Cantrill Review – Hybrid Environments and Migration

Border Gateway Protocol (BGP) 101

* Autonomous System (AS) – Routers controlled by one entity – a network in BGP
* Autonomous System Number (ASN)
  + Unique and allocated by IANA
    - 0-65535
    - 64512 – 65534 are private
      * Can be used in private peering arrangements without officially allocating
* BGP operates over tcp/179 – it’s reliable
* Not automatically – peering is manually configured
* BGP is a path-vector protocol
  + It exchanges the best path to a destination between peers
    - The path is called the ASPATH
* iBGP = Internal BGP
  + Routing within an AS
* eBGP = External BGP
  + Routing between AS’s
* In a route table, An ASPATH of “I” is the origin – the network was learned from a locally connected network
* BGP exchanges the shortest ASPATH between peers. In example, Brisbane -> Alice Springs would default to the satellite link (5 mbps), even though the longer fibre (1 gbps) would provide better performance.
* Shortest Route is chosen, not the most performant route or route with higher throughput capabilities.
  + To override, use a technique called ASPATH Prepending
* ASPATH Prepending can be used to artificially make the satellite path look longer, making the fibre path preferred
  + Add additional ASN to the path to make it appear longer/less desirable
    - Ex. ASPSATH – 202,202,202,I (satellite) vs. 201,202,I (fibre) vs. 202,I (satellite w/out ASPATH prepending)
* An AS will advertise all the shortest paths it knows to all its peers.
  + The AS prepends its own AS number onto the path
    - This creates a source to destination path which BGP routers can learn & propagate

IPSEC VPN Fundamentals

* IPSEC is a group of protocols
* It sets up secure tunnels across insecure networks
* Between two peers (local and remote)
* Provides authentication
* Provides encryption
* Visibility to any modifications
* IPSEC Tunnels
  + Created as required –
  + “Interesting Traffic” – traffic which matches certain rules – ex. Network prefixes or complex traffic types
    - If traffic matches the rules defined as interesting traffic, a VPN tunnel is created to carry traffic through to its destination
    - No interesting traffic – tunnels are torn down and rebuilt if/when interesting traffic is detected
* IPSEC has two main phases
* IKE Phase 1 – (Slow & Heavy)
  + Authenticate – Pre-shared key (password) / Certificate
  + Using Asymmetric encryption to agree on, and create a shared symmetric key
  + IKE SA (Security Association) – Created as phase 1 tunnel
* IKE Phase 2 – (Fast & Agile)
  + Uses the keys agreed in phase 1
  + Agree encryption method, and keys used for bulk data transfer
  + Create IPSEC SA – Phase two tunnel (architecturally running over phase 1)
  + Allows IPSEC VPN Tunnels to be torn down and rebuilt if/when interesting traffic is detected
* IKE Phase 1 – How does it work?
  + Diffie-Hellman Key Exchange (DH) –
    - Each side creates a DH private key
    - And derives a public key
    - Used to decrypt data and sign things
    - Public keys are exchanged over the public internet
    - Each side takes their private key, and the remote peers public key – independently generate the “same” shared Diffie-Hellman Key
    - Using DH Key – exchanges key material and agreements
    - Each side generates a Symmetrical key – using DH key and exchanged material
      * Used to encrypt anything going across the Phase 1 tunnel.
        + IKE Security Association (Phase 1 Tunnel)

Can be used at Phase 2

* IKE Phase 2 – How does it work?
  + DH Key on both sides
  + Same symmetric phase 1 tunnel
* Policy Based VPN
  + Rule Sets match traffic => a pair of Sas
  + Different rules/security settings for different types of traffic
* Route-Based VPNs
  + Target matching (prefix)
  + Matches a single pair of SAs
  + Less functionality, but simpler setup

