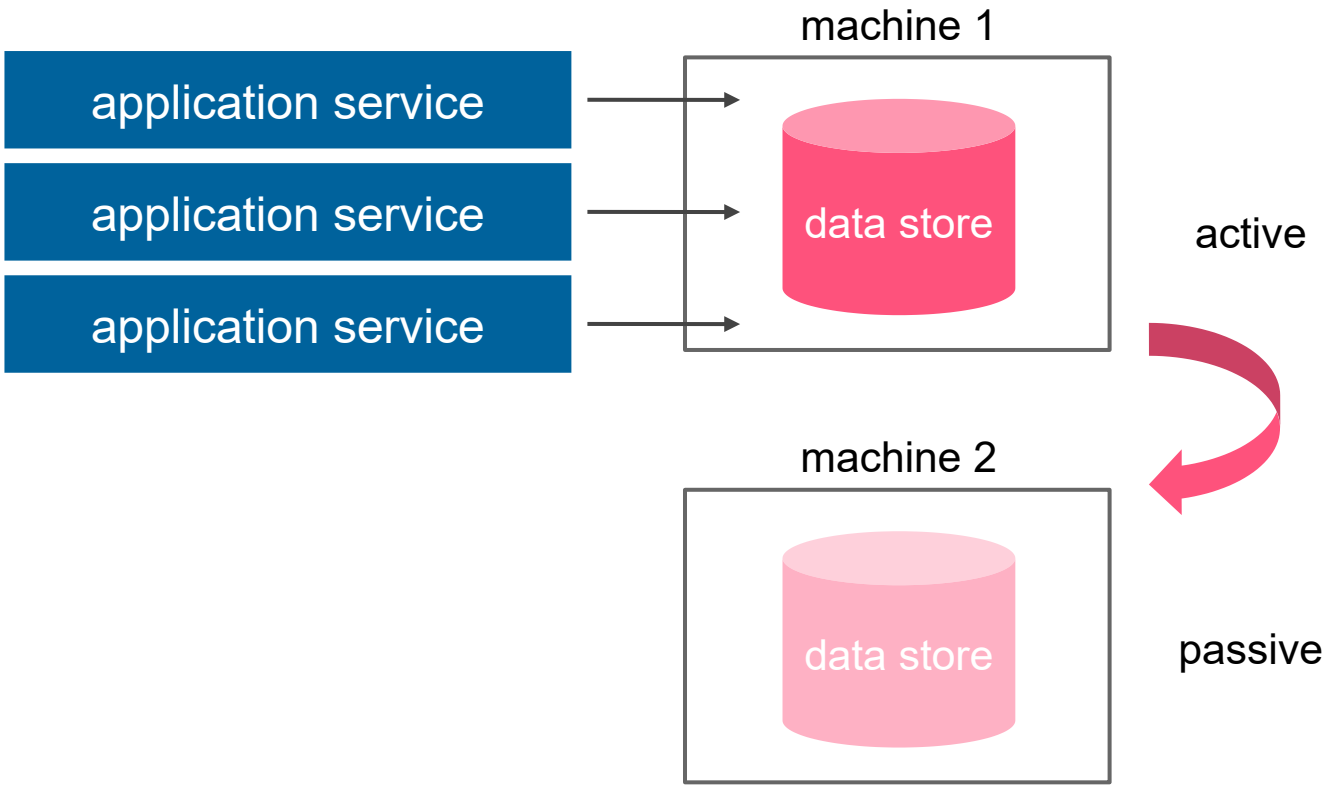


Data replication

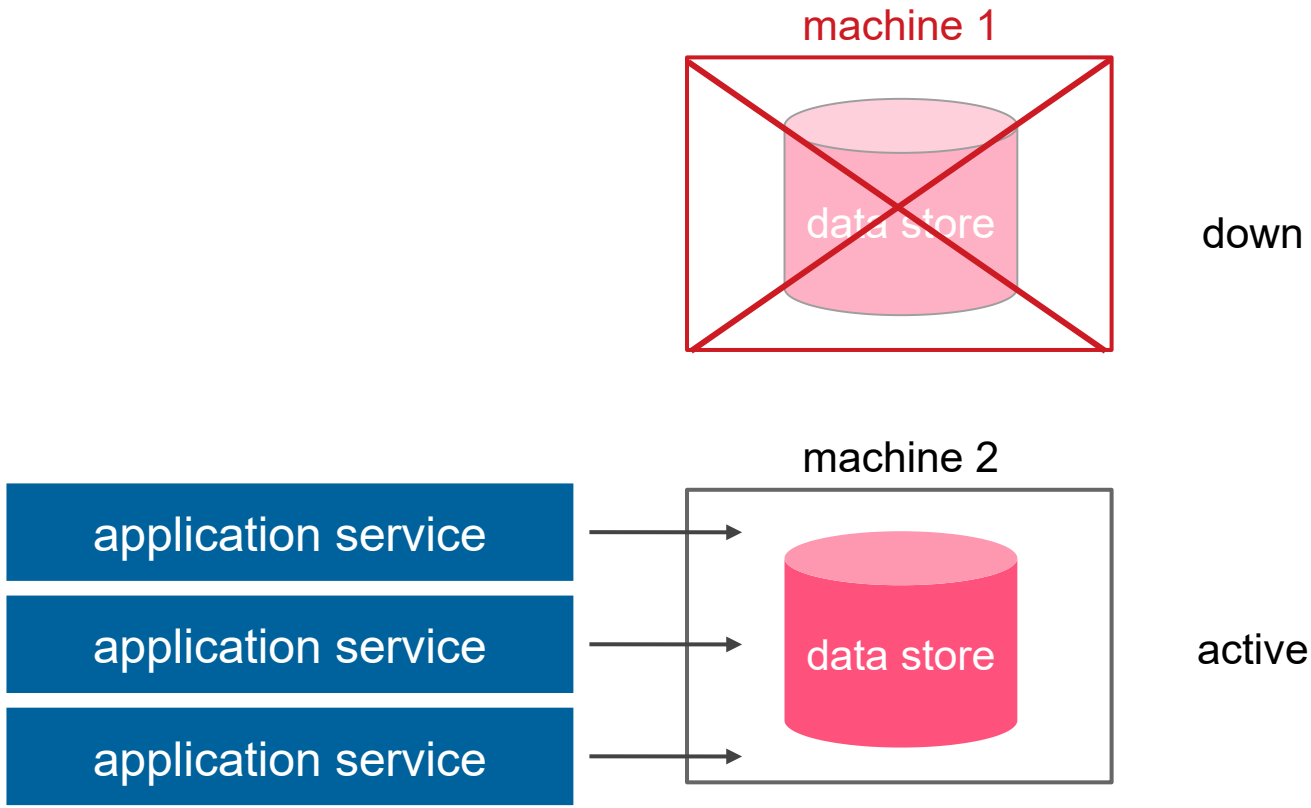
Option 1: Active-Passive (Master-Slave) Replication



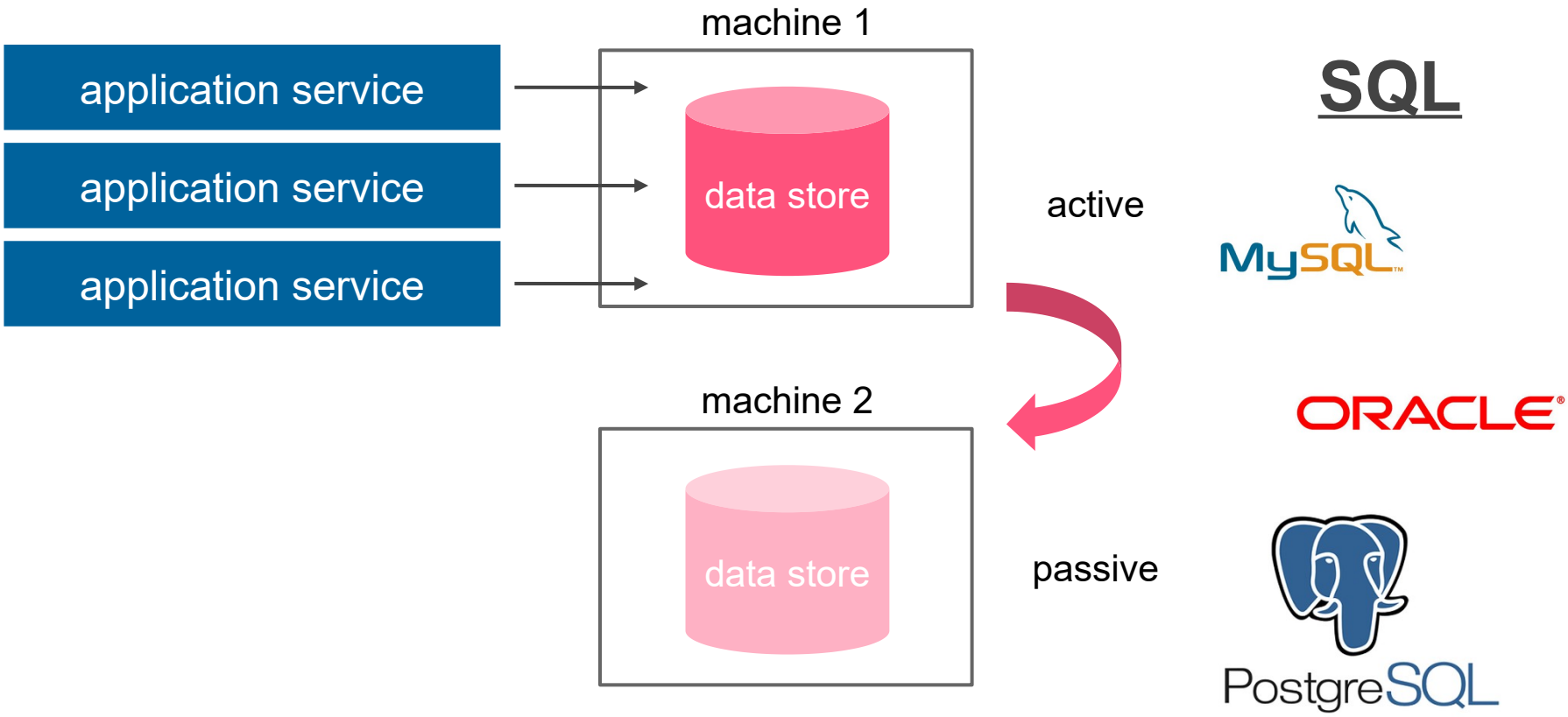
Option 1: Active-Passive (Master-Slave) Replication



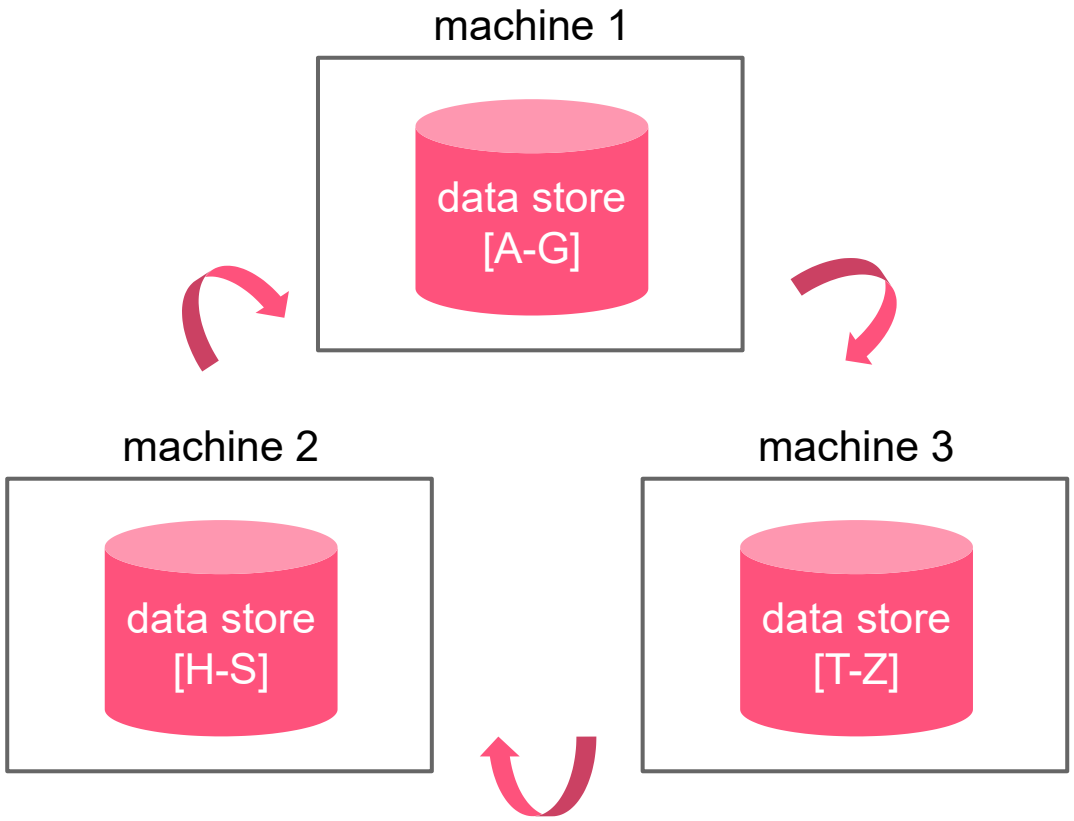
Option 1: Active-Passive (Master-Slave) Replication



