

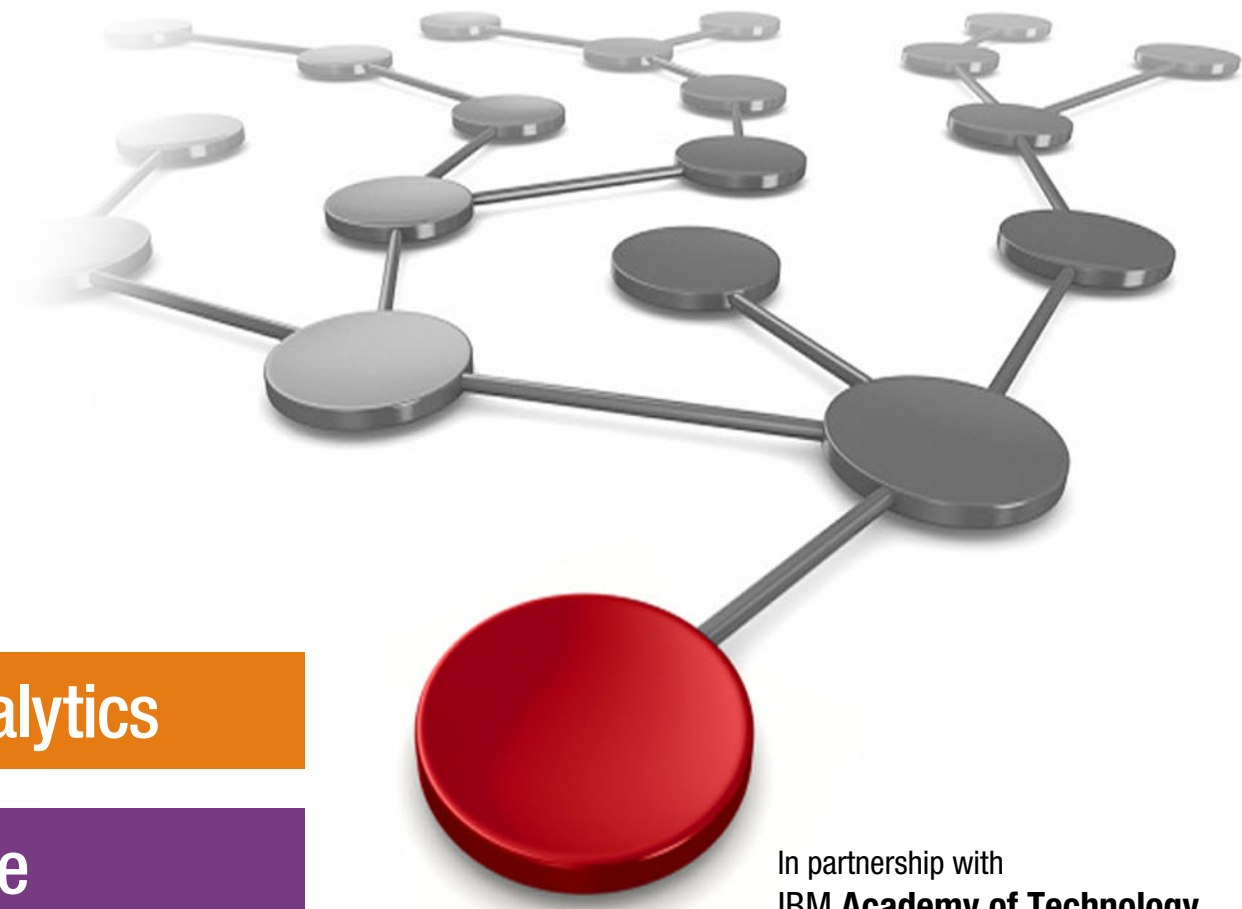
Hortonworks Data Platform with IBM Spectrum Scale

Reference Guide for Building an Integrated Solution

Sandeep R. Patil

Wei G. Gong

Larry Coyne



 Analytics

Storage

In partnership with
IBM Academy of Technology



Hortonworks Data Platform with IBM Spectrum Scale: Reference Guide for Building an Integrated Solution

This IBM® Redpaper™ publication provides guidance on building an enterprise-grade data lake by using IBM Spectrum™ Scale and Hortonworks Data Platform for performing in-place Hadoop or Spark-based analytics. It covers the benefits of the integrated solution, and gives guidance about the types of deployment models and considerations during the implementation of these models.

Hortonworks Data Platform

Hortonworks Data Platform (HDP) is a leading Hadoop and Spark distribution. HDP addresses the complete needs of data-at-rest, powers real-time customer applications, and delivers robust analytics that accelerate decision making and innovation.

IBM Spectrum Scale

IBM Spectrum Scale™ is flexible and scalable software-defined file storage for analytics workloads. Enterprises around the globe have deployed IBM Spectrum Scale to form large data lakes and content repositories to perform high-performance computing (HPC) and analytics workloads. It can scale performance and capacity both without bottlenecks.

Integrated solution overview

HDP brings business value to enterprises by analyzing their disparate data sources and deriving actionable insights from them. This analytics journey typically starts with consolidation of different data silos to form an *Active Archive*. The Active Archive is then used to get a single view of the customer and perform further predictive analytics on them.

With IBM Spectrum Scale, clients can build highly scalable and globally distributed data lakes to form their Active Archives. IBM Spectrum Scale becomes the storage layer for your HDP environment as an alternative to native Hadoop Distributed File System (HDFS). It supports the access of the data by using HDFS APIs and is not apparent to the applications that use HDP. With IBM Spectrum Scale, you get more flexible deployment models for your storage system that help you optimize infrastructure costs.

IBM Spectrum Scale and HDP were first certified with IBM Spectrum Scale V4.2.3 and HDP 2.6. Since the first certification, Hortonworks and IBM have signed an agreement to certify both the products on an ongoing basis for their new releases and keep the certification current. See Table 1 on page 21 for certified software levels. This certification is for IBM Spectrum Scale software and applies to all deployment models of IBM Spectrum Scale, including IBM Elastic Storage™ Server. This certification includes a paper certification for using Hortonworks DataFlow (HDF) V3.0 with IBM Spectrum Scale.

Benefits of integration

The following are the top benefits of using IBM Spectrum Scale with HDP:

- ▶ Extreme scalability with parallel file system architecture

IBM Spectrum Scale is a parallel architecture. With a parallel architecture, no single metadata node can become a bottleneck. Every node in the cluster can serve both data and metadata, enabling a single IBM Spectrum Scale file system to store billions of files. This architecture enables clients to grow their HDP environments seamlessly as the data grows. Additionally, one of the key value propositions of IBM Spectrum Scale, especially with IBM Elastic Storage Server, is running diverse and demanding workloads, plus the ability to tier down to Active Archive.