AWS Site-to-site VPN

* A logical connection between a VPC and on-premises network
  + Encrypted using IPSEC – running over the public internet
* Full HA – if you design and implement it correctly
* Quick to provision – less than an hour
* Key Components
  + VPC
  + Virtual Private Gateway (VGW)
  + Customer Gateway (CGW)
    - Logical – AWS
    - On-prem – device
  + VPN Connection between the VGW and CGW
* Static vs. Dynamic VPN (BGP)
  + Static
    - Routes for remote side (AWS) must be added to route tables as static routes
    - Networks for remote site statically configured on the VPN connection (CGW)
    - No load balancing and multi-connection failover
  + Dynamic
    - Routes CAN be added to remote side (AWS) route tables as static routes OR enable route propagation
      * Route Propagation – if enabled, means routes are added to RT’s automatically on AWS side
    - BGP is configured on both the customer and AWS side using ASN
      * Networks are exchanged via BGP
    - Multiple VPN connections provide HA and traffic distribution
* VPN Considerations
  + Speed limitations – 1.25 Gbps limit
  + VPN Gateway – 1.25 Gbps limit
  + Latency Considerations – inconsistent, public internet
  + Cost – AWS hourly cost, GB out cost, data cap (on premises)
  + Speed of setup – hours – all software configuration
  + Can be used as a backup for Direct Connect (DX) –
    - DX = physical hardware
  + Can be used with Direct Connect (DX)
    - Use while waiting for Direct Connect to setup – up to month

Direct Connect (DX) Concepts

* A physical connection
  + 1, 10, or 100 Gbps
* Business Premises => DX Location => AWS Region
* Port Allocation at a DX location
* Port Hourly Cost & Outbound Data Transfer
  + DX Location and Speed of the port
  + Inbound is free
* Provisioning time – physical cables & no resilience
  + Potentially weeks for physical cable to be installed
  + Add multiple DX in meantime
* Low & consistent latency + High Speeds
* Access AWS Private Services (VPCs) and AWS Public Services – NO INTERNET
* DX Location –
  + Not owned by AWS
  + Large Regional Centre
  + AWS have space and equipment
* If you don’t have space/equipment at a DX location – a COMMS provider can extend the DX port into your business premises
* Cross Connect – AWS Port to Customer/Partner Port link

Direct Connect (DX) Resilience and HA

* AWS Regions have multiple Direct Connect (DX) locations
  + Normally located in Major Metro Datacenters
* DX Locations are connected to AWS Region via redundant high speed connections
* By default, a single cross-connect links a DX port with a customer or provider router
* A DX is extended from the DX location to a Customer Premises
* Single points of failure –
  + DX Location
  + DX Router
  + Cross Connect
  + Customer DX Router
  + Extension
  + Customer Premises
  + Customer Router
* By default, Direct Connect (DX) is not HA – requires additional configuration
* For HA using DX provision –
  + 2 DX Routers
  + 2 Customer DX Routers
  + 2 Customer Premises Routers
  + This architecture uses multiple devices at both the Direct Connect and customer premises locations. It’s resilient against hardware failure in 1 path.
  + Single Points of Failure when using the HA configuration –
    - DX Location, Customer Premises, and potentially extension link path (between AWS DX Router and Customer DX Router)
    - If a telco (Customer DX Router) uses the same cable path between locations – it’s a hidden SPOF
  + For added resiliency, provision in two separate customer premises with two DX Customer Premises Routers, connecting to 2 DX Locations - (2) Customer DX Routers, using 2 extension links to connect to 2 AWS DX Routers, through to the Virtual Private Gateway (Private AWS Resources) or to Public AWS Resources (IP/Range)
    - This architecture can protect against the failure of a location OR hardware – and will continue to operate at a reduced resiliency level in the event of an outage.
    - Risk of outage if an entire location fails AND a router in the other location OR corresponding customer premises fails
* Key takeaways –
  + DX is a physical component – not HA unless you architect/implement it redundantly
  + Can use S2S VPN for DX Backup