Option 1: Active-Passive (Master-Slave) Replication



Option 2: Clustering



NoSQL

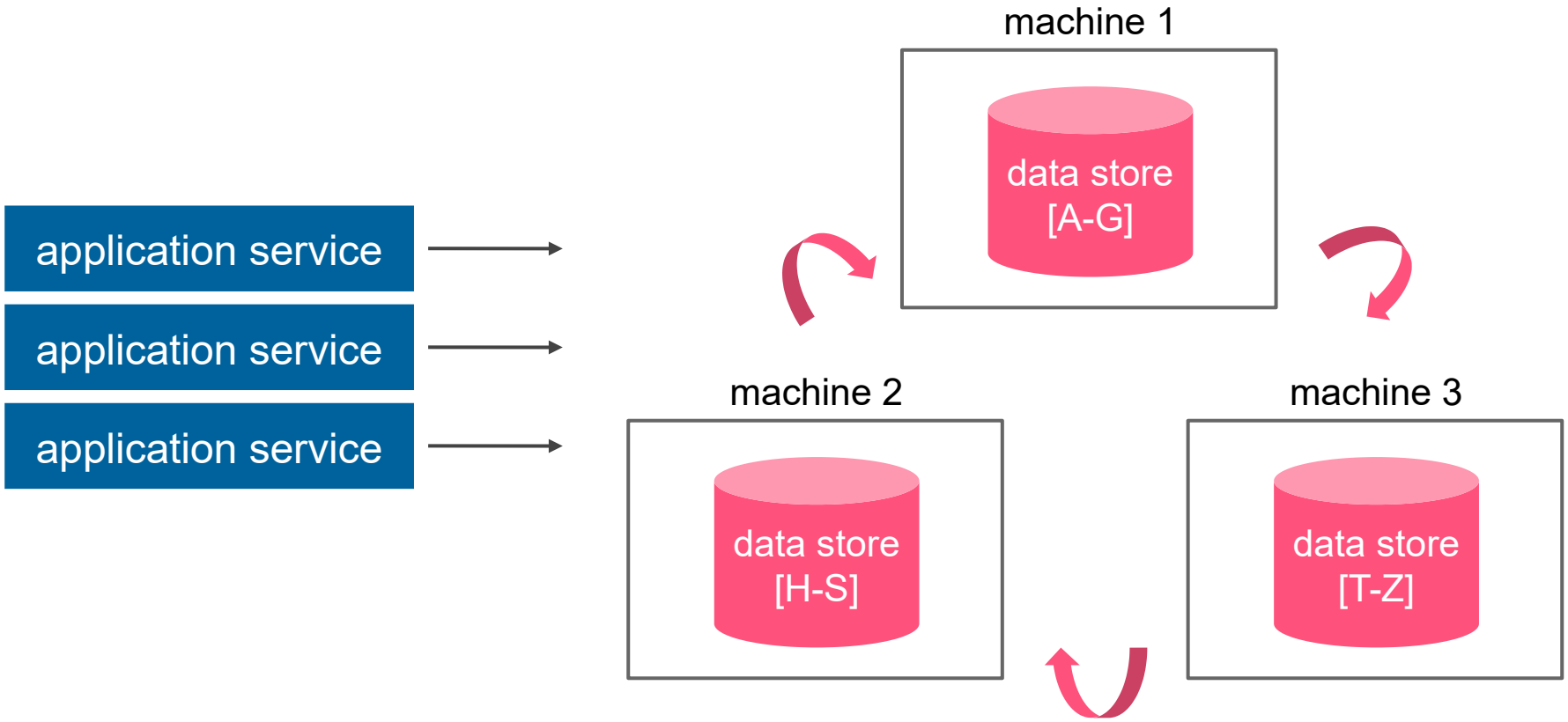
cassandra

hazelcast

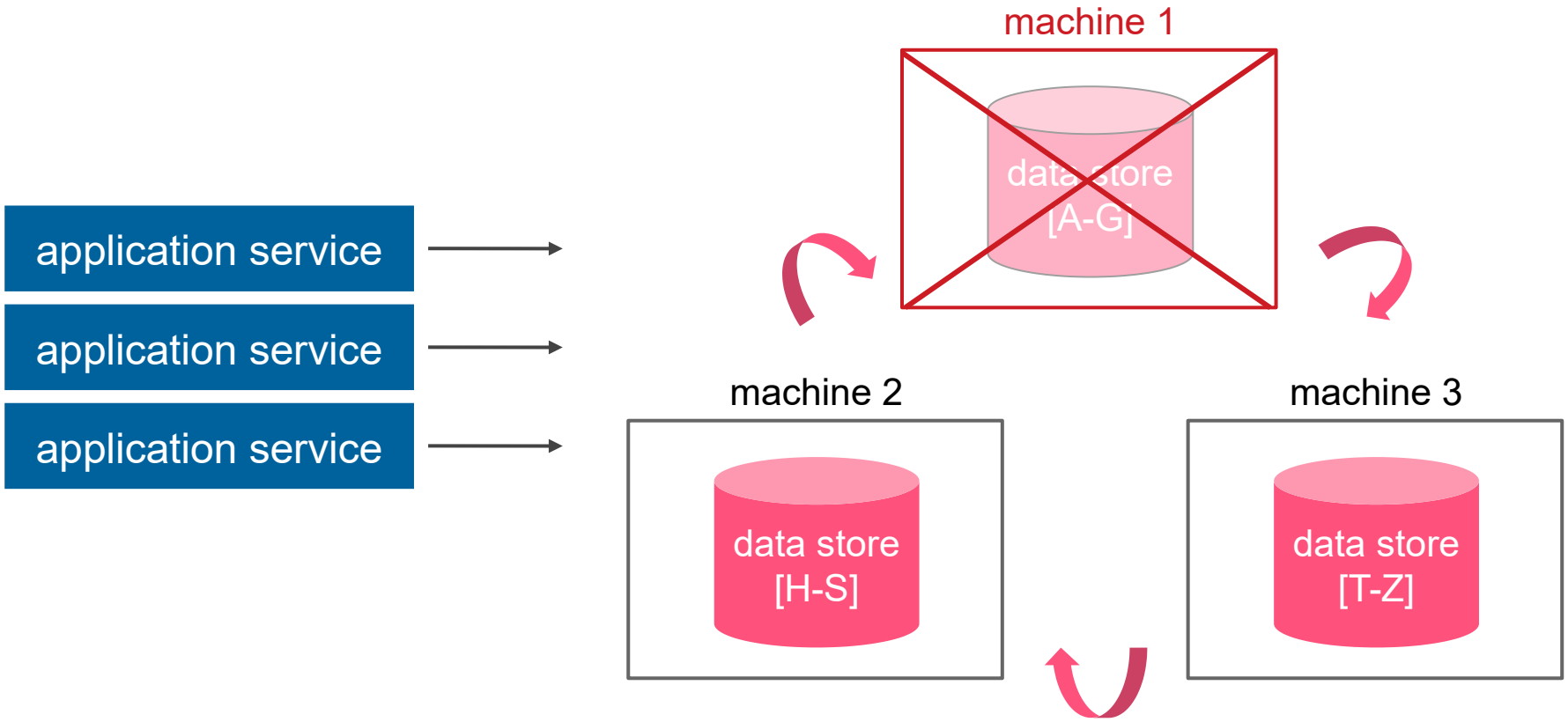
redis

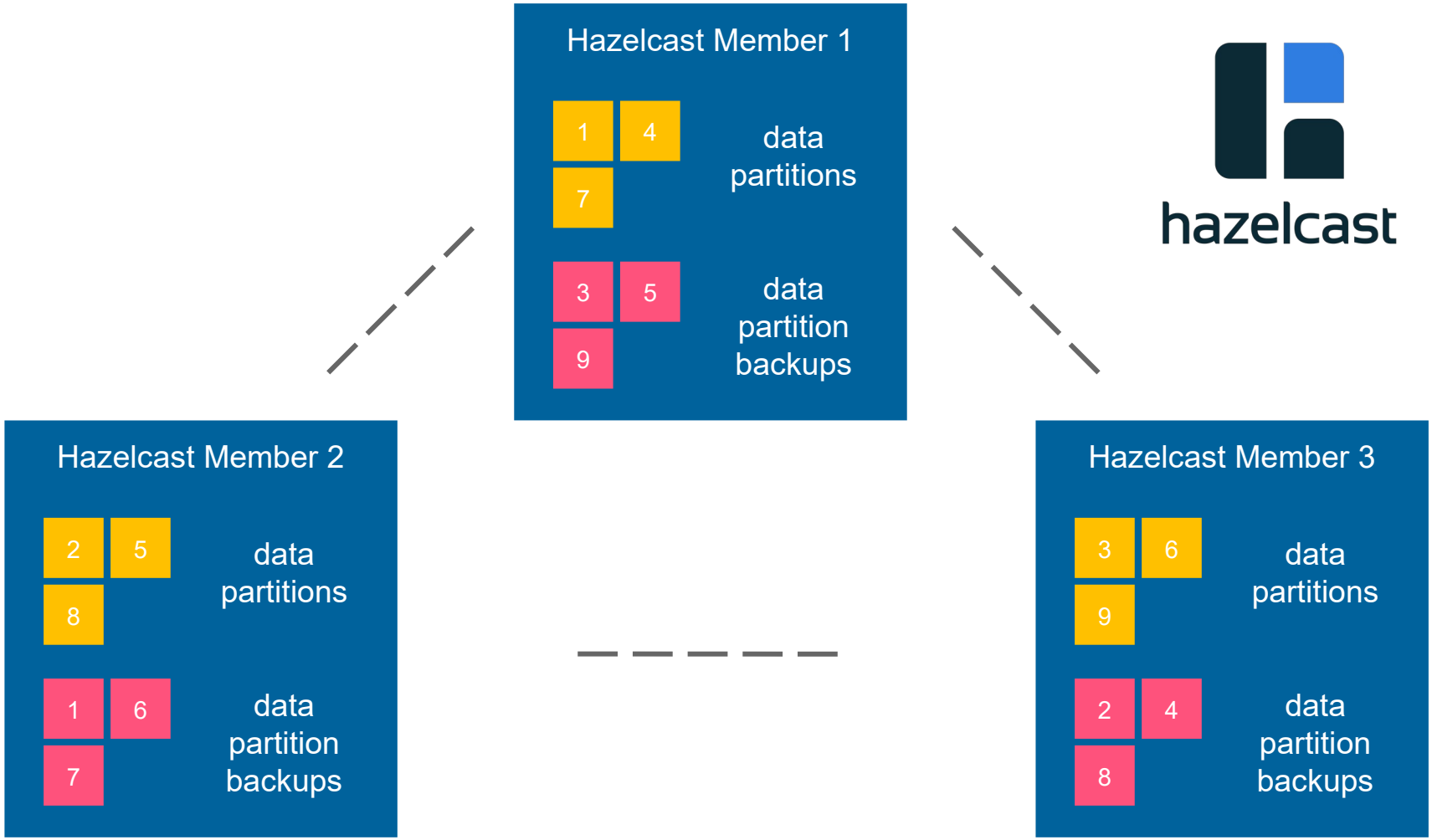
mongoDB.

Option 2: Clustering

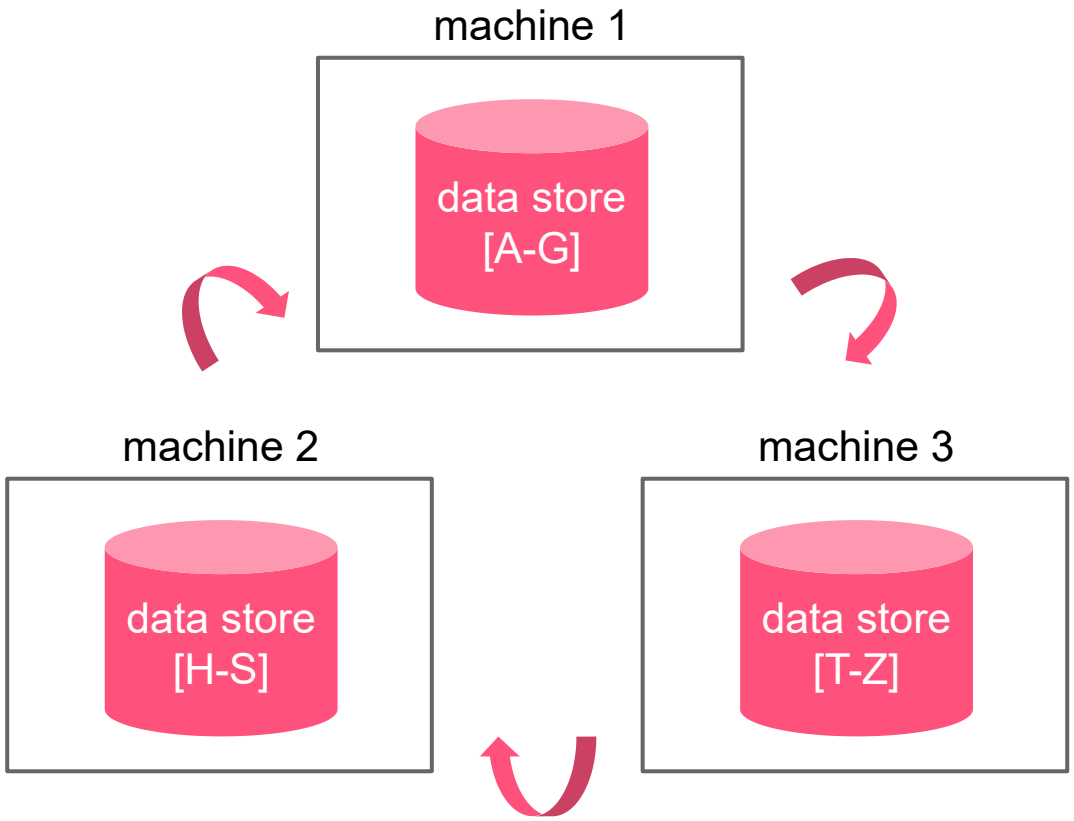


Option 2: Clustering




Hazelcast Member 1
1
4
7
data partitions
3
5
9
data partition backups
hazelcast
Hazelcast Member 2
2
5
8
data partitions
1
6
7
data partition backups
Hazelcast Member 3
3
6
9
data partitions
2
4
8
data partition backups

Option 2: Clustering



NoSQL

cassandra

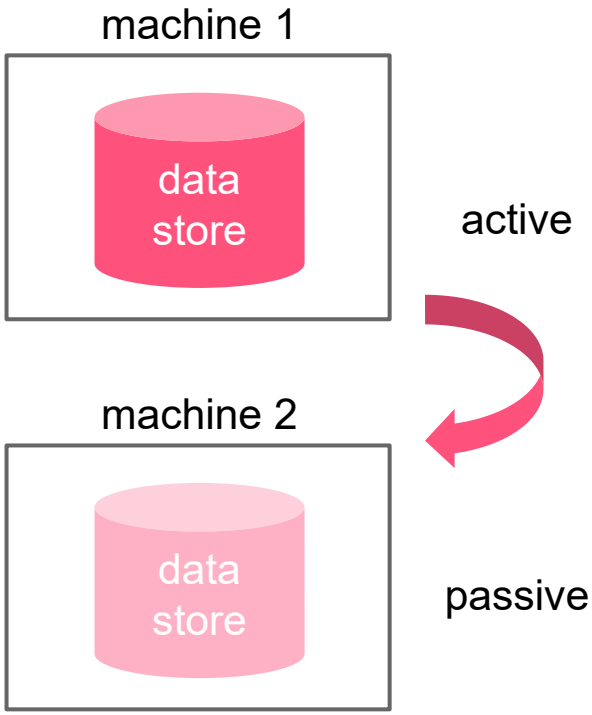
hazelcast

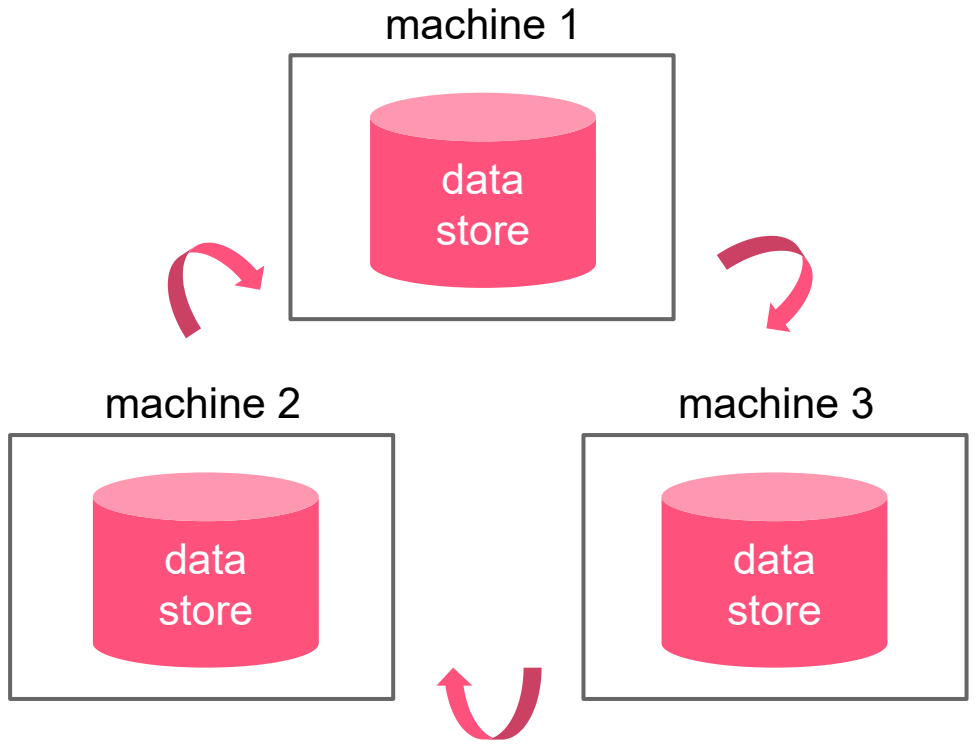
redis

mongoDB.

Synchronous vs Asynchronous

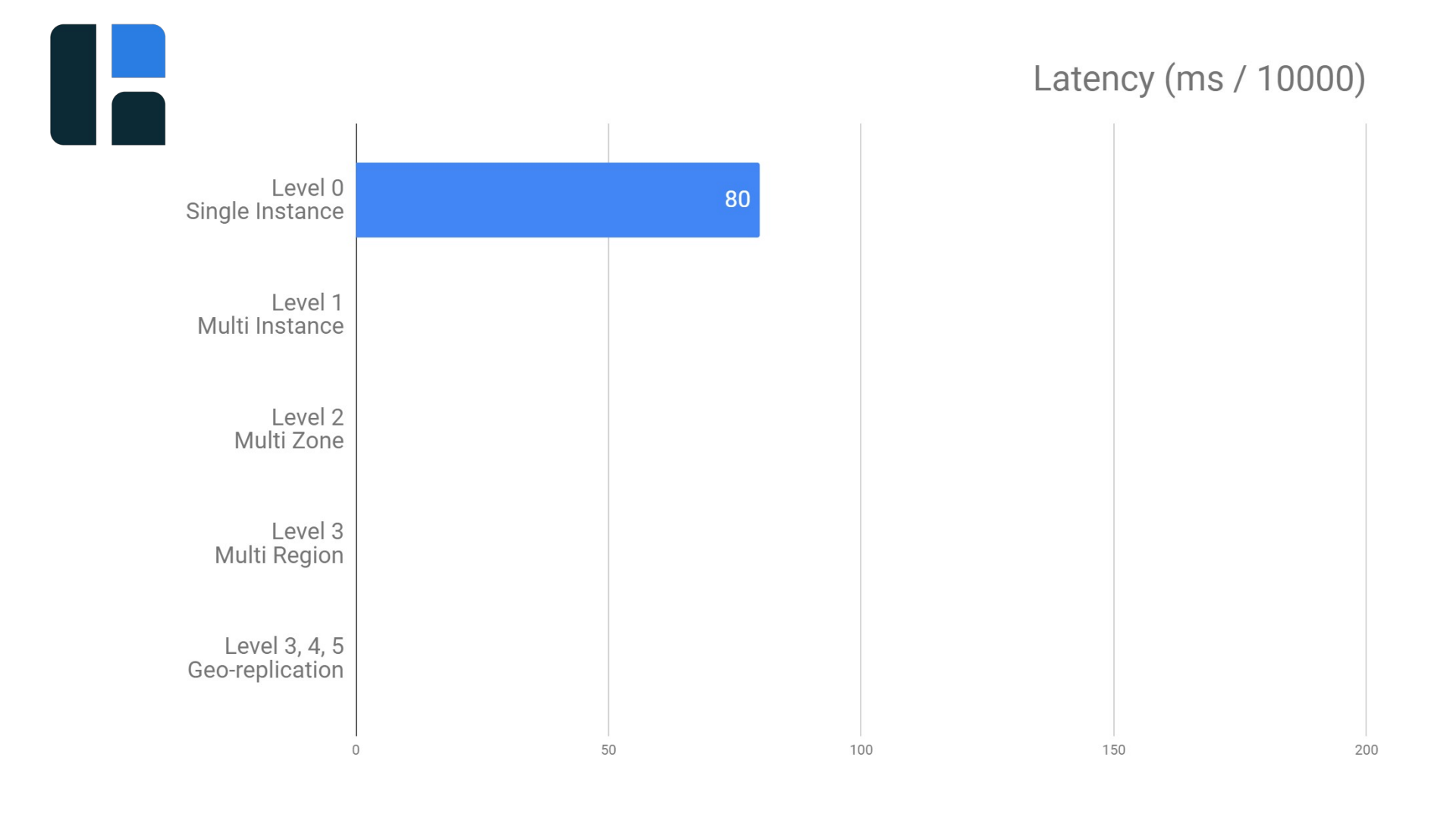
Synchronous (Consistency) or Asynchronous (Latency)?

