

Amazon's Availability Zones

November 2011

A major step forward in achieving high availability in the cloud is Amazon's Availability Zones, introduced in 2008. Availability Zones allow a company to run multiple instances of its critical applications in different data centers so that the applications can survive even a data-center failure

There have been several spectacular cloud failures recently, ranging from hours to days, due to a wide variety of causes – power, storage, networks, and people. These outages cut across all cloud service providers, large and small – Amazon and Google have both contributed their share. A lesson to be learned from these outages is that the root cause of the next cloud failure is probably unimaginable.

Consequently, for critical applications that just cannot go down, it is imperative to run instances of the application in two or more geographically separated data centers so that the application survives any destructive event. To this end, Amazon operates data centers in multiple regions throughout the world. Within each region are multiple fault-isolated Availability Zones. Amazon's Availability Zones are a major move towards achieving high availability in the cloud.¹

The Amazon cloud, through its Amazon Web Services (AWS) that were introduced in 2006, provides a suite of IT infrastructure resources that can be used by businesses to the extent that they are needed. The processing, storage, and network capacity provided by the AWS services can expand and contract as the application needs change. A customer pays only for what his applications use. Amazon's retail service is now running in the Amazon cloud.

The primary AWS services from an availability viewpoint include:

- Elastic Compute Cloud (EC2) – provides resizable compute, storage, and network capacity.
- Elastic IP Addresses – static IP addresses that can be migrated between EC2 instances.
- Simple Storage Service (S3) – stores and retrieves data via a simple web interface (URLs).
- Elastic Block Store (EBS) – provides persistent storage for an EC2 instance.
- Relational Database Service (RDS) – rescalable MySQL and Oracle database support.
- Elastic Load Balancing - distributes traffic across multiple EC2 instances that can be in different Availability Zones.
- Availability Zones – fault-isolated virtual data centers within a regional data center.

In this article, we describe the AWS services that make up the Amazon cloud and how they interoperate to provide high availability of cloud-based applications.

¹ Availability Zones are not geographically dispersed, but they are fault-isolated.

Regions

Amazon operates multiple data centers in several *regions*. It currently operates data centers in six regions:

- US East (northern Virginia)
- U.S. West (northern California)
- U.S. West (Oregon)
- Europe (Ireland)
- Asia Pacific (Singapore)
- Asia Pacific (Tokyo)

In addition, it runs a special region, GovCloud, for the U.S. government. Within each region are multiple Availability Zones.

A region has an availability SLA of 99.95%, representing about four hours of downtime per year.

Elastic Compute Cloud (EC2)

The Elastic Compute Cloud (EC2) is the primary infrastructure for the Amazon cloud. It provides resizable compute facilities in the AWS cloud. Amazon operates its data centers under virtualization. It uses the Xen hypervisor to support multiple virtual machines (VMs) on a single server.

The fundamental building block of the Elastic Compute Cloud is an *EC2 instance*. An EC2 instance is an application running in the AWS cloud as a virtual machine (VM). A VM is physically implemented as an Amazon Machine Image (AMI). The AMI is the boot image for an EC2 instance. The AMI includes the applications, libraries, data, transient storage, and other configuration parameters for the EC2 instance.

An EC2 instance is created simply by the virtualization facility running the AMI boot file. At run time, the CPU size, memory size, operating system (various versions of Linux, Windows, and Solaris), and other parameters are specified for the EC2 instance. Once running, the applications provided by the EC2 instance can be accessed over the Internet.

Multiple EC2 instances can be run from the same AMI in the same or different Availability Zones or regions.

Elastic IP Addresses (EIP)

Early EC2 instances were limited by having just one IP address. Therefore, it was difficult to run multiple instances of the same AMI unless the instances used dynamic IP addressing. The introduction of Elastic IP addresses changed that.

