

AIX Backup Performance Best Practices and Guidelines on IBM Power Systems Virtual Server

IBM Power Systems Performance Team

Authors
Saritha Vinod
Gayathri M

04 March 2022

Table of Contents

1. Introduction	2
1.1 File level Backup for AIX VMs in PowerVS Using Spectrum Protect	2
2. IBM Spectrum Protect Configuration and Tunings for Backup	3
2.1 Spectrum Protect Configuration	3
2.2 Spectrum Protect Server Tunings	3
3. Backup Performance Guidelines	4
3.1 Backup Performance Measurements with 1Gbps DL 2.0 Link	4
3.1.1 Performance Guidelines for Spectrum Protect Server-side Deduplication and Compression with 1Gbps DL 2.0 Link	4
3.1.2 Performance Guidelines for Spectrum Protect Client-side Deduplication and Compression with 1Gbps DL 2.0 Link	5
3.1.3 Settings and Best Practices for Spectrum Protect Client-side Deduplication and Compression	5
3.2 Backup Performance Measurements with 10Gbps DL 2.0 Link	6
3.2.1 Performance Guidelines for Spectrum Protect Server-side Deduplication and Compression with 10Gbps DL 2.0 Link	7
3.2.2 Performance Guidelines for Spectrum Protect Client-side Deduplication and Compression with 10Gbps DL 2.0 Link	7
4. Acknowledgements	8

1. Introduction

This document specifies the best practices, guidelines, and tunings for optimum performance of file level backup for AIX virtual machines on Power Systems Virtual Server (PowerVS) using IBM Spectrum Protect. Based on benchmarking results with the best practices and tunings discussed in this document, below table provides a summarized view of the guidelines for backup rates that are achievable on different network link bandwidths. The sections below detail each of these different configs and some of the trade-offs to be considered while using these backup settings.

Direct Link (DL 2.0) Bandwidth	Backup rates achievable	Settings Used
1 Gbps	~300-310 GB/hour	Server-side dedup (~50%), 250KB Extentsize
1 Gbps	~330-530 GB/hour	Client-side dedup (~50%), 250KB Extentsize, dedupcache off
10 Gbps	~560-615 GB/hour	Server-side dedup (~50%), 250KB Extentsize
10 Gbps	~680-725 GB/hour	Server-side dedup (~50%), 750KB Extentsize
10 Gbps	~780-880 GB/hour	Client-side dedup (~50%), 750KB Extentsize, dedupcache off

1.1 File level Backup for AIX VMs in PowerVS Using Spectrum Protect

For file level backup, IBM Spectrum Protect Server is installed and configured on an x86 Linux VSI located in the IBM Cloud Classic Infrastructure area of IBM Cloud. PowerVS AIX Virtual Server Instance connects to the Spectrum protect Server using the Direct Link 2.0 (DL 2.0) for backup-restore operations and can access Cloud Object Storage using an internal endpoint for data storage/retrieval as depicted in Figure 1 below. The installation details of Spectrum protect is outside the scope of this document. Refer to this link for a tutorial on the details of installation of Spectrum Protect for PowerVS AIX backup.

https://cloud.ibm.com/media/docs/downloads/power-iaas-tutorials/PowerVS_AIX_Backups_Tutorial_v1.pdf

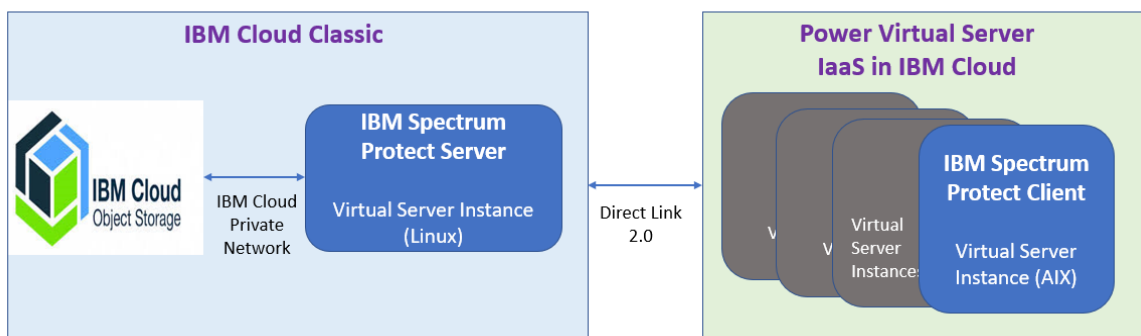


Figure 1: Backup environment using Spectrum Protect for AIX Virtual Machines on PowerVS

2. IBM Spectrum Protect Configuration and Tunings for Backup

This section provides the configurations and tunings applied for IBM Spectrum Protect Server and Client.