Direct Connect (DX) + Public VIF + VPN

* When combined, provide end to end (IPSEC encrypted tunnel) access to private VPC resources
* Public VIF + VPN
  + Encrypted & Authenticated tunnel
  + Over DX (low latency & consistent latency)
  + Uses a Public VIF + VGW/TGW Public Endpoints
  + Transit Agnostic (DX/ Public Internet)
  + End-2-End – CGW ⬄ TGW/VGW – MACsec is single hop based
  + Wider vendor support
  + VPN has more CRYPTOGRAPHIC overhead vs MACsec – limits speeds
  + Can be used while DX is being provisioned and/or as a DX backup
* IPSEC doesn’t compete with MACsec

AWS Transit Gateway

* The AWS Transit Gateway is a network gateway which can be used to significantly simplify networking between VPCs, VPN, and Direct Connect
* It can be used to peer VPCs in the same account, different account, same or different region, and supports transitive routing between networks
* Network Transit Hub to connect VPC’s to on-premises networks
* Significantly reduces network complexity
* Single network object – HA and Scalable
* Attachments to other network types
  + Supported types
    - VPC
    - S2S VPN
    - Direct Connect Gateway
* VPC Attachments are configured with a subnet in each AZ where service is required
  + Specify a subnet in each AZ in each of the VPC that you want to use the GW with
* When connected to Customer Gateway, Transit GW allows bi-directional transitive routing between on-premises and multiple AWS VPCs
* TGW can be peered with other TGWs in other accounts, in other regions
  + Peering Attachment –
    - Cross-Region
      * Any cross-region data uses the AWS Global Network instead of public internet
    - Same/Cross-Account
* TGW can integrate with direct connect gateway using a transit VIF
* TGW comes with a default RT but supports multiple RT’s allowing complex routing architectures
* TGW Considerations
  + Supports transitive routing
  + Can be used to create global networks
  + Can share TGW between accounts using AWS RAM
    - RAM – share products and services between AWS accounts
  + Peer with different regions – same or cross account
  + Less complexity vs w/o TGW
    - Ideal solution vs VPC peering

Storage Gateway

* Storage Gateway is a product which integrates local infrastructure and AWS storage such as S3, EBS Snapshots, and S3 Glacier
* Virtual Machine (or hardware appliance\*)
* Presents storage using iSCSI, NFS, or SMB
* Integrates with EBS, S3, and Glacier within AWS
* Use cases
  + Migrations
  + Extensions
  + Storage Tiering
  + DR
  + Replacement of backup systems
* For the exam – picking the right mode
* Volume Stored
  + All data stored locally
  + Data is copied into S3 in the form of EBS snapshots
    - Constantly
  + Use cases
    - Great for full disk backups of servers
    - Assists with DR – create EBS Volumes in AWS
    - Low latency access to data
  + Doesn’t improve on-premises datacenter capacity – Main copy of data is stored on the gateway
  + Limitations –
    - 32 Volumes per GW
    - 16 TB per volume
    - Total of 512 TB per GW
* Volume Cached
  + Instead of Storage GW – Volume Stored Mode storing all data on-premises – Data is stored in AWS S3
  + AWS Managed area of S3 – only visible from Storage GW console – not visible from S3 directly – raw block storage
  + Can be used to create EBS snapshots
  + Storage appears on-premises
  + Benefits
    - Datacenter capacity extension
    - Data stored in AWS – Cached on-premises
  + Limitations –
    - 32 Volumes per GW
    - 32 TB per volume
    - Total of 1 PB per GW

