Things I got wrong:

Monday, July 13, 2020 10:10 PM

- EC2 cloudwatch detailed monitoring enables 1 rate per minute metrics
- High resolution custom metrics have a min resol of 1 second
- SSO scales to adding accounts in the future
- ECR (elastic container registry) stores images, ECS (elastic container service) deploys them
- CodeDeploy helps deploy code to EC2 instances.
- Gp2 is for burst, io2 is consistent iops, yes you can set iops rate within a range based on disk size.
- ASG launch config is immutable, so a new one needs to be changed
- SSE-C allows client to provide encryption key
- **EMR is big data platform for spark, hive, hbase, etc, petabase scale analytics on hadoop.**
- SSM doesn't rotate keys, I think.. Secrets manager does though.
- IAM auth works with MySQL and PostgreSQL (RDS)
- RDS changes cname to point at standby, so failover is automated.
- NLB gives 2 public ips static
- To encrypt an unencrypted RDS database snapshot, restore database from snapshot as encrypted, and terminate previous database.
- EFS IA allows for mounting, S3 intelligent tiering does not.
- SQS FIFO needs group id if you want multiple consumers.
- **ASG launch template supports mix of on-demand and spot, unlike ASG launch config that doesn't do the mix and spot fleet.**
- If users can attach and detach policies (have access to AWS Managed Policies), you need to specify IAM permission boundary that restricts managed policies they can attach to themselves.
- S3 and DynamoDB are gateway endpoints, everything else is interface.
- Use snowmobile for 10PB or more, up to 100TB each.
- $.1 per gig PROVISIONED on EBS, $.023 per gig on S3 standard, $.3 for EFS
- SQS FIFO batching allows up to 3000 messages per second, otherwise its 300 without batching.
- Kinesis shards need provisioning so it's not truly serverless and automated scaling.
- For S3 lifecycle rules, you cannot go from bottom of waterfall (deep archive) to top (s3 standard)
- ASG: step scaling and simple is for manual CW alarms (cannot do CPU), target tracking for auto CW alarms for CPU and other metrics).
- Redshift can query data from files in S3 without having to load data into it like athena.
- Database upgrade for RDS causes both secondary and primary upgraded at same time, downtime until complete during maintenance window
- Fargate is charged differently than ECS, it charges by vcpu and memory resources instead of EC2 launch type.
- Lifecycle hooks are for putting ASG instances in wait state, scheduled action is for time based actions.
- EFS can be used with an inter-region VPC peering connection for other regions and on prem
- **AWS global accelerator provides static ips that act as fixed entry points for your app endpoints in single or multiple AWS regions for ALB, NLB, and ec2 instances, but not S3**.
- S3 transfer accelerator is free if transfers are not accelerated.
- ASG standby mode is for troubleshooting instances or upgrades and does not cause an increment in quantity.
- For Aurora failover, RDS will promote read replicate with highest priority (lowest tier), and then if same tier the largest size.
- enableDnsSupport and enableDnsHostnames for internal custom domains.
- Go, C#/NET, Java, NodeJS python and ruby are supported by lambda. R is not, nor is php. Think high level programming languages.
- ASG first terminates unhealthy then launches a new instance to replace terminated.

