

How to design an IBM Spectrum Scale solution

Tomer Perry
Spectrum Scale Dev.

Outline

- Scope: Designing Spectrum Scale correctly
- Gathering requirements
- Gathering environment characteristics
- Translating requirements into Scale configurations
- Summary



Session Scope

- This session is based on multiple customer cases where solution design workshops were performed
- On some cases, the workshop was done “after the facts” :-)
- The session will discuss how to design the software part of Spectrum Scale
- Basic knowledge of Spectrum Scale is required (education...)



Gathering requirements

- At the end of the day, the purpose of any infrastructure solution is to serve the business needs
- The problem is that in most cases, the IT guys don't know the business needs – but at least can define what are the application requirements
- In some cases, meeting with the application owners is required in order to better understand what they need
- We don't always need “exact” requirements. In many cases they are unknown....
- One of the major challenges is that most customers don't understand the concept of SDS
- Still, collecting the requirements is crucial part of the solution design, Thus, we might need to work together with the customer to define them.



Gathering requirements - “workload families”

- In many cases, there are many applications. So, we can’t do a “per application design”.
- Thus, we will try to divide them into “workload families”
- For each workload, we need to find out the following (again, as best as we can):
 - Dataset type (large number of small files, small number of large files, large directories etc.)
This will allow us to decide what is the right data/metadata block size. Can we share filesystem/s between different workloads etc.
 - Access pattern (random/seq, metadata intensive)
This will help us to define the right media to use (disks, flash etc.)
 - Access type (read/write, parallel, batch/interactive)
Will help us defining the right media, and can we combine different workloads
 - Performance requirements (GB/s or OP/s, file creates per second)
This is one of the most difficult parts. As most customers just don’t know. Be careful. IOPs are usually not an interesting factor at this stage. Filesystem ops are.



Gathering requirements - “workload families”

- The next stage, is to define the per workload functional requirements:
 - Snapshot policy:
While we can create a per fileset snapshot. Creating a snapshot require quiescing the whole filesystem. It might make sense to try and group filesets with similar snapshot requirements into the same filesystem
 - Quotas:
Spectrum Scale support many different types of quotas. Its important to understand what is the user/group quota policy that the customer require. But, also, how can we use fileset quotas in order to manage “space allocation” between different groups/applications
 - Namespace requirements:
In many cases, the customer is migrating from existing system. Changing application paths might be difficult. So, we might need to use different techniques to “virtualize” the namespace (bind mounts/Symlinks)
 - Security requirements:
What level of isolation is required? Are ACLs enough? What type of ACLs?. Do we need to implement encryption? Do we need to use different filesystems? Do we need to use multiple client clusters ?
 - Capacity requirements:
Not the overall capacity. Rather the per workload. How to manage it. Is it per workload or per group?



Gathering requirements - “workload families”

- Per workload functional requirements – continued:
 - Replication/DR:
Do we need sync replication? Async replication? Disaster recovery ? What are the required RPO/RTO?
 - Backup:
What are the backup requirements? What backup application is being used?
 - Information Lifecycle Management (ILM)
Does it make sense to implement ILM for the workload? Which type? Do we have the required software (Spectrum Archive? Spectrum Protect? TCT ?)
- Again, gathering those requirements, will later allow us to design the overall solution correctly.



Understanding the customer's environment

- As any software, Spectrum scale rely on the customer's environment. We need to make sure that the environment can satisfy Spectrum Scale requirements:
 - Network topology and vendor:
As a distributed software, Spectrum Scale highly depends on the network capabilities. Most of the issues we're facing are due to faulty networks. This is why its extremely important to understand how the network is designed: Who maintain the network? Is it a dedicated network or shared with other workloads? Is the network oversubscribed? Do we have any special network virtualization in place? In some cases, it might make sense to get some expert advice about the network.
 - Name Resolution:
How is name resolution performed? Is there HA in place for name resolution? Do we want to use the existing DNS servers or create a dedicated DNS caching servers ?
 - Time services:
Do the customer has reliable NTP servers ?
 - Authentication servers:
What are the authentication requirements? Do Spectrum Scale support those requirements ?
 - Key servers infrastructure:
Usually relevant when encryption is required. Is the key server infrastructure is supported by Spectrum Scale? HA?



Understanding the customer's environment

- WAN environment (if relevant):
What is the WAN environment looks like? What is the latency (sync vs. async)? What equipment is being used? What is the available bandwidth?
- Monitoring infrastructure:
What the customer is using for health monitoring? Performance monitoring? How to integrate Scale monitoring with the customer's environment ?
- Migration requirements:
What are the requirements to migrate existing data? Existing applications ?



Translating requirements into Scale configurations

- So now we (hopefully) have enough information about what are the requirements and how the environment looks like. How do we translate that into Scale configurations....
- Essentially, we need to start going through the various Scale layers:
 - Clusters design
 - Filesystem/s design
 - Storage pools design
 - Fileset design
 - Replication design
 - Monitoring design



Cluster design

- How many clusters do I need?
 - Cluster is an “administrative domain” (ssh)
 - Cluster nodes needs to be able to talk to each other and to filesystem owning cluster nodes (but not to other client clusters...) - Network structure
 - Filesystem access Isolation
 - But..what about licensing? DM should solves that problem
 - Preferably, separate storage and compute
- For each cluster, carefully select quorum nodes, manager nodes, helper nodes etc.



Filesystem design

- Do I need more than one?
- Reasons for having more than one filesystem:
 - Block size (performance and allocation)
 - Recovery
 - Snapshots impact
 - Quotas impact
 - Cross Cluster isolation
- Reasons for having small number of filesystems
 - Management overhead
 - Resource isolation (performance)



Storage Pools design

- Storage virtualization
- Reasons for having multiple storage pools:
 - Data vs. Metadata
 - Different media for different workloads
 - Colder data
 - Resource isolation (poor man QoS)



Fileset Design

- Namespace virtualization
- Fileset is “mini filesystem” (dependent vs. independent)
- Namespace virtualization
- Immutability
- Policy entity
- Snapshot policy
- Space allocation (quotas, inodes)
- Always think if fileset can replace a filesystem



Replication design

- Sync vs. Async
 - Sync:
 - Latency impact
 - ReadReplicaPolicy (default, local, fastest)
 - Carefully define failure groups
 - Async:
 - AFM
 - What mode to use? (RO, LU, SW, IW, DR)
 - AFM GW: collocating or dedicated?
 - WAN tuning: tcp buffers
 - RPO for DR (can we make it on time...)



Monitoring design

- Health monitoring vs. Performance monitoring
- “Isolated” vs. “Integrated” monitoring (SMB vs. Enterprise)
- Custom monitoring vs. standard monitoring
- SNMP or not...
- Callbacks



TWENTY QUESTIONS



VAN



DICK



HERB



FLORENCE



ALDO RAY



Session reference links

- Spectrum Scale Docs - <https://ibm.biz/Bdiaq5>
- Spectrum Scale wiki - <https://ibm.biz/BdHA6N>
- Spectrum Scale FAQ - <https://ibm.biz/BdiaqN>
- Spectrum Scale User Group - <https://ibm.biz/BdiaqN>