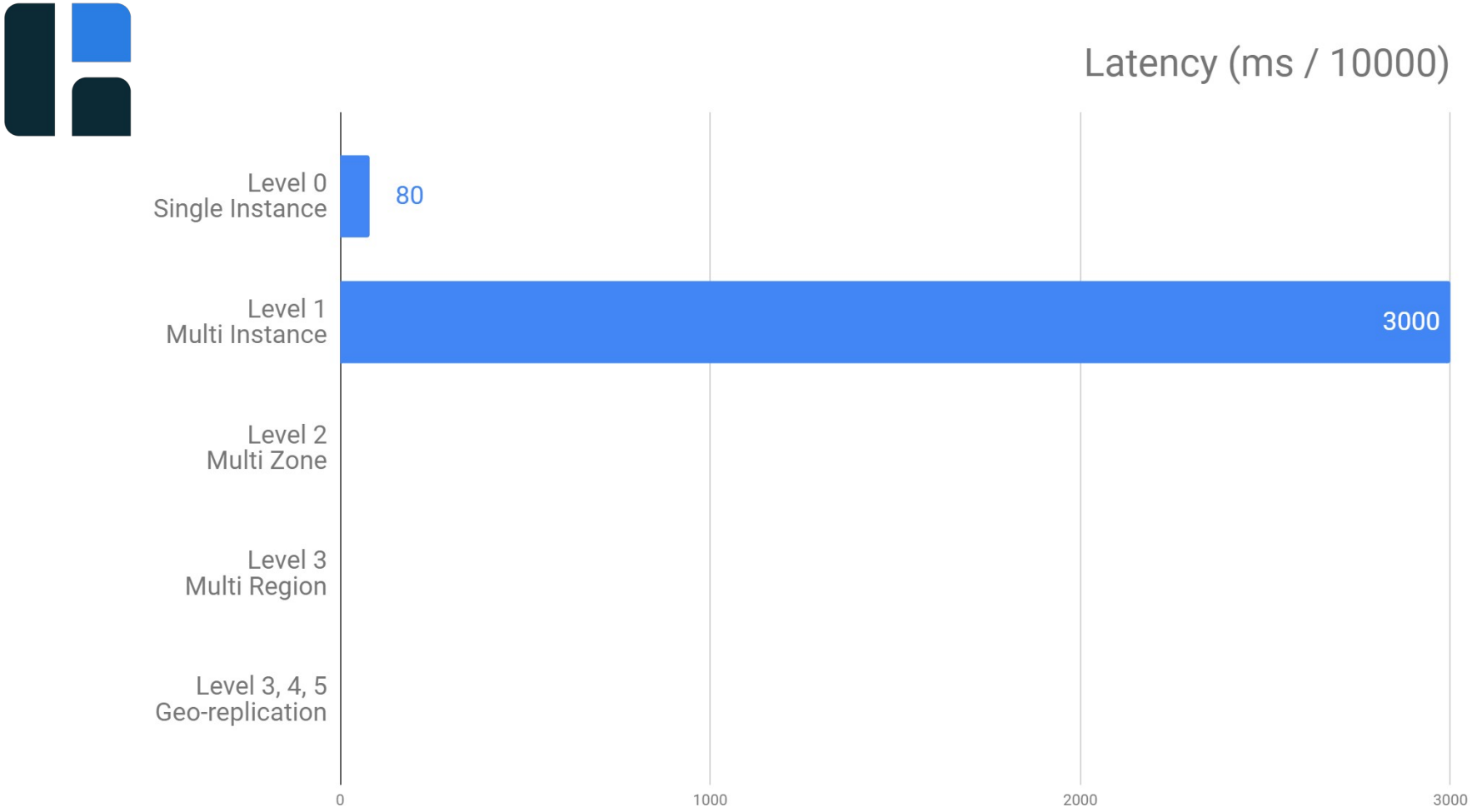




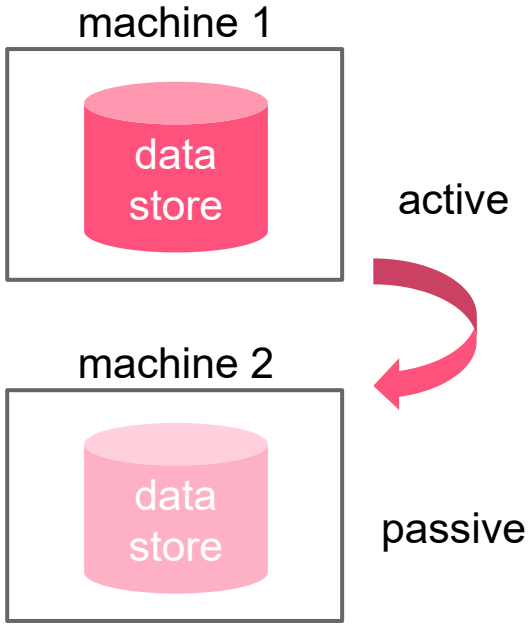
LATENCY EXPERIMENT

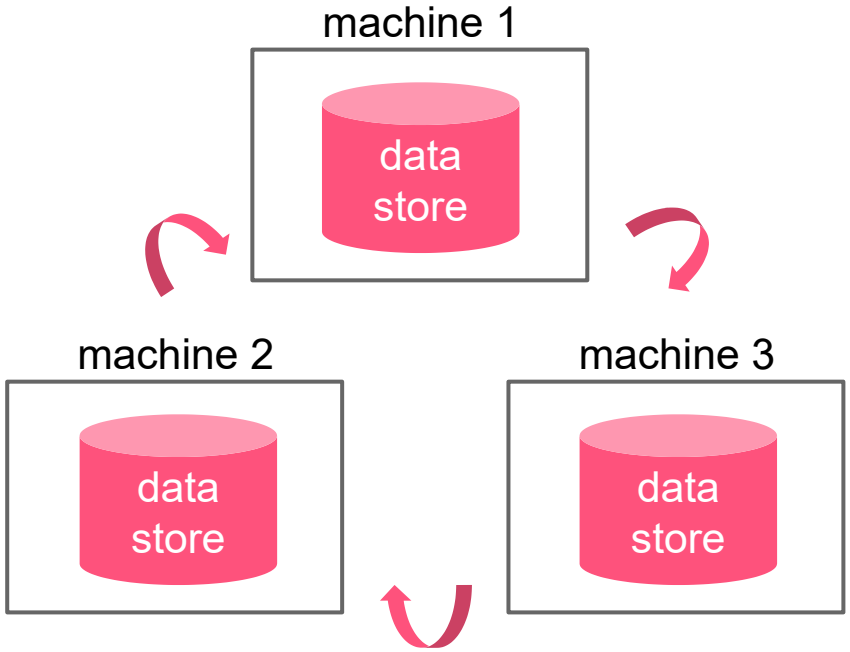
hazelcast®


Latency (ms / 10000)
Level 0
Single Instance
80
Level 1
Multi Instance
Level 2
Multi Zone
Level 3
Multi Region
Level 3, 4, 5
Geo-replication
0
50
100
150
200


Latency (ms / 10000)
Level 0
Single Instance
80
Level 1
Multi Instance
3000
Level 2
Multi Zone
Level 3
Multi Region
Level 3, 4, 5
Geo-replication
0
1000
2000
3000

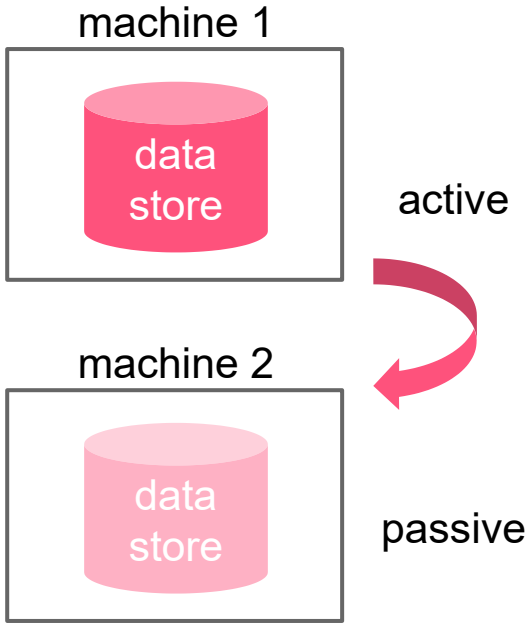
Synchronous (Consistency) or Asynchronous (Latency)?



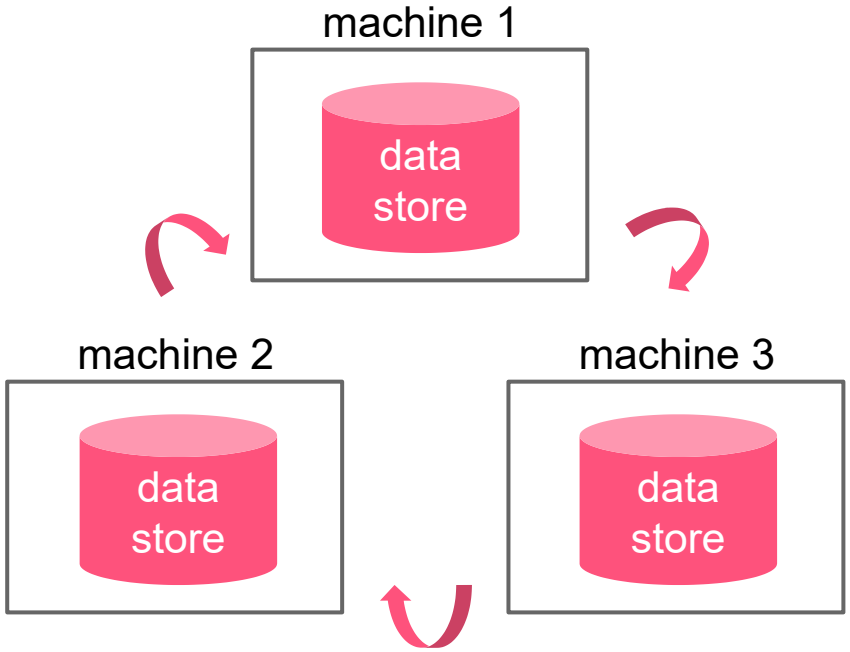


synchronous

Synchronous (Consistency) or Asynchronous (Latency)?



?



synchronous

What does "Level 1: Multi Instance" mean to You?

- Data consistency!
- Most tools supported
- Cloud-specific toolkit (e.g. AWS SQS)
- Simple setup (even on-premises)
- High latency if accessed multi regions

hazelcast®

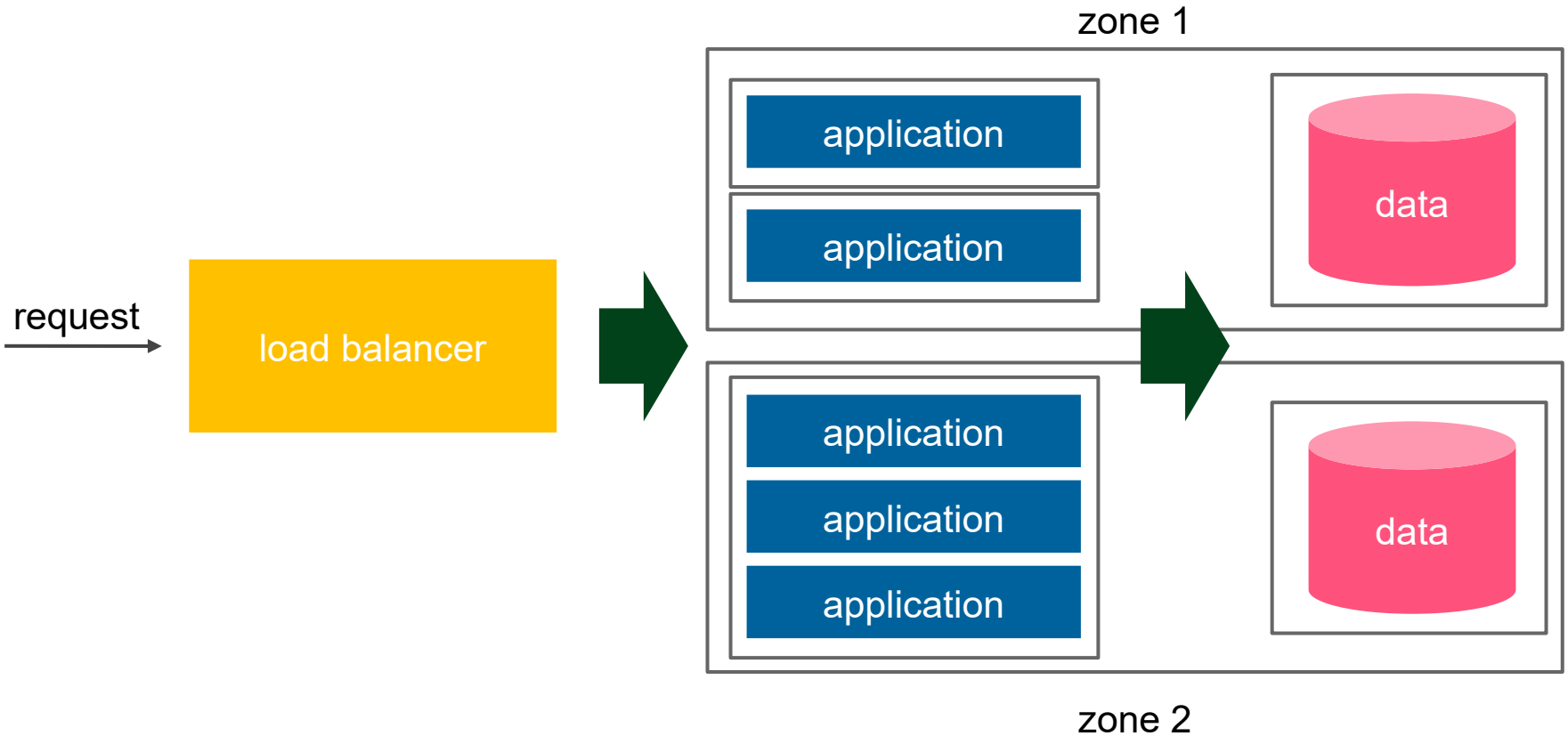
Agenda

- Introduction ✔
- High Availability Levels
  - Level 0: Single Instance ✔
  - Level 1: Multi Instance ✔
  - Level 2: Multi Zone
  - Level 3: Multi Region
  - Level 4: Multi Cloud
  - Level 5: Hybrid Cloud
- Summary

Level 2: Multi Zone

If one availability zone is down, the system is still available

Level 2: Multi Zone



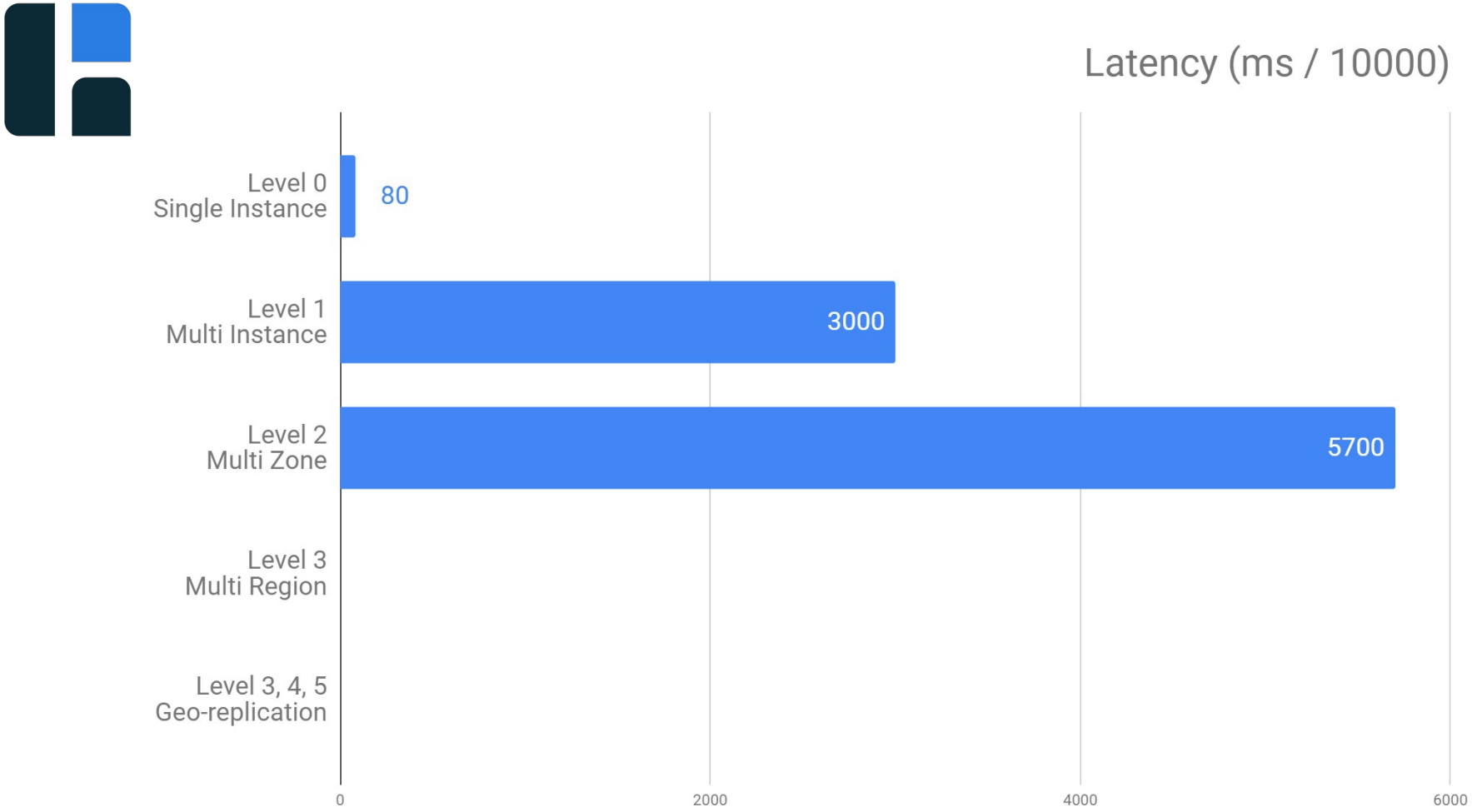
Is multi-zone deployment any different?

No

No... but Yes

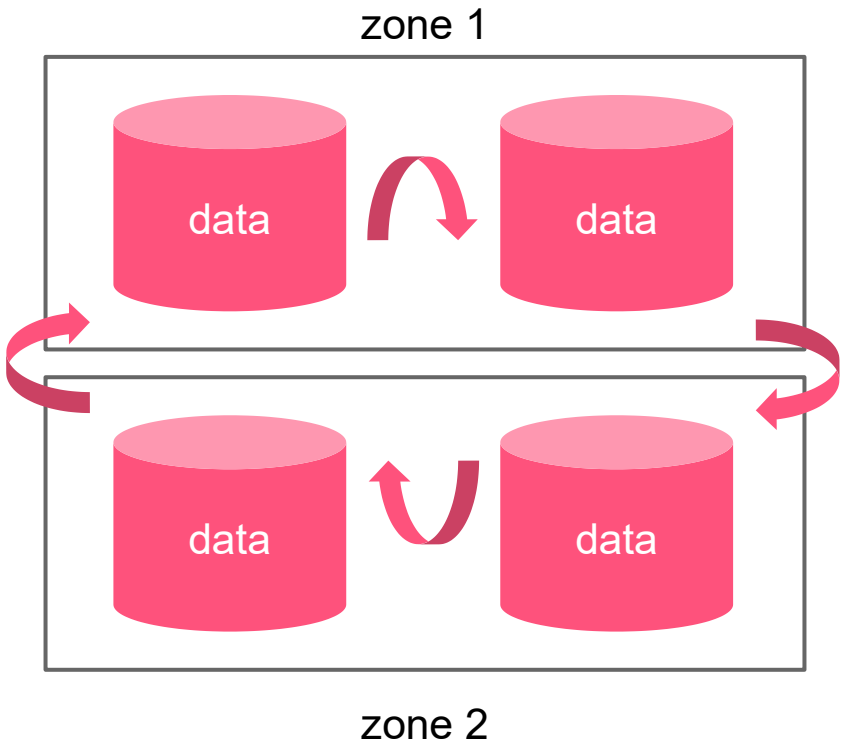
LATENCY EXPERIMENT

hazelcast®


Latency (ms / 10000)
Level 0
Single Instance
80
Level 1
Multi Instance
3000
Level 2
Multi Zone
5700
Level 3
Multi Region
Level 3, 4, 5
Geo-replication
0
2000
4000
6000

