
AWS
re:Invent

EMB013

Powering Amazon EBS: Infrastructure deep dive

Marc Olson

Principal Software Engineer

AWS

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Definitions

The evolution of Amazon Elastic Block Store (Amazon EBS)

EBS Block Express

What’s next



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

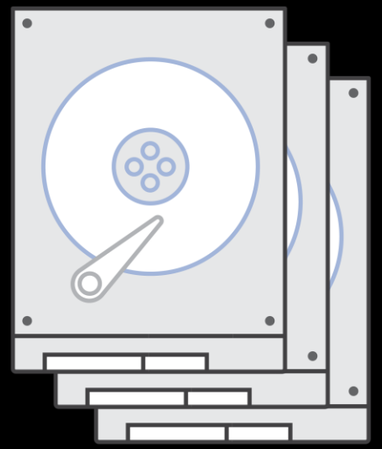
Definitions



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What is EBS?

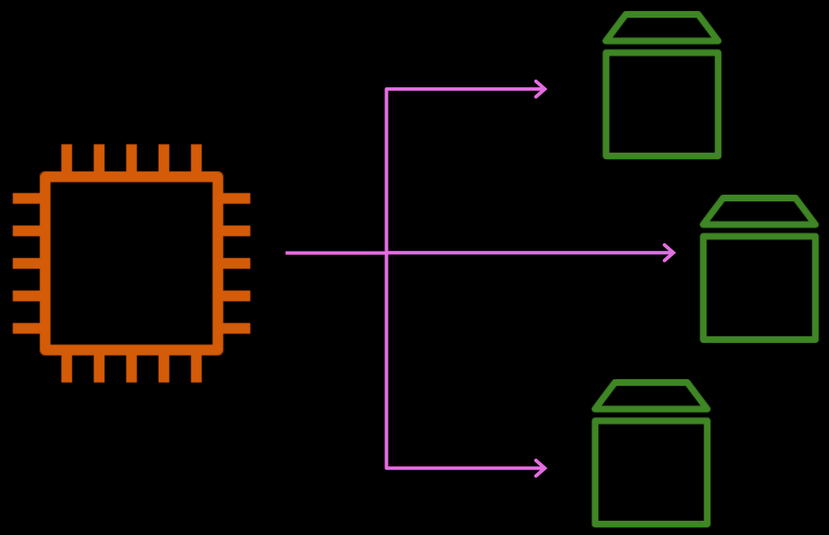
EBS ⊜



What is EBS?

- Network block storage as a service for Amazon EC2 instances
- Data services features
- Large distributed system

