



Maximizing uptime and eliminating outages for Oracle databases on AWS, Azure, and GCP clouds

Art Danielov
FlashGrid, CEO
info@flashgrid.io

June 17 – 20, 2024 • Caesars Palace • Las Vegas, NV

The logo for ASCEND 2024 features the word "ASCEND" in a bold, black, sans-serif font. To the left of the "A" is a stylized graphic consisting of three parallel diagonal lines in blue, red, and blue. Below "ASCEND" is the year "2024" in a smaller, black, sans-serif font.

ASCEND
2024

Presenter



FlashGrid Inc.

HQ in Sunnyvale, CA

Global customer base

Serving Oracle customers on AWS,
Azure, and GCP since 2017



Art Danielov
CEO, FlashGrid

The Challenge



- Oracle Database powers many mission-critical applications and services
 - Database downtime means the service is down, especially for processing transactions
 - Some services must be up 24x7, no “maintenance window”
 - Even short database downtime may take down the application tier, causing longer service recovery times
-
- Maintenance is required
 - Failures happen

Designing for Maximum Database Uptime



- High availability architecture
- Reliable and compatible components
- Systems sized properly for the workload
- Systems and software configured properly for the workload and for the HA architecture
- Controlling and limiting the workload – prevent self-produced DoS attacks
- Continuous testing / analyzing / improving – by internal engineering and/or vendor
- Planning for maintenance: planned downtime or active-active architecture

Important, but not covered in this session:

- DR
- Backup
- Security

Multi-AZ vs. Single-AZ



Multi-AZ advantages

- ✓ Survive failures affecting entire data center
- ✓ Avoid the need for DR failover during shorter outages (e.g., up to 1 day)
- ✓ Low inter-AZ latency => synchronous data replication => Zero RPO

Multi-AZ considerations

- Higher latency than within single AZ (see FlashGrid KB for all inter-AZ latencies)
- More careful planning of resource placement needed
- Entire application stack must be deployed as Multi-AZ

Multi-AZ vs. Multi-Region



	Multi-AZ	Multi-Region
Active-Passive HA	Yes	No
Active-Active HA	Yes	No
Local Disaster Recovery	Yes	Yes
Major Disaster Recovery	No	Yes
Recovery From Regional Cloud Service Outage	No	Yes
Latency	<1 ms	10-100 ms
Data replication	Synchronous	Asynchronous
RPO	Zero	Non-Zero

Multi-AZ HA Options for Oracle DB



	Engineered and supported by	Managed by	Target Uptime SLA	Anticipated Downtime Per Year
Amazon RDS for Oracle	Cloud vendor	Cloud vendor	99.95%	<4 hours
Build-Your-Own Oracle DB Servers + DG FSFO	Self	Self	99.95% (?)	<4 hours (?)
FlashGrid Server + DG FSFO	Vendor + Self	Self	99.95%	<4 hours
FlashGrid Cluster for Oracle Failover HA	Vendor	Self	99.95%	<4 hours
FlashGrid Cluster for Oracle RAC: 2 nodes	Vendor	Self	99.99%	<1 hour
FlashGrid Cluster for Oracle RAC: 3+ nodes	Vendor	Self	>99.99%	< 0.1 – 1.0 hour

Multi-AZ HA Options for Oracle DB



	Engineered and supported by	Managed by	Target Uptime SLA	Anticipated Downtime Per Year
Amazon RDS for Oracle	Cloud vendor	Cloud vendor	99.95%	<4 hours
Build-Your-Own Oracle DB Servers + DG FSFO	Can you ensure the reliable operation of each server and proper failure isolation?			
FlashGrid Server + DG FSFO	Vendor + Self	Self	99.95%	<4 hours
FlashGrid Cluster for Oracle Failover HA	Vendor	Self	99.95%	<4 hours
FlashGrid Cluster for Oracle RAC: 2 nodes	Vendor	Self	99.99%	<1 hour
FlashGrid Cluster for Oracle RAC: 3+ nodes	Vendor	Self	>99.99%	< 0.1 – 1.0 hour

Oracle Database Failure Modes



What failures are usually anticipated

- Cloud instance failure
- Dead disk
- Database crash

What often happens in reality

- Out-of-memory and swapping
- Connection storms
- Disk i/o freezing (no errors returned!) temporarily
- Network disruptions
- Various failures due to configuration errors

Fault Tolerance and Failure Isolation



Clear errors/failures are easier to handle

- Disk i/o returns an error
- The network becomes fully unavailable
- System rebooted