- ▶ A global namespace that can span multiple Hadoop clusters and geographical areas

Using IBM Spectrum Scale global namespace, clients can create active, remote data copies and enable real-time, global collaboration. This namespace enables global organizations to form *data lakes* across the globe, and host their distributed data under one namespace.

IBM Spectrum Scale also enables multiple Hadoop clusters to access a single file system while still providing all the required data isolation semantics.

The IBM Spectrum Scale Transparent Cloud Tiering feature can archive data into a S3/SWIFT compatible cloud object storage system, such as IBM Cloud™ Object Storage or Amazon S3, by using the powerful IBM Spectrum Scale information lifecycle management (ILM) policies.

- ▶ A reduced data center footprint with the industry's best in-place analytics

IBM Spectrum Scale has the most comprehensive support for data access protocols. It supports data access by using NFS, SMB, Object, POSIX, and the HDFS API. This feature eliminates the need to maintain separate copies of the same data for traditional applications and for analytics.

- ▶ True software-defined storage that is deployed as software or as a pre-integrated system

You can deploy IBM Spectrum Scale as software directly on commodity storage-rich servers running the HDP stack, or deploy it as part of a pre-integrated system by using the IBM Elastic Storage Server. Clients can use software-only options to start small, while still using enterprise storage benefits. With IBM Elastic Storage Server, clients can control cluster sprawl and grow storage independently of the compute infrastructure. IBM Elastic Storage Server uses erasure coding to eliminate the need for the three-way replication for data protection that is required with other solutions.

► IBM hardware advantage

A key advantage for IBM Elastic Storage Server is to lower capacity requirements. IBM Elastic Storage Server requires 30% extra capacity to offer similar data protection benefits. IBM Power Systems™ servers along with the IBM Elastic Storage Server offer the most optimized hardware stack for running analytics workloads. Clients can enjoy up to three times reduction of storage and compute infrastructure by moving to IBM Elastic Storage Server compared to commodity scale-out x86 systems.

To support the security and regulatory compliance requirements of organizations, IBM Spectrum Scale offers Federal Information Processing Standard (FIPS) compliant data encryption for secure data at rest, policy-based tiering/ILM, cold data compression, disaster recovery, snapshots, and backup and secure erase. The HDP Atlas and Ranger components provide additional data governance capabilities and the ability to define and manage security policies.

Component diagram

The component diagram in Figure 1 shows the relationships between IBM Spectrum Scale and the HDP components.

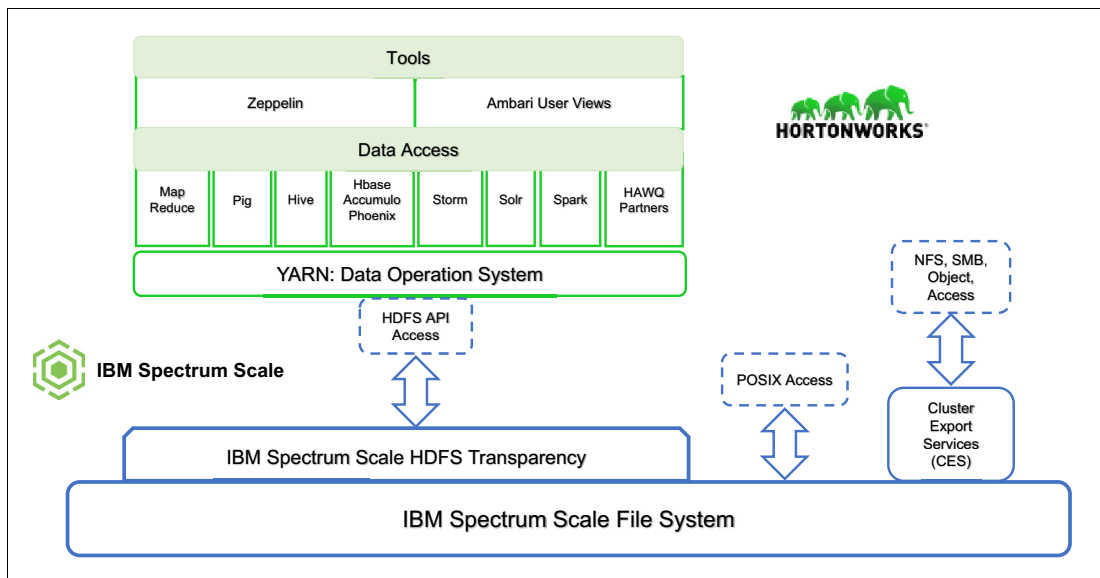


Figure 1 HDP and IBM Spectrum Scale component diagram

IBM Spectrum Scale file system supports the direct POSIX access interface. The IBM Spectrum Scale Cluster Export Services (CES) provides high availability NFS, SMB, and Object protocol services. You can use these different data access protocols for direct data access, such as importing data into IBM Spectrum Scale Data Lake for Hadoop analytics and retrieving the results.

IBM Spectrum Scale HDFS Transparency

IBM Spectrum Scale HDFS Transparency enables applications to use a standard HDFS client to access IBM Spectrum Scale through native Remote Procedure Call (RPC) requests. All data transmission and metadata operations in HDFS use the RPC mechanism and are processed by NameNode and DataNode services within HDFS. IBM Spectrum Scale HDFS Transparency integrates both the NameNode and the DataNode services, and responds to the requests from HDFS clients. HDFS clients can continue to access IBM Spectrum Scale seamlessly, just as it does with HDFS.

Figure 2 shows the IBM Spectrum Scale HDFS Transparency component.