2.1 Spectrum Protect Configuration

All guidelines and tunings mentioned in this document is using the **Spectrum Protect Direct-to-Cloud** Configuration, where backup data is stored directly to cost-effective IBM Cloud Object Storage (hosted within IBM Cloud).

For the direct-to-cloud architectures, backup data is ingested directly into a cloud-container storage pool with a performant cloud accelerator cache disk location tuned for a system's ingestion workload as the initial "landing spot". Data is then immediately, asynchronously transferred to cloud object storage while further data is also ingested into the disk staging area.

Below provided configuration is used to deploy an extra-small system for **IBM Spectrum Protect Server** as defined in the IBM Spectrum Protect Cloud Blueprints for IBM Cloud.

- Spectrum Protect Server Version 8.1.12
- 4 cores Intel Xeon CPU (M1.4x32 Virtual Server Instance in IBM Cloud Classic Infrastructure)
- 32 GB RAM
- Operating System – RHEL 8.3
- Block Storage (SAN Storage Volumes)
 - 100GB Volume (boot volume)
 - 300GB Volume (database and active log disk)
 - 500GB Volume (database backup and archive log disk)
 - 1TB Volume (Cloud accelerator cache disk)
- Network
 - 1Gbps public and private network
- Private IBM Cloud Object Storage endpoint (Regional)

Spectrum Protect Client (Version 8.1.11) is installed on an AIX 7.2 virtual machine in PowerVS in the same Cloud region as the virtual server instance for Spectrum Protect Server. PowerVS AIX VM used for performance measurements is allocated 4 cores and 64GB memory and the CPU, memory, disk and network resources on AIX VM was mainly consumed by Spectrum protect during the backup operation. The required CPU/memory allocations on AIX VM can vary based on the workload running on VM. Backup measurements used Oracle Database type data from a workload generator on the AIX VM.

PowerVS AIX VM can use either 1Gbps or 10Gbps Direct Link Connect 2.0 (DL 2.0). The Direct Link Connect 2.0 service creates a seamless connection that allows access to Spectrum Protect Server virtual instance on IBM Cloud from Power Systems Virtual Server instance.

2.2 Spectrum Protect Server Tunings

Below provided are the recommended tunings in general on the Spectrum Protect Server. It is recommended that below settings should be applied as soon as possible after a node is defined and before initial backups take place.

- Kernel tunings
 - kernel.randomize_va_space = 0
 - vm.swappiness = 5
- OS ulimit tunings
 - nofile = 65536
- Network Settings
 - TCPWINDOWSIZE of 0 which is the default.
- Spectrum Protect server tunings
 - Tune MINIMUMExtentsize parameter value from 50KB (default) to 250KB (recommended for Oracle database type workloads)
 - Note that, when there is a change to the extent size parameter, the subsequent first backup (irrespective of incremental or not) can take slightly longer, as further backups would not dedup well against older backups.

3. Backup Performance Guidelines

The following sections outline the guidelines, tunings and best practices based on the backup throughput results that were measured in the IBM Cloud and PowerVS environments with the configurations described above.

In some of the scenarios, multiple instances of IBM Spectrum Protect client backup sessions were initiated in parallel as mentioned in respective sections. In such case, the results presented here for backup represent the aggregate throughput experienced with the largest number of sessions tested against the system. Data sizes used for measurements varied between 250GB to 750GB.

All throughput values are obtained for the actual data size specified for backup. These are the data rates reported by Spectrum Protect Client as the average aggregate throughput when backing up data to the IBM Spectrum Protect server. On ingestion, the actual quantity of data that makes it to accelerator cache disk and onwards to cloud object storage will be less, depending on the data deduplication and compression rates.