An EIP is assigned to an account, not to an EC2 instance. An EIP address is essentially a static IP address that can be assigned to an EC2 instance. However, unlike a static address, it can easily be migrated to another instance. When an EC2 instance is created, it is given an EIP address. All traffic routed to that EIP address then goes to the instance to which the EIP address is assigned. If the EIP address is reassigned to another EC2 instance, all further traffic will be routed to the new instance. It takes about three minutes to reassign an EIP address.

EIP addressing brings a significant availability advantage to EC2 in that it provides a simple mechanism to reroute traffic to a backup EC2 instance should the primary instance fail. All that needs to be done is to issue a command to point the EIP address to the new instance that should take over processing.

EIP addressing also allows planned downtime to be eliminated. A new upgraded EC2 instance can be created with its own test IP address and exercised until it has gained the developer's confidence. The EIP address of the primary EC2 instance can then be switched to the upgraded EC2 instance and processing continues. If the upgraded EC2 instance subsequently exhibits problems, traffic can easily be switched back to the old EC2 instance.

Simple Storage System (S3)

Introduced with EC2, the Simple Storage System (S3) provides storage services that can be used to store any amount of data. Data is stored in *buckets*. A bucket contains multiple data objects, typically files. A data object's size can be anywhere from one byte to 5 GB.

A bucket can be stored in any region. It is identified by a unique URL specified by the user. Objects within the bucket are specified by user defined keys.

All versions of a data object are saved, and the user can recover any version upon request.

Access to a bucket is via the Internet over HTTP. Therefore, the contents of a data object in a bucket may not be consistent as it is written, since data may be flowing over different Internet routes. The bucket will be consistent when its writing is complete. Amazon calls this "eventual consistency."

When Availability Zones were introduced, the capability to synchronously copy S3 buckets across multiple AZs was introduced. Up to six replicas can be maintained. The copy does not complete until the bucket has been successfully copied to all specified AZs. S3 will detect corrupted copies and will repair them. With sufficient redundancy, Amazon claims a bucket durability of eleven 9s and a bucket availability of four 9s.

Elastic Block Store (EBS)

The Elastic Block Store is another major availability improvement. Up until it was introduced, an EC2 instance had only its instance storage and S3 storage. The instance storage was nonpersistent. Though it would survive a reboot, it disappeared when the EC2 instance was terminated.

S3 storage was persistent, but it had high access latency since it was accessed over the Internet. Therefore, it was not suitable for applications that required high performance.

EBS changed all that. EBS is persistent block-level storage. An EBS volume can be sized from 1 GB to 1 TB. It can be mounted by any EC2 instance (just one at a time) and can be used just as any block storage device. Typically, a file system is run in an instance that uses EBS for its underlying storage. EBS also serves as a root device for holding the AMI for an EC2 instance. In this way, an AMI does not have to be first loaded before bringing up an EC2 instance. In addition, an EC2 instance can be stopped and started with its state maintained if EBS is its boot device.

EBS is designed to be highly available. A logical EBS volume is synchronously replicated to multiple EBS physical volumes in the same AZ. One of the physical volumes is the primary volume and is the one used by the EC2 instance that has it mounted. The other volumes are used only for redundancy to ensure that data is not lost due to a volume failure. If a volume should fail (primary or backup), EBS searches the physical volumes to find one with enough space to use as a replacement redundant volume. It then makes a copy of the logical volume onto the new physical volume. This is called *remirroring*. During remirroring, the logical volume is not accessible. Applications attempting to use this volume are said to be "stuck."

An EC2 instance can mount several logical EBS volumes. For instance, it could mount multiple volumes and create a RAID 0 array by striping data across them to improve read performance.

Backup EBS logical volumes can be created in other Availability Zones and kept synchronized with the primary volume using synchronous replication. In this way, if the primary AZ should fail, a new EC2 instance can be rapidly created in another AZ and can mount the backup volume. As soon as the Elastic IP address is redirected and remirroring is completed, the applications hosted by that EC2 instance are back up and running.

EBS volumes can also be launched in multiple regions for additional protection. However, in this case, replication to the backup copies in other regions is the responsibility of the application.