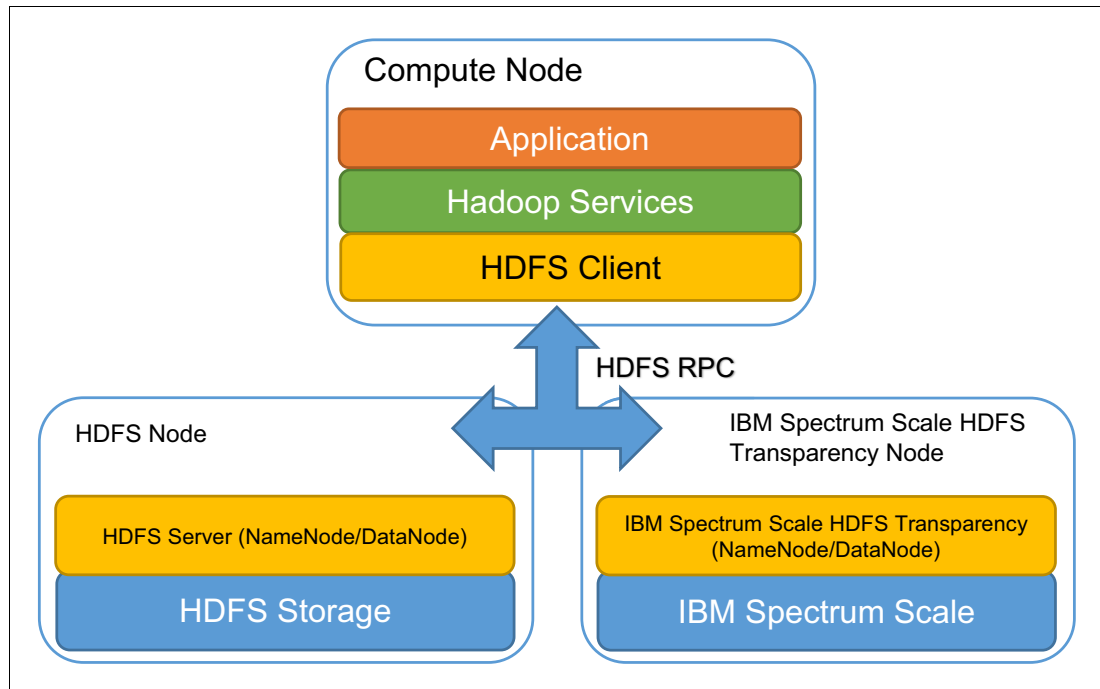


Figure 2 IBM Spectrum Scale HDFS Transparency

The following are the key advantages of IBM Spectrum Scale Transparency Connector:

- ▶ An IBM Spectrum Scale Client is not needed on every Hadoop node. The HDFS client can access data on IBM Spectrum Scale as it does with HDFS storage.
- ▶ Full Kerberos support.
- ▶ Improved security management by Kerberos authentication and encryption for RPCs.
- ▶ Support for more Hadoop components or HDFS-compliant APIs/commands (for example, **discp** and **webhdfs**).
- ▶ Use the HDFS client cache.

The following are other key HDFS Transparency and IBM Spectrum Scale differences to note:

- ▶ If one file is set with Access Control List (ACL) (POSIX or NFSv4 ACL), IBM Spectrum Scale HDFS Transparency does not provide the interface to disable the ACL check at the IBM Spectrum Scale HDFS Transparency layer. If you want to disable the ACL for one file, the only way is to remove the ACL.
- ▶ IBM Spectrum Scale HDFS Transparency does not support HDFS encryption. IBM Spectrum Scale provides its own encryption mechanism that can protect against attacks that target the disks (for example, theft or acquisition of improperly discarded disks) and attacks performed by unprivileged users of a IBM Spectrum Scale node in a multi-tenant cluster.
- ▶ IBM Spectrum Scale provides its own caching mechanism that does not support HDFS caching. Caching that is done by IBM Spectrum Scale is more optimized and controlled, especially when you run multiple workloads. The interface `hdfs cacheadmin` is not supported by IBM Spectrum Scale HDFS Transparency.
- ▶ NFS Gateway from native HDFS is not supported by IBM Spectrum Scale HDFS Transparency. IBM Spectrum Scale provides multiple protocol interfaces, including POSIX, NFS, and SMB. Customers can use IBM Spectrum Scale Protocol for NFS to access the data.
- ▶ The option `distcp -diff` is not supported for snapshot over IBM Spectrum Scale HDFS Transparency. Other options from `distcp` are supported.
- ▶ The interface from `hdfs dfs` is supported, whereas others (such as `hdfs fsck`) are not needed for IBM Spectrum Scale HDFS Transparency.

Apache Ambari integration

Apache Ambari is an open source management platform for provisioning, managing, monitoring, and securing Apache Hadoop clusters. Apache Ambari, as part of the Hortonworks Data Platform, enables enterprises to plan, install, and securely configure HDP. It makes it easier to provide ongoing cluster maintenance and management no matter the size of the cluster.

The Ambari Management Pack for IBM Spectrum Scale service extends the capabilities of Apache Ambari to rapidly provision, deploy, manage, and monitor IBM Spectrum Scale and the IBM Spectrum Scale HDFS Transparency Connector.

The Ambari Management Pack for IBM Spectrum Scale module can either provision an IBM Spectrum Scale shared-nothing (File Placement Optimizer (FPO)) cluster from scratch, or detect an existing IBM Spectrum Scale cluster (including IBM Elastic Storage Server) for use in a Hadoop environment. In addition, the Ambari Management Pack installs and configures the IBM Spectrum Scale HDFS Transparency Connector based on customer input through the Ambari UI. The Ambari Management Pack ensures that all IBM Spectrum Scale components are configured based on preferred practices for deployment within a Hadoop environment.

When provisioned, the Ambari Management Pack can monitor IBM Spectrum Scale components through metrics that are integrated into the Ambari Metrics System (AMS) framework and available through the Ambari UI, as shown in Figure 3.