Storage Gateway Tape – VTL Mode

* VTL – Virtual Tape Library
* Traditional Tape Backup
  + Large Backups => Tape
    - E.g., LTO-9 Media can hold 24 TB Raw Data
      * Up to 60 TB Compressed
  + 1 Tape Drive can use 1 tape at a time
  + Loaders (Robots) can swap tapes
  + A Library is 1+ drive(s), 1+ loader(s), and slots
  + Components
    - Drives
    - Library
    - Robots – Media Changer
    - Tape Shelf (anywhere but the Library) – different physical location
  + Offsite Tape Storage – Often Managed by a 3rd Party
  + Only tapes physically in the library can be used for backups and restores. Tapes are moved offsite when not in active use. Transport between locations take time and often has a cost
* AWS Configuration
  + On prem –
    - Backup Server
    - Media Changer
    - Tape Drive(s)
    - Upload Buffer
    - Local Cache
  + On prem connects to AWS owned/managed
    - Virtual Tape Library (VTL) – stored in S3
    - Tape Shelf (VTS) – stored in S3 Glacier or Deep Archive
  + Capacity
    - Virtual Tape – 100GB => 5 TiB
    - Storage GW can handle 1PB of data across 1500 virtual tapes (S3)
      * When not in use, stored on Tape Shelf (S3 Glacier or Deep Archive)
    - Virtual Tape Shelf – Unlimited Storage
* Storage Gateway (VTL) – pretends to be an iSCSI tape library, changer, and drive
  + Capabilities
    - Maintain existing backup system, but replace the expensive parts with AWS storage
    - Extend an existing backup system by using AWS Storage
    - Migration – while migrating the backups to AWS Storage
* Storage Gateway (File)
  + File Mode
  + Bridges local file storage over NFS and SMB with S3 Storage
  + Supports multi-site, maintains storage structure, integrates with other AWS products and supports S3 Object Lifecycle Management
  + Mount points (shares) available via NFS (Linux) OR SMB (Windows)
    - Map directly on an S3 bucket
    - Visibility to the bucket
  + Files stored into a mount point, are visible as objects in an S3 bucket
  + Read and Write Caching ensure LAN-like performance
  + On Prem – create File shares that are each linked with a single S3 bucket
    - A Bucket Share = AWS S3 & On-premises file share
  + S3 objects are visible in the file share on-premises
  + File structure (flat object storage system) created using the s3 standard for file naming
  + 10 Bucket Shares per Storage GW
  + Primary Data is held in S3
  + Integrate with other AWS Services
    - Athena
    - S3 Events
    - Lambda
    - Etc.
  + Multiple Contributors (On-premises locations) supported
  + Use the NotifyWhenUploaded API to notify other gateways when objects are changed
  + File Gateway doesn’t support Object Locking
    - Use read only mode on other shares or tightly control file access
    - New files of same name overwrite previous version
  + Multiple Contributors + Replication
    - Cross region replication
    - 2 on prem File GWs
    - 2 AWS Regions with 2 S3 Buckets
  + File Gateway + S3 Lifecycle Management
    - Specify default storage class
    - Add Lifecycle Policies on a bucket to move data between storage classes

Snowball / Snowball Edge / Snowmobile

* Move large amounts of data IN and OUT of AWS
* Physical storage – suitcase or truck
* Ordered from AWS Empty, Load up, Return
* Ordered from AWS with Data, empty, and Return
* For the exam know which to use

Snowball

Data Sync

* A product which can orchestrate the movement of large scale data (amounts or files) from on-premises NAS/SAN into AWS or vice-versa
* Data Transfer service TO and FROM AWS
* Migrations, Data Processing Transfers, Archival/Cost Effective Storage, or DR/BC
* Designed to work at huge scale
* Keeps metadata (e.g., permissions/timestamps)
* Built in data validation
* Key Features
  + Scalable
    - 10 Gbps per agent (~100TB per day)
  + Bandwidth Limiters – avoid link saturation
  + Incremental and scheduled transfer options
  + Compression and Encryption
  + Automatic recovery from transit errors
  + AWS Service Integration – S3, EFS, FSx
  + Pay as you use – per GB cost for data moved
* The DataSync Agent (installed locally) runs on a virtualization platform such as VMWare and communicates with the AWS DataSync Endpoint
* DataSync Endpoints – Locations define the SRC or DST for the sync of data TO or FROM AWS
  + S3, EFS, FSx, NFS, SMB
* Encryption in transit – TLS
* Schedules can be set to ensure the transfer of data occurs DURING or AVOIDING specific time periods
* Customer impact can be minimized by setting a bandwidth limit in (MiB/s)
* Uses DataSync Agent on-prem SAN/NAS storage migration over NFS or SMB
* DataSync Components
  + Task – A Job within DataSync, defines what is being synced, how quickly, FROM where TO where.
  + Agent – Software used to read or write to on-premises data stores using NFS or SMB
  + Location – every task has two locations FROM and TO
    - Network File System (NFS) – Linux
    - Server Message Block (SMB) – Windows
    - Amazon EFS
    - Amazon FSx
    - Amazon S3