Repeating intermittent errors may create double failures and require timely isolation

- Disk keeps producing non-fatal i/o errors
- Network unstable or turns on/off repeatedly
- System keeps crashing/rebooting

Brown-outs are dangerous, and must be prevented

- Swapping as a result of low memory
- Connection storms
- Running low on critical system resources

Handling and Isolating Disk I/O Failures



- Disk i/o is critical for database functioning.
- Disk i/o failures happen.
- How are your disk timeouts configured?
- What happens when a disk i/o is stuck for 1 minute?
- What happens when a disk i/o is stuck for 1 hour?
- Have you tested it?
- Have you tested it for different disk types – data, voting, db home, OS root?
- Have you re-tested it after changing cloud instance type or upgrading OS?

Tip: use *AWS Fault Injection Simulator for EBS*

HA Solution: Build Your Own vs Buy Pre-Integrated



Building your own HA

- Significant and continuous investment of engineering resources
- Highly skilled DB and cloud engineers must be available 24x7
- If not done right, may result in extended downtime

Advantages of a pre-integrated HA solution

- ✓ Shorter deployment schedule
- ✓ Reduced risk of configuration errors: extensive testing, experience with other customers
- ✓ 24x7 support

Downtime for Maintenance and Patching?



Maintenance is unavoidable and can reduce uptime SLA considerably

- Security/reliability patches for database, OS
- Resizing
- Reconfiguration
- Physical host retirement

Plan for maintenance:

- What will be the impact on your application/service?
- How much time can be allocated?
- How often?
- How soon can be scheduled?
- Who will need to approve?

Zero-downtime maintenance with active-active HA/RAC

- ✓ Zero-downtime “rolling” patching
- ✓ Zero/minimal impact on applications
- ✓ Simplified maintenance planning

RAC + Multi-AZ enabling 99.999% uptime



- ✓ Near-zero RTO with RAC
- ✓ Survive a double failure with 3+ RAC nodes
- ✓ Survive an outage affecting entire data center
- ✓ Avoid the need for DR failover during DC outages (e.g., up to 1 day)
- ✓ Node-level reliability and failure isolation with Active Reliability Framework

99.999% Oracle database uptime is attainable

Basic Steps to Achieve Required Uptime SLA



- Decide on service uptime SLA
- Decide on database uptime SLA
(must be higher than the service uptime SLA)
- Choose database HA architecture that enables the required uptime SLA
(plus DR and backup strategies)
- Choose between Build-Your-Own vs. Pre-Integrated HA implementation
- Ensure that the database and the app tier are sized and used correctly
- Continuously monitor and adjust configuration as needed

Multi-AZ HA Options for Oracle DB



	Engineered and supported by	Managed by	Target Uptime SLA	Anticipated Downtime Per Year
Amazon RDS for Oracle	Cloud vendor	Cloud vendor	99.95%	<4 hours
Build-Your-Own Oracle DB Servers + DG FSFO	Self	Self	99.95% (?)	<4 hours (?)
FlashGrid Server + DG FSFO	Vendor + Self	Self	99.95%	<4 hours
FlashGrid Cluster for Oracle Failover HA	Vendor	Self	99.95%	<4 hours
FlashGrid Cluster for Oracle RAC: 2 nodes	Vendor	Self	99.99%	<1 hour
FlashGrid Cluster for Oracle RAC: 3+ nodes	Vendor	Self	>99.99%	< 0.1 – 1.0 hour



Q&A

info@flashgrid.io

 **ASCEND**
2024



Thank You For Attending!

Please complete the session
survey in the conference app.

