

Bob Yeh, Solutions Architect, Amazon Web Services

The reliability pillar focuses on the ability to prevent, and quickly recover from failures to meet business and customer demand. Key topics include foundational elements around setup, cross project requirements, recovery planning, and how we handle change.



Resources:

Website:

https://aws.amazon.com/architecture/well-architected/





Resources:

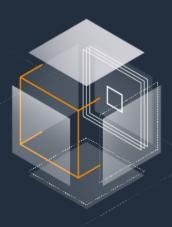
HTML Version of Performance Pillar:

https://wa.aws.amazon.com/wat.pillar.reliability.en.html



Background

How to calculate Availability? How many 9?



Availability	Max Disruption (per year)	Application Categories
99%	3 days 15 hours	Batch processing, data extraction, transfer, and load jobs
99.9%	8 hours 45 minutes	Internal tools like knowledge management, project tracking
99.95%	4 hours 22 minutes	Online commerce, point of sale
99.99%	52 minutes	Video delivery, broadcast systems
99.999%	5 minutes	ATM transactions, telecommunications systems



Background





Definitions:



Foundation - Networking

Application Design for Availability

Understand Availability Needs



Foundation - Networking

- Allow <u>IP address space</u> for <u>more than one VPC</u> per Region.
- Consider <u>cross-account connections</u>. For example, each line of business might have a unique account and VPCs. These accounts should be able to connect back to shared services.
- Within a VPC, allow space for <u>multiple subnets</u> that <u>span multiple AZ</u>.
- Always leave unused CIDR block space within a VPC.
- How are you going to be resilient to failures in your topology?
- What happens if you misconfigure something and remove connectivity?
- Will you be able to handle an unexpected increase in traffic/use of your services?
- Will you be able to absorb an attempted DoS attack?



Foundation - Networking

- Key Services for Network Topology
 - Amazon VPC
 - AWS Direct Connect
 - Amazon EC2
 - Amazon route53
 - Elastic Load Balancing
 - AWS Shield



Definitions:



Foundation - Networking

Application Design for Availability

Understand Availability Needs



Application Design for Availability

- Fault Isolation Zones
 - Multiple independent component in parallel
 - Multi-AZ
- Redundant components
- Micro-service architecture
- Recovery Oriented Computing
- Distributed systems best practices
 - Throttling
 - Retry with exponential fallback
 - Fail fast
 - Use of idempotency tokens → assume an action must occur exactly once
 - Constant work
 - Circuit breaker
 - Bi-modal behavior and static stability

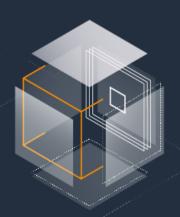


Detective Controls

- Capture and Analyze Logs:
 - Capture: CloudTrial, AWS Config
 - Store: CloudWatch Logs, S3, Glacier
 - Analyze: Elasticsearch Service, EMR, Athena
- Integrate Auditing Controls with Notification and Workflow:
 - CloudWatch, CloudWatch Events
 - AWS Config Rules
 - CloudWatch API & AWS SDKs
 - Inspector



Definitions:



Foundation - Networking

Application Design for Availability

Understand Availability Needs



Operational Consideration for Availability

- Automate Deployments to Eliminate Impact
 - Canary deployment
 - Blue-Green deployment
 - Feature toggles
 - Failure isolation zone deployments
- Testing
- Monitoring and Alarming

Generation \rightarrow Aggregation \rightarrow Real-time processing and alarming \rightarrow Storage \rightarrow Analytics

- Operational Readiness Reviews
- Auditing



- Automate Deployments to Eliminate Impact
 - AWS Code Deploy
- Testing
- Monitoring and Alarming
 - Amazon Cloudwatch
 - AWS X-Ray
 - Amazon S3
 - Amazon EMR
- Operational Readiness Reviews
- Auditing
 - Amazon Cloudwatch Logs
 - AWS Config
 - AWS CloudTrail



Q: How are you managing AWS Service Limits for your Account(s)?

Unaware

Aware but not Tracking

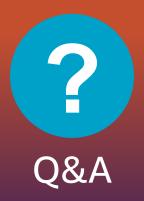
Monitor and Manage Limits

Aware of Fixed Service Limits

Sufficient Buffer in Service Limits to Accommodate for Failover

Service Limits are Considered







Thank you!