No... but Yes

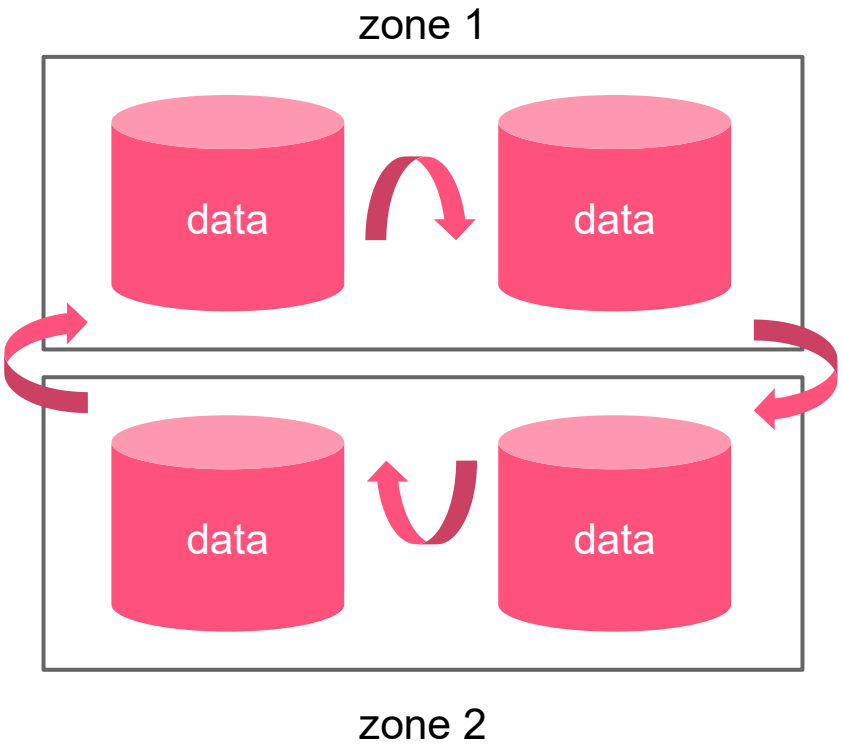
Level 2: Multi Zone



Level 2: Multi Zone

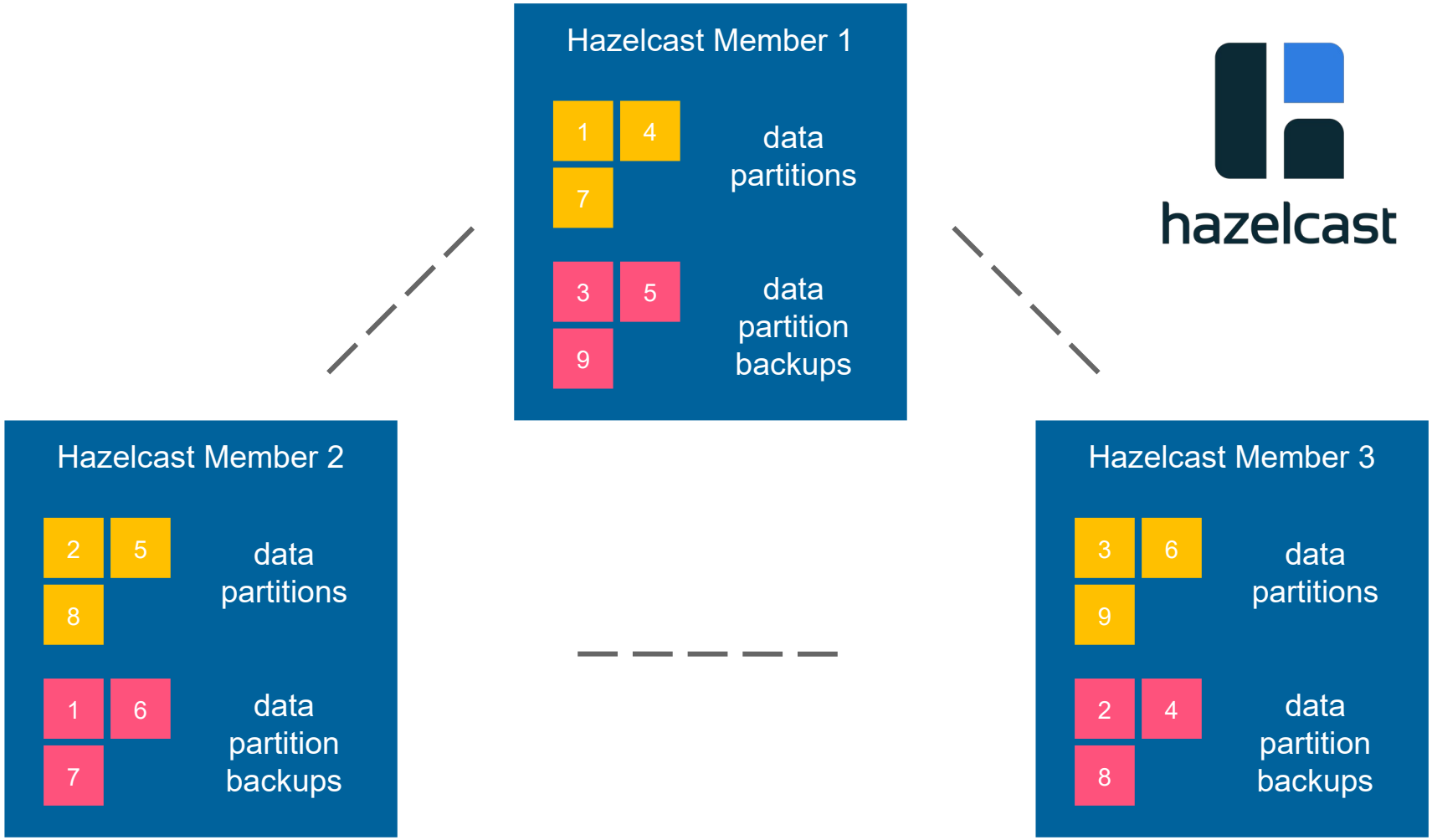
Assumptions:

- Machines in at least 2 AZ
- Fast and reliable network

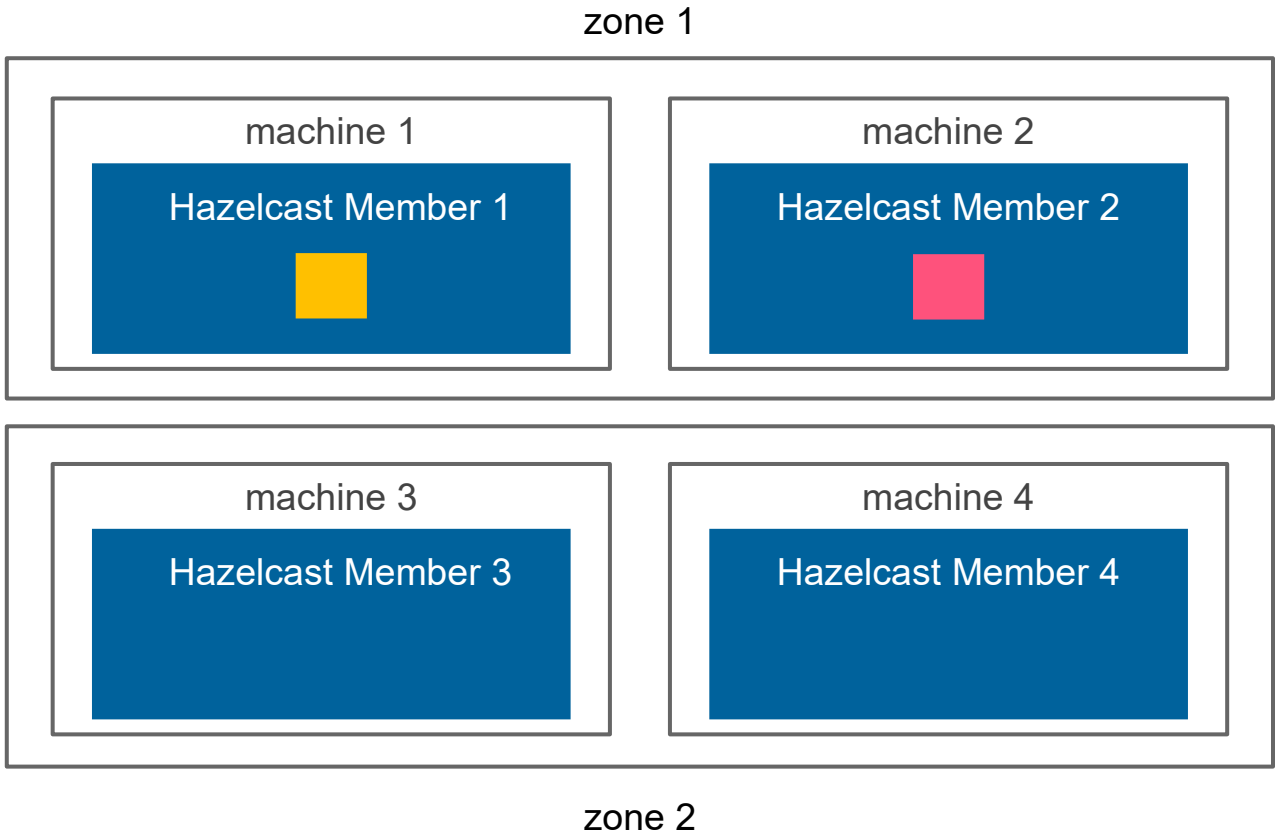


For example:

- EC2 Instances in 2 AWS Availability Zones
- Azure VM instances in 2 Availability Sets


Hazelcast Member 1
1
4
7
data partitions
3
5
9
data partition backups
hazelcast
Hazelcast Member 2
2
5
8
data partitions
1
6
7
data partition backups
Hazelcast Member 3
3
6
9
data partitions
2
4
8
data partition backups

Level 2: Multi Zone



hazelcast

Hazelcast Zone Aware Feature

Hazelcast configuration:

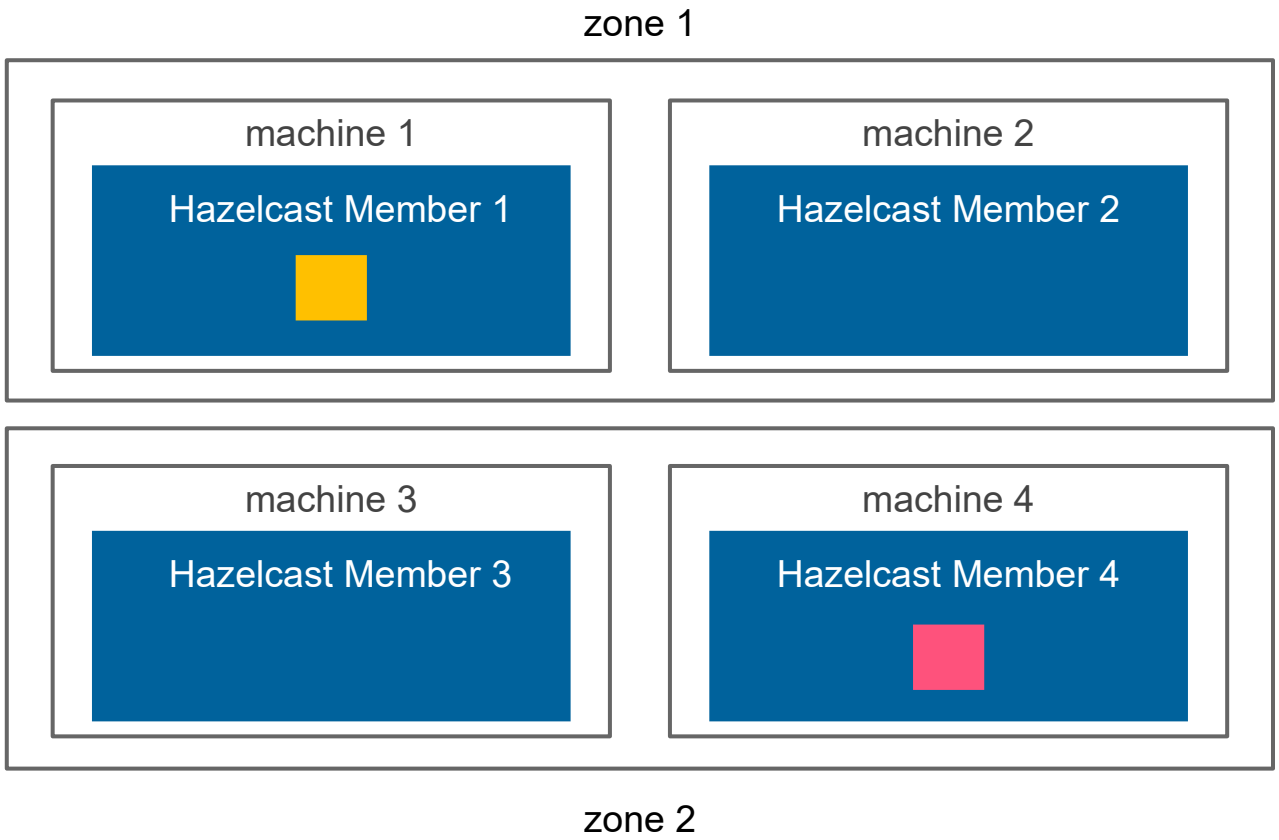
```
hazelcast:
    partition-group:
        enabled: true
        group-type: ZONE_AWARE
```

Hazelcast Zone Aware Feature

Google Cloud Platform

kubernetes

Level 2: Multi Zone



hazelcast

No... but Yes

Make sure your data store is ZONE AWARE

What does "Level 2: Multi Zone" mean to You?

- Currently top 1 choice!
- Data consistency!
- Cloud-specific toolkit (e.g. AWS SQS)
- High latency if accessed multi regions
- Not all tools are "zone aware"

hazelcast®

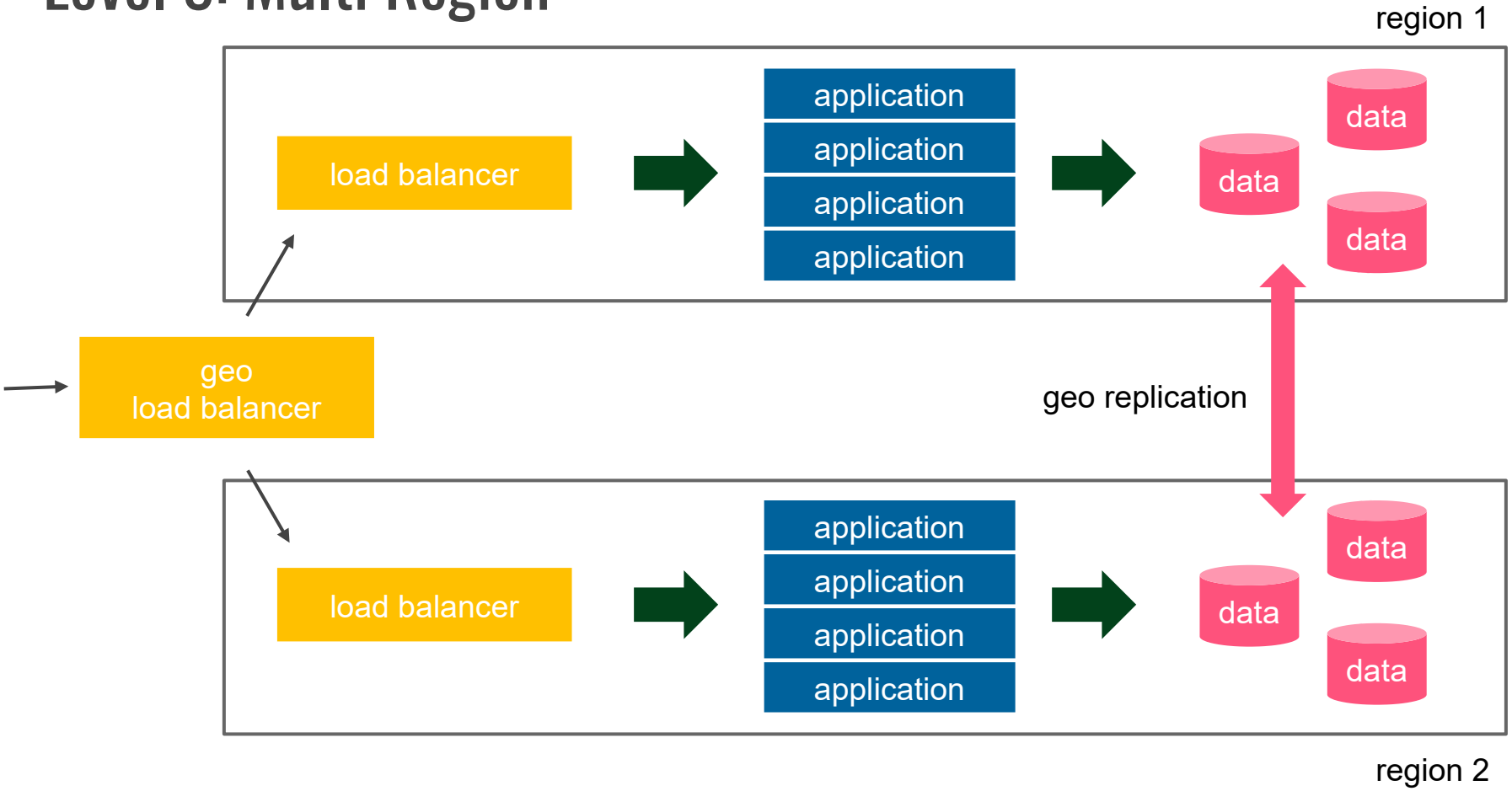
Agenda

- Introduction ✔
- High Availability Levels
  - Level 0: Single Instance ✔
  - Level 1: Multi Instance ✔
  - Level 2: Multi Zone ✔
  - Level 3: Multi Region
  - Level 4: Multi Cloud
  - Level 5: Hybrid Cloud
- Summary