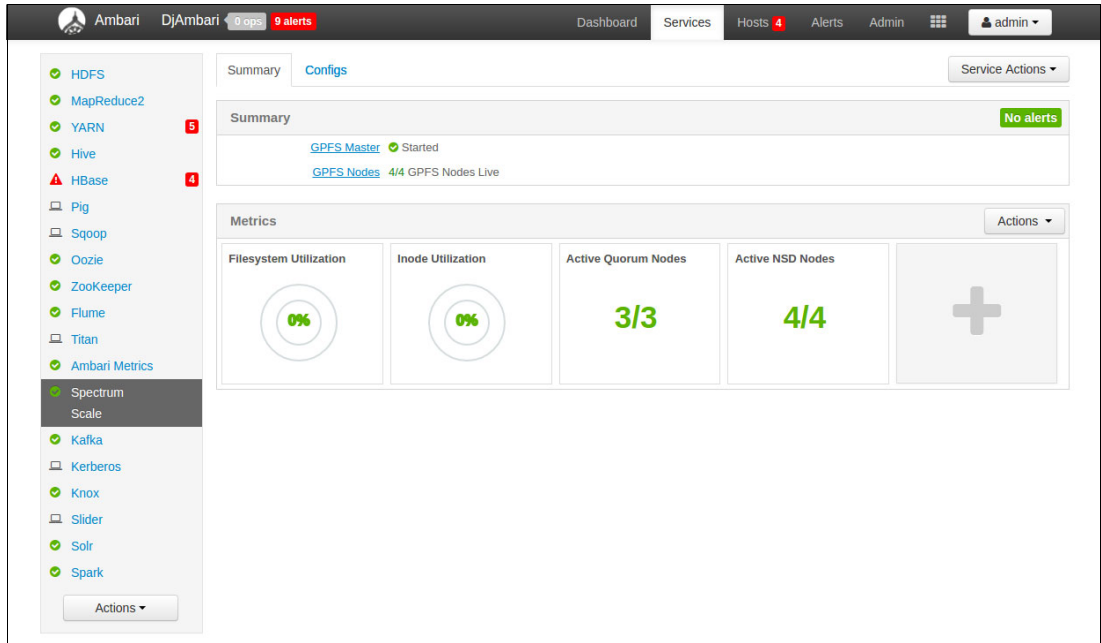


Figure 3 IBM Spectrum Scale Ambari integration

The Ambari Management Pack supports ongoing maintenance of both the IBM Spectrum Scale cluster and the IBM Spectrum Scale HDFS Transparency Connector, including component lifecycle management (Start, Stop, Add, Remove, and Move nodes) and the capability to upgrade these components.

Multiple Hadoop clusters over the same file system

By using HDFS transparency, you can configure multiple Hadoop clusters over the same IBM Spectrum Scale file system. For each Hadoop cluster, you need one HDFS transparency cluster to provide the file system service.

As shown in Figure 4, an IBM Spectrum Scale file system services Hadoop cluster 1 and Hadoop cluster 2 at the same time through HDFS Transparency cluster 1 and HDFS Transparency cluster 2.

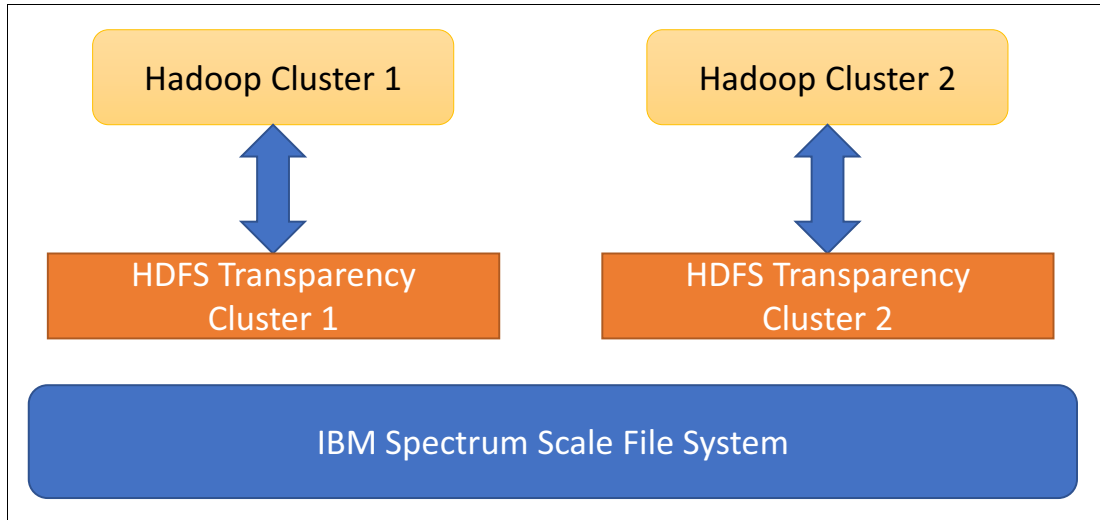


Figure 4 Two Hadoop Clusters over the same IBM Spectrum Scale file system

Hadoop Storage Tiering with IBM Spectrum Scale

For customers adopting IBM Spectrum Scale and IBM Elastic Storage Server (pre-integrated solution powered by IBM Spectrum Scale software) with Hortonworks Hadoop/Spark solution, a key requirement is being able to add IBM Elastic Storage Server (ESS) into an existing HDP cluster. This feature eliminates the need to set up a separate HDP cluster to gain the benefits of IBM ESS.

Hadoop Storage Tiering with IBM Spectrum Scale addresses this requirement. Enterprises that already have a standard HDP cluster with native HDFS can now add ESS as a storage tier in the same HDP cluster. See Figure 5. This configuration will help enterprises manage cluster sprawl by adding ESS-based shared storage to their existing HDP clusters.

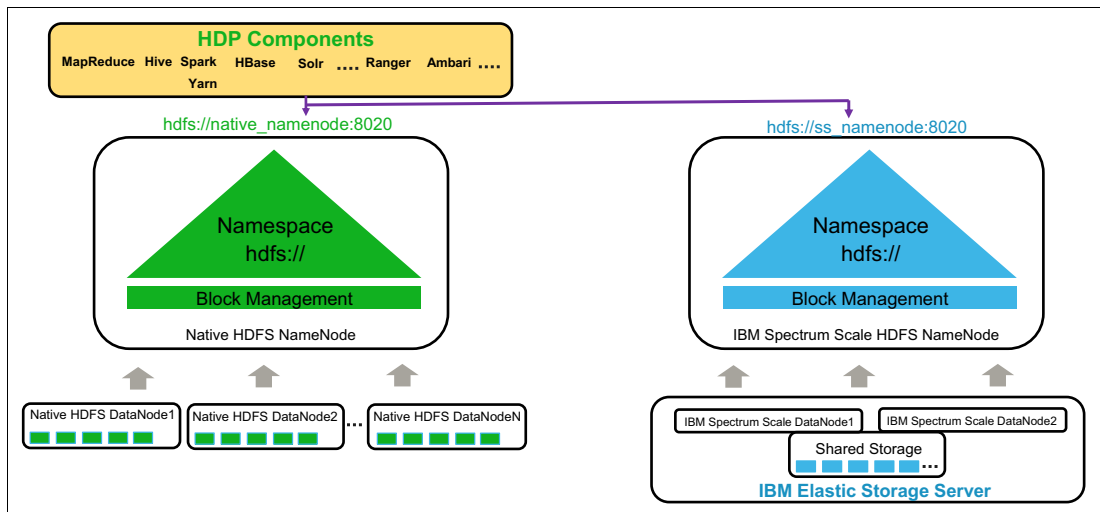


Figure 5 Hadoop storage tiering with IBM Spectrum Scale

The following are some ways to use this feature:

