**VPC/ Network 관련**

VPN: 네트워크와 네트워크 간 보안 연결을 설정하는 서비스이다.

Site to Site VPN: VPC와 On-Prem 연결해주는 네트워크이다. Site to Site VPN이라고 하며 IPsec(인터넷 프로토콜 보안)을 따른다.

Client VPN: 사용자 로컬에서 On-Prem이나 AWS 리소스로 액세스할 수 있도록 연결한다.

PrivateLink: 인터넷을 통과하지 않고 VPC에 연결할 수 있는 VPC endpoint를 제공한다. 이 endoint는 다른 VPC, AWS 서비스, On-Prem등에 연결될 수 있다.

VPN vs PrivateLink: VPN은 비공개 연결 과정을 의미하고, Privatelink는 Endpoint를 열어주는 관점이다. 원래 On-Prem에 연결하는 PrivateLink는 DC만 가능했는데, 이제는 VPN까지 지원된다.

NAT Gateway: Private Subnet의 인스턴스가 VPC외부의 서비스에 연결할 수 있지만, 외부 서비스가 인스턴스에 연결할 수 없게, 단방향적인 연결을 위해 사용하는 게이트웨이다. 디폴트로 Public subnet에 위치시키며, 인터넷 뿐 아니라 다른 VPC나 On-Prem네트워크에도 연결할 수 있다. NGW로 연결하는 순간 나가는 IP는 더 이상 인스턴스의 IP가 아니라 NGW의 elastic IP가 된다.

Bastion: EC2 연결하기 전, primary access point가 되는 special purpose server이다. SSH와 RDP연결을 지원하고, 주로 public subnet에 위치한다. Jump server로서의 역할을 한다. NGW는 Bastion host가 될 수 없다.

Virtual Private Gateway(VGW): DC를 각 VPC에 연결하는 게이트웨이이다. DC Private Virtual Interace(VIF)를 사용한다. 이와 비슷하지만, Direct Connect Gateway는 DC를 multi Region의 VPC들에 연결할 때, Transit Gateway는 DGW와 동일하지만 VPN 연결까지 가능하다.

NLB: 3,4계층 UDP, TCP 기반 부하분산 서비스이다. Static IP를 가지며, Within Region이다. GA와 잘 맞는다. CloudFront와 사용되지 않는다.

ALB: 7계층 HTTP/HTTPS 기반 부하분산 서비스이다. 가변적인 IP를 가지며, within Region이다. CloudFront와 잘 맞는다. Static IP가 필요한 경우 GA와 사용될 수 있다. (static ip를 위해 NLB를 거치기도 한다.)

TLS termination: 프론트 connection 끝내고 온 리퀘스트를 decrypt 하는 과정. 여기서부터 백엔드로 넘어간다고 보면 된다. 예전엔 ALB만 지원했는데, 이제는 NLB도 지원한다.

A record vs CNAME vs Alias:

A record는 도메인 주소를 ip로 연결하는 레코드이다. 한 번의 요청으로 원하는 IP로 찾아갈 수 있음은 장점이지만, IP가 자주 바뀌는 환경이라면 상당히 귀찮아진다.

CNAME은 도메인 주소를 또 다른 도메인 주소로 연결하는 별명 레코드이다. 한번에 두 단계를 거쳐 IP를 찾아가는 것은 불편하지만, IP가 자주 바뀌는 환경이라면 맵핑 하나만 바꾸면 되니 편하다. 같은 이름의 record가 없어야 한다. 절대 root domain name으로 쓰여서는 안된다.

Alias는 CNAME과 동일하지만 같은 이름 중복될 수 있으며, Zone Apex로 쓰일 수 있다.

Route53 active-active vs active-passive failover: active active는 완전히 동등한 수준(name, type, 둘다 고르게 사용)의 두 resource에 대해 failover되는 것이다. Active passive는 하나는 primary, 하나는 secondary(standby)로 두고 primary에서 standby로 failover되는 것이다.

**EC2, Storage, DB 관련**

Dedicated Host: 내가 쓰고싶은대로 쓰게 해주는 physical server를 제공한다. EC2에서 ms나 oracle같은 3rd party의 라이선스를 “소켓/코어”단위로 사용하는데 사용된다. 이 경우 AWS License Manager와 찰떡궁합이다. Dedicated instance보다 좀 더 physical server환경의 관리 권한을 부여받는다고 볼 수 있다.

EFS vs FSX: 리소스 공유하는 파일시스템. 기본적으로 Concurrent한 공유는 EFS이다. 거기에 추가적으로 윈도우나 MS, SMB 강조하면 FSX windows이고, HPC 강조하면 FSX Lustre이다. + S3에 곧바로 integrated되는 것은 FSX이다. EFS는 그런 기능 없다.