EBS in two parts: Data plane



- Compute instance
- Nitro controller
- Network
- Storage servers

EBS in two parts: Control plane



- API front ends
- Volume metadata
- Data placement services
- Repair workflows

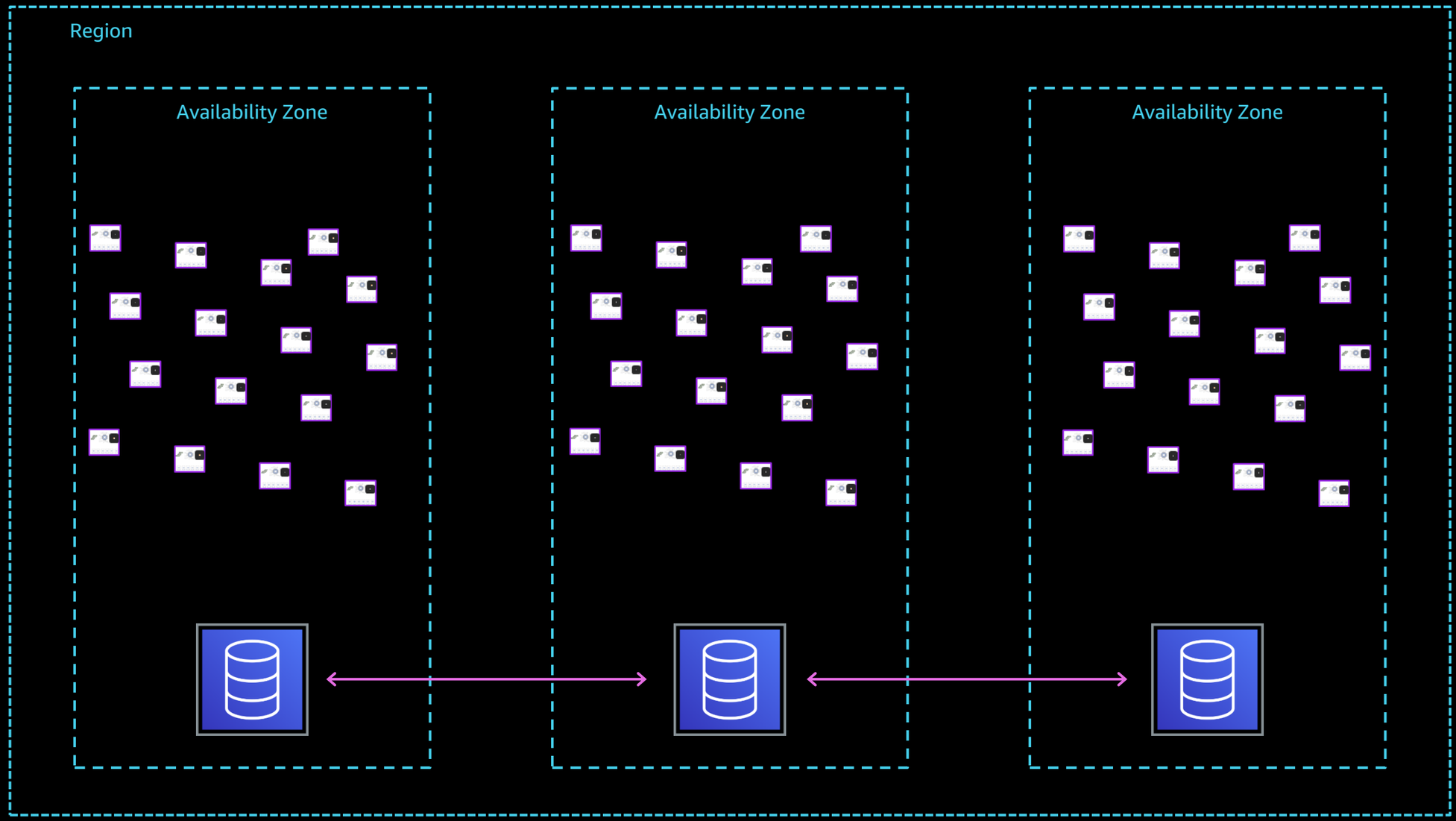
The evolution of EBS



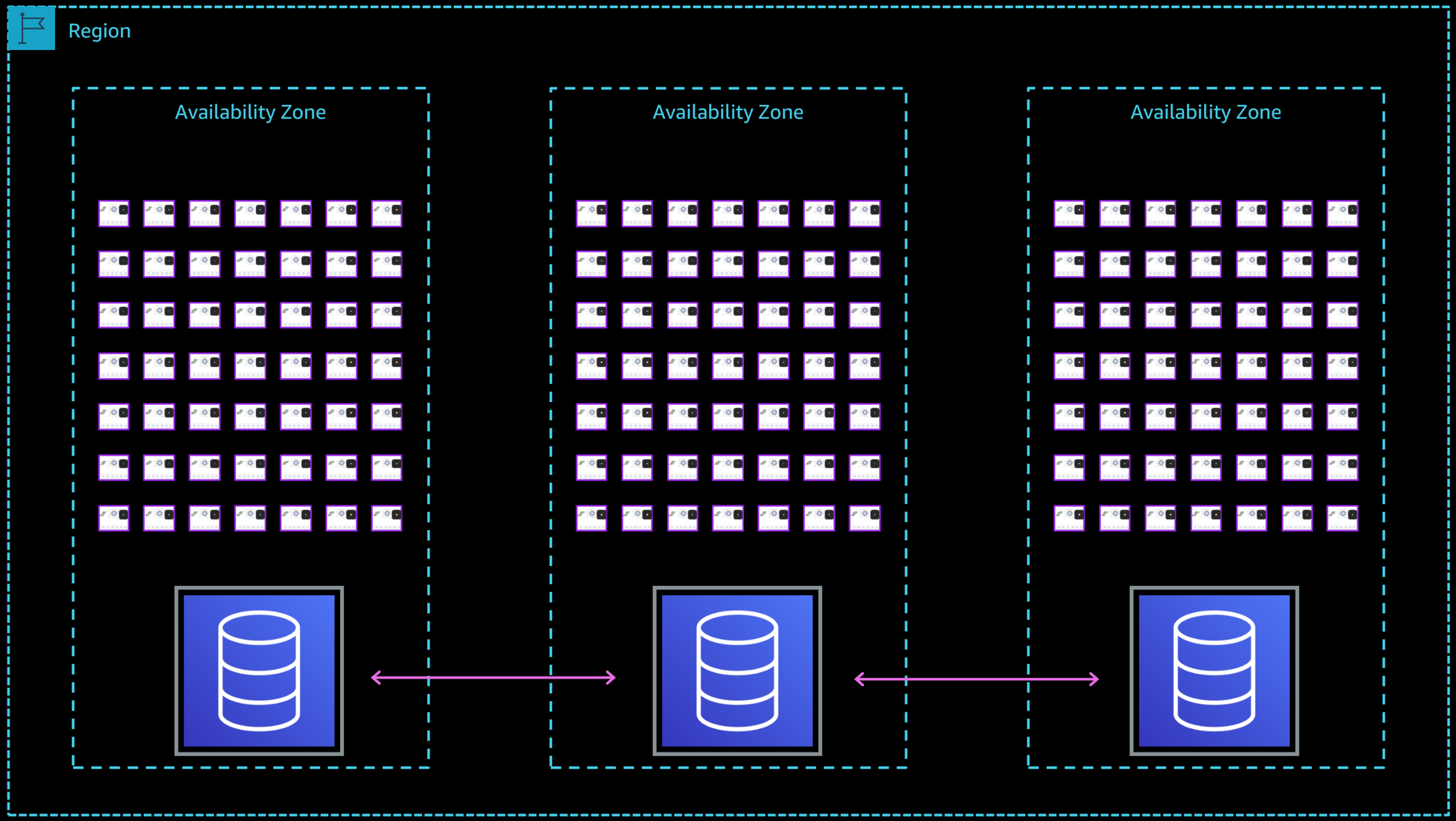
AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

In the beginning . . .