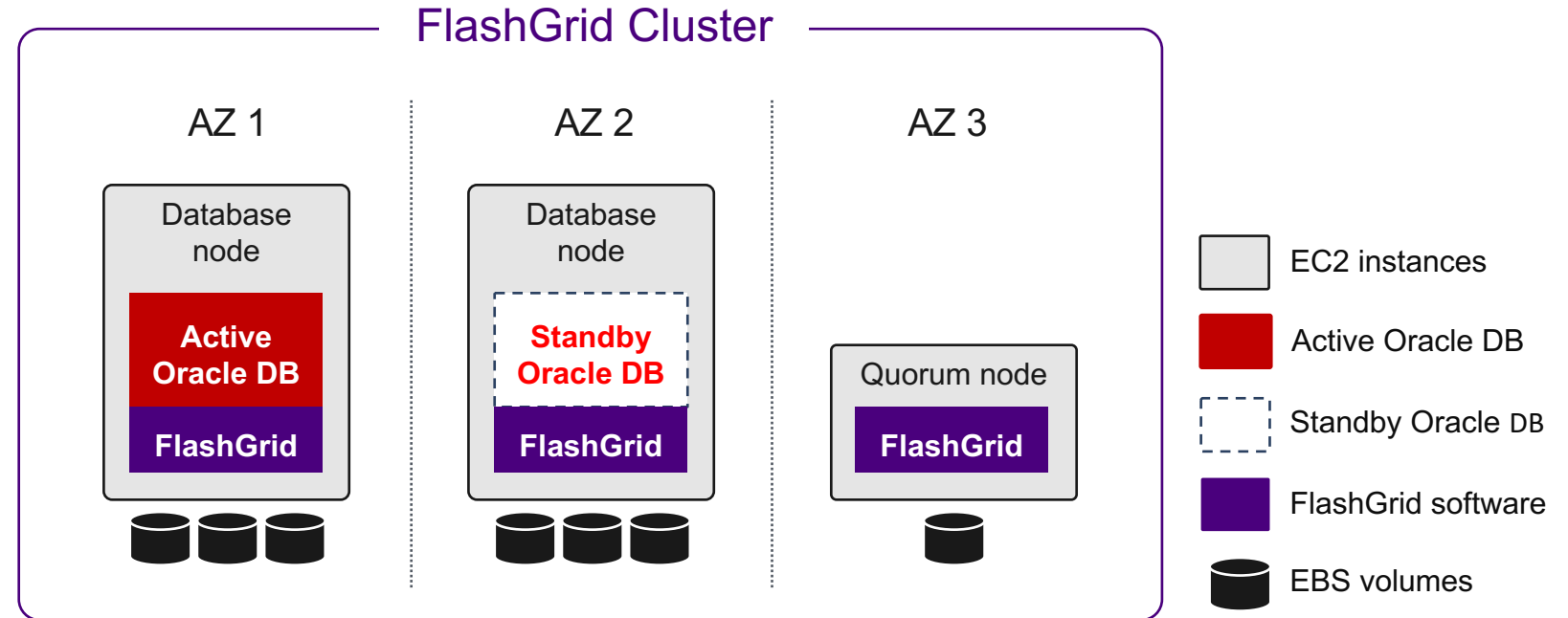


Backup slides

Oracle Failover HA on AWS



- ❖ Active-Passive HA
- ❖ Automatic Failover
- ❖ Multi-AZ
- ❖ Suitable for **99.95%** target uptime