- HDDs cannot be used as boot, so ST1 and SC1. Instance store can.
- Deleting CMK in KMS has a waiting period, from 7 - 30 days where it pends deletion.
- Cloudfront signed urls and signed cookies control content, but https is not needed.
- Io1 is a iops SSD, good for database workloads.
- Use RDS multi-az for availability, read replicas for scaling
- Long polling allows a server to wait to get a message. Visibility timeout causes duplicate messages if set too short, and long delays for retries if too long. However, a config issue can cause it too.
- The two formats for s3 website endpoints is bucket-name.s3-website-Region.amazonaws.com and the same thing but s3-website.Region
- Fifo SQS is not allowed as S3 event notification destination
- Centrallized storage = cannot make multiple buckets.
- Cloudfront and s3 transfer acceleration can speed up uploads and downloads internationally
- Kinesis agent cannot write to kinesis firehose for which the data stream source is already set as kinesis data streams.
- Cloudwatch alarm is a thing, and alarm action is a thing too, especially to reboot instances
- For elastic ip to be attached to ec2, you need access to api calls.
- Vpc with public subnet only and site-to-site vpn access is not available with vpc console wizard
- Cloudwatch recovery process is identical to original instance, except memory, even ipv4 public address.
- Cloudfront only does HTTP/RTMP protocols and not UDP. Global Accelerator does UDP, IoT(MQTT) or VoIP.
- SGs can be associated with a NAT instance but not NAT gateway
- Delay queues for delaying when components need more time to process (consumers)
- SCPs affect all users and roles including root user, but they do not affect service-linked roles.
- Basic monitoring is enabled when you use AWS console, but detailed is enabled when you use CLI to launch config..
- If you need to delay certain messages, use message timers.
- Step functions are useful when you have multiple functions and need something to manage end to end workflow
- AD connector will not allow you to run directory-aware workloads in AWS, only to login to AWS apps and services.
- Kinesis data streams store data for up to 7 days and can keep ordered data.
- S3 glacier supports encryption by default for data at rest and in transit.
- Tenancy.. You can change it from dedicated instance to host, and dedicated host to instance.
- EBS to instance traffic is encrypted.
- Scale-up for vertical scalability, scale-out for horizontal
- EFA is better than ENA, oops.
- Route53 outbound for forwarding queries to resolves on prem, inbound for vice versa.
- Vpc sharing is to share subnets, not vpcs.
- AWS Glue takes a lot of development, and is really meant for batch. Use aws data migration service for easy migrations.
- Spot blocks are designed to not be interrupted, but rarely can be due to lack of resources
- For launch configs, vpc tenancy set to dedicated overrides pretty much anything, same with launch config set to dedicated.
- "server bound licenses" are best used with dedicated hosts.
- Glue is for loading data for analytics, not for database migrations. (to redshift)
- AWS schema conversion tool is good for converting schemas.
- Use ALB to route based on port, not route 53.
- NLB doesn’t do outbound ips, right.
- You can assign IAM policies to ECS tasks using TaskRoleARN parameter
- Relational databases don't store session data well.
- RDS multi-az is not good for improving query performance, only for failover.
- The default SCP is FullAWSAccess which grants Allow of * on *.

- Amazon FSx may refer to either windows or lustre.
- Glacier deep archive has a min storage of 180 days. 12-48 hr retrieval. Glacier is 1 min to 12 hrs
- SNS can invoke lambda function, SQS cannot. (SQS consumers pull, but SNS pushes).
- Use dynamodb for active-active. Aurora global does not do active-active replication in multiple regions.
- Cloudfront can do both dynamic and static
- EFS connects to on prem!
- Privatelink does not use SGs
- Redshift can run complex queries in addition to analytical ones
- Cloudfront can do custom error pages
- S3 transfer acceleration only works for uploads
- Parallel requests and byte ranges can help download files faster
- To block all non-cloudfront traffic you would need to use AWS lambda to update ips of a secruity group for the ELB
- Cross-region replication is S3, not ec2
- Api gateway decouples ui from underlying services
- Rds snapshots in different regions can use different keys
- Aurora standbys can be used as read replica
- Pre-signed urls can be used to upload and download
- S3 is object based, efs is file based
- Use encryption helpers to encrypt things in lambda
- For direct connect you use VPG to setup encryption, not IPSec
- Firehose is for transformation not processing of data - that's for kinesis streams.
- Hdd throughput optimized is cheaper than gp2 good for small requests large data
- Use placement groups to put things in clusters, etc.
- GA is better at reducing latency than route 53 geo routing.
- Cloudwatch logs can send notification of alerts utilizing sns.
- Internet Gateway has no restriction on bandwidth, and is redundant. NAT gateway is 45GBPS limit.
- You can use MFA with cognito user pool
- Cloudformation can make subnets, and it can leverage the aws organizations api to make accounts
- DynamoDB is a key-value database (noSQL) non-relational.
- **SQS cannot invoke things!!! REMEMBER THIS**
- SWF (Simple Workflow Service) coordinates tasks across components
- You can migrate public ips to global accelerator
- PTR are reverse dns records
- Dynamodb streams will capture item-level modifications
- 500 iops max for ST1 (throughput optimized hdd)
- Suspend ASG processes instead of disabling them.
- AWS Batch multi-node parallel jobs let you run jobs that span multiple EC2 instances. Good for MPI, apache MXNet, tensorflow, and caffe2 HPC related comms.
- NAT Gateways are highly avaialbe within each AZ, whatever that means.
- Nat gateways cannot be assigned SGs.
- Opworks stacks is for on prem and aws, opworks for chef automate is just within aws
- Tolerate delays means you are able to let your job fail a few times with spot instances.
- Grant programmatic access to apps using iam policy, not bucket policy.
- EMR processess data, uses hadoop, good for analyzing logs.
- Identity pools in cogniity is used to grant temp creds to AWS services
- Instance can terminate if you reach your ebs limit, or snapshot is corrupt, or lack of permission for key, or missing part of ami file.
- EBS Volume with CloudFormation needs logical ids to provision.
- BGP Prefix advertising both vpn and direct connect is cheaper than using transit gateway
- Query API and keys are used to manage IAM over https.
- NAT gateways are for outbound connection, not inbound.

- Connection draining is for ELBs and not RDS.