Level 3: Multi Region

If one region is down,
the system is still available

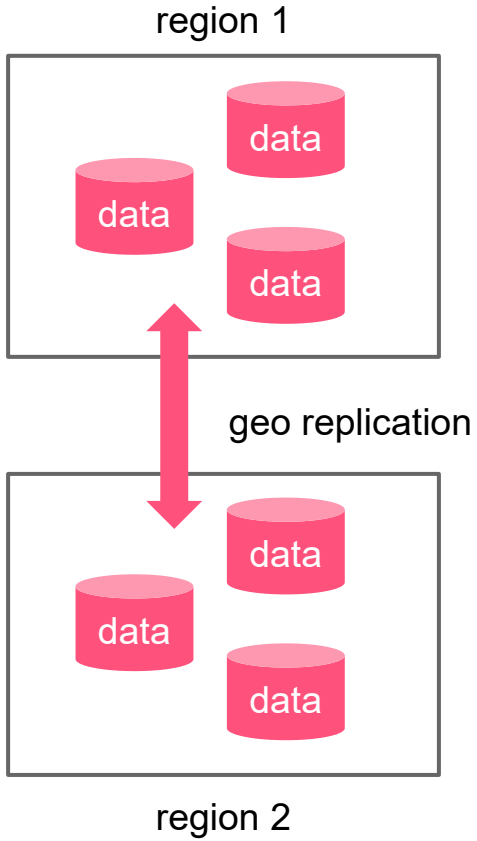
Level 3: Multi Region



Level 3: Multi Region

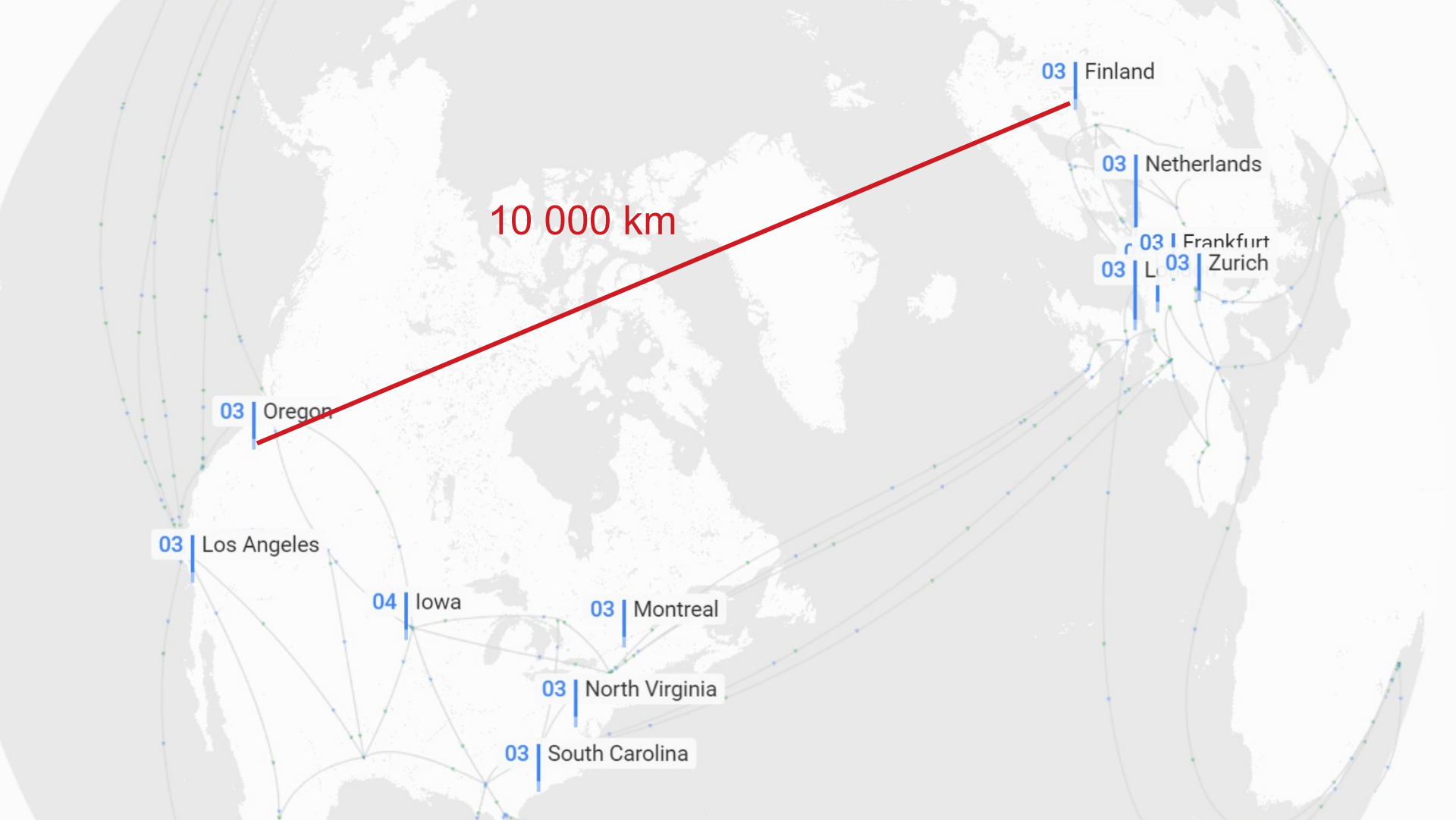
Assumptions:

- Machines in at least 2 geographical regions
- Network may be slow and unreliable



For example:

- EC2 Instances in regions: eu-central-1 and us-west-2


03 Finland
03 Netherlands
10 000 km
03 Frankfurt
03 L
03 Zurich
03 Oregon
03 Los Angeles
04 Iowa
03 Montreal
03 North Virginia
03 South Carolina

Level 3: Multi Region

Speed of light: 300 000 km/s

Distance: 10 000 km

RTT (Round Trip Time) = 60 ms

Geo-replication

Level 3: Multi Region (Geo-replication)

hazelcast®

Geo-replication

- It's **asynchronous**
- Your data store **must support it**
- You must be prepared for **data loss**
- Two modes:
  - Active-Passive
  - Active-Active

hazelcast®

Active-Passive Geo-replication

active

passive

hazelcast®

Active-Passive Geo-replication

- **data loss** possible
- (eventual) **consistency**

active

passive

hazelcast®

Active-Active Geo-replication

active

active

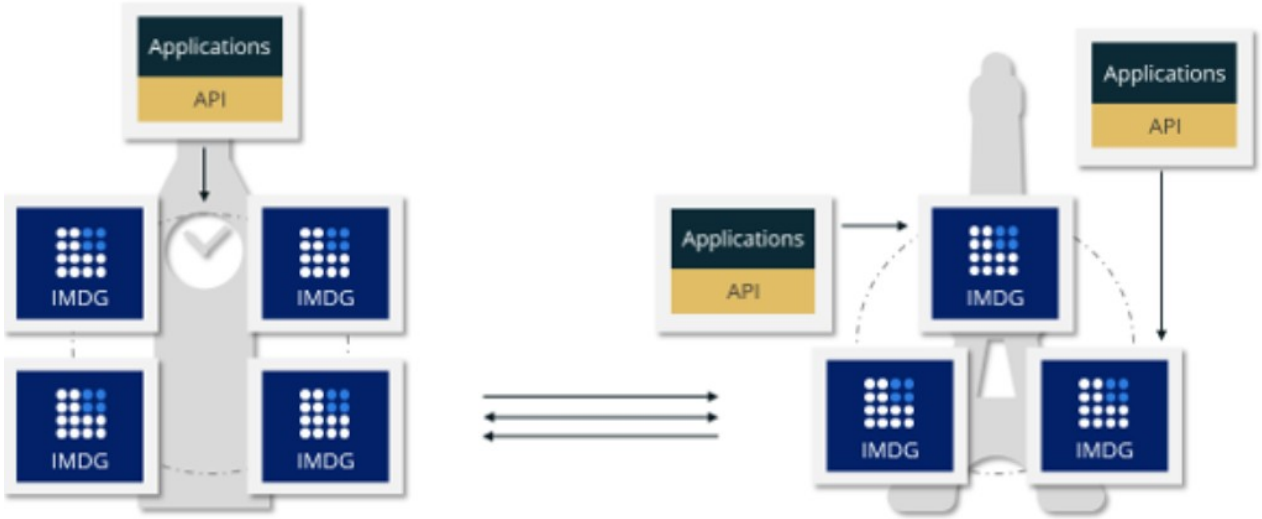
hazelcast®

Active-Active Geo-replication

- **data loss** possible
- **eventual consistency**
- **conflict resolution**

hazelcast®

Hazelcast WAN Replication



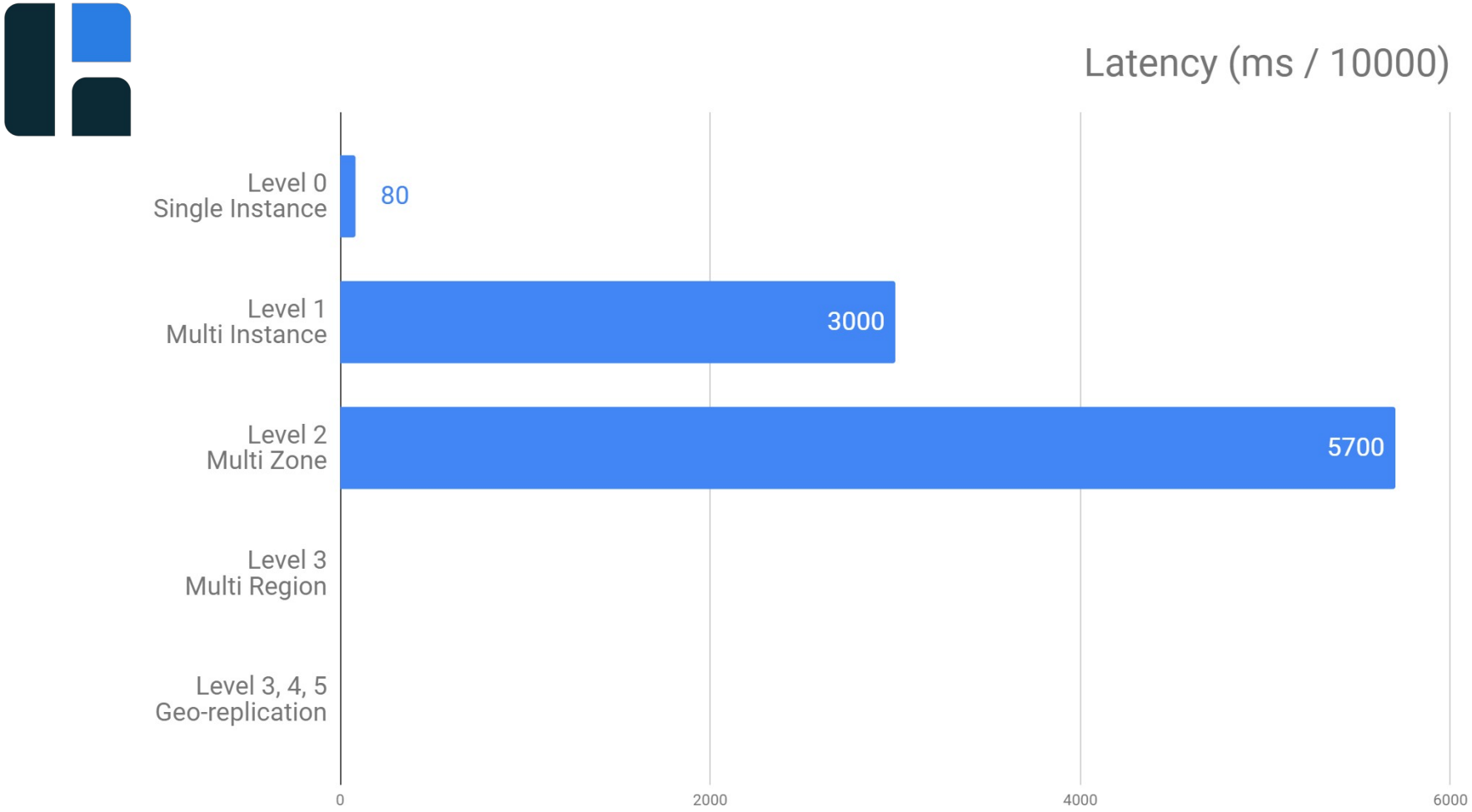

```
hazelcast:
  wan-replication:
    batch-publisher:
      target-endpoints: 35.184.122.109
```

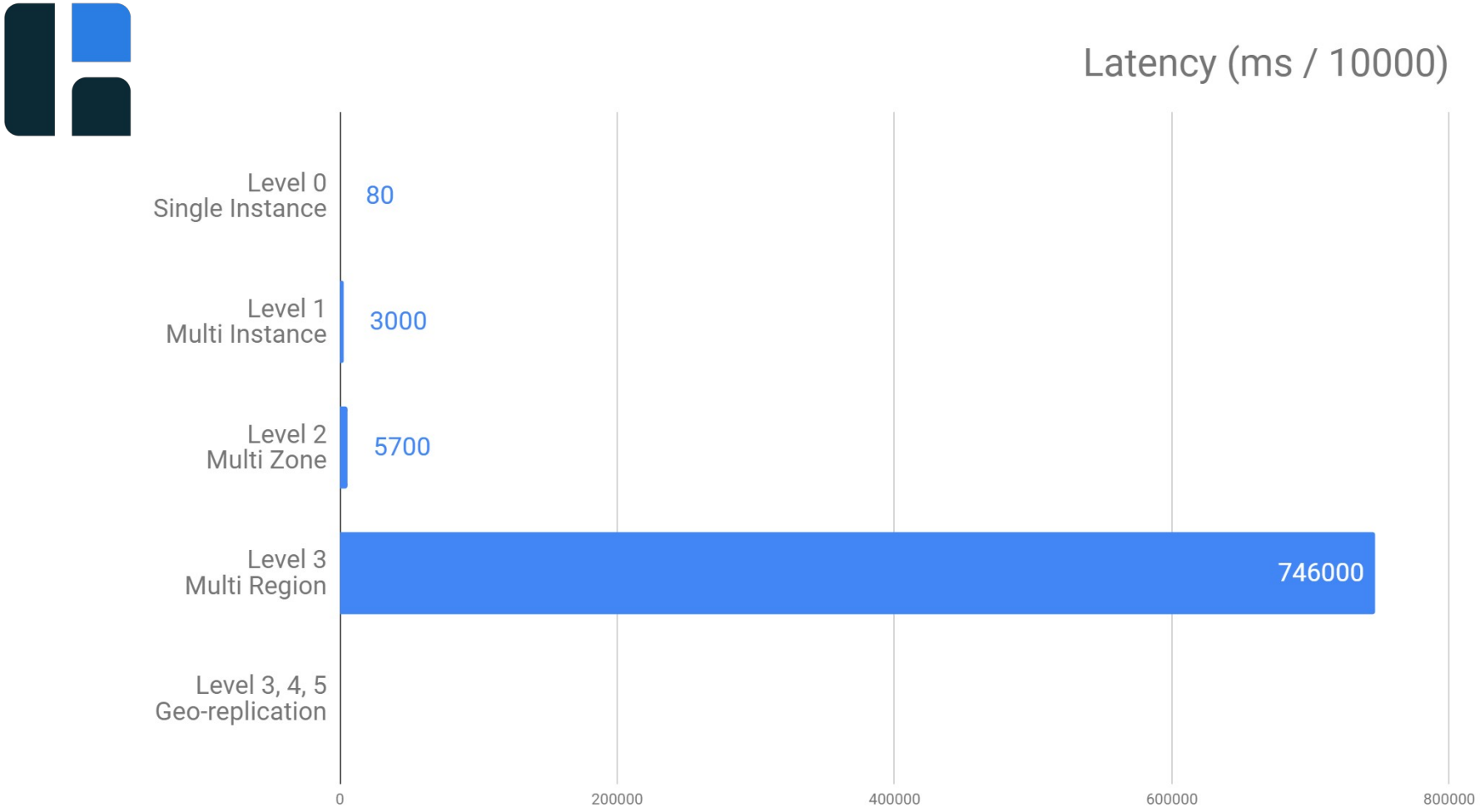
hazelcast®

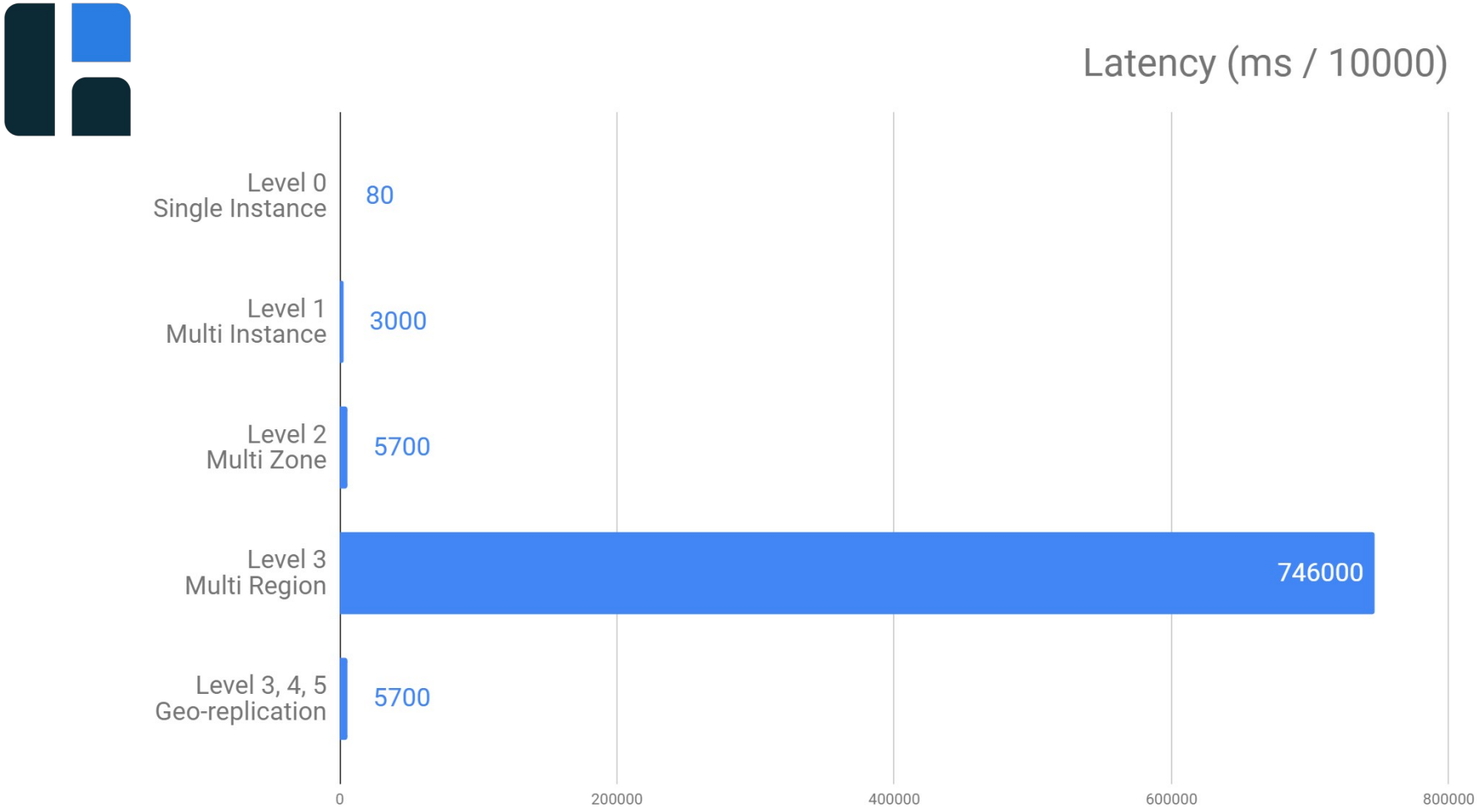
Do I really need to lose consistency?

LATENCY EXPERIMENT

hazelcast®


Latency (ms / 10000)
Level 0
Single Instance
80
Level 1
Multi Instance
3000
Level 2
Multi Zone
5700
Level 3
Multi Region
Level 3, 4, 5
Geo-replication
0
2000
4000
6000


Latency (ms / 10000)
Level 0
Single Instance
80
Level 1
Multi Instance
3000
Level 2
Multi Zone
5700
Level 3
Multi Region
746000
Level 3, 4, 5
Geo-replication
0
200000
400000
600000
800000


Latency (ms / 10000)
Level 0
Single Instance
80
Level 1
Multi Instance
3000
Level 2
Multi Zone
5700
Level 3
Multi Region
746000
Level 3, 4, 5
Geo-replication
5700
0
200000
400000
600000
800000

Strong Consistency in Multi Region

- **NewSQL** (Spanner, CockroachDB)
- Multi-region distributed transactions
- **Consensus** algorithms (Paxos, Raft)
- Always a **trade-off**: consistency vs latency

What does "Level 3: Multi Region" mean to You?