Storage Gateway vs Datasync: 둘 다 On-Prem에서 Cloud로 데이터를 옮기는 서비스이다. Storage Gateway는 hybrid한 성격이다. Backup 중, 혹은 후에 migrated data에 On-Prem에 low latency로 접근할 수 있다.

Elasticache가 S3에 쓰일 수 있나? Elasticache는 실시간 transaction, 게임 리더보드에 자주 쓰이는 in-memory 캐싱 기능이다. 특정 데이터나 DB앞에서의 캐싱 용도이지, S3와는 같이 쓰일 수 없다.

DynamoDB secondary index: NoSQL은 관계가 없으니까 데이터 찾기에 시간이 오래 걸린다. 그걸 효율적으로 찾을 수 있도록 사용하는 검색 지표가 secondary index이다. local과 global은 단순히 secondary index에서 hash key 사용해서 찾는 방식의 차이인데, 이것까진 몰라도 될 듯하다.

RAID: 디스크를 얼마나 배열로 분배하는지. EBS에서 세부 설정으로 켤 수 있으며, RAID 0은 performance용(IOPS증가), RAID1은 Fault Tolerant용으로 사용하기 좋다.

Snowball: in-transit encrypted. server-side encryption 가능하며 Client 단위에서 알아서 encrypt 해준다.

Storage Gateway: in-transit encrypted. SSE-S3 방식으로, SSL/TLS 로 encrypt된다. option으로 KMS CMK사용할 수 있는데, 이때는 무조건 symmetric CMK만 지원한다. Asymmetric은 지원하지 않는다.

DataLake vs DataWarehouse: DataLake는 미리 정해진 schema가 없고, NoSQL, RDS 등등 가리지 않고 모든 데이터를 수집한다. 주로 DataScience 용도이다. DataWarehouse는 미리 Schema를 정하며, 주로 RDS나 operational DB가 대상이다. Business analysis 용도이다.

**S3 Storage 비교**

(1) 최소 보관 기간 / Retrieval time

Standard & RRS 0 days immediately

Standard IA 30 days immediately

Intelligent 0 days n/a

One-Zone IA 30 days immediately

Glacier 90 days 5분~n시간, Expedited 기준 5분 이내.

Glacier Deep Archive 180 days 12시간 이내

(2) 이동 시 주의할 점

(1)의 최소 보관 기간보다 빠르게 lifecycle에서 transition을 설정할 수 없다.

{ Standard, Standard IA } → {Standard IA, OneZone IA} 무조건 Standard에서 30일 지나야 한다.

(3) 이동 가능 여부 = (1)에서 아래로

Standard & RRS → Standard IA, Intelligent, OneZone IA, Glacier, Glacier Deep Archive

Standard IA → Intelligent, OneZone IA, Glacier, Glacier Deep Archive

Intelligent → OneZone IA, Glacier, Glacier Deep Archive

OneZone IA → Glacier, Glacier Deep Archive

Glacier → Glacier Deep Archive

(4) Storage Gateway에서 올 때

File Gateway → S3에 저장 후, Lifecycle 설정을 통해 Any storage class로 갈 수 있다.

Cached & Volume Gateway → S3에 EBS snapshot 형식으로 백업한다. 이를 읽으려면 snapshot을 복구해야 한다.

Tape Gateway → 특별히 Glacier, Glacier Deep Archive에 바로 저장 지원한다. 다만 읽으려면 Storage GW api 사용해서 해당 storage의 retrieve 시간 따라야 한다.

정리하면, File Gateway만 lifecycle 설정 통해 S3 storage에서 바로 읽을 수 있고, 나머지는 snapshot형태로 backup이 저장되는 방식이다.

**Application 관련**

Kinesis DS vs Firehose: 기능만으로 완벽히 구분하기는 어렵다. 다만 Datastream은 후에 analysis 관련 서비스(Kinesis Data Analytics, EMR, EC2, Lambda)로, Firehose는 후에 저장이나 search 관련(S3, Redshift, Elastisearch, Splunk) 등으로 연결되는 경향이 있다.

Throttling: 특정 기간동안 유저가 보낼 수 있는 리퀘 량 제한하는 것이다. API에 쓰면 API 리퀘량 제한한다.

SaaS: 클라우드를 통해 제공되는 SW. Slack, ms365 등이 해당된다.