3.1 Backup Performance Measurements with 1Gbps DL 2.0 Link

This section in particular talks about backup performance guidelines when using the 1Gbps DL 2.0 link between Spectrum Protect Client in PowerVS and Spectrum Protect Server in IBM Cloud. This includes two configurations – (1) with default Spectrum Protect Server-side Deduplication and Compression, (2) enabling Spectrum Protect Client-side Deduplication and Compression

3.1.1 Performance Guidelines for Spectrum Protect Server-side Deduplication and Compression with 1Gbps DL 2.0 Link

By default, inline data deduplication, compression and encryption are enabled on the Spectrum Protect server. This occurs during client backup operations, data is automatically deduplicated and compressed as it is ingested into the cloud-container storage pool at the server. However, the actual

backup data size gets transferred over the network (through DL 2.0 link) from the AIX virtual machine to Spectrum Protect server on IBM Cloud.

With the settings discussed above, for a full backup operation peak network utilization of ~80% and aggregate network utilization of ~72% was measured on a 1Gbps DL 2.0 link. With these settings, full backup data transfer rate of up to **~300-310 GB/hour** can be attained considering 50% deduplication rates for the data being backed up using a 1Gbps DL 2.0 link. If the data deduplication rates for backed up data is higher, the backup data transfer rate per hour also can increase. The above rate is achieved with “MINIMUMExtentsize” parameter value of 250KB (mentioned as tunings in Section 2.3).

3.1.2 Performance Guidelines for Spectrum Protect Client-side Deduplication and Compression with 1Gbs DL 2.0 Link

Client-side data deduplication and compression can be enabled to work together with server-side deduplication. In this case deduplication and compression of data happens at the client before the data gets transferred to the server over the network (via the DL 2.0 link). This can reduce the size of actual data transferred over network, depending on the deduplication and compression rates achieved.

The steps to enable **client-side deduplication and compression** and best practices to be followed while doing so are discussed in Section 3.1.3 below. Following those best practices, for a full backup operation peak network utilization up to ~76% was measured on a 1Gbps DL 2.0 link with multiple parallel instances (up to 4 instances), with 50% data deduplication rate and dedup cache turned off. With these settings, full backup data transfer rate of up to **~330-530 GB/hour** can be attained considering 50% deduplication rates for the data being backed up using a 1Gbps DL 2.0 link. The above rate is achieved with “MINIMUMExtentsize” parameter value of 250KB (mentioned as tunings in Section 2.3).

3.1.3 Settings and Best Practices for Spectrum Protect Client-side Deduplication and Compression

Enable client-side de-duplication to work together with server-side de-duplication as below:

– On the Server: use below commands (replace nodename with name of the client node)

update node <nodename> compression=yes

update node <nodename> deduplication=clientorserver

– On the Client: update the dsm.sys file to add these lines

compression yes

deduplication yes

For a detailed description of these commands refer to Spectrum Protect documentation links below.

<https://www.ibm.com/docs/en/spectrum-protect/8.1.12?topic=commands-update-node-update-node-attributes>

<https://www.ibm.com/docs/en/spectrum-protect/8.1.12?topic=reference-compression>

<https://www.ibm.com/docs/en/spectrum-protect/8.1.12?topic=reference-deduplication>

Below provided are some **best practices to use client-side data deduplication**

- Use compression along with deduplication at client side.
 - On the Client: update the dsm.sys file to add these lines
 - ***compression yes***
 - ***deduplication yes***
- Use multiple parallel sessions to improve backup time and network utilization. This can be done by initiating multiple dsmc incr command and splitting the subdirectories between these commands. Eg :
 - dsmc incr /dir/subdir1/ -subdir=yes
 - dsmc incr /dir/subdir2/ -subdir=yes
- Use the default resource utilization (RU = 2)
- Use de-dup cache turned off with lower data deduplication rates.