Point-in-time snapshots can be taken of an EBS volume and stored in S3 storage. Snapshots are incremental and can be taken as rapidly as every five minutes. If an EBS volume is lost, it can be reconstructed from its last snapshot; and only the data that was created since the snapshot is lost. Snapshots can also be used to restore data integrity following file corruption such as the accidental deletion of a file.

When restoring a volume from an S3 image, the EC2 instance does not have to wait for the load to complete. "Lazy loading" is used. If the application requests a block that has not been loaded, the block is immediately fetched from the snapshot.

Depending upon file size, an EBS volume has an availability ranging from .995 to .999, with downtime being predominantly the remirroring time. Larger volumes have lower availabilities (the .995 availability is for a 20 GB volume).

Relational Database Service (RDS)

RDS sets up, manages, and scales a relational database in the cloud. Both MySQL and Oracle are supported. RDS automatically applies database patches and performs periodic backups of the database.

RDS uses EBS as its underlying block store. Users can take snapshots of RDS databases and store them in S3 storage in a manner similar to EBS snapshots.

An RDS database can be deployed across multiple AZs either for backup or for read replicas. If used for backup, the backup databases are kept synchronized with synchronous replication. In this case, the backup databases cannot be used for reading.

RDS databases can also be deployed in multiple AZs as read replicas. In this case, they are kept synchronized via asynchronous replication and will lag the primary database.

Elastic Load Balancing

AWS Elastic Load Balancing distributes traffic across multiple EC2 instances that can be in multiple Availability Zones. It detects unhealthy EC2 instances and ceases to route traffic to them.

Availability Zones (AZ)

Early on, applications running in a regional data center were subject to failure and data loss if the data center suffered an outage. Amazon took a major step in curing this problem with its introduction of Availability Zones (AZs). A region now can support multiple Availability Zones.

Availability Zones are distinct locations within a region that are engineered to be insulated from failures in other Availability Zones and that provide inexpensive, low latency connectivity to other Availability Zones in the same region. An Availability Zone is in effect a separate data center in a region. It has its own

power, cooling, and communication infrastructure that is separate from the other Availability Zones in its region. Therefore, a failure of one AZ will not affect another AZ except for a regional disaster.

An AZ is not necessarily in a different data center from the other AZs. It is only necessary that the data center be built to ensure the independence of the AZs.

Elastic IP addresses allow users to run backup EC2 instances in other AZs so that if an AZ fails, the backup instance can rapidly take over.

Intra/Inter Availability Zone Communications

The EC2 instances in a region communicate with their EBS storage arrays via the redundant EC2 Control Plane. All requests for creation, deletion, and read/write access between an EC2 instance and an EBS volume is carried by the Control Plane as shown in Figure 1.