SQS visibility timeout: SQS는 스스로 메세지를 지우지 못한다. Consumer가 지우라고 요청해야한다. 처음 Consumer가 받은뒤, 처리하는 동안 다른 Consumer가 가져가지 못하게 막아두는 시간을 Visibility timeout이라 한다. 디폴트 30초, 최대 12시간까지 설정 가능하다. 만약 이 timeout이 너무 짧으면 삭제가 되지 않아 계속 같은 작업만 반복될 수도 있고, 아예 처리 자체가 안될 수도 있다. ChangeMsgVisibility를 통해 직접 Consumer가 작업 못끝냈으면 추가시간 요청하게 할 수도 있다.

Stateless two-tier application: 단순히 과거의 요청 정보, 트랜젝션 정보가 저장되지 않는 two-tier app

RedShift Spectrum vs Athena: 둘 다 Serverless S3 query 서비스이다. Spectrum이 성능 면에서는 Athena를 능가할 여지가 많다. Redshift 이미 사용중이면 Spectrum, S3만 사용중이고 historical이라면 Athena쓴다.

EMR: 빅 데이터 플랫폼이다. 각종 오픈소스 프레임웤 생성용이지, 이 자체로 instance나 storage는 아니다.

QuickSight vs Kibana: 둘 다 Visualization tool이며, 기능 자체로는 구분 안된다. QuickSight는 Redshift나 S3, Athena와 같이 쓰는 경향 있고, Kibanan는 Elaticsearch와 같이 쓰는 경향 있다.

Rolling update deploy: 하나의 인스턴스가 업데이트중이어도 서비스가 원활(HA보장)하게 이루어지는 배포 방식이다. 해당 방식은 [HA를 위해 업데이트중인 인스턴스 하나 + Fail한 인스턴스 하나]의 경우를 생각하고 HA구조를 짜야한다.

**권한/계정/보안 관련**

Service Control Policy: Organization에 한해 account의 권한을 제어하는 기능이다.

Resource Access Manager: AWS계정 간 리소스 공유하는 가장 기본 기능이다. 계정 뿐 아니라 Organization, Organization Unit도 RAM의 대상이다.

Macie: S3한정, 머신러닝과 패턴매칭 알고리즘 사용하여 보안이 필요한 데이터(PII)를 보호한다. 이후 Cloudwatch Events로 신호를 주는 식으로 진행된다.

GuardDuty vs Shield vs Shield Advanced vs WAF

GuardDuty: 계정, S3, 데이터 등 종합적으로 intelligent하게 Threat을 탐지하고 모니터링, ‘사전차단’한다. Cloudtrail, vpc flow logs, dns logs 등을 활용한다. = active antivirus 정도로 볼 수 있다.

Shield: DDoS 보호 용도이다. Standard version은 CF, Route53과 함께 3, 4계층의 공격을 막는다.

Shield Advanced: WAF 통합 가능하다. EC2, ELB, GA에 대해 DDoS이상의 보호를 제공한다.

WAF: CF나 ALB, API Gateway등에 DDoS외에 SQL주입이나 스크립팅 같은 웹 공격 패턴을 막는다.

VPC Flow Log: VPC단위(or 서브넷)에서 IP 트래픽 로그를 수집하는 기능이다. S3나 CloudWatch logs에 제공된다.

S3의 encryption: client side encryption(S3로 들어오기 전 in transit에서 encryption)과 3개의 server-side encryption으로 구성된다. SSE-S3(AWS가 다 관리), SSE-KMS(AWS가 data관리, 내가 master 관리), SSE-C(내가 다 관리)로 나뉜다.

KMS: 키 관리해주는 서비스이다. Secrets Manager이더라도 key의 생성, encryption을 위해서는 KMS가 있어야 한다. AWS managed CMK (AWS가 직접 CMK 생성 및 관리한다.), Customer-Managed CMK(고객이 직접 관리하고 모든 권한을 갖는다+key rotation 가능하다), AWS owned key(우리가 접근도 못하는)으로 나뉘며, CloudHSM을 활용한 방식도 있다.

Symmetric vs Asym. CMKs: Sym은 encrypt, decrypt모두 같은 키로 진행되며 AES-256을 따른다. Asym.은 public, private key pair를 생성한다. RSA를 따르며 KMS 사용하는 AWS services는 asym. 사용하지 않는다.

CloudHSM: Hardware Security Module이다. Dedicated hardware로 아주 강력하게 키를 관리하는 서비스이다. Single-tenancy이고 fips lv 3으로 KMS(lv2)보다 강력하다. AWS의 서비스는 HSM에 접근할 수 없다. AWS차원에서의 API도 없고, PKCS, JCE, Microsoft CryptoNG같은 Api를 따른다.