FlashGrid Cluster with active and failover Oracle databases

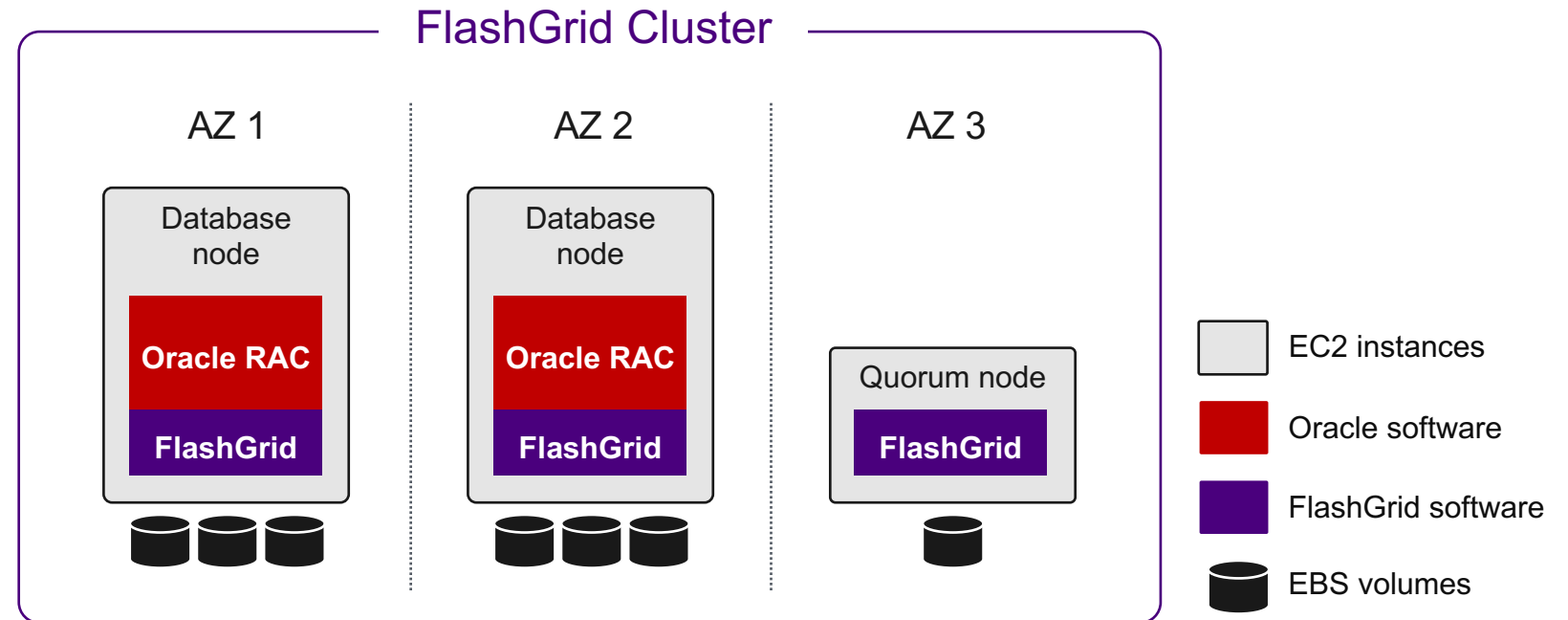
Additional benefits

- ✓ Fully controlled by customer
- ✓ All Oracle features and tools available
- ✓ Your own patching / upgrade schedule
- ✓ Larger instances up to 448 vCPU
- ✓ Larger storage capacities >64TB
- ✓ Higher storage throughput up to 10 GBPS
- ✓ Bare-metal instances

Oracle RAC on AWS: 2 db nodes



- ❖ Active-Active HA
- ❖ Multi-AZ
- ❖ Can survive loss of **one node**
- ❖ Suitable for **99.99%** target uptime