- Super high available!
- Low latency if accessed from multi region
- Sometimes possible to use Cloud-specific toolkit (e.g. Google Spanner - yes, AWS Elasticache - no)
- Geo-replication (asynchronous)!
- Eventual consistency (conflict resolution)

hazelcast®

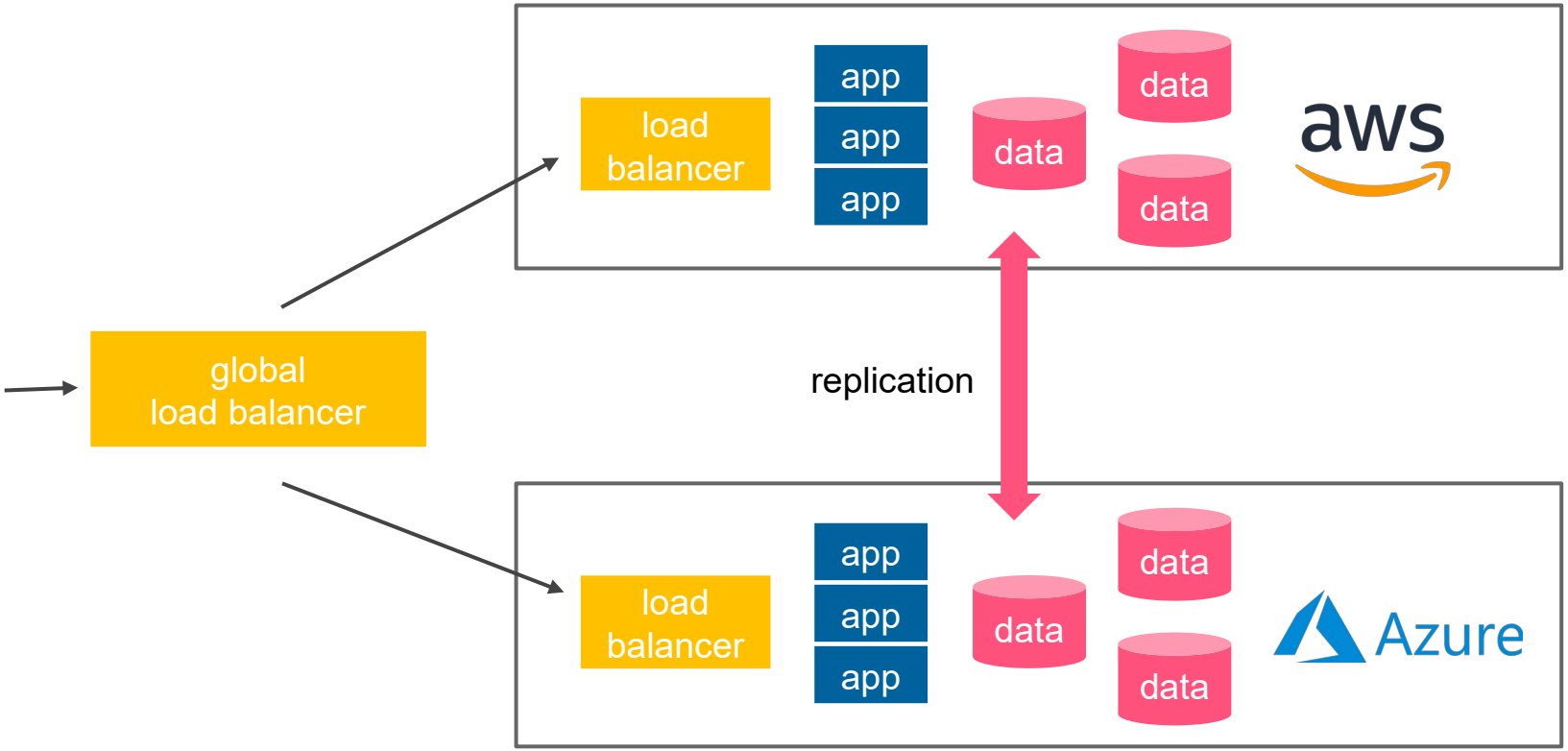
Agenda

- Introduction ✔
- High Availability Levels
  - Level 0: Single Instance ✔
  - Level 1: Multi Instance ✔
  - Level 2: Multi Zone ✔
  - Level 3: Multi Region ✔
  - Level 4: Multi Cloud
  - Level 5: Hybrid Cloud
- Summary

Level 4: Multi Cloud

If one cloud provider is down, the system is still available

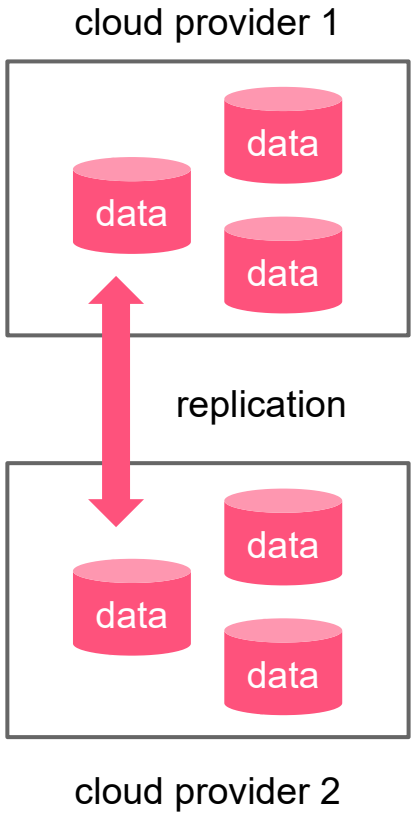
Level 4: Multi Cloud



Level 4: Multi Cloud

Assumptions:

- Machines in at least 2 cloud providers
- Network may be slow and unreliable
- Machines may be in different geo regions



For example:

- EC2 Instances in eu-central-1 and GCP VM Instances in us-west1-a

What's different from multi-region?

Level 4: Multi Cloud

- **No Cloud-specific** tools
- **No VPC Peering** across Cloud providers
  - Latency
  - Security
- **Cost**

AWS SQS

Cloud Spanner

Is High Availability the only reason for Multi-Cloud?

Reasons for Multi-Cloud

- High Availability / Disaster Recovery
- Avoiding vendor lock-in
- Cloud cost optimization
- Risk Mitigation
- Low latency
- Data Protection / Regulations / Compliance
- Best-Fit Technology (Cloud-specific portfolios)

What does "Level 4: Multi Cloud" mean to You?

- No vendor lock-in!
- Cloud cost negotiations
- Low latency if accessed from multi-cloud
- Complex setup!
- No Cloud toolkit (e.g. AWS SQS)

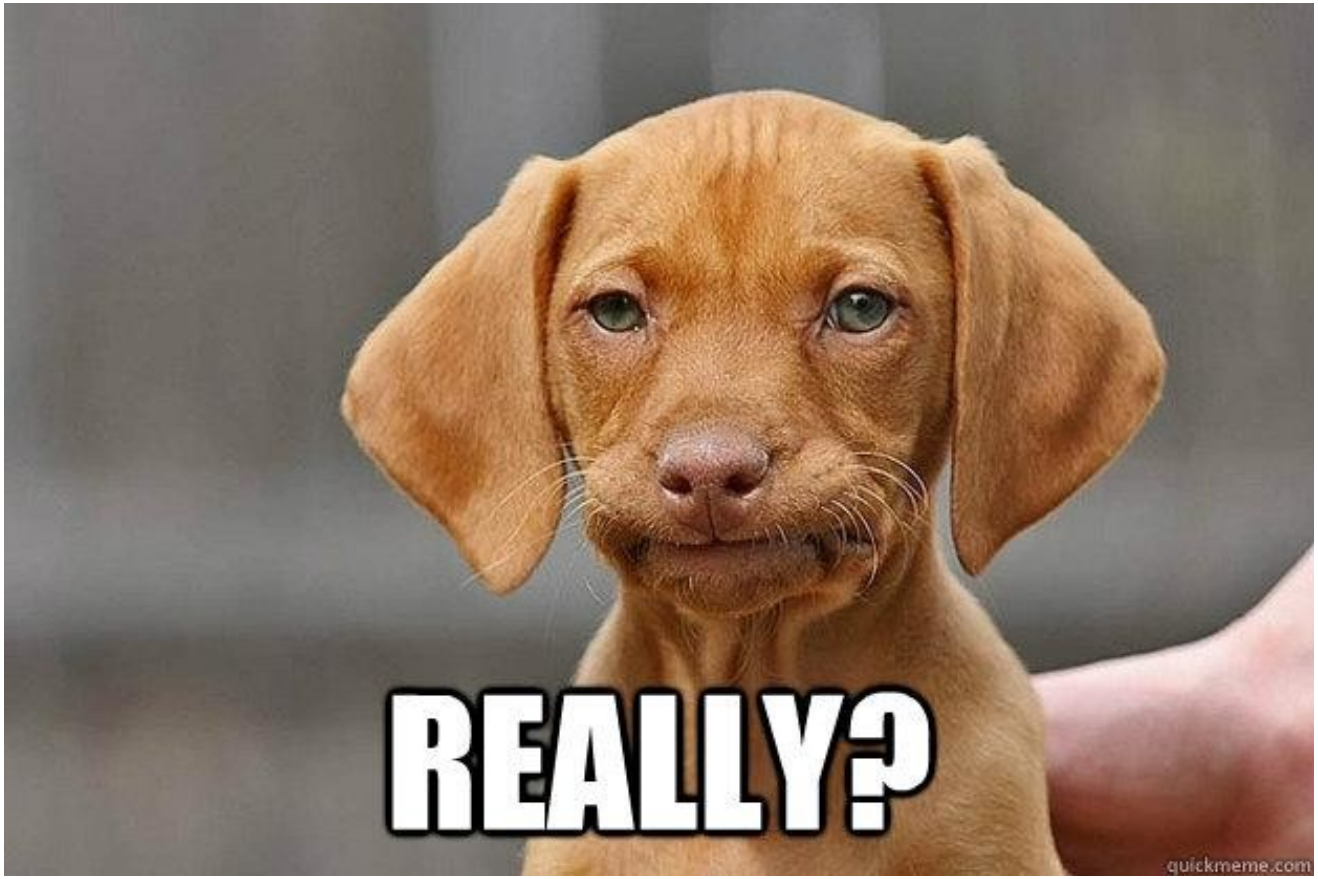
hazelcast®

Agenda

- Introduction ✔
- High Availability Levels
  - Level 0: Single Instance ✔
  - Level 1: Multi Instance ✔
  - Level 2: Multi Zone ✔
  - Level 3: Multi Region ✔
  - Level 4: Multi Cloud ✔
  - Level 5: Hybrid Cloud
- Summary

Level 5: Hybrid Cloud

If all cloud providers are down, the system is still available


REALLY?
quickmeme.com

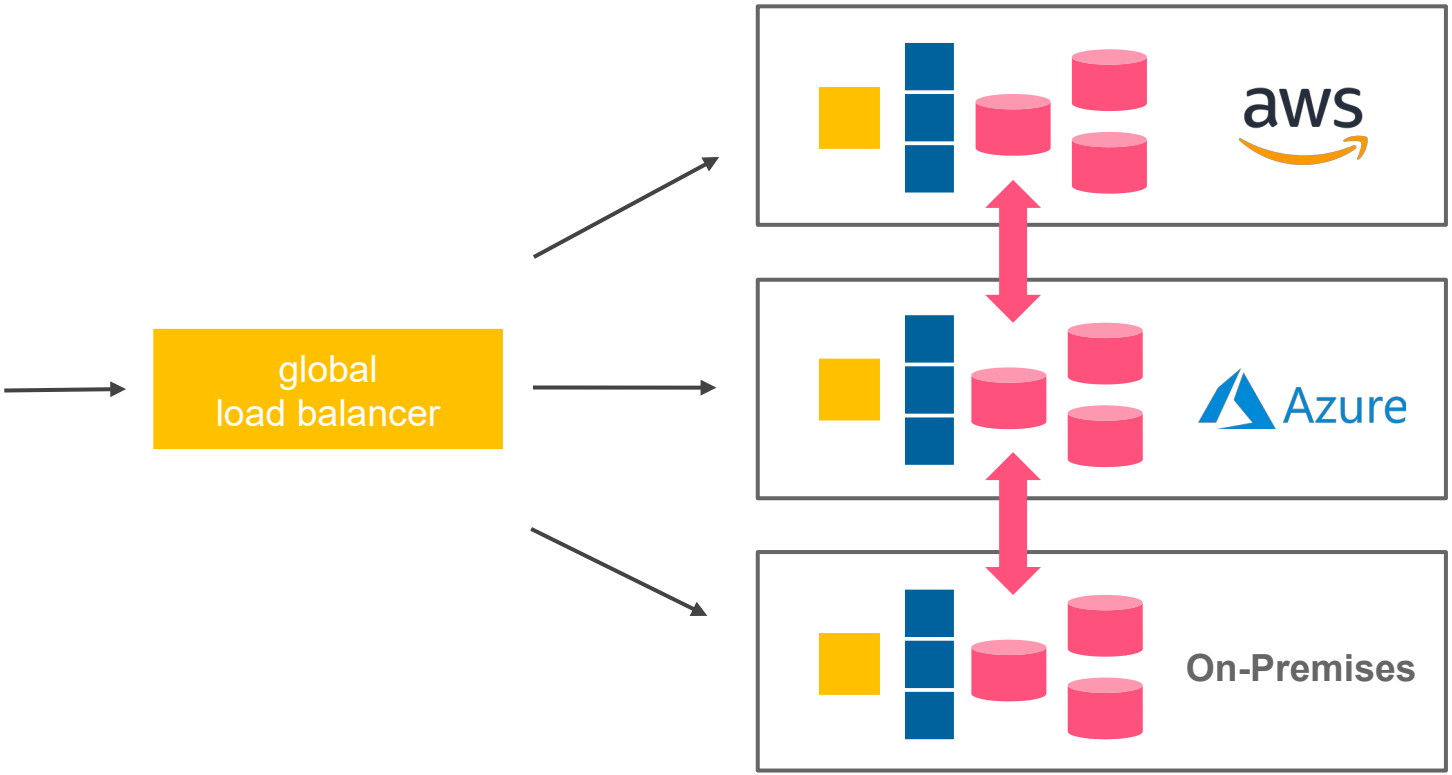
Is it possible that all cloud providers are down?

No!

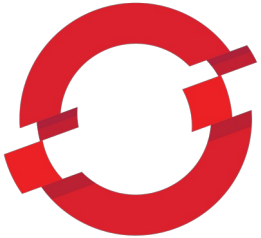
Reasons for Hybrid Cloud

- Data requirements / regulations
- Data security
- Moving to Cloud
- Cost reduction
- All mentioned already in Multi-Cloud

Level 5: Hybrid Cloud



kubernetes



OPENSHIFT

IBM Cloud

What does "Level 5: Hybrid Cloud" mean to You?

- No Cloud lock-in!
- Low latency if accessed from custom networks
- Super complex setup!
- Usually extra layer needed (e.g. Kubernetes, OpenShift)
- Costs a fortune!