- ▶ As an ingest tier for faster ingest
Enterprises can use IBM Spectrum Scale POSIX support with flash-based IBM ESS to get super-fast ingests for their existing Hadoop data lakes.
- ▶ As a secondary tier with shared storage
Enterprises can use IBM ESS as a secondary tier in their existing Hadoop data lakes. This configuration enables them to grow storage independently of compute and also eliminates the need for three-way replication. The key benefit is the ability to run analytics directly on the secondary tier without having to bring the data into the primary tier.
- ▶ For data sharing between clusters
If an enterprise wants to build a new analytics workflow on a new HDP cluster, but also needs access to the data from an existing HDP cluster, the tiering feature can enable this without having to create data copies. IBM ESS can be used as a secondary tier for the existing HDP cluster. And the same ESS can act as the storage for a new HDP cluster. For example, some IBM customers are considering this scenario to introduce new IBM Power-based HDP clusters for demanding next generation analytics workflows.

Hadoop Storage Tiering with IBM Spectrum Scale is supported by IBM Spectrum Scale 4.2.3+ and HDP 2.6+ versions, and with all deployment models of IBM Spectrum Scale, which include IBM ESS and non-ESS-based deployments. See Table 1 on page 21 for more information. Hadoop Storage Tiering with IBM Spectrum Scale is planned to be enhanced after HDP 3.0 to allow this tiering within the clusters enabled for HDFS federation as well.

Short Circuit Read and Short Circuit Write

Short Circuit Read means reading data directly from the local file system when HDFS client and target data nodes are collocated on the same physical node. When the HDFS client requests the DataNode to read a file, the DataNode reads that file off the disk and sends the data to the client over a TCP socket. The short circuit read obtains the file descriptor from the DataNode and then reads the file directly.

For Short Circuit Write, if HDFS Client and HDFS Transparency DataNode are on the same node, when writing file from HDFS client, Short Circuit Write writes data directly into the IBM Spectrum Scale file system instead of writing data through RPC. This process might reduce the RPC latency through the local loop network adapter, and therefore enhance the write performance.

Deployment models

The following two deployment models are supported for IBM Spectrum Scale in an HDP environment:

- ▶ Shared Storage model
- ▶ Shared Nothing Storage model

Shared Storage model

IBM Spectrum Scale can be deployed in a Shared Storage model behind an HDP cluster by using IBM Elastic Storage Server. IBM Elastic Storage Server is a pre-integrated file storage solution that is powered by IBM Spectrum Scale software on IBM Power Systems and disk enclosures (JBOD), as shown in Figure 6.

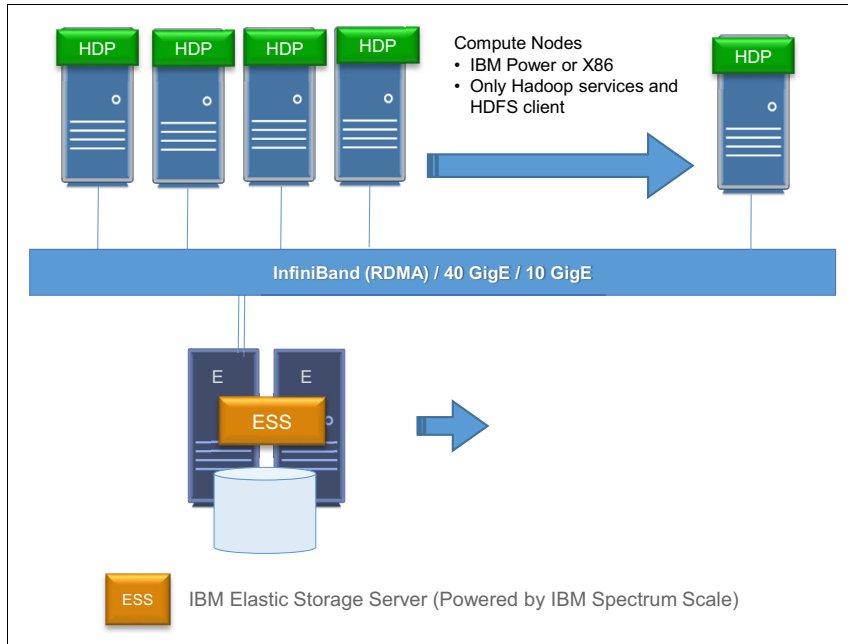


Figure 6 IBM Spectrum Scale as a shared Storage deployment

Additionally, IBM Spectrum Scale can be deployed with any SAN-attached Shared Storage and is a supported configuration for HDP environments. This publication focuses on IBM Elastic Storage Server based deployments for the Shared Storage model.

Shared Nothing Storage model

IBM Spectrum Scale can be deployed in a Shared Nothing Storage model directly on storage-rich servers running an HDP stack, as shown in Figure 7. This type of IBM Spectrum Scale deployment is called FPO deployment. This deployment is similar to native HDFS deployment.

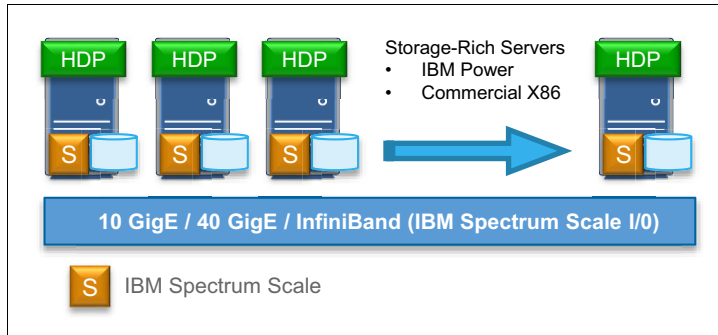


Figure 7 IBM Spectrum Scale as a Shared Nothing Storage deployment

Shared Storage model

This section contains the details of deploying HDP with IBM Elastic Storage Server as a Shared Storage model.

IBM Elastic Storage Server

IBM Elastic Storage Server is a high-performance, pre-integrated IBM Spectrum Scale based network storage disk solution. This technology combines the CPU and I/O capability of the IBM POWER® architecture and matches it with 2U and 4U storage enclosures. Sustained data streaming performance can reach 25 GBps (40 GBps with all-flash storage IBM Elastic Storage Server models) in each building block of IBM Elastic Storage Server, growing as more blocks are added.