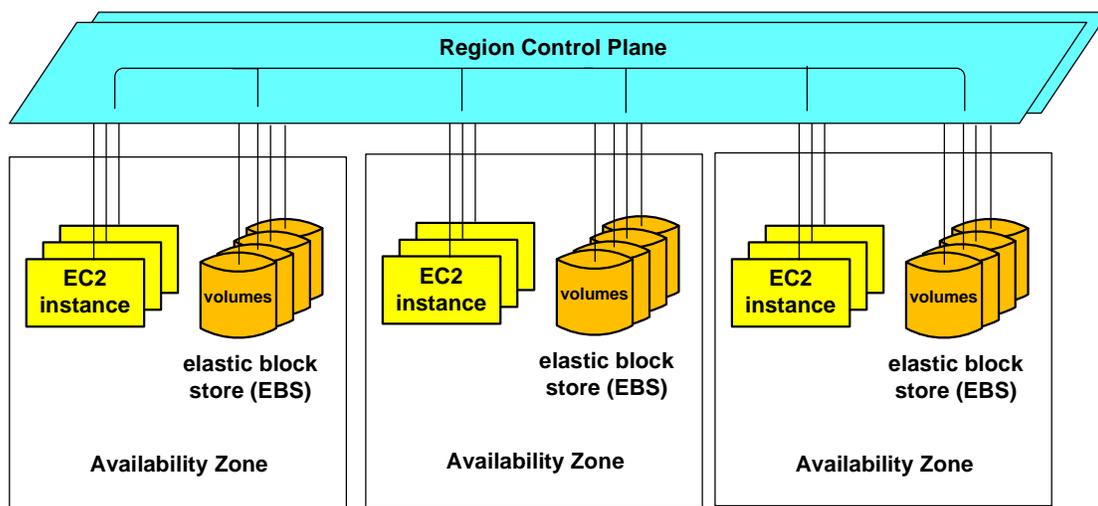


Figure 1: An Amazon EC2 Region

A recent (April, 2011) fault in Amazon's EU² region overloaded the control plane as multiple volumes tried to remirror simultaneously. Amazon plans to improve the Control Plane architecture by having an independent Control Plane in each Availability Zone to handle intra-AZ traffic, as shown in Figure 2. Inter-AZ traffic will be carried by the Regional Control Plane.

² [Lightening Downs Amazon – Not!](http://www.availabilitydigest.com/public_articles/0609/amazon_lightning_strike.pdf), *Availability Digest*, September 2011.
http://www.availabilitydigest.com/public_articles/0609/amazon_lightning_strike.pdf.

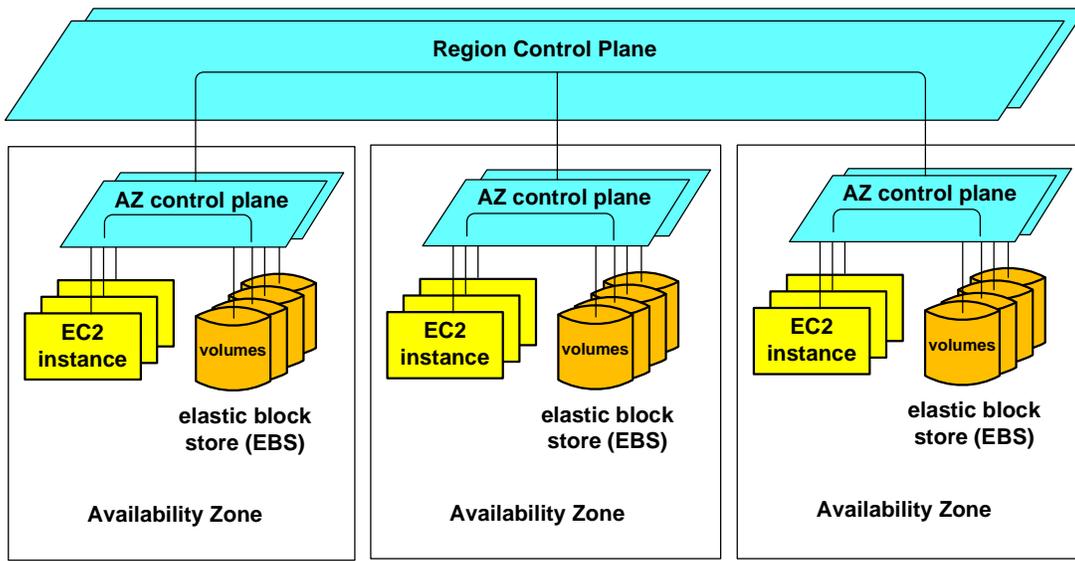


Figure 2: Upgraded Amazon EC2 Region

High-Availability Fault-Tolerant AWS Architectures

There are a variety of ways that Availability Zones can be used to achieve various levels of availability. The decision is a cost/availability tradeoff.