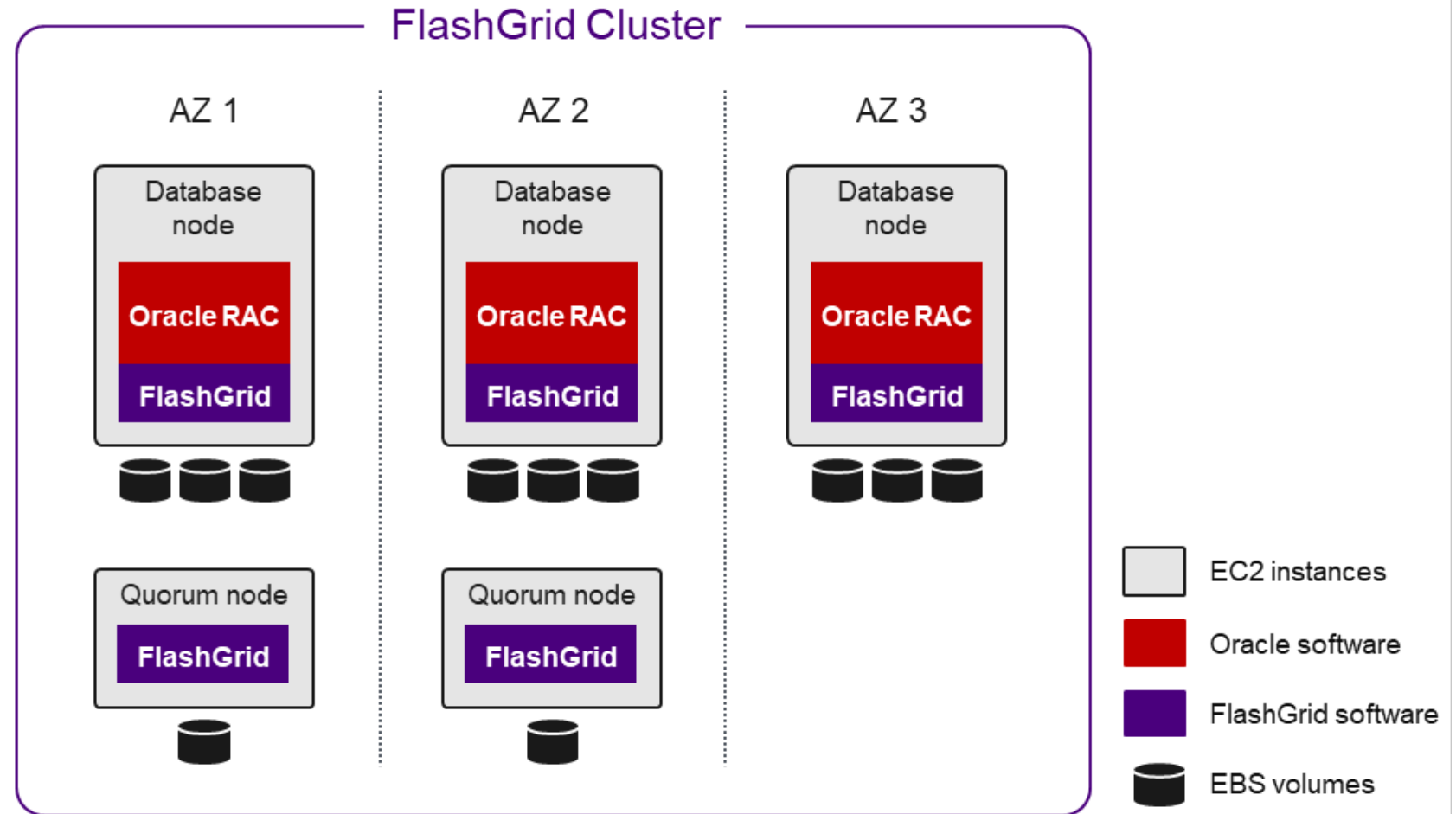


FlashGrid Cluster with two RAC database nodes

Oracle RAC on AWS: 3 db nodes

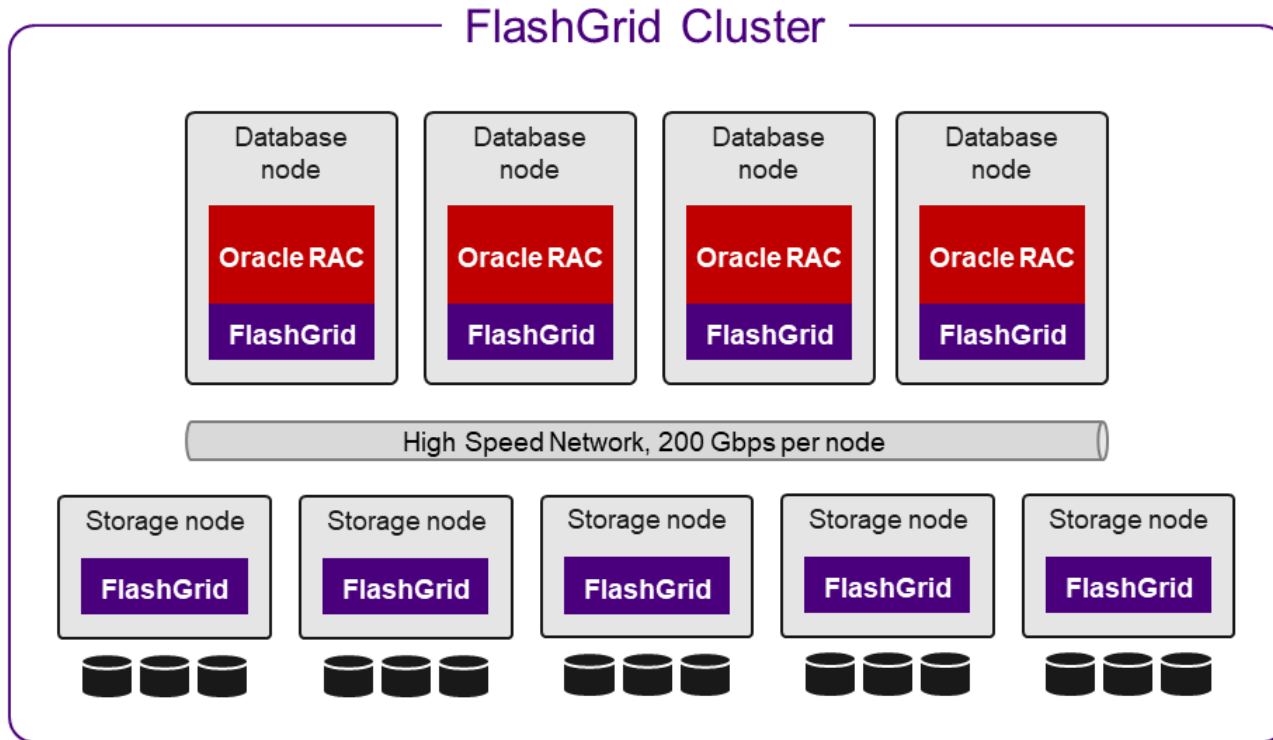


- ❖ Active-Active HA
- ❖ Multi-AZ
- ❖ Can survive loss of **two nodes**
- ❖ Suitable for **>99.99%** target uptime