IBM Elastic Storage Server employs IBM Spectrum Scale de-clustered RAID software for data protection for performance and fast recovery. This software RAID function eliminates the three-way replication for data protection that alternative storage solutions require. Instead, IBM Elastic Storage Server requires just 30% extra capacity to offer similar data protection benefits.

Scenarios for Shared Storage model

The Shared Storage model architecture is beneficial in the following scenarios:

- ▶ Medium to large capacity deployments.
- ▶ When you should grow storage independent of compute resources. This requirement is typically true when your data growth outpaces the need for additional compute resources yet you continue to invest in storage-rich servers (compute + storage). There is no benefit in increasing your investment in compute resources in this case.
- ▶ In-place analytics with mixed workloads (Hadoop and traditional applications) on the same storage.

- ▶ Storage consolidation for unstructured data with options for File (POSIX/NFS/SMB) and Object (S3/SWIFT) access.
- ▶ High scalability and performance needs.

Implementation guidelines

The following sections describe architecture and implementation guidelines when implementing HDP with IBM Elastic Storage Server.

Cluster configuration

In an HDP + IBM Elastic Storage Server deployment model, IBM Elastic Storage Server serves as central back-end storage and a set of Hadoop compute nodes are network-connected to the IBM Elastic Storage Server systems, as shown in Figure 8. Each of the compute nodes running HDP also runs IBM Spectrum Scale client along with IBM Spectrum Scale Hadoop Transparency Connector. Replacing HDFS with IBM Spectrum Scale clients has many advantages.

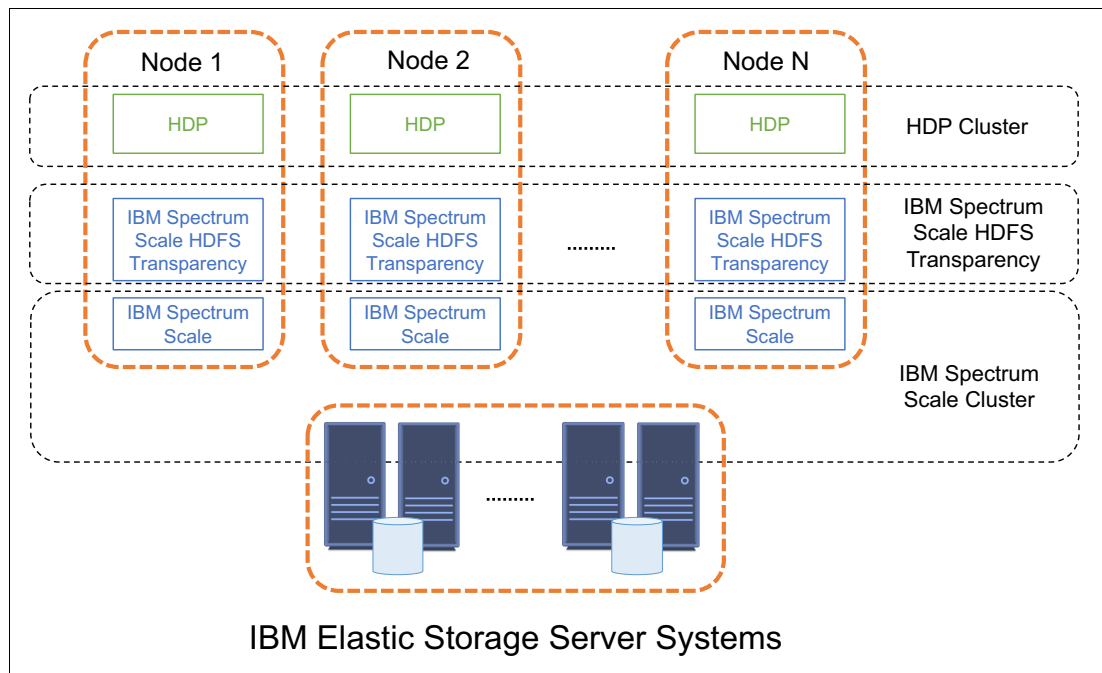


Figure 8 HDP + IBM Elastic Storage Server cluster configuration

The HDP + IBM Elastic Storage Server deployment model does not require storage-rich servers as your compute nodes because IBM Spectrum Scale clients access the data over a high-performance network. IBM Spectrum Scale client data access is spread across all of the storage nodes (IBM Elastic Storage Server I/O nodes in this case).

Although IBM Elastic Storage Server and all the IBM Spectrum Scale clients can be in one IBM Spectrum Scale cluster, as shown in Figure 8 on page 11, you can deploy IBM Spectrum Scale clients/HDP compute nodes in a different IBM Spectrum Scale cluster than the IBM Elastic Storage Server. In this case, the clients can access the data on IBM Elastic Storage Server through Remote Mount, as shown in Figure 9.

Using multiple clusters is preferable because it helps to separate out the IBM Elastic Storage Server nodes from the HDP cluster for better manageability. For more information about remote mount, see the [Accessing a remote GPFS file system](#) section in IBM Knowledge Center.