Scaling


Region
Availability Zone
Availability Zone
Availability Zone

Blast radius reduction



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

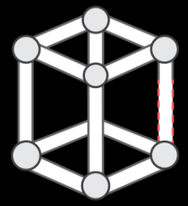
Blast radius reduction: Configuration manager

Region

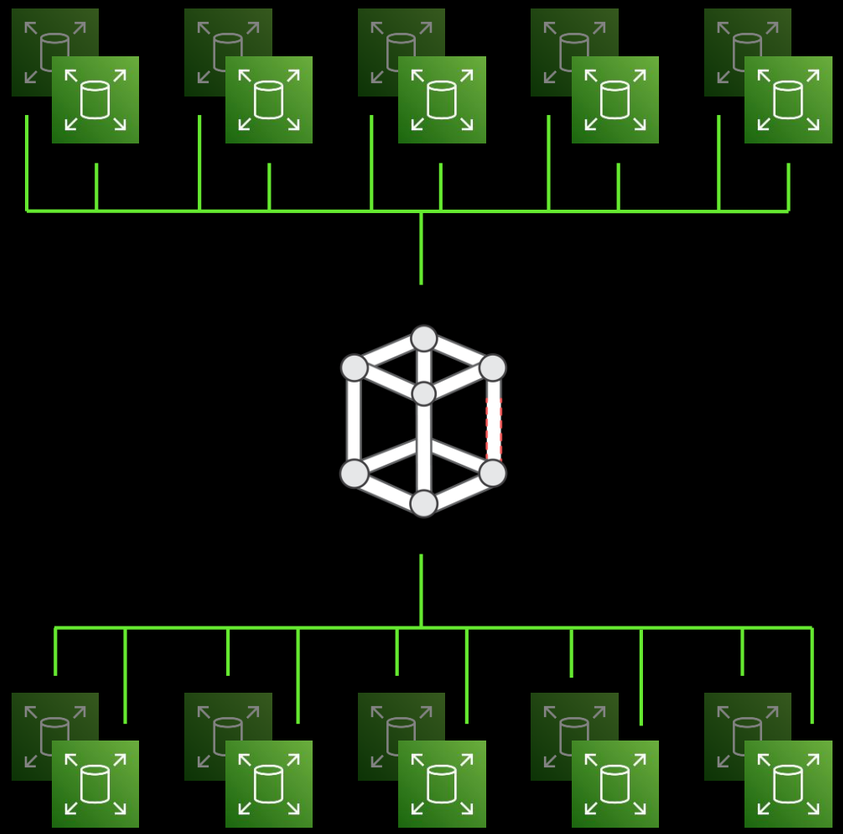




