Best Practices for AWS and IT Audit

●●●

Andrew Clark

Data Economist

BlockScience

BLOCKSCIENCE

About me

- B.S. in Business Administration with a concentration in Accounting, Summa Cum Laude, from University of Tennessee at Chattanooga.
- M.S. in Data Science from Southern Methodist University.
- Ph.D. Candidate in Economics at the University Reading.
- American Statistical Association Graduate Statistician (GStat), INFORMS Certified Analytics Professional (CAP) and AWS Certified Solutions Architect – Associate.
- Experienced in designing, built and deployed numerous machine learning and continuous auditing solutions using open source technologies.
- Have worked in IT Audit for two publicly traded companies, one of them a Fortune 500 financial institution.
- Working as a Data Economist creating ecosystem economic design specifications by simulating the designed ecosystem using Python-based methods. Employing mathematical engineering technologies, I create novel solutions by utilizing time-tested systems engineering practices to solve business problems.

About BlockScience

- BlockScience is an engineering, research and development, and analytics firm focused on the design and analysis of complex networks. We apply mathematical engineering technologies associated with time-tested systems engineering practices to solve business problems. Whether identifying systemic risks in company operations or guiding expansion into a new line of business, BlockScience provides thoroughly researched, mathematically engineered solutions.

Outline

- Overview of AWS and it’s key services
- Discussion of the unique risks that are present in a cloud computing environment.
- Remediation strategies for dealing with cloud computing risks
- Describe best practices for an enterprise AWS deployment.
- Understand the potential for completely flexible and scripted computing environments that AWS enables.

AWS: The current king of cloud

- AWS is the current king of cloud computing. I like to think of AWS’s infrastructure as a box of Lego bricks, that a skilled architect can assemble together to make something remarkable.
- Microsoft Azure has a solid offering as well, and may be a better option for small businesses, as it is a little easier to plug and play. AWS is the gold standard for customization however.
- Google Cloud has a respectable entry as well, although lagging between AWS and Azure in many regards, one of which is maturity

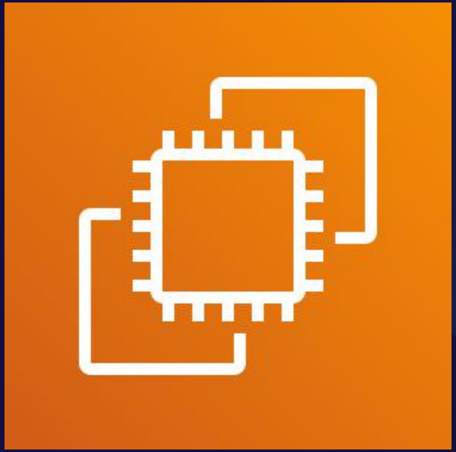
Key components and concepts of AWS

- VPCs
- Regions
- Availability Zones
- Pay as you go

BLOCKSCIENCE

Elastic Compute Cloud (EC2)

- EC2 is flexible computing in the cloud.
- It is an unmanaged service where the user provision virtual machines of the required size and accesses them through SSH.
- It takes only minutes to provision and access an EC2 instance to deploy an application. It is possible to run an EC2 instance as your main work computer, when using something like a Google Chromebook to access it.



Elastic Block Store (EBS)

- EBS is the hard drive for your EC2.
- It can be solid state drive (SSD) or hard disk drives (SDD) and range in sizes from 4GB to 16 TB https://aws.amazon.com/ebs/features/?nc=sn&loc=1

Simple Storage Service (S3)

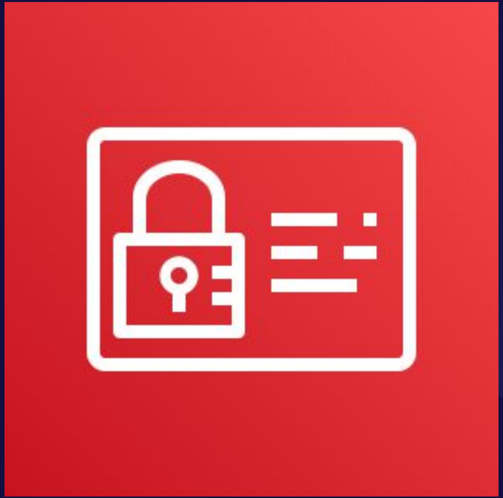
- S3 is an object based storage service.
  - You can think of it as Dropbox (the first 8 years of Dropbox’s existence it was essentially a wrapper around AWS S3 - https://www.wired.com/2016/03/epic-story-dropboxs-exodus-amazon-cloud-empire/
- S3 comes in three tiers (S3, S3-I, and Glacier), and can store data between 0 bytes and 5 Tbs in an single bucket -https://aws.amazon.com/s3/faqs/

Relational Database Service (RDS)