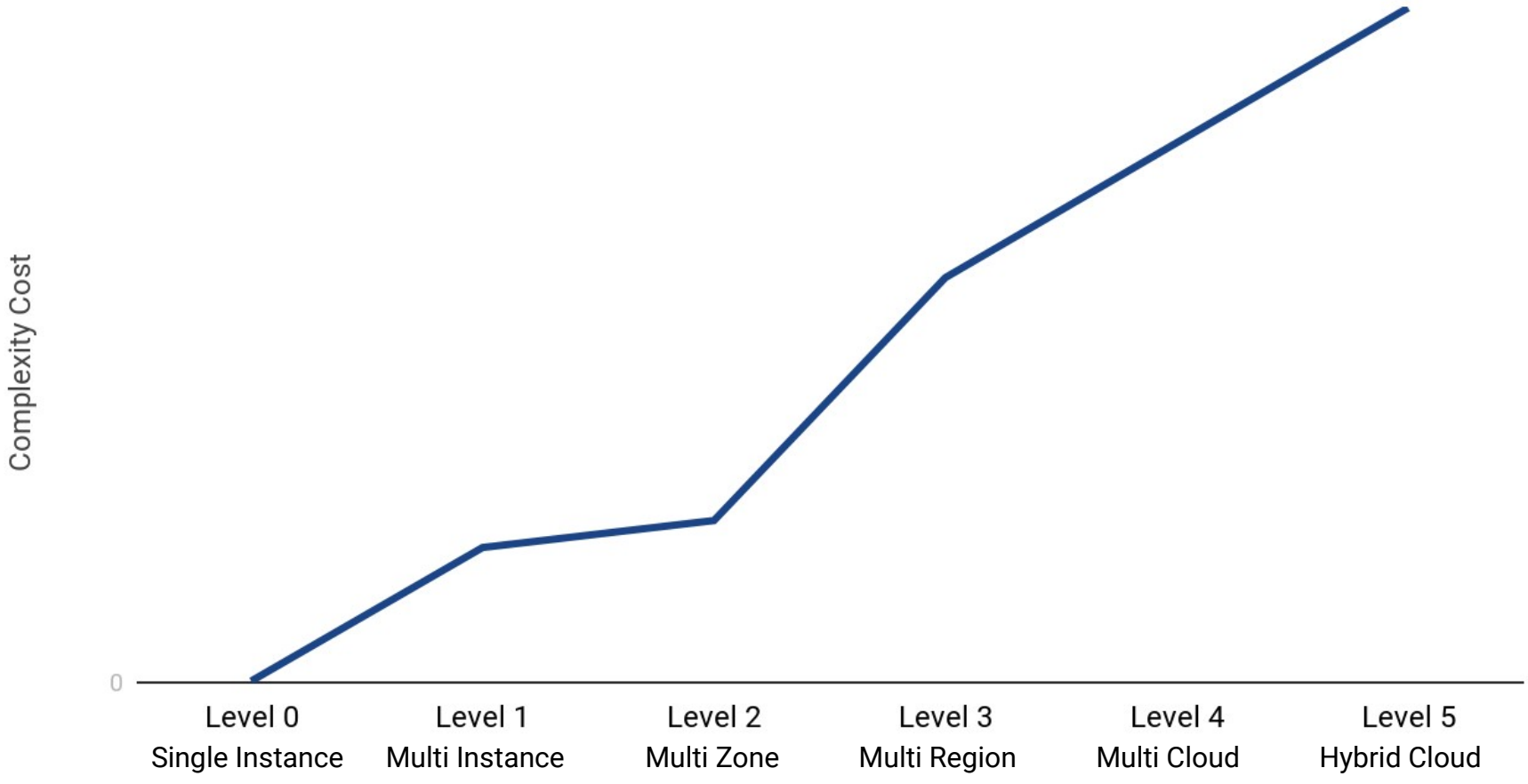
hazelcast®

Agenda

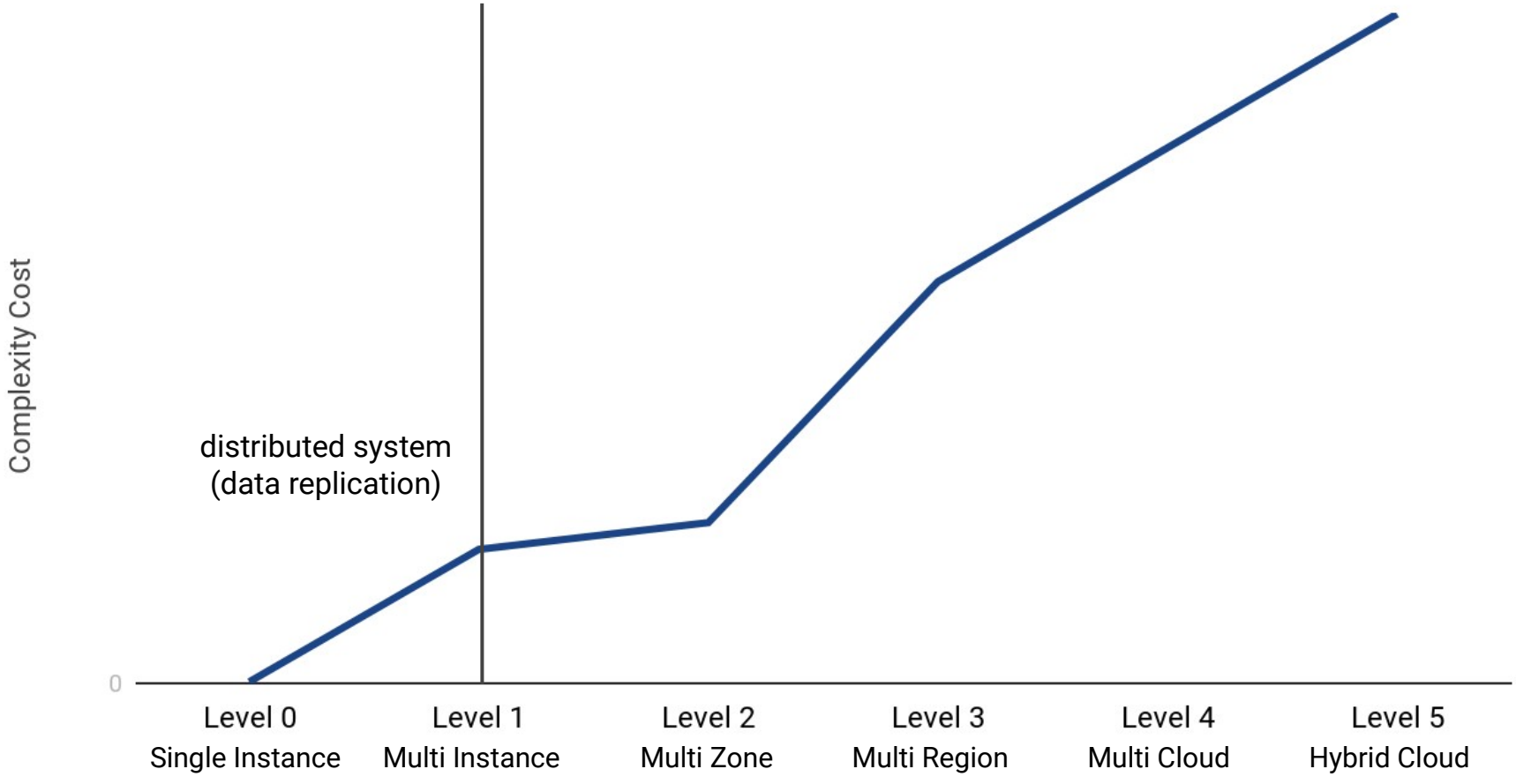
- Introduction ✔
- High Availability Levels
  - Level 0: Single Instance ✔
  - Level 1: Multi Instance ✔
  - Level 2: Multi Zone ✔
  - Level 3: Multi Region ✔
  - Level 4: Multi Cloud ✔
  - Level 5: Hybrid Cloud ✔
- Summary

Summary

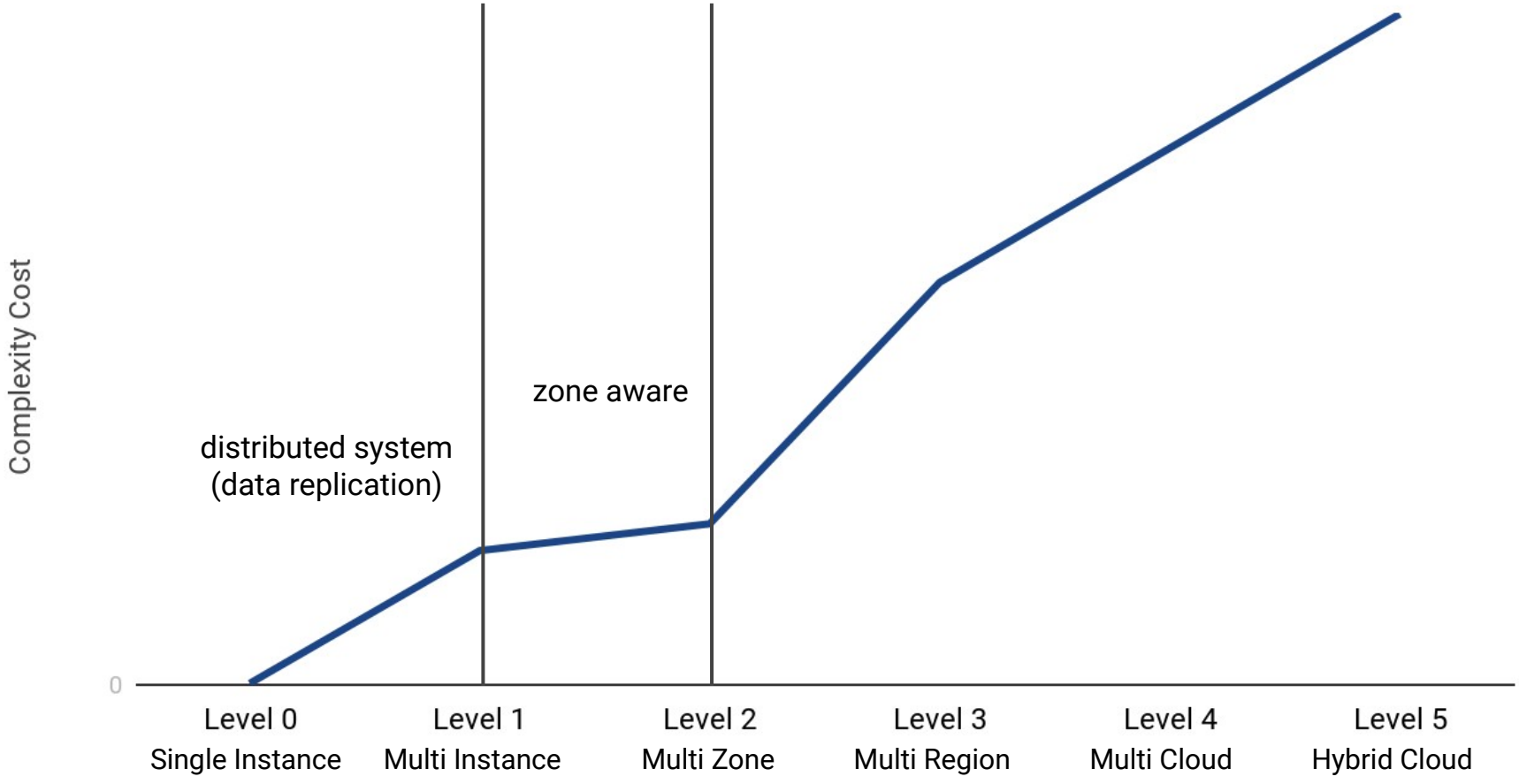
Complexity Cost / High Availability Level



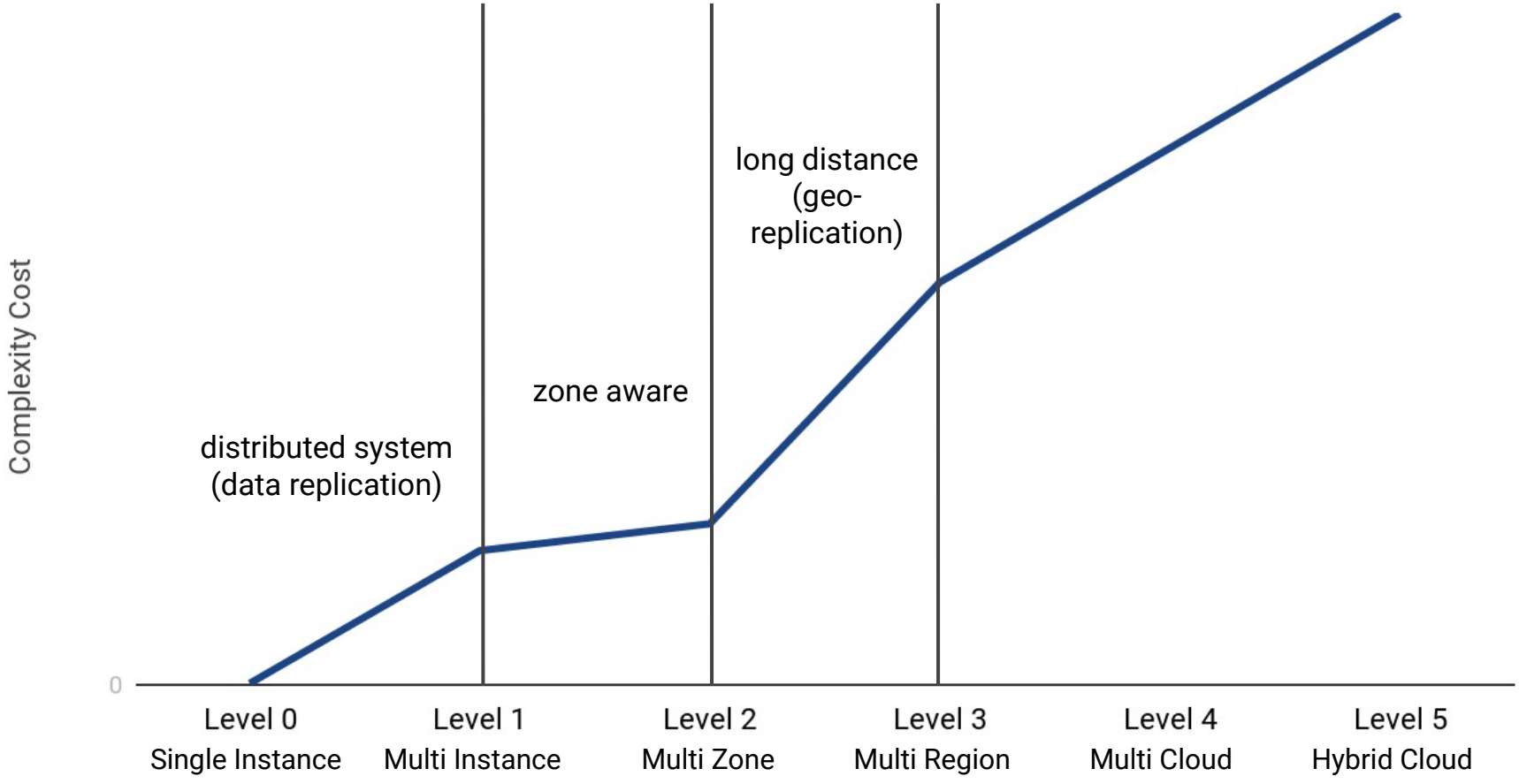
Complexity Cost / High Availability Level



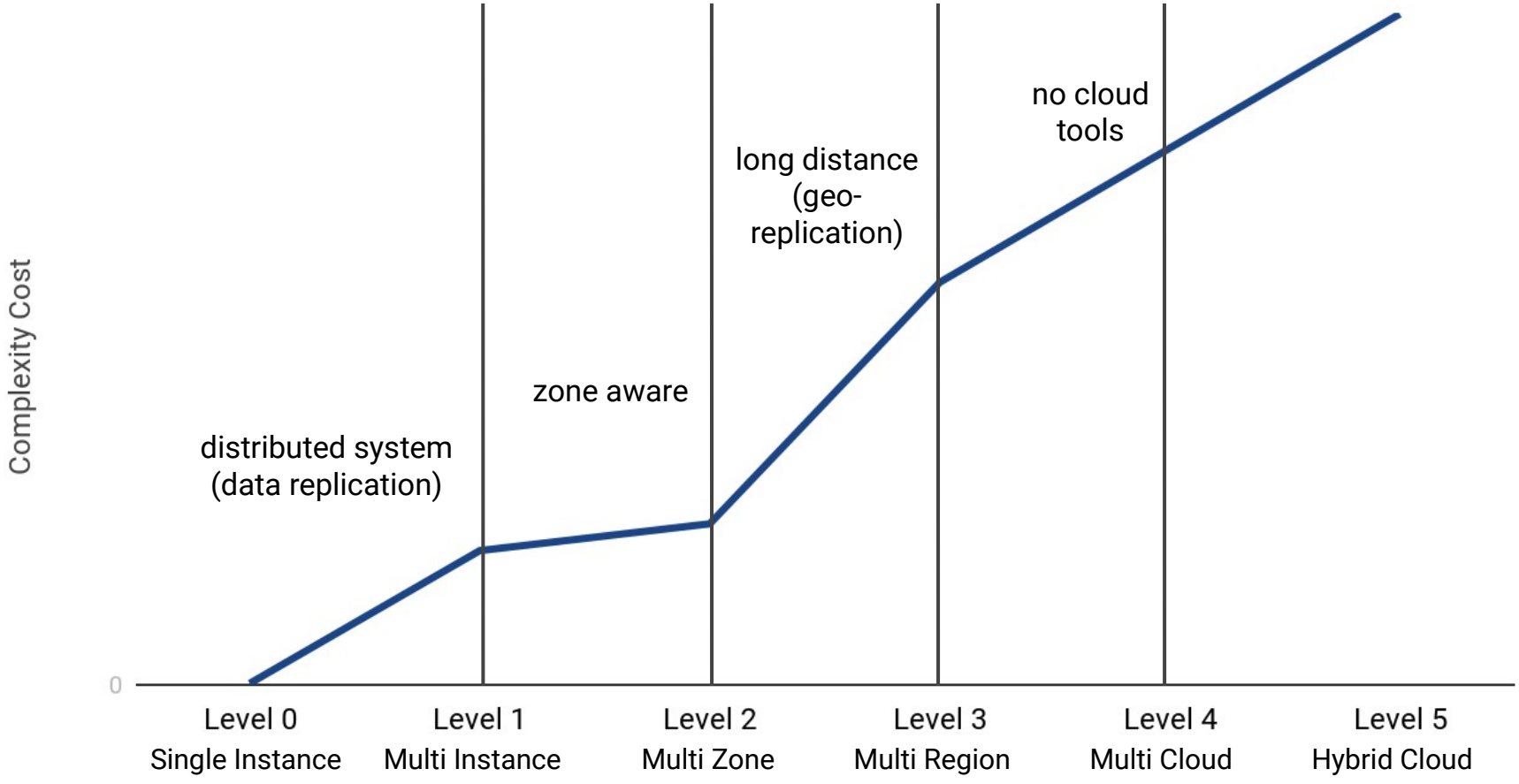
Complexity Cost / High Availability Level