Availability Zone

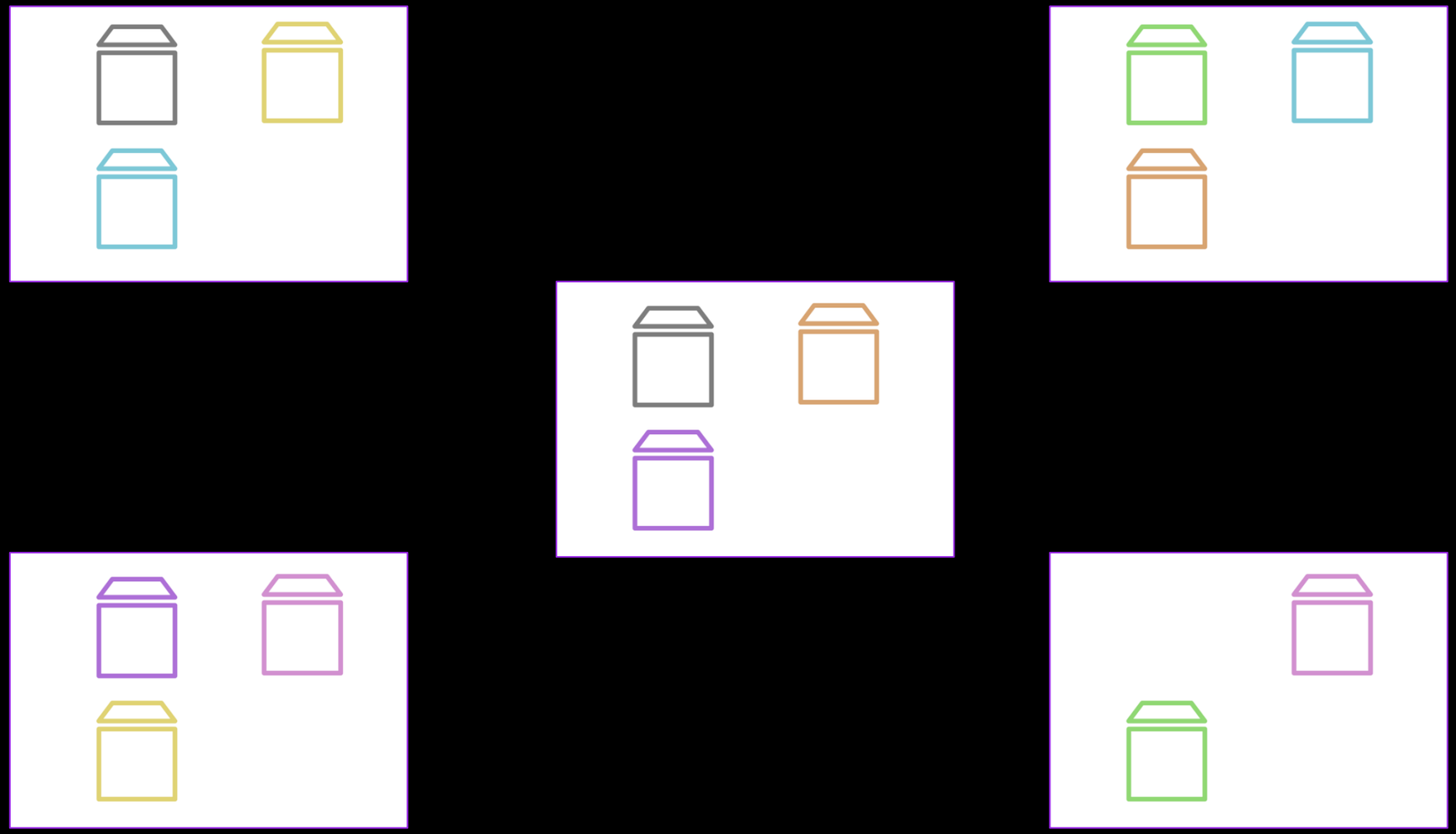


Configuration manager

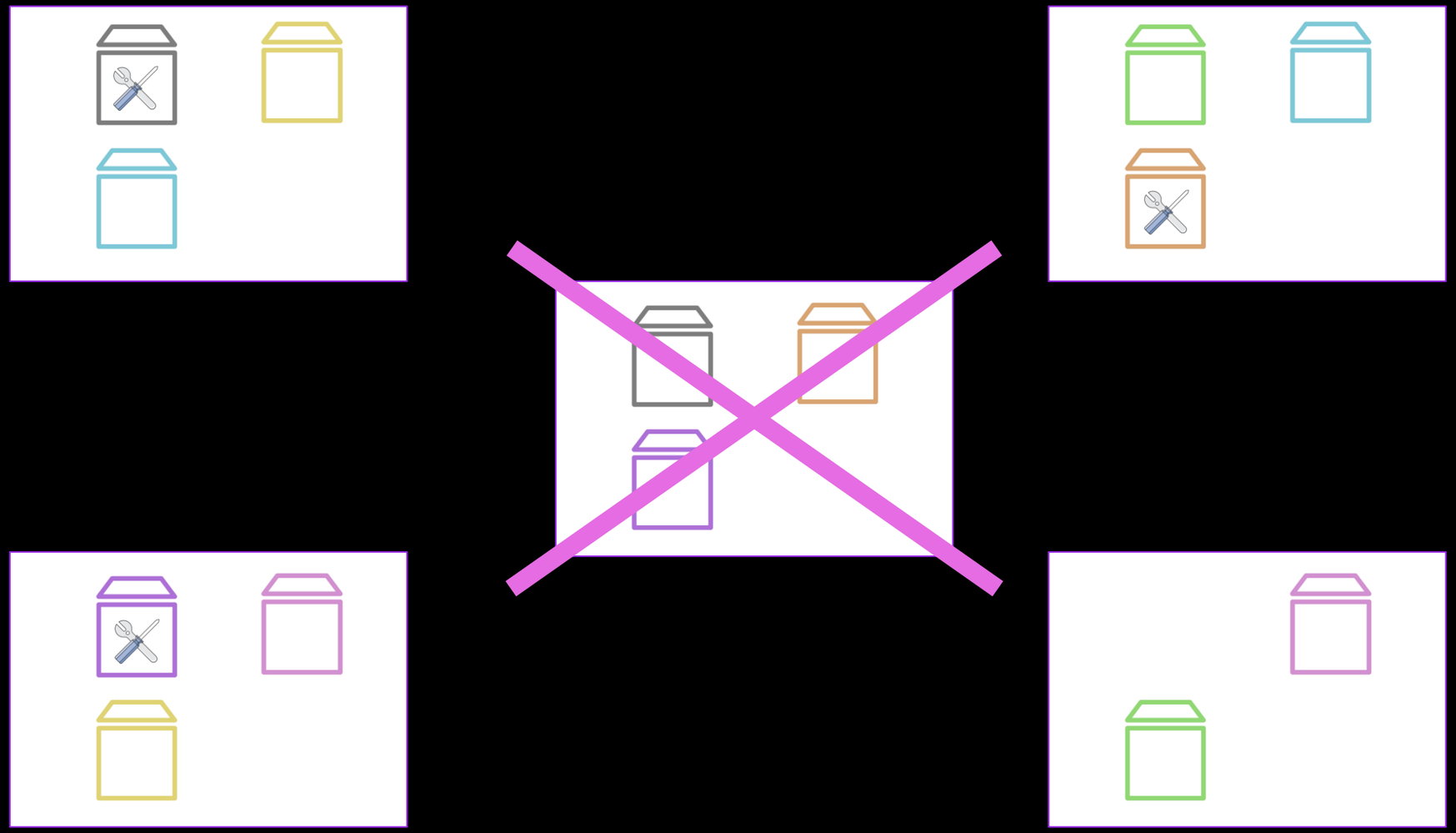


- Authority map for host and volume
- Wholly contained in Availability Zone
- Standalone service
- Paxos-based replication