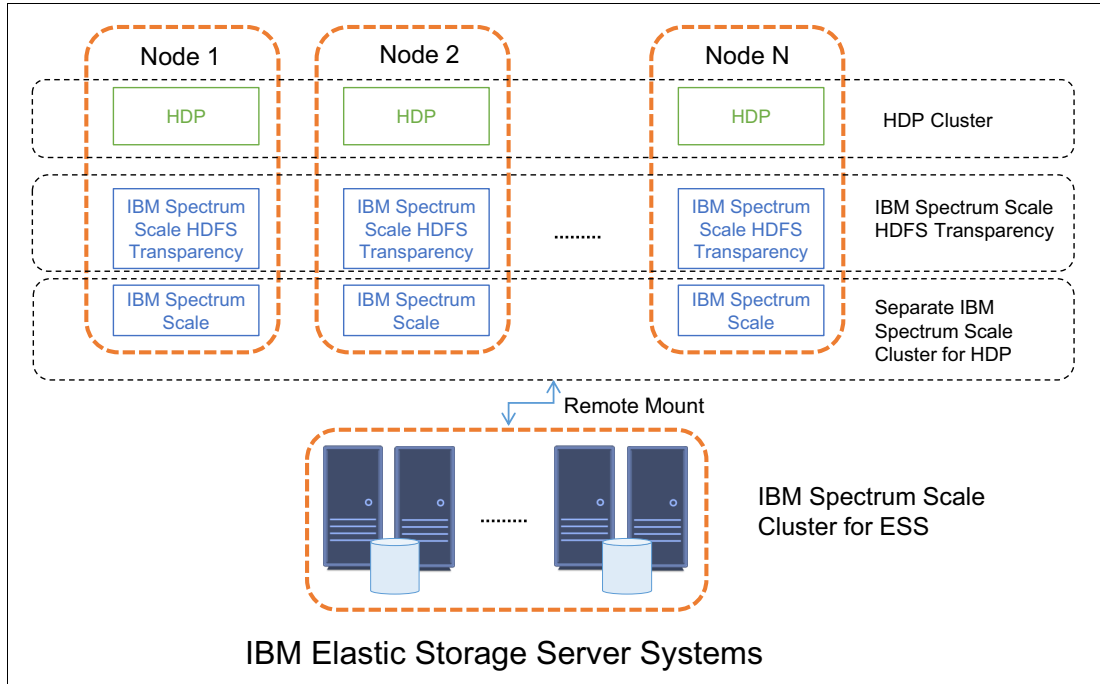


Figure 9 HDP + IBM Elastic Storage Server cluster configuration with Remote Mount (preferred)

System design

In the architecture shown in Figure 10, the IBM Elastic Storage Server is connected to the HDP compute nodes and set of CES nodes (a minimum of two, which are optional based on your requirement of having NFS, SMB, or Object access to data). For better performance, have all the HDP compute nodes run IBM Spectrum Scale Client with the IBM Spectrum Scale HDFS Transparency Connector. Figure 10 shows the IBM Spectrum Scale configuration types that are needed on the nodes. For more information about IBM Spectrum Scale configurations, see [IBM Knowledge Center](#).

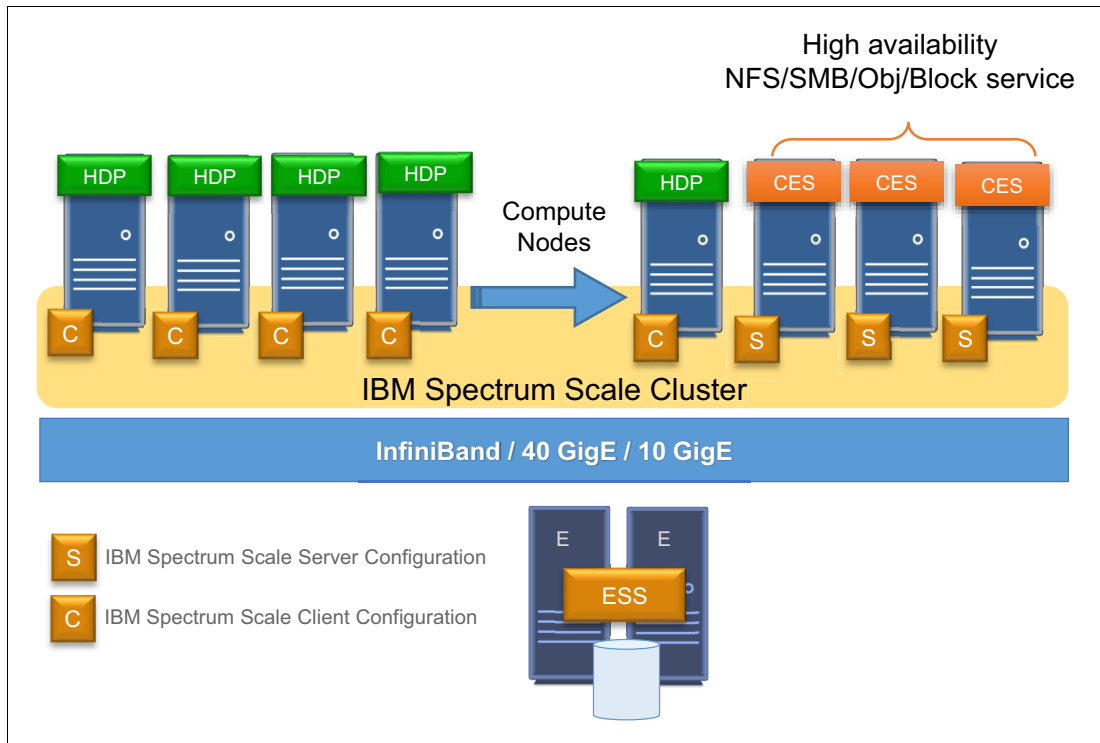


Figure 10 HDP + IBM Elastic Storage Server with protocol nodes

IBM Elastic Storage Server models

IBM Elastic Storage Server supports many high capacity and high IOPS model variations to fit your workload. The models are shown in Figure 11. Select the model that best supports your overall capacity, performance, and availability requirements.

