IR states or greenfield build.

* DataSync can help you rapidly migrate your data to AWS
* DataSync was built to help solve issues with size of data sets. More data sets to migrate today and DataSync can help them do that
* Application of file data is one use case for DataSync
* Transfer data into the cloud for processing. Using the cloud to process and analyze ;large amounts of data
* Archive to free up on-premises storage capacity. This helps free up or shut down legacy systems
* Replication for data protection and recovery in case of disaster. Could be one time or on-going replication
* It takes significant amount of work and money to do data transfer

***WHAT IS AWS DATASYNC?***

* It is an *online transfer service* that simplifies, automates, and accelerates *moving data between on-premises storage and AWS.*
* It characterized by fast data transfer, easy to use, secure and reliable, cloud integrated and cost effective
* DataSync is built for speed so performance is one of its strong points
* Its very secure and reliable. It encrypts data in flight and also supports encryption at rest as well on S3 and EFS. It provides option for validating your data making sure that all data was transfer successfully
* It integrates with CloudWatch, cloud trail, IAM etc.

***HOW AWS DATASYNC WORKS***

* Start by deploying an agent on-premises in a VM environment and you connect the agent to the shared file system and decide where you want your data to be transferred. The data can be copy to other s3 storage classes as well. Once source and destination are decided you create a task which will automatically take care of moving the data from on-prem fiule system to the cloud
* <<Datasyncimage.png>>
* <<AWS DataSync Implementation Run book (2).pdf>>

***WHAT ARE NFS AND SMB***

* NFS is network file system.

***NFS****stands for Network File Sharing*. It is an easy-to-use and relatively affordable protocol. *NFS provides a solution for remote file sharing between the servers*, by using the existing internet protocol infrastructure. NFS comes in a variety of flavors, with NFSv3 being the most popular. Some of the interesting protocol options include caching options, security support, shared file locking characteristics and much more. NFS has progressed a lot to facilitate more security, file sharing and better performance as well.

*From <*[*https://www.serveradminz.com/blog/introduction-to-nfs-vs-smb/*](https://www.serveradminz.com/blog/introduction-to-nfs-vs-smb/)*>*

* ***WHAT IS SMB***

***SMB****stands for  Server Message Bloc*k. It is a network file sharing protocol that deals applications to read and write files on a system. An application (or a user of an application) can use this protocol to access files or other resources on a remote server, which allows them to automatically read, create, and update files on the distant server. It can also communicate with any server application configured to handle SMB client requests.

*From <*[*https://www.serveradminz.com/blog/introduction-to-nfs-vs-smb/*](https://www.serveradminz.com/blog/introduction-to-nfs-vs-smb/)*>*

***HANDS-ON***

* Miscellaneous folder should be copied to EFS as their destination
* The cold folder are files that are not accessed regularly so I want to transfer them to s4 glacier deep archive to save some money
* The many folder contains files that need to be accessed a little more frequently and I want to copy them to s3 standard
* The on-prem was deployed in us-wests-1 (Northern California) and the AWS region was deployed in us-east-2(Ohio)
* *Showmount -e* shows you all the NFS exports that you have in your NFS server
* We created an s3 and EFS file system in the AWS region
* We don’t want to copy the junk data in the miscellaneous folder
* To find out how many files are in a folder la into that folder and | wc -l
* So it should look like *ls data | wc -l*
* Before you start copying files you need to activate the agent
* Deploy your agent as close to your source data as possible. Activate the agent in the region where you are copying the agent to.
* All data in DataSync is encrypted

***Type of endpoints agent will connect to***

* The first option which is public service endpoint in which the agent will connect to the DataSync service is done the internet. Here you don’t have a dedicated service
* The second option which is FIPS service endpoint is one that is commonly used by American government agencies.
* The last option is for people who have a dedicated connection either via VPN or direct connect
* To activate the agent you get the IP address of the agent to do this. This helps the browser connect to the agent through its public IP over port 80 and running a quick activation script. Once the activation is ready the agent will close the port and tell us that its ready.
* Next step is creating a task in which I create a source and destination location
* Here you provide information for the location (NFS in our case), the Region, the agent, the NFS server IP if we picked NFS and the mount path which in this case is the export path
* DataSync always performs checks during transfers
* An agent can only run one task at a time but if you select queuing then DataSync automatically starts the next tasks if you are trying to run multiple tasks
* The task also gives you the ability to filter out files you don’t want to copy
* You can also configure logging
* Every time you run a task you have the option to run it with different settings
* *Agent are activated in the region data will be transferred to*
* When the location is an S3 bucket you create an IAM role in which DataSync will use to communicate with S3
* DataSync communicates with S3 using VPC endpoints
* You can have multiple DataSync agents pointing to the same location or multiple agents pointing to the same location or one agent pointing to multiple locations

***DATASYNC ARCHITECTURAL GUIDANCE***

* One agent can fill a 10Gbps network link
* Supports private link - don’t have to use internet to move data
* Transfers to /from AWS
* It can copy directly to low cost S3 storage classes
* Transfer data via multiple DataSync agents to the same target
* Use DataSync for EFS-EFS migration