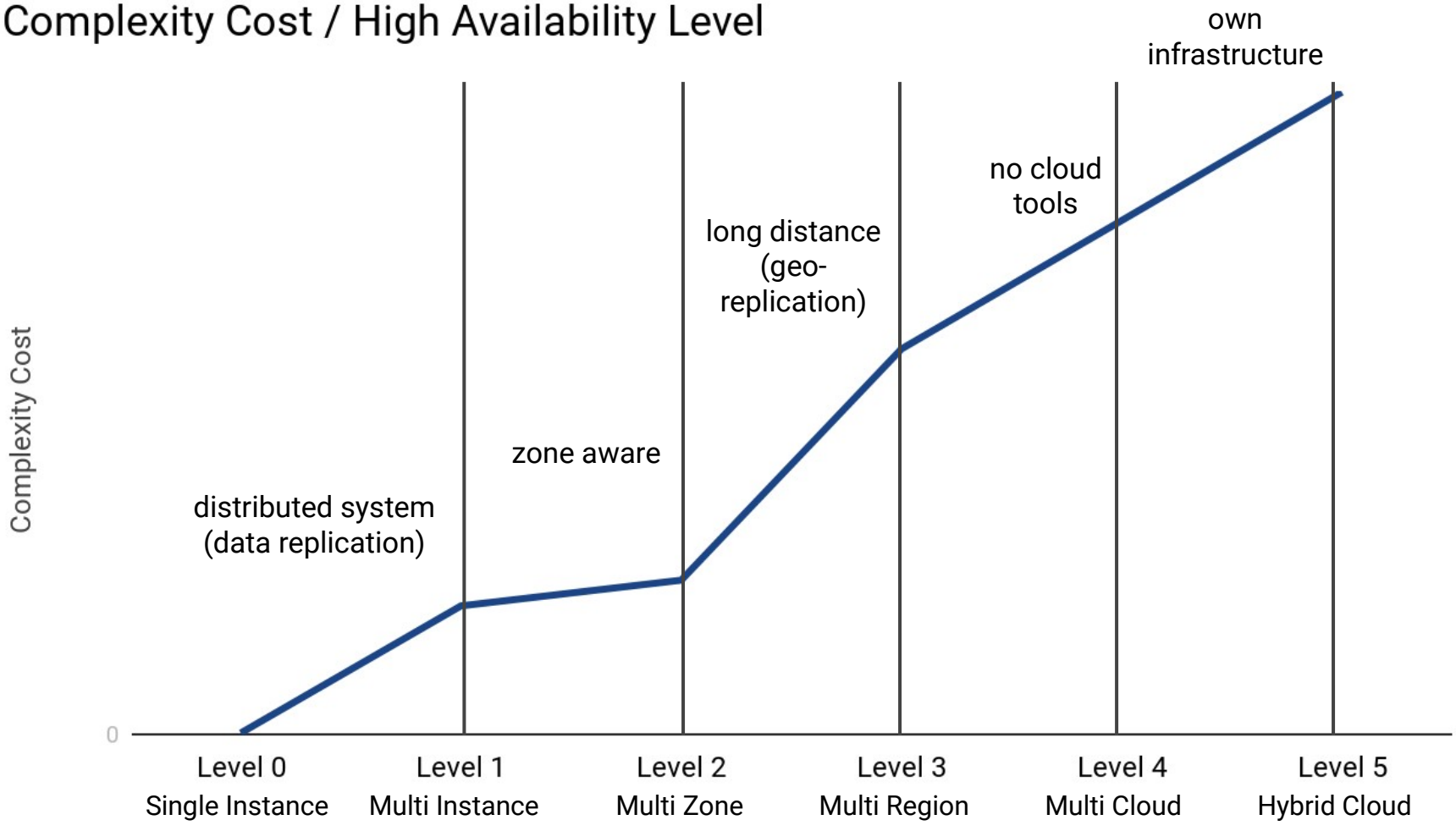


Complexity Cost / High Availability Level



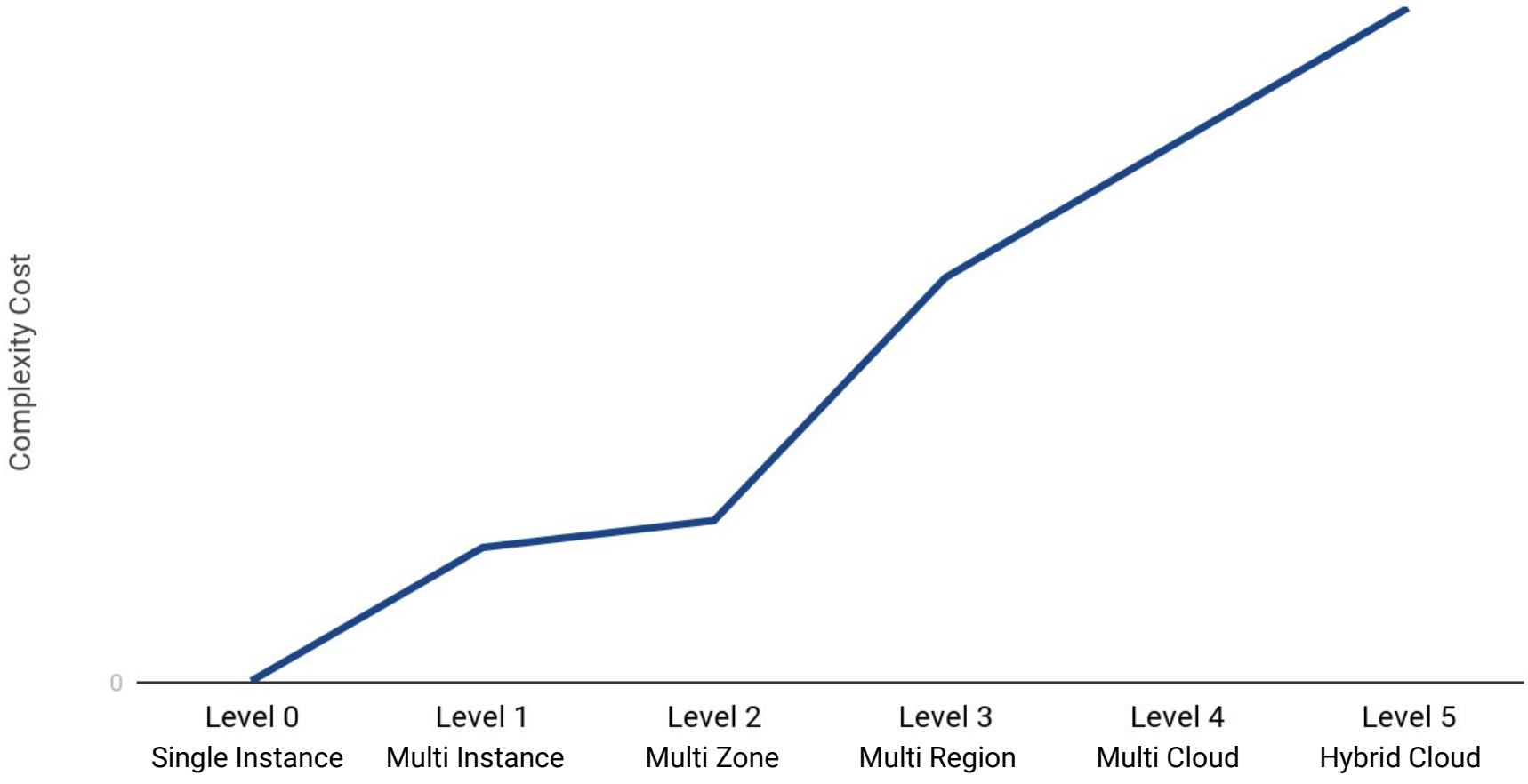
Complexity Cost / High Availability Level



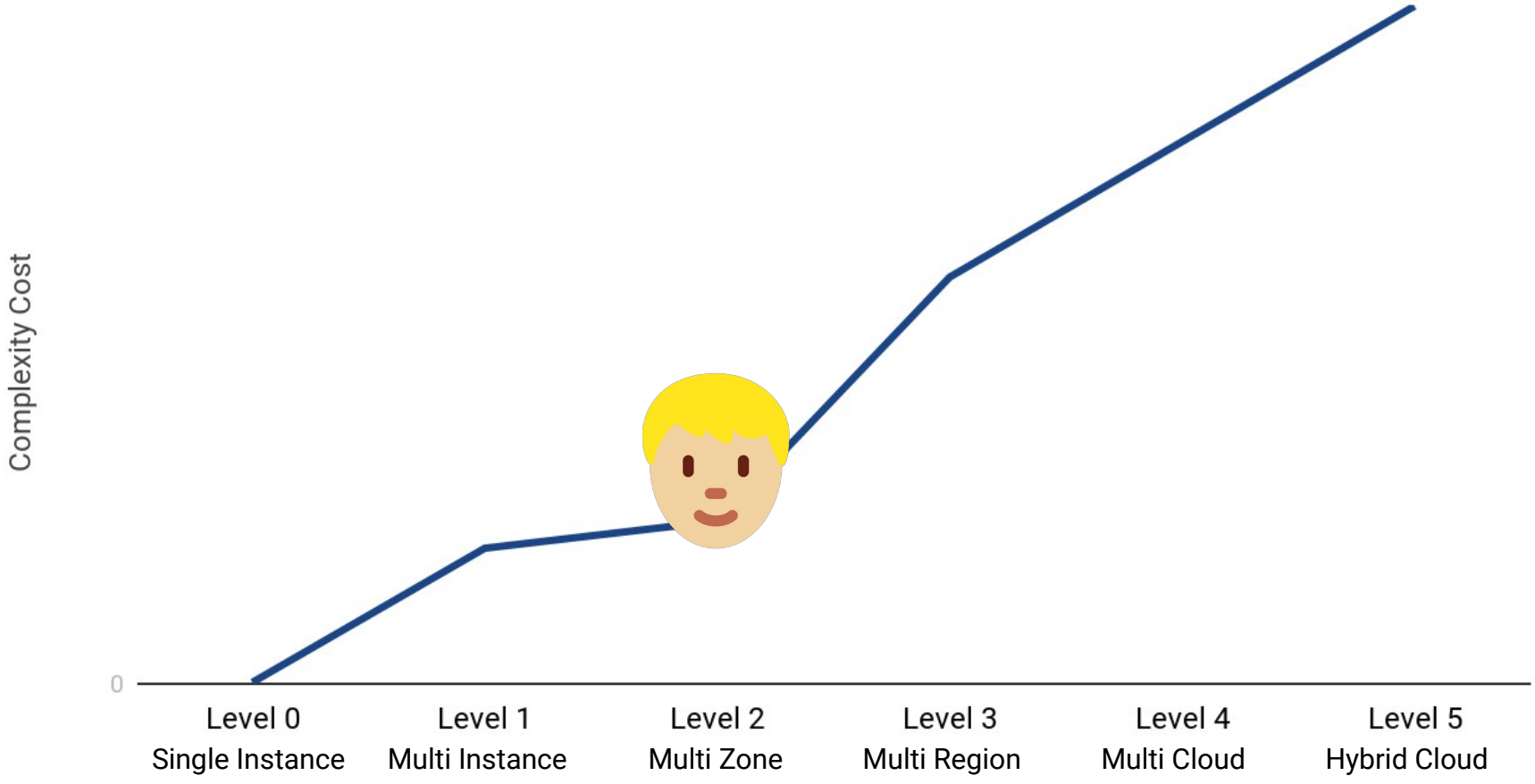

Complexity Cost / High Availability Level
Complexity Cost
0
distributed system
(data replication)
zone aware
long distance
(geo-
replication)
no cloud
tools
own
infrastructure
Level 0
Single Instance
Level 1
Multi Instance
Level 2
Multi Zone
Level 3
Multi Region
Level 4
Multi Cloud
Level 5
Hybrid Cloud

Which High Availability Level is for me?

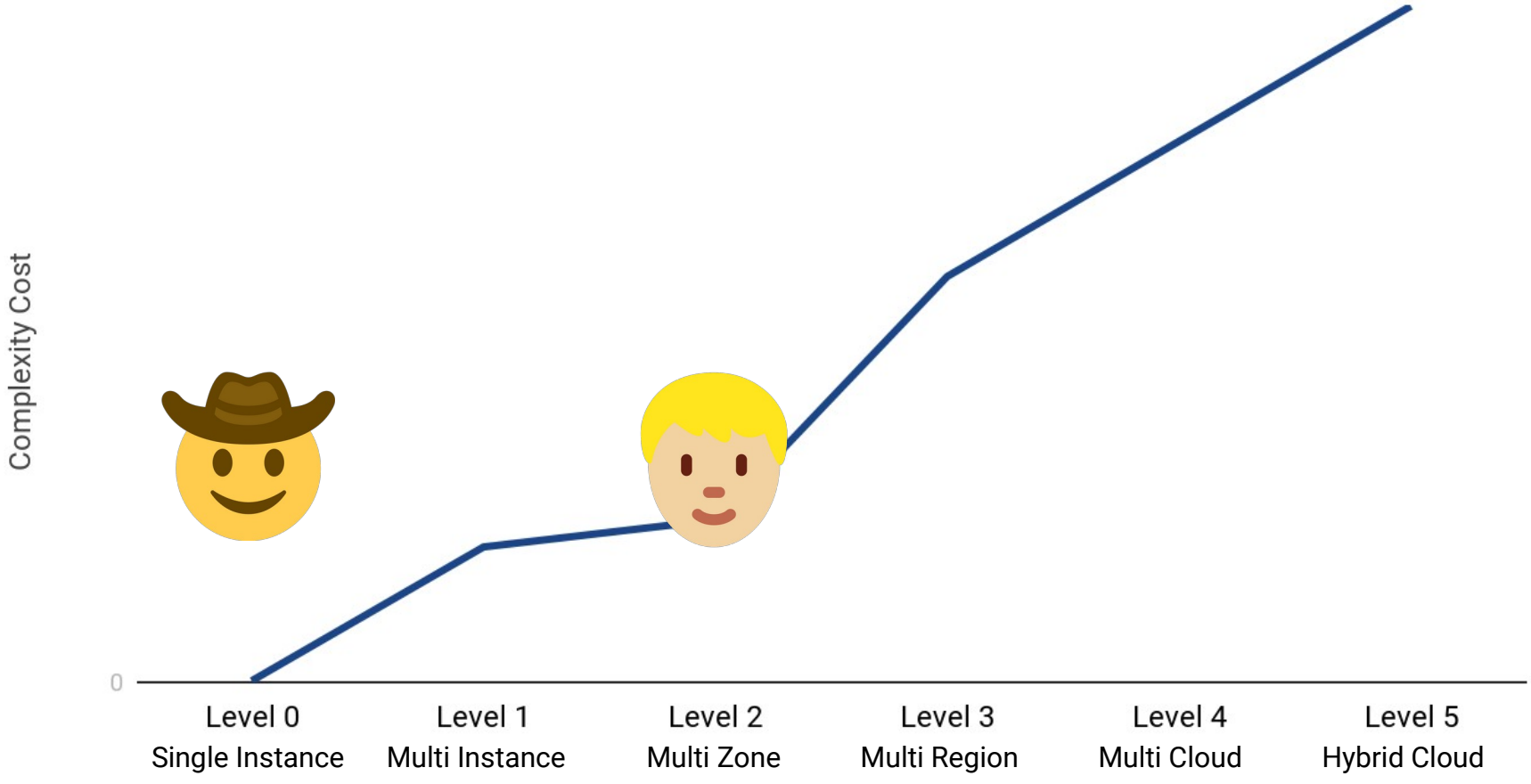
Complexity Cost / High Availability Level



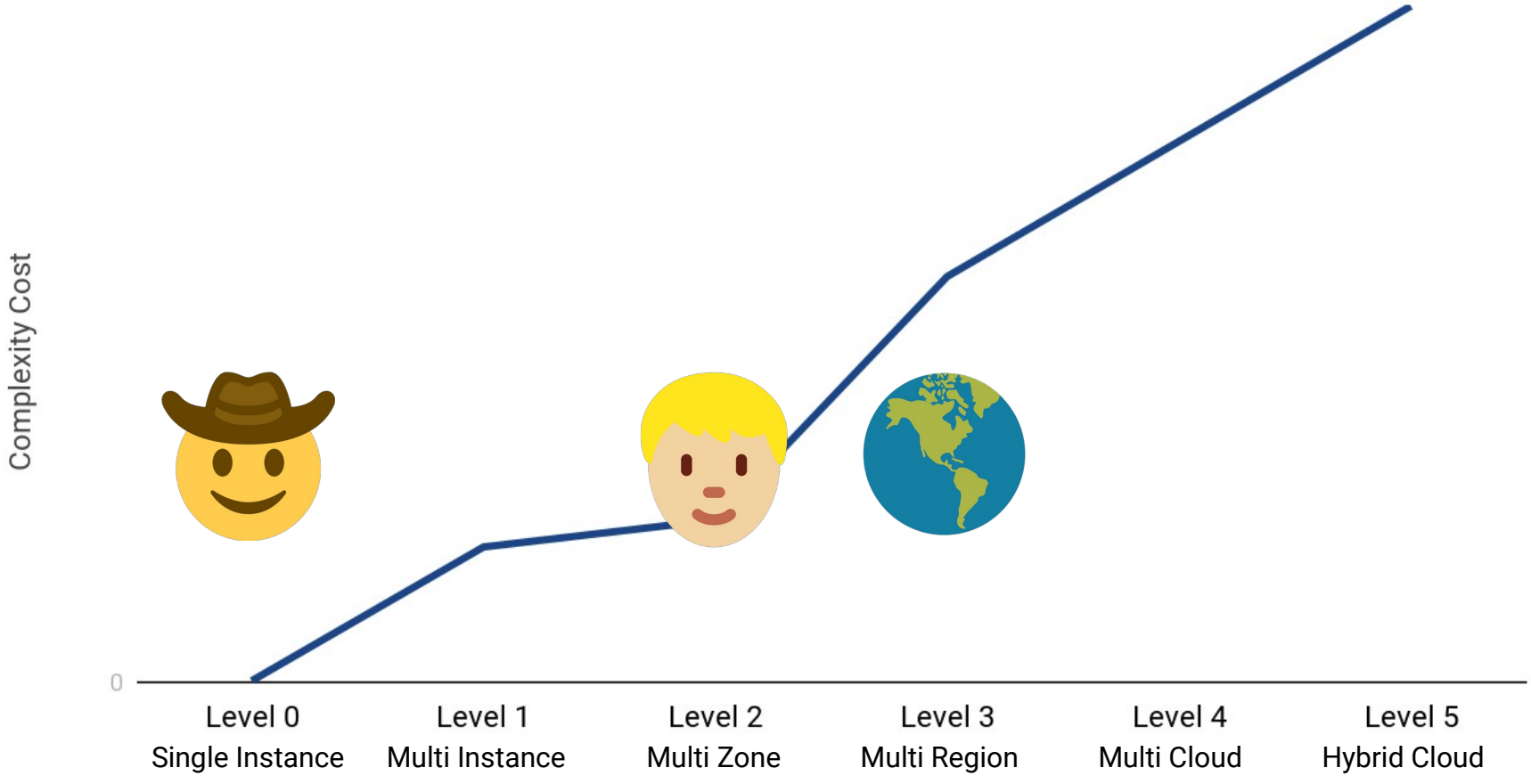
Complexity Cost / High Availability Level



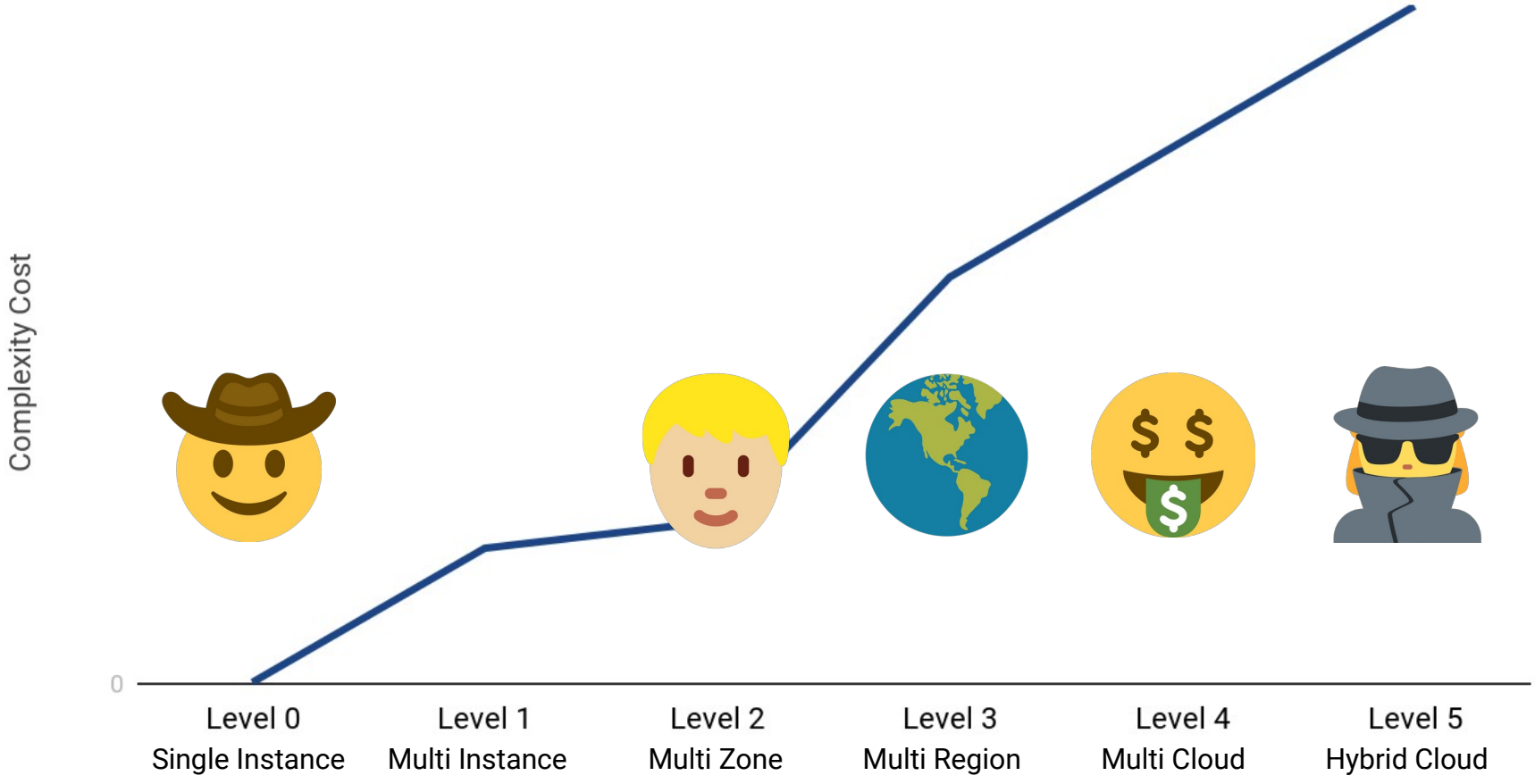
Complexity Cost / High Availability Level



Complexity Cost / High Availability Level



Complexity Cost / High Availability Level



Thank You!

Rafał Leszko
@RafalLeszko
rafalleszko.com