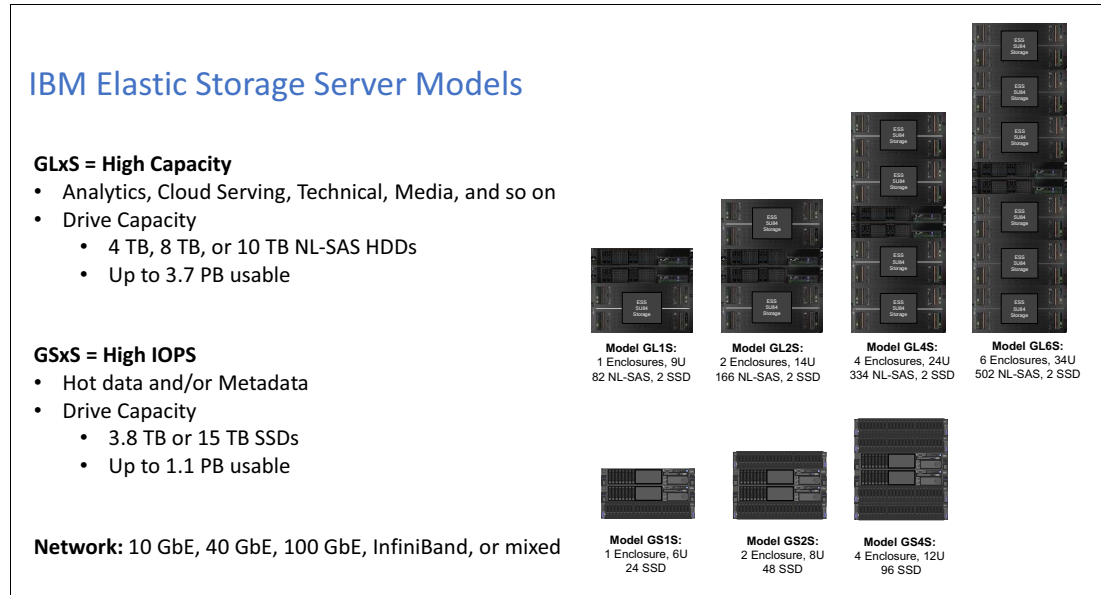


Figure 11 IBM Elastic Storage Server models

Network

IBM Elastic Storage Server offers network adapter options. Three PCI slots are reserved for SAS adapters and one PCI slot is configured by default with a 4-port 10/100/1000 Ethernet adapter for management. Three other PCIe3 slots are available to configure with any combination of Dual-Port 10 GbE, Dual-Port 40 GbE, or Dual-Port InfiniBand PCI adapters. For more information about updates to the 100 GbE or Enhanced Data Rate (EDR) InfiniBand adapters that are based on Mellanox ConnectX-4 cards, see the [IBM Knowledge Center](#).

Which network you choose depends upon your performance requirement and networking infrastructure. In a 10 GbE network topology with a single building block (IBM Elastic Storage Server), the maximum theoretical bandwidth per client cannot exceed the bandwidth of two network sockets, which provides a throughput of approximately 2 GBps. In comparison, a 40 GbE network can scale up to 8 GBps. For all GL4 and GL6 models, consider RDMA/InfiniBand or a 40 GbE or 100 GbE topology. Otherwise, the performance benefits from an IBM Elastic Storage Server building block are limited by the connectivity between the IBM Elastic Storage Server server and its clients.

Note: It is important to have a reliable network for IBM Spectrum Scale to work optimally.

Data protection

IBM Elastic Storage Server implements IBM Spectrum Scale erasure coding RAID software. IBM Spectrum Scale RAID implements sophisticated data placement and error-correction algorithms to deliver high levels of storage reliability, availability, and performance with cost-effective JBOD storage. For more information about IBM Spectrum Scale RAID and its components, see [IBM Spectrum Scale RAID Administration Guide](#).

IBM Spectrum Scale RAID supports 2- and 3-fault-tolerant Reed-Solomon erasure codes and 2, 3, and 4-way replication. These configurations detect and correct up to one, two, or three concurrent faults, depending on the chosen RAID level.

Note: The level of protection from drawer or enclosure failure depends on the IBM Elastic Storage Server model and RAID level. For enclosure protection on a GLS6 or GS6 models, you can choose either 8+2P or 8+3P. If you choose an IBM Elastic Storage Server GL4 or GS4, you must use an 8+3P protection scheme.

Scaling

A primary advantage of this deployment model is its ability to grow storage performance and capacity independent of the compute infrastructure. If storage capacity or storage performance is not sufficient, you can add more storage into your cluster dynamically. At the same time, when the compute capacity is not sufficient, you can add more compute nodes without investing in capacity. This granularity enables investment of resources based on your need, as shown in Figure 12.

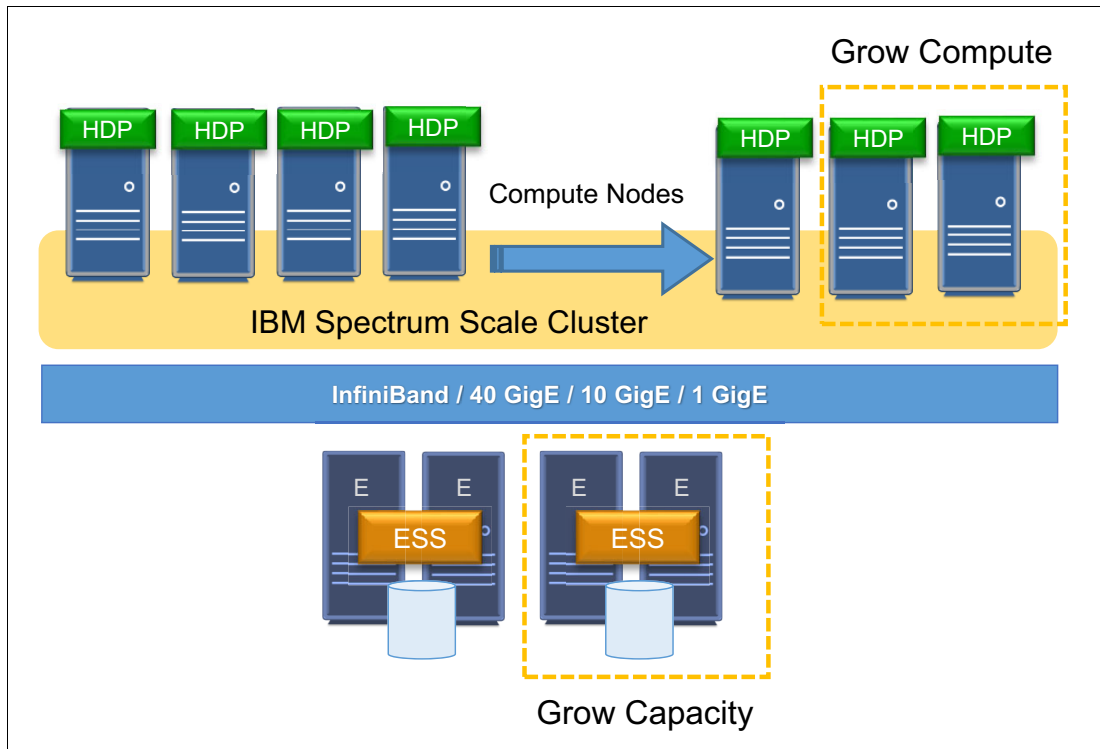


Figure 12 IBM Elastic Storage Server scaling

Other preferred practices

Consider the following preferred practices while planning the deployment of HDP with IBM Elastic Storage Server:

► Tiering

IBM Spectrum Scale supports policy-based tiering and the ability to place metadata on separate storage from data. For performance-sensitive workloads, it is common to use solid-state storage for the file system metadata. For data, you can write policies to move file data to the flash tier for faster access. Policies can use many file attributes, including file heat, which enables you to create a policy based on how often the file is accessed, and not just on the last access.

For more information about IBM Spectrum Scale tiering, see [IBM Knowledge Center](#).

► File system block size

When creating a file system, design for two types of parameters: Those parameters that can be changed after the file