To turn off the deduplication cache from the client, edit the dsm.sys file to add this line

enablededupcache no

- Turn on de-dup cache when the data deduplication is higher (de-dup close to 100%)

To turn on the dedupe cache from the client, edit the dsm.sys file to add this line

enablededupcache yes

NOTE on client-side deduplication cache:

- For applications that use the IBM Spectrum Protect API, do not use the client data deduplication cache (enablededupcache option). The dedup cache can be used with backup-archive clients. This is a restriction in IBM spectrum Protect. Please refer to this Spectrum Protect documentation link for more details on this:

<https://www.ibm.com/docs/en/spectrum-protect/8.1.12?topic=throughput-tuning-client-side-data-deduplication>

3.2 Backup Performance Measurements with 10Gbps DL 2.0 Link

This section in particular talks about backup performance guidelines when using the 10Gbps DL 2.0 link between Spectrum Protect Client in PowerVS and Spectrum Protect Server in IBM Cloud. This includes two configurations – (1) with default Spectrum Protect Server-side Deduplication and Compression, (2) enabling Spectrum Protect Client-side Deduplication and Compression

Also note that the extra-small configuration used for Spectrum Protect Server (described in Section 2.2) uses a x86 Virtual Server Instance in IBM Cloud Classic Infrastructure that has a 1Gbps non-rate limited link for public and private network. Hence the 10Gbps DL 2.0 link connectivity from PowerVS AIX VM to Spectrum Protect server cannot be fully utilized.

With that limitation, the **best practices used for backup with 10Gbps DL 2.0 link** are as follows:

- For server-side as well as client-side deduplication, use multiple parallel instances (below guidelines are with 4 instances) to maximize utilization of the 10Gbps link with jumbo frames enabled at client and server
 - To enable jumbo frames at Spectrum Protect Server (RHEL OS):
 - Check the current setting by running the command: *“ip link show dev eth0”*
 - Change the current setting to support 9000 MTU by running the command: *“ip link set mtu 9000 dev eth0”*
 - Change the setting so that it persists after the system is restarted. Edit the file *“/etc/sysconfig/network-scripts/ifcfg-eth0”*, and add *“MTU=9000”*
 - To enable jumbo frames at Spectrum Protect Client (AIX OS): (Note: change interface name below based on the actual interface used)
 - **chdev -Pl ent2 -a jumbo_frames=yes**
 - **chdev -Pl en2 -a mtu=9000**
- Client-side deduplication and compression:
 - Use client-side deduplication and compression for data that has higher deduplication rates (50% and above).

3.2.1 Performance Guidelines for Spectrum Protect Server-side Deduplication and Compression with 10Gbps DL 2.0 Link

With server-side deduplication and compression and the above settings, full back up data transfer rate of up to **~560-615 GB/hour** can be attained considering 50% deduplication rates for the data being backed up using a 10Gbps DL 2.0 link. The above rate is achieved **with “MINIMUMExtentSize” parameter value of 250KB** (mentioned as tunings in Section 2.3).

If the **“MINIMUMExtentSize” parameter value is increased to 750KB**, full back up data transfer rate can increase up to **~680-725 GB/hour** considering 50% deduplication rates for the data being backed up using a 10Gbps DL 2.0 link.

Note that using 750KB extent size could increase the network bandwidth utilization, thereby reducing the backup time taken. However, it may have other trade-offs and may not suite all types of workloads (one implication could be that de-dup rates achieved might decrease). Hence setting 250KB extent size is recommended for workloads like Oracle databases, where higher de-duplication is expected.

3.2.2 Performance Guidelines for Spectrum Protect Client-side Deduplication and Compression with 10Gbps DL 2.0 Link

With client-side deduplication and compression, with de-dup cache turned off, and **with “MINIMUMExtentSize” parameter value of 750KB**, full back up data transfer rate of up to **~780-880 GB/hour** can be attained considering 50% deduplication rates for the data being backed up using a 10Gbps DL 2.0 link.

4. Acknowledgements

We hereby acknowledge the following people for their review feedback, suggestions and support while creating these best practices guide.

- Rakesh Sharma, Power Systems Virtual Server Networking
- James Damgar, Spectrum Protect Systems
- Jeff Tenner, Power Systems Performance
- Kokil Deuri, Power Systems Performance