- Relational database in the cloud.
- Available in many different flavors, such as SQL Server, Oracle, MySQL, PostgresSQL, MarioDB, and Amazon Aurora.
- Fully managed service, meaning you only have to worry about the database maintenance, not running the underlying server.
- Note: You can run a database off of an EC2, there is no reason you must use RDS. - https://aws.amazon.com/rds/faqs/

Identity and Access Management (IAM)

- One of the key components of the AWS.
- The ‘logical access engine’
- Create users, roles, etc - https://aws.amazon.com/iam/?nc2=h_m1



Virtual Private Cloud (VPC)

- One of the key components of the AWS.
- How ‘your environment in the cloud’ is constructed.
- Define security groups specifying who can access which resources.
- Define networking, availability joins, etc. - https://aws.amazon.com/vpc/?nc2=h_m1

Route 53

- DNS service
- Scalable
- Flexible options for directing traffic to reduce latency -

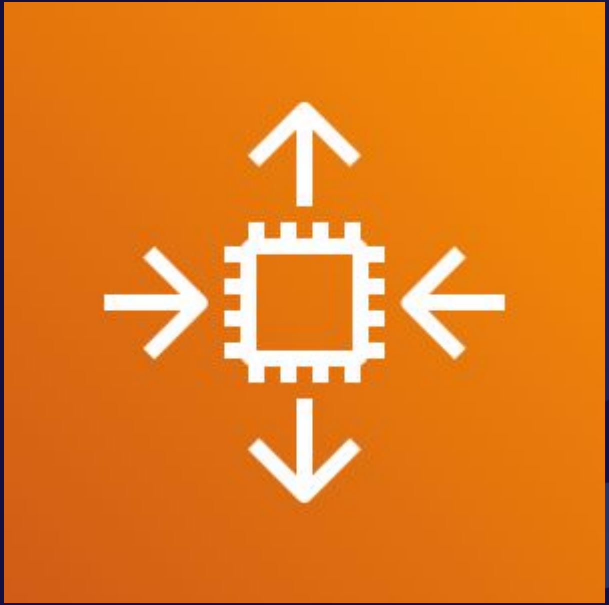
https://aws.amazon.com/route53/?nc2=h_m1

Elastic Load Balancing

- Load balancing service that can handle multiple targets across multiple availability zones.
- Secure
- Elastic - https://aws.amazon.com/elasticloadbalancing/?nc2=h_m1

EC2 AutoScaling

- Provides the ability to add or remove EC2 instances based on defined user criteria. I.e, if CPU usage goes above 70%, add another EC2 (and an Elastic Load Balancer can balance the traffic out between the instances)
- Scheduled scaling - https://aws.amazon.com/ec2/autoscaling/?nc2=h_m1



CloudFormation

- Infrastructure as code.
- Provides flexibility not traditionally possible
- Potential for increase in security
- Don’t patch, update the latest operating system versions in your CloudFormation script and rebuild your infrastructure. “Rehydration”
- Similar to Chef, Ansible, etc. -

https://aws.amazon.com/cloudformation/?nc2=h_m1

Honorable mentions:

- Lambda
- Fargate
- ECS
- Redshift
- There are hundreds of additional services, it is extremely overwhelming to try and stay on top of.

Security Considerations

- Confidentiality
- Access control - root accounts
- Resilience / availability
- Compliance
- Security
- Vendor lock-in
- Insufficient visibility

Security Considerations Cont.

- Inadvertently exposing data - S3 bucket moved to public
- Access control, access control, access control. Root accounts should be very limited, and use MFA when necessary. Strong passwords, should have MFA for all. Least privileged access, use IAM religiously. Limit IP address in security groups.
- Key management - Rotate or change keys every 90 days or so.
- For some of the AWS managed services, Machine Learning tools, etc, Amazon has some fine print about the right to view your data, so be care which services you use, if using managed.
- Use CloudTrail for audit history

Best Practices (selected items)

- Create a single “pane of glass” for viewing your environment.
- Use tagging religiously
- Keep an inventory of all instances (CloudWatch and Config, or custom)
- Rehydrate every 60 days
- Conduct penetration tests, both social engineering and technical
- Deploy to multiple regions and availability zones, with load balancing for failover.
- Conduct tests of removing services in one region and see if the system is resilient enough to withstand it. i.e., simulating a region going offline.

BLOCKSCIENCE

CloudFormation and the scripted environment

- Follow this tutorial and resources to see the power of CloudFormation and automated infrastructure:
- https://github.com/aclarkData/NACACS-2019/blob/master/README.md

To learn more about AWS:

- https://github.com/aclarkData/NACACS-2019
- https://aws.amazon.com/
- A Cloud Guru – Paid membership, but the best on the market I’ve found

Questions?

BLOCKSCIENCE

Thank you!

BLOCKSCIENCE

Andrew Clark

Data Economist

BlockScience

LinkedIn: https://www.linkedin.com/in/andrew-clark-b326b767/

Email: andrewtaylorclark@gmail.com

Website: https://aclarkdata.github.io/