FlashGrid Cluster with three RAC database nodes

Oracle RAC on AWS: Extra Large Databases



Extra-large database cluster on AWS with 4+ RAC nodes and separate storage nodes

20 read + 20 write GBPS per DB node (with r6in.metal)

Up to 10 GBPS per storage node (with c6in.metal)

Any number of db and storage nodes

Up to 400 TB (25 x 16TB) raw per storage node

Each storage node in a separate failgroup

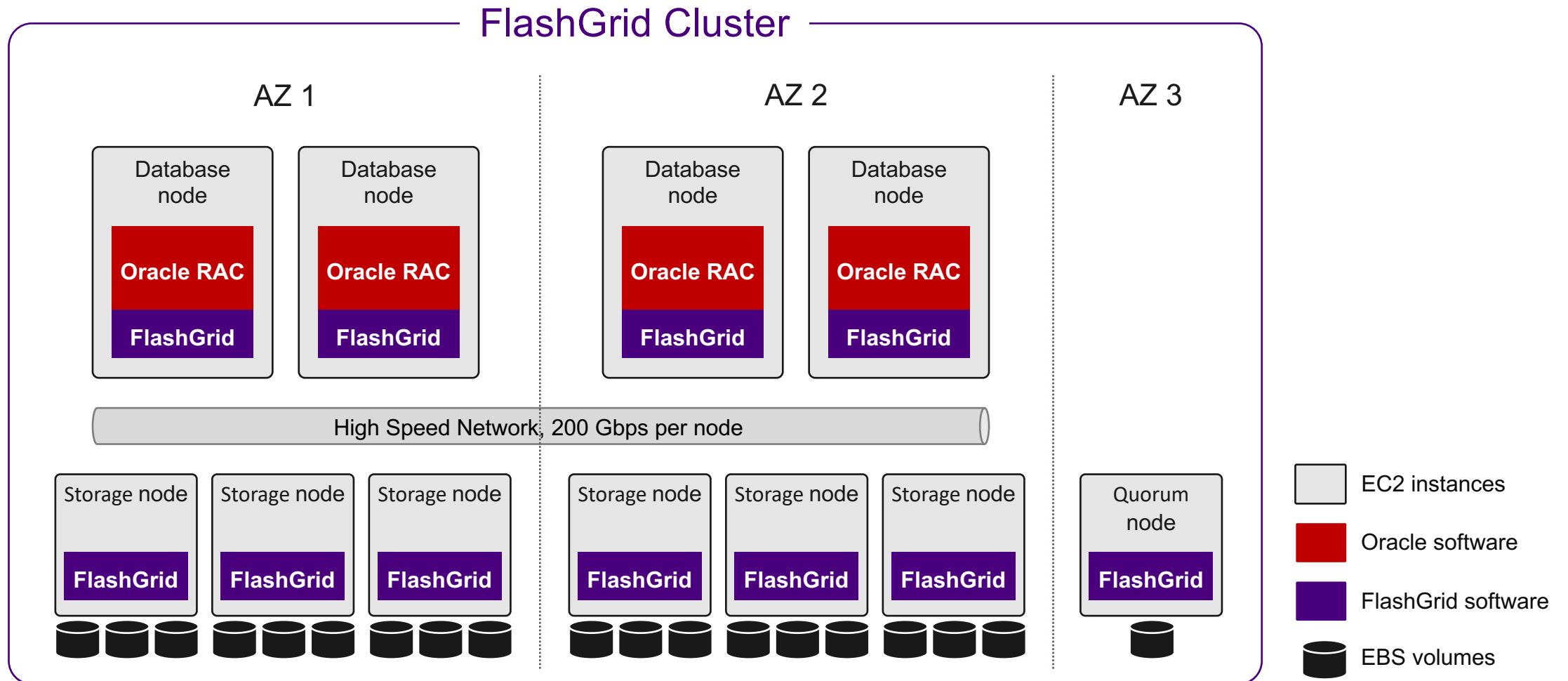
Normal Redundancy

- ❖ 2x data mirroring
- ❖ usable capacity = 1/2 raw capacity
- ❖ Can survive loss of one storage node