EBS volume repair



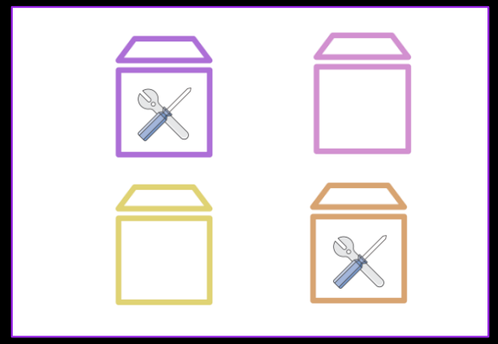
EBS volume repair

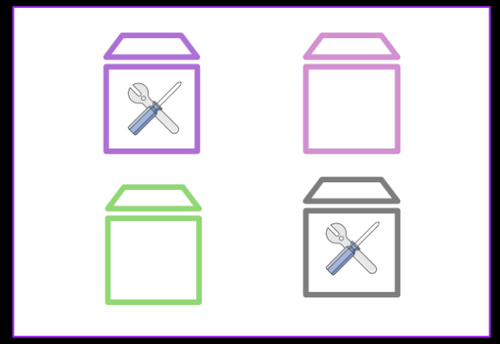


EBS volume repair









EBS volume repair

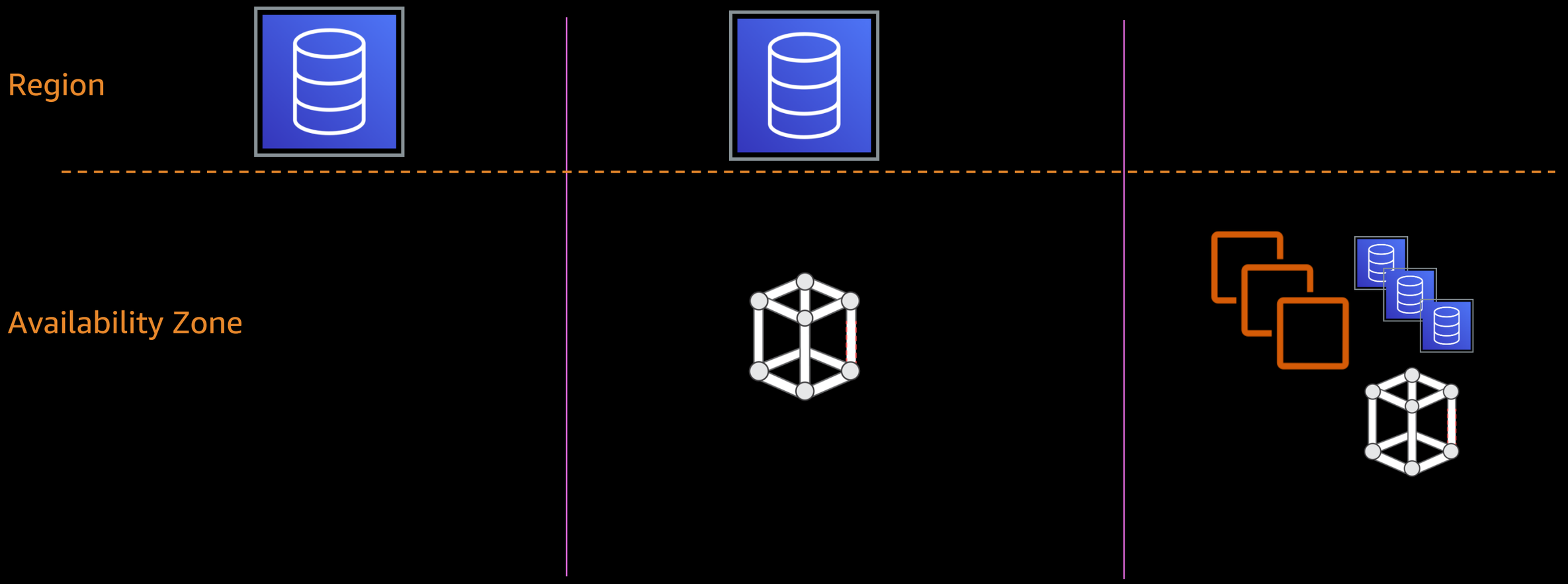




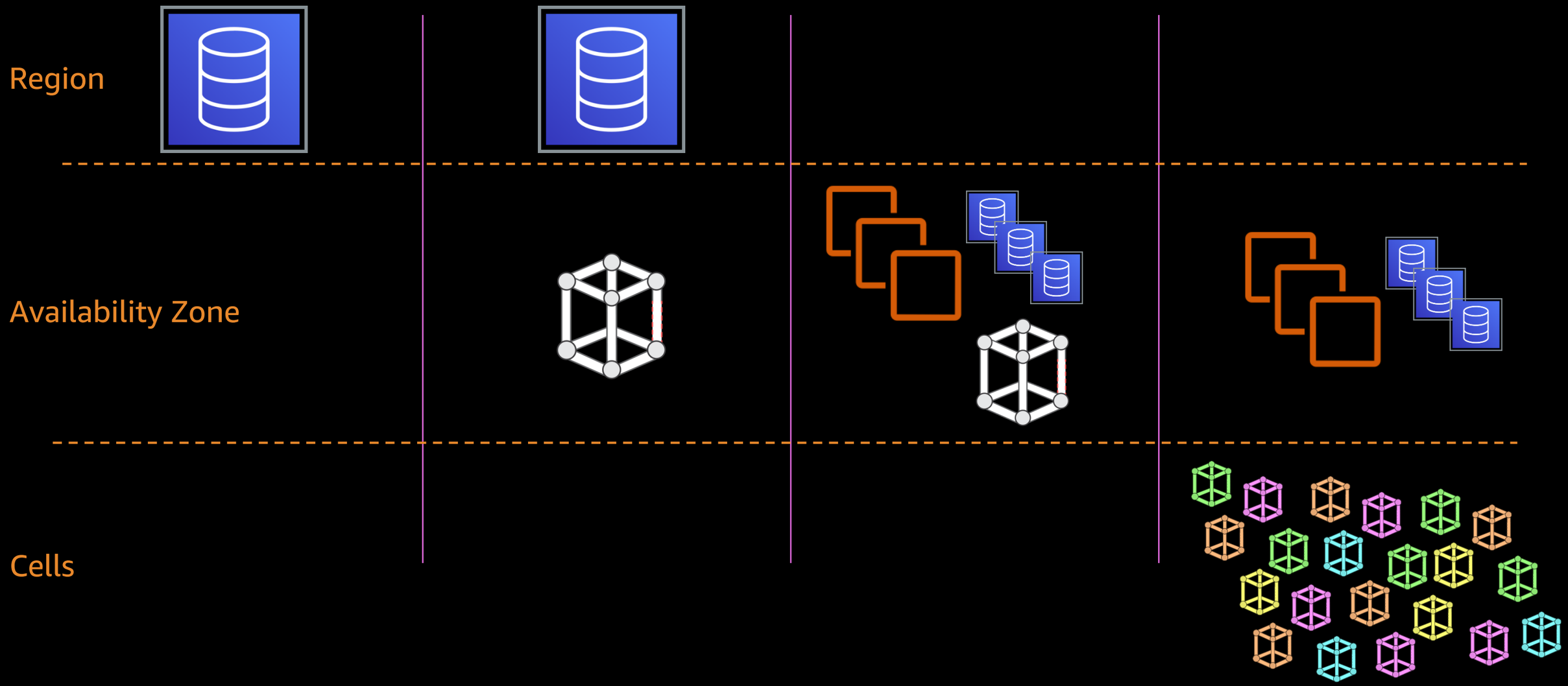




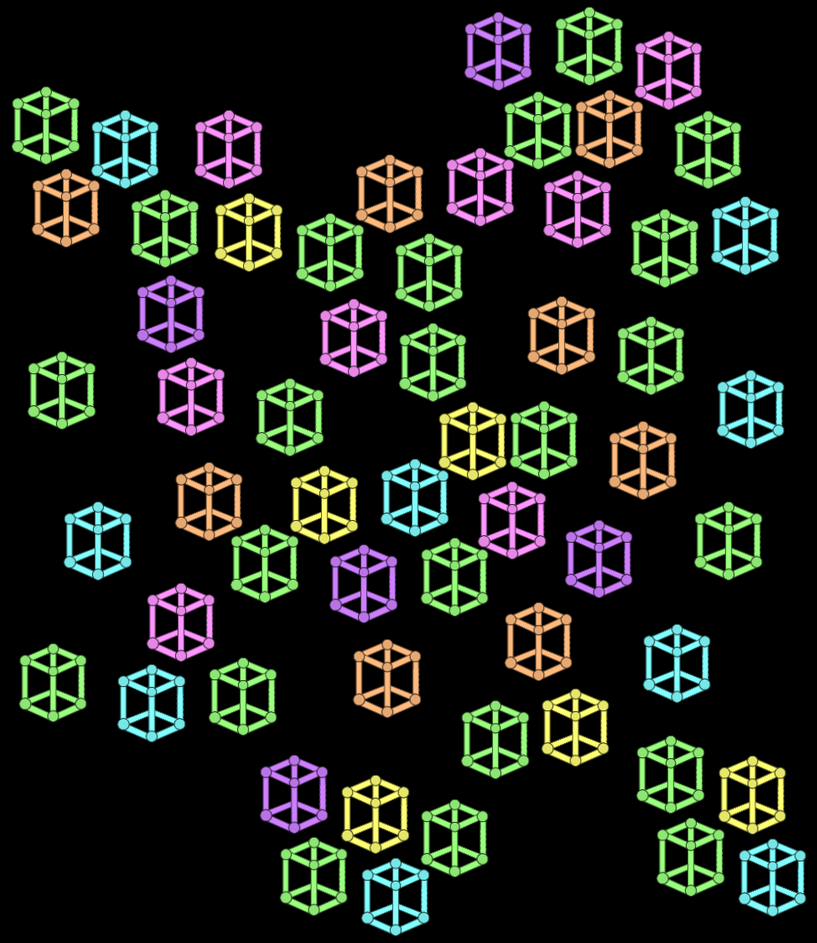
Blast radius reduction: AZ independence



Blast radius reduction: Physalia



Physalia: Millions of tiny databases



- Small number of clients per database
- Simple to manage
- Cell can reside near clients
- Reduced blast radius
- Each cell is a 7-node Paxos cluster

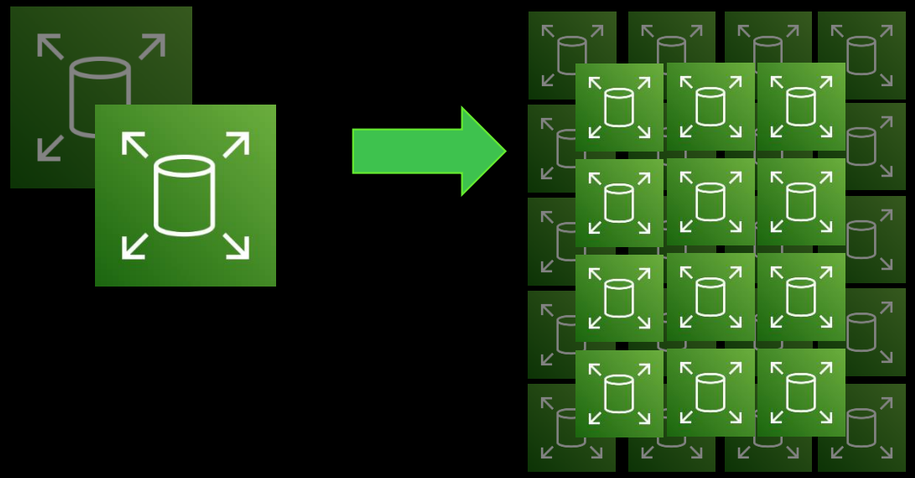
Read more: assets.amazon.science/c4/11/de2606884b63bf4d95190a3c2390/millions-of-tiny-databases.pdf

Static stability is the ability for systems to return to steady state after small disturbances