Basic Availability Configuration

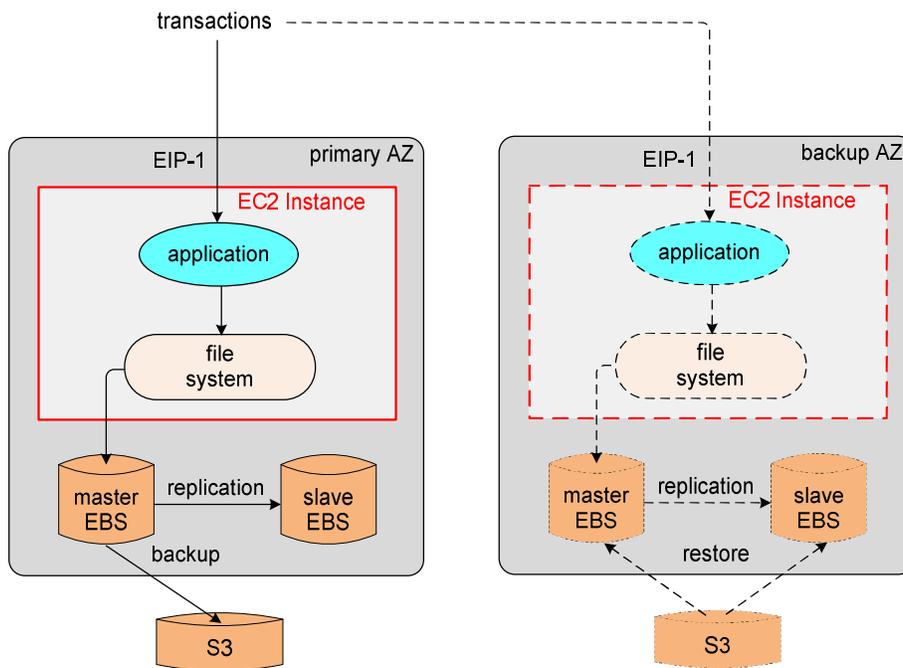


Figure 3: Basic Availability Configuration

In a basic availability configuration, the production application runs in the primary AZ. In Figure 3, the production EC2 instance comprises the application and a file system, with transactions received on the Elastic IP address EIP-1. The file system uses an EBS volume as its block store. A slave copy of the master EBS volume is configured in the same Availability Zone and is kept synchronized with the master volume via synchronous replication. A snapshot is periodically taken of the master volume and is stored in S3 storage (which itself is replicated across Availability Zones).

A backup clone is ready to be launched in a backup AZ. If the primary AZ should fail, the master and slave EBS volumes in the backup AZ are rebuilt from the last snapshot. The backup EC2 instance is then launched, and the EIP address is switched to the backup EC2 instance. Recovery time depends upon the boot size for the instance and is typically about 10 minutes plus the time to restore the master and slave databases (remember – lazy loading allows the applications to begin before database restoration is complete).

Data since the last snapshot will be lost.

Intermediate Availability Configuration

A configuration with higher availability is shown in Figure 4. Two EC2 instances operating in different Availability Zones are given different EIP addresses, and they share the transaction load. All applications use the master database via the file system in EC2 Instance 1 (remember that only one EC2 instance can mount an EBS volume). Communication between the EC2 instances is via the Control Plane.