- Only use dedicated host when there are hardware requirements
- For lambda to access inner-vpc resources, you need to provide the subnet id and security group ids
- Iam Groups
    - Collections of users with policies attached
    - Not an identity and cannot be idenified as principal in iam policy
    - Use groups to assign permissions to users
    - Iam groups cannot be used to group ec2 instances
    - Only users and services can assume a role to take on permissions, groups cannot.
- critical for pilot light vs core for warm standby
- Dynamodb best practices
    - Keep items small
    - Use seeperate tabes for serial data
    - Store frequenty and less frequently accessed data in speerate tables
    - Compress larger attribute values
    - Store larger than 400kb objects in s3 and use pointers
- Transit virtual interface is used to access vpc transit gateways
- RDS multi-az uses synchronous replication, but read replicas use asynch.
- Cloudwatch AWS Lambda tracks latency per request and total number of requests.
- Elasticache can help with rds read performance issues
- **FIFO queues ensure messages are delivered only once**
- VPN CloudHub is for vpn so it is slower than a direct connect gateway with VIFs.
- RDS public access requires SG to allow access and be put into a public subnet, with the option to allow RDS to be public accessible
- Route 53 A records can be used to map resources including ALB, API Gateways, CloudFront distrubitons, Beanstalk, S3 buckets, VPC Interface. Cannot do public ip addresses for on prem
- COPY command allows copying of data in databases.
- Mysql cant use STS, but it can use the AWS Auth plugin with IAM
- ElastiCache with MemCached allows multiple CPU cores or threads, redis does not support that.
- IAM policy to apply folder level permissions and create IAM Group to attach policy for allowing access for users to their own directories in S3.
- Step functions are for serverless workflows, cloudformation is for describing and provisioning.
- NAT gateway still needs internet gateway to function. Duh.
- ElastiCache has in-transit encryption as an optional feature.
- Instance metadata query tool is a thing, lol. EC2 Config performs tasks when you start or stop instances.

- 5 elastic ips per region, better to use bastion host
- Cloudfront is not region specific
- Bulk is lowest cost to retrieve from glacier
- Replatform = core is the same some components change
- Lambda cannot listen on ports
- Provisioned capacity
- Immutable and rolling with additional batch will update app with full capacity via elastic beanstalk
- Origin servers need to be publicly accessible
- Can't make VPC with elastic beanstalk
- ACL are processed in order
- Lambda can access internet by default
- Cloudwatrch supports custom metrics that work with alarms
- Lambda is not instant on and may have some startup delay
- Autoscaling automatically replaces unheahlthy instances

- Lambda allows you to choose amount of ram allocated
- ELB requires private ip not public
- Redshift is 1/10 of cost of other warehouse tech but not good at very large number of read or write requests, think analytics
- Route 53 uses TTL settings configured in service that alias record points to
- Dynamo db capacity reservation is for regions
- Sqs can hold for 14 days
- Vpc flow log for audting inbound and outband traffic in vpc
- Repurchasing is to move from on-prem software stack to SaaS

- SNS does not have DLQ. Lambda does have dead letter config tho
- Infrastructure event mgmt is a capabiltiy AWS offers.
- Spot instances that are terminated within the hour by AWS is not charged to customer
- Eventual consistency does not double write throughput
- Cloud formation does not support auto checks of acconut limits.
- Don't confused reserved with scheduled instances
- ECS Task Role limits permissions of containers running tasks. Task execution role is fargate specific, similar to ECS Instance role which is granted to all containers.
- Metadata can query user data, who knew.
- SNS -> SQS fan out for multiple subscribers to poll when needed
- Cloudwatch only aggregates metrics within a region
- Automated backups are deleted when DB instance is deleted
- When daily backup is enabled, DB trxn logs are also backed up more often
- SNS is meant as a message pusher not buffer, kinessis streams is for buffering
- Lambda can attach ENI in vpc to talk, also need NAT to work
- Cross region replication needs versioning enabled for S3
- When RDS automatic backup is configured, you can initiate a point-in-time restore and specify any second during your retention period, up to the Latest Restorable Time
- --------
- **Key-value pairs is nosql like dynamodb.**
- MapReduce, Kafka, log processing, data warehouse, and ETL workloads for HDD.
- 2xlarge gives more cpu than just large, even if its one tier lower.
- When using ELB use elb health checks instead of instance ones for ASG.
- New SG has inbound deny and outbound allow all
- EFS allows for posix permissions or EFS SGs
- S3 select helps analyze and process data within s3 buckets.
- S3 with elasticache is good for high reads and consistent throughput
- Run command to launch commands on windows
- Custom NACL has inbound rule deny all, outbound rule deny all. Default is opposite
- Put vpc flow log on network interfaces not subnets, it's more secure?
- AWS batch is for monitoring and managing batch computing jobs.