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Rethinking our volumes



- Split volumes into shards
- No longer constrained by individual storage nodes
- Faster to replicate in failure

Building a better network protocol



- Throughput limited by incast
- Difficult to offload in hardware
- Slow to recover in failure

Building a better network protocol



- Understand your environment
- Design for failure
- Recognize properties of your data stream

Nitro card for EBS



NVMe controller

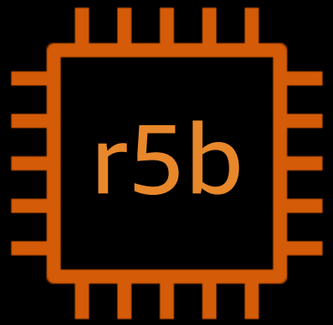
- Standard drivers broadly available

EBS data plane

- Encryption support
- NVM to remote storage protocol
- EBS optimized by default

NEW

New instance family: r5b



- Same VCPU:memory ratio as r5
- Up to 260,000 IOPS
- Up to 60 Gbps EBS optimized BW (7,500 GiB/s)
- Built on Nitro + SRD

Moving to services-based architecture improves control plane agility. Can we do the same in a high-performance data plane?



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Disaggregated data plane

- Reducing code in the fast path improves performance

  Think critically about what absolutely needs to be there!

- Smaller code is easier to reason about security and correctness
- Can move faster with defined APIs

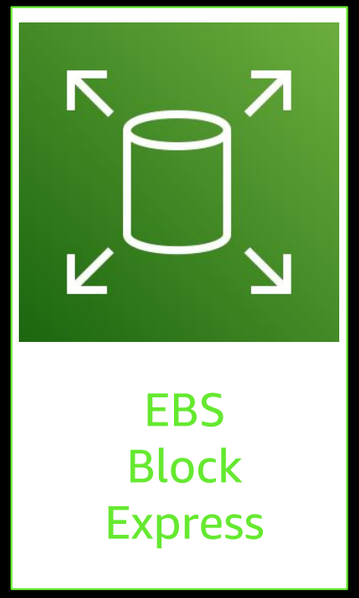
EBS Block Express



AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

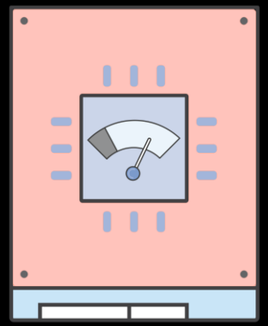
EBS Block Express



- Disaggregated infrastructure
- Disaggregated data plane and data services
- Nitro and SRD
- More consistent I/O latency