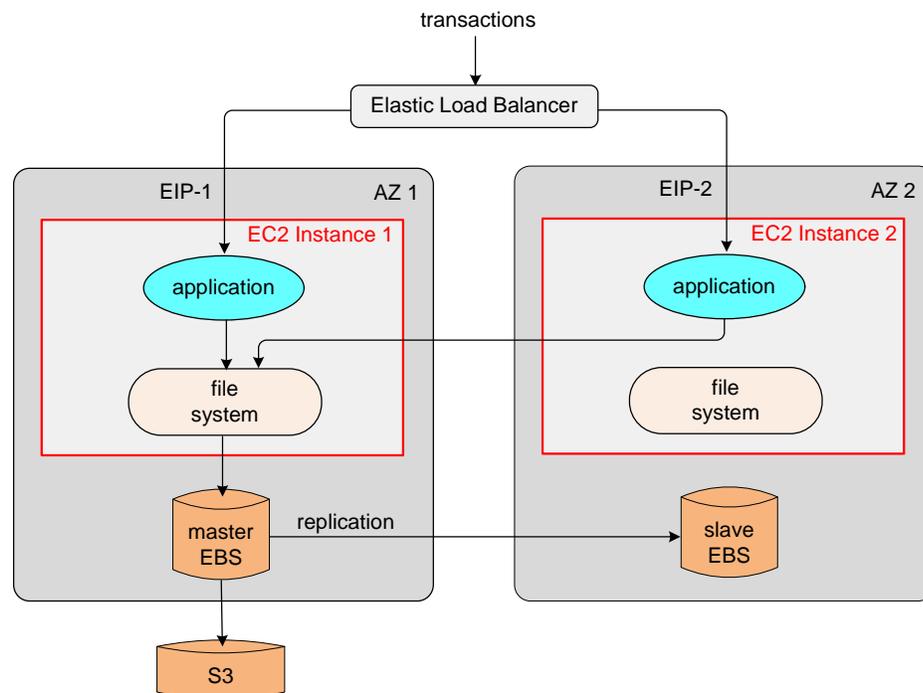


Figure 4: Intermediate Availability Configuration

The master and slave EBS volumes are located in different Availability Zones so that there is no delay in rebuilding the database nor is there any data lost should an AZ fail. If an AZ fails, all further transactions are routed to the surviving AZ. If the AZ containing the master database is the one that failed, then the slave database is promoted to master before the applications connect to it.

In this case, both EC2 instances must be configured to handle the entire transaction load. Alternatively, upon an EC2 failure, the surviving EC2 can be reconfigured to handle the additional load. This is a cost tradeoff that accepts slower response time until the surviving EC2 instance can be reconfigured.

Advanced Availability Configuration

Availability of this system can be further extended by creating another backup instance of the application in another AZ should one of the AZ's fail, as shown in Figure 5. In normal operation, the configuration of the application is that shown in Figure 4.

Assume that AZ 1 containing the master copy of the database fails. The slave database in AZ 2 is promoted to master, and the application in AZ 2 reconnects to the new master database in AZ 2. Once service has been restored, another EC2 instance is created in another Availability Zone. It is given the IP address EIP-1 to restore redundancy. The configuration has now been returned to that of Figure 4.

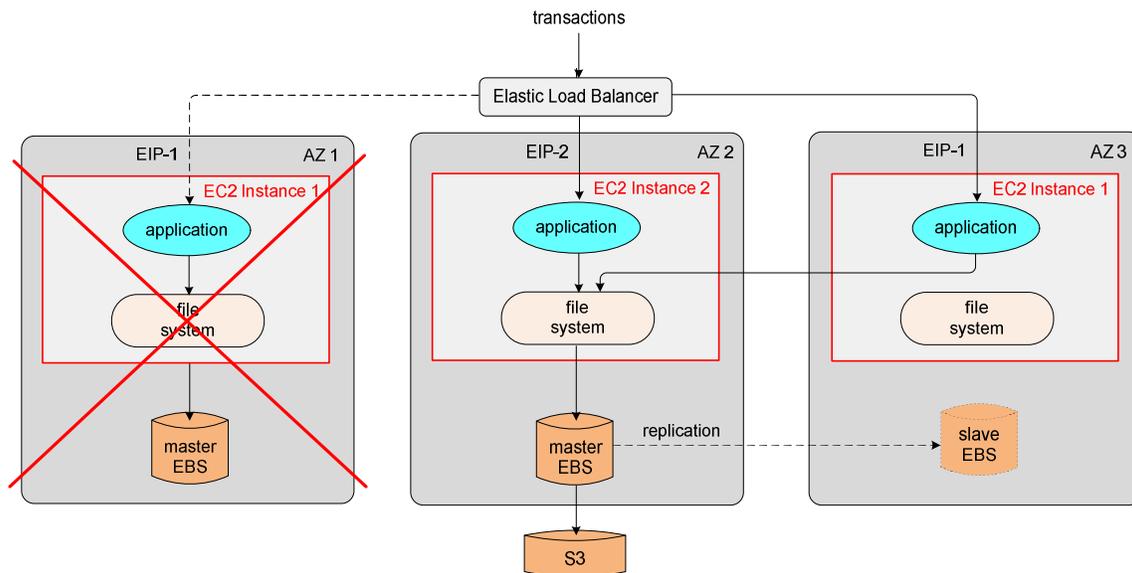


Figure 5: Advanced Availability Configuration

Summary

It is often advised not to run critical applications in the cloud. Cloud failures have become all too commonplace.