FSx for Windows File Server

* Advanced shared file system accessible over SMB, and integrates with Active Directory (either managed, or self-hosted)
* Fully managed native windows file servers/shares
* Designed for integration with windows environments
* Integrates with Directory Service or Self-Managed AD
* Single or Multi-AZ within a VPC
* On-demand and Scheduled Backups
* Accessible using VPC, Peering, VPN, Direct Connect
* Know when to use vs. EFS
* Uses AD for user store
* Supports managed or self-managed AD (On-premises)
* [\\fs-xxx123.animals4life.org\catpics](file:///\\fs-xxx123.animals4life.org\catpics)
  + Example access path
  + Workspaces for example
* Native windows file system, supports de-duplication (sub file) Distributed File System (DFS), KMS at-rest encryption, and enforced encryption in transit
* Accessed via SMB
* Support volume shadow copies – file level versioning
* Performance
  + 8MB/s to 2GB/s
  + 100K’s IOPS
  + <1ms latency
* FSx Key Features and Benefits
  + VSS – User Driven Restores
  + Native file system accessible over SMB
  + Windows permission model
  + Supports Distributed File System (DFS) – scale-out file share structure
  + Managed – no file server admin
  + Integrates with DS AND your own directory

FSx for Lustre

* Designed for HPC
* Delivers extreme performance for scenarios such as Big Data, Machine Learning, and Financial Modeling
* Managed Lustre – Designed for HPC – LINUX Clients (POSIX)
* 100s GB/s throughput & sub millisecond latency
* Deployment types
  + Persistent
    - Longer Term
    - HA (in one AZ)
    - Self-healing
  + Scratch
    - Highly optimized for Short term
    - No replication
    - Fast
* Accessible over VPN or Direct Connect
* Data resides in the File System while processing occurs
  + Data is Lazy Loaded from S3 (S3 Linked Repository) into the file system as it’s needed
  + Data can be exported back to S3 at any point using hsm\_archive
* Metadata stored on Metadata Targets (MST)
* Objects are stored on Object Storage Targets (OSTs) (1.17 TiB)
* Baseline performance based on size
* Size – min 1.2 TiB then increments of 2.4 TiB
* For Scratch – Base 200 MB/S per TiB of storage
* For Persistent – 50 MB/s, 100 MB/s, and 200 MB/s per TiB of storage
* Both Scratch and Persistent modes support burst capability of up to 1,300 MB/s per TiB (Credit System)
* Lustre runs from 1 AZ – 1 ENI - 1 VPC
* Key features
  + Scratch is designed for pure performance
    - Short term or temp workloads
    - NO HA
    - NO Replication
  + Larger file systems means more servers, more disks, and more chance of failure
  + Persistent has replication within ONE AZ only
    - Auto-heals when hardware failure occurs (1 AZ)
  + You can backup both modes to S3 – Manual or Automatic 0-35 retention

AWS Transfer Family

* Managed file transfer service – supports transferring TO or FROM S3 and EFS
* Provides managed “servers” which support protocols
  + FTP – Unencrypted file transfer
  + FTPS – File transfer with TLS encryption
  + SFTP – File transfer over SSH
  + Applicability Statement 2 - AS2 – Structured B2B data
* Identities – Service Managed, Directory Service, Custom (Lambda/API Gateway)
* Managed File Transfer Workflows (MFTW) – serverless file workflow engine
* Endpoint Types – Public, VPC – Internet, and VPC – Internal
* Key features –
  + Multi AZ – Resilient and Scalable
  + Provisioned Server per hour $ + data transferred $
  + FTP and FTPS – Directory Service or Custom IDP only
  + FTP – VPC ONLY (cannot be public)
  + AS2 – VPC Internet / Internal ONLY
  + If you need to access S3/EFS, but with existing protocols
    - Integrating with existing workflows
    - OR Using MFTW to create new ones