PREVIEW

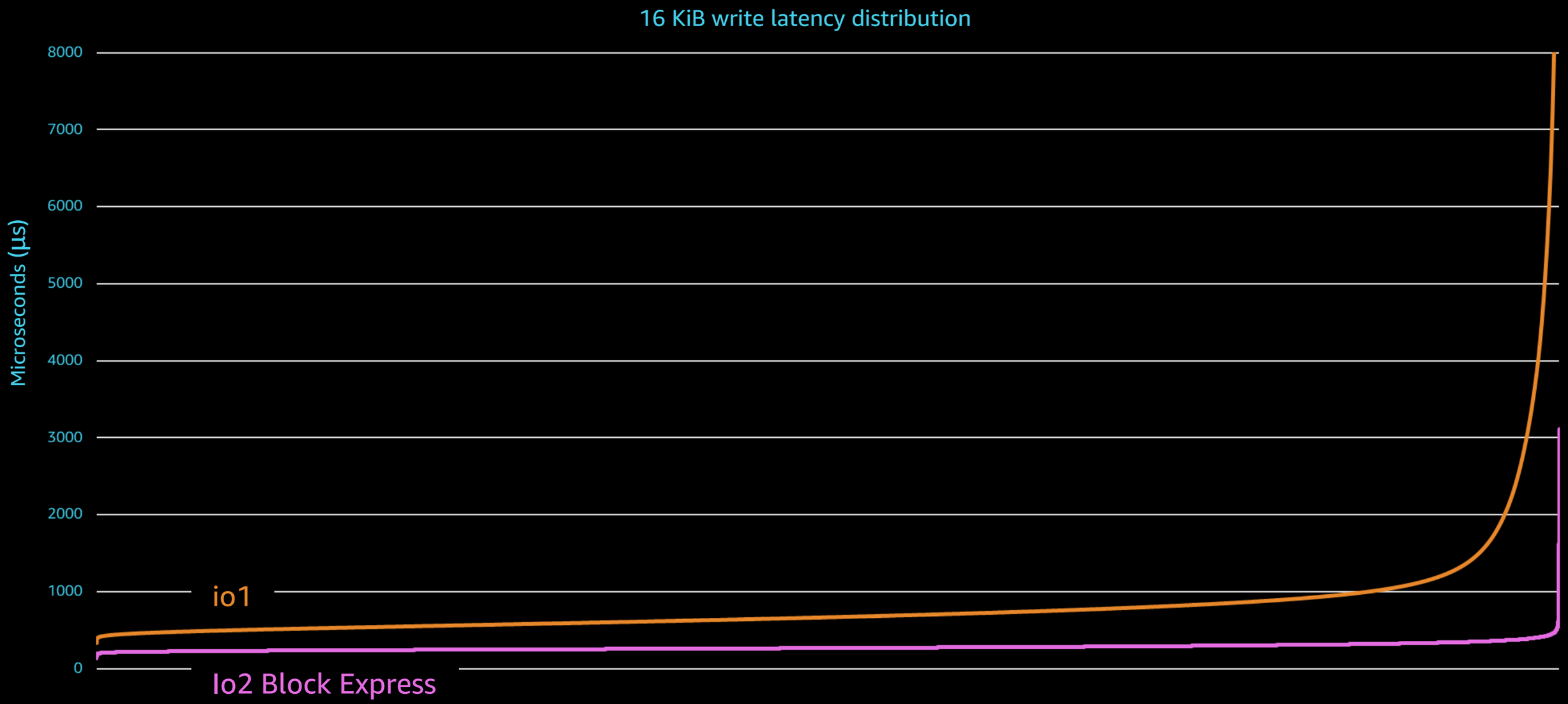
Preview: io2 Block Express



io2
Block
Express

- IOPS: up to 256,000
- Bandwidth: up to 4,000 MiB/s
- Latency: <1 ms average
- Capacity: 4 GiB–64 TiB
- Durability: 99.999%

Io2 Block Express: Performance



Application compatibility

All applications will work, including SAN workloads

Best for applications requiring low-latency I/O

Preview limitations

- Supports encrypted volumes* and snapshots
- Can only be used with the r5b instance family
- Cannot modify volume to change size, performance, or type
- Cannot attach volumes to multiple instances

* Does not currently support EBS encryption by default

What’s next?



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

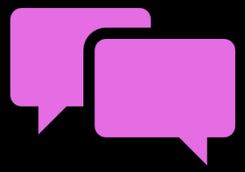
What’s next?



All EBS features



All Nitro instances



Feedback?

Thank you!

Marc Olson

Principal Software Engineer
AWS
@molson

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Please complete the session survey

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.