Amazon's Availability Zones provide a powerful approach to guarantee survivability of critical applications even if an entire Availability Zone should fail. Each Availability Zone is an independent data center that is fault-isolated from other Availability Zones. Application instances can be run in two or more Availability Zones either as multiple operational instances or as active/backup pairs. Should an Availability Zone fail, an instance in another Availability Zone can take over the processing of the application instance in the failed Availability Zone.

References

Amazon Web Site

What is AWS? <http://aws.amazon.com/what-is-aws/>
Amazon's Elastic Compute Cloud www.amazon.com/ec2
Amazon Elastic Block Store <http://aws.amazon.com/ebs/>
Amazon Simple Storage Service (Amazon S3) <http://aws.amazon.com/s3/>
Amazon Relational Database Service (Amazon RDS) <http://aws.amazon.com/rds/>
Announcement: Amazon EC2 Release: Introducing Elastic IP Addresses, Availability Zones, and new public AMIs/Kernels <https://forums.aws.amazon.com/ann.jspa?annID=295>

Wikipedia

Amazon Web Services
Amazon Elastic Compute Cloud
Amazon Elastic Block Store
Amazon S3
Amazon Relational Database Service

RightScale

EC2 Architecture Diagrams http://support.rightscale.com/12-Guides/EC2_Best_Practices/EC2_Site_Architecture_Diagrams
RightScale supports the new amazon EC2 Elastic IP addresses and availability zones
<http://blog.rightscale.com/2008/03/26/rightscale-supports-the-new-amazon-ec2-elastic-ip-addresses-and-availability-zones/>
Setting up a fault-tolerant site using Amazon's Availability Zones
<http://blog.rightscale.com/2008/03/26/setting-up-a-fault-tolerant-site-using-amazons-availability-zones/>
Best Practices for using Elastic IPs (EIP) and Availability Zones
[http://support.rightscale.com/09-Clouds/AWS/02-Amazon_EC2/Designing_Failover_Architectures_on_EC2/00-Best_Practices_for_using_Elastic_IPs_\(EIP\)_and_Availability_Zones](http://support.rightscale.com/09-Clouds/AWS/02-Amazon_EC2/Designing_Failover_Architectures_on_EC2/00-Best_Practices_for_using_Elastic_IPs_(EIP)_and_Availability_Zones)
DNS, Elastic IPs (EIP) and how things fit together when upgrading a server
<http://blog.rightscale.com/2008/03/26/dns-elastic-ips-and-how-things-fit-together-when-upgrading-a-server/>

Availability Digest

Amazon's Cloud Downed by Fat Finger
http://www.availabilitydigest.com/public_articles/0605/amazon_ebs.pdf
Lightning Downs Amazon – Not!
http://www.availabilitydigest.com/public_articles/0609/amazon_lightning_strike.pdf

Miscellaneous

On the Road to Highly Available EC2 Applications (Werner Vogel Amazon CTO)
http://www.allthingsdistributed.com/2008/03/on_the_road_to_highly_availabl.html
Introduction to EBS Volumes http://1201restart.com/Articles/Introduction_to_EBS_Volumes
Differences between S3 and EBS <http://www.cloudiquity.com/2009/03/differences-between-s3-and-ebs/>
Amazon Web Service EC2 – Part 1: Introduction and Availability Zones
<http://clouddb.info/2009/02/26/amazon-web-services-ec2-part-1-introduction-and-availability-zones/>
Amazon EC2 Adds Elastic IP Addresses and Availability Zones
<http://blog.adsdevshop.com/2008/03/27/amazon-ec2-adds-elastic-ip-addresses-and-availability-zones/>