High Redundancy

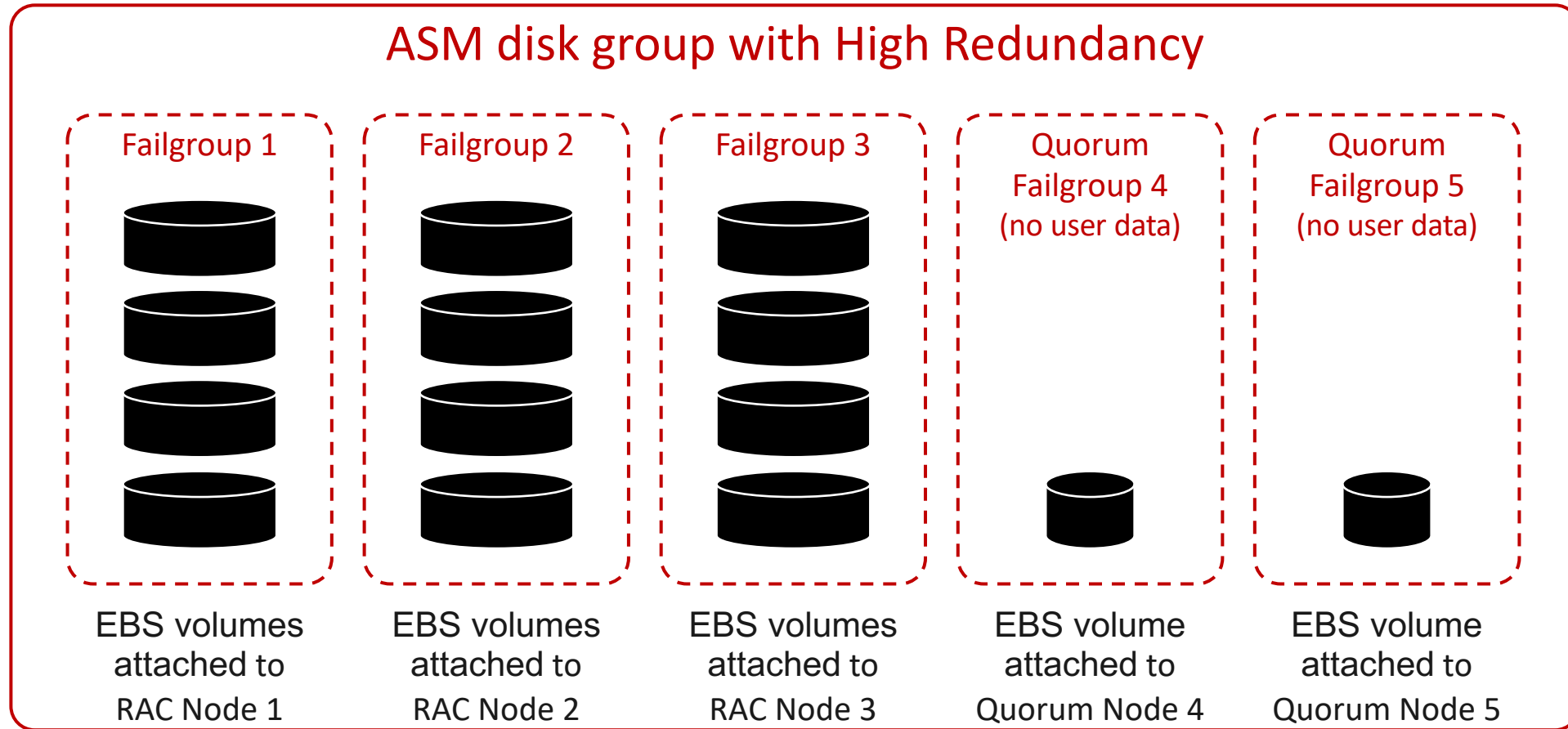
- ❖ 3x data mirroring
- ❖ usable capacity = 1/3 raw capacity
- ❖ Can survive loss of two storage nodes

Oracle RAC on AWS: Extra Large Databases, Multi-AZ



Extra-large database cluster with multi-AZ

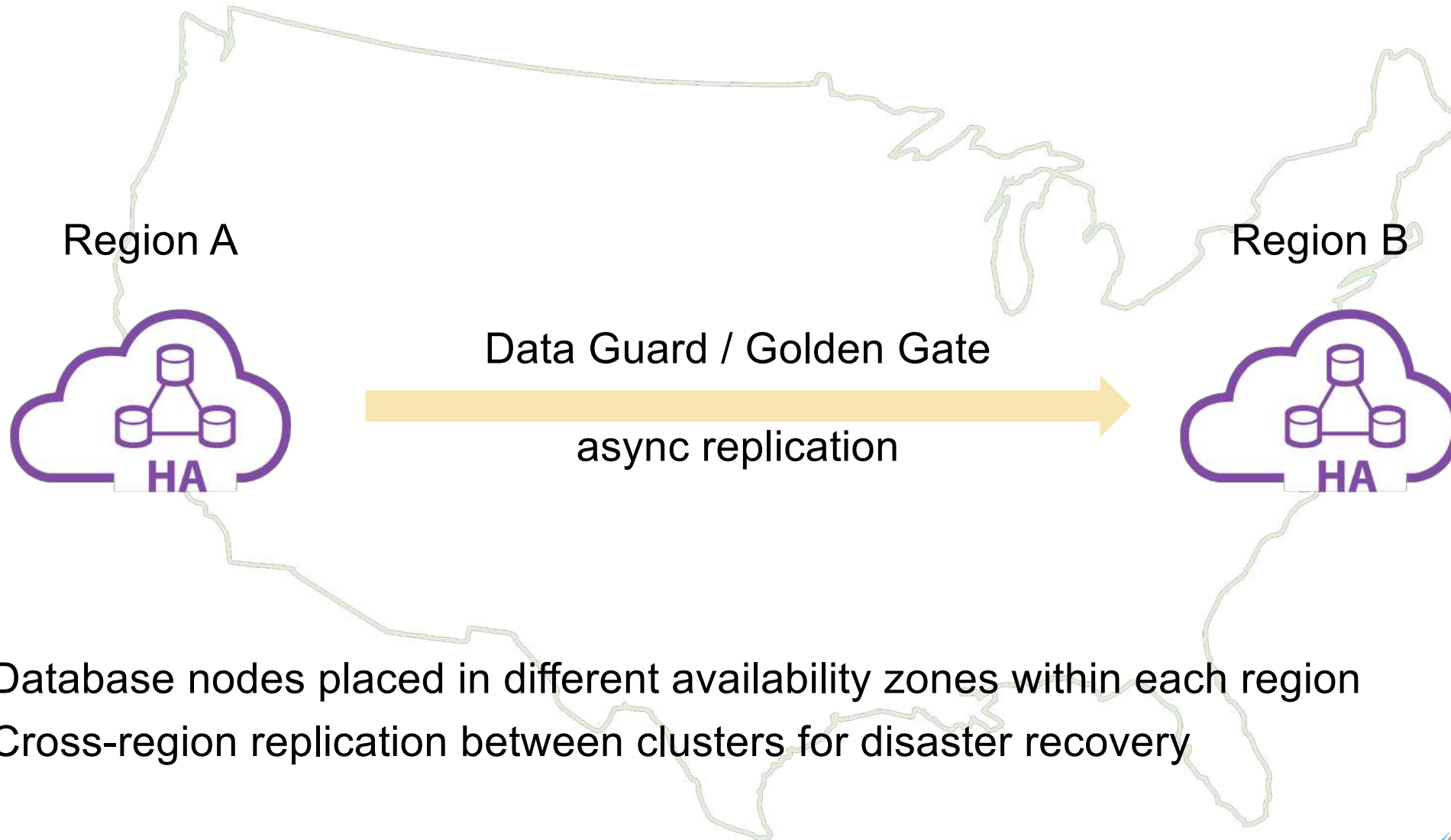
ASM Disk Group configuration example: 3 Nodes, 3-way mirroring



Additional Quorum Failure Groups recommended for better HA

Quorum Disk: 1 GiB disk on the quorum node, very low performance needed (~1 IOPS)

Maximum Availability: HA + DR



- ❖ Database nodes placed in different availability zones within each region
- ❖ Cross-region replication between clusters for disaster recovery



To learn more, visit:

www.flashgrid.io

The logo for ASCEND 2024. It features the word "ASCEND" in a bold, black, sans-serif font. To the left of the "A" is a stylized graphic consisting of three parallel diagonal lines in blue, red, and white. Below "ASCEND" is the year "2024" in a smaller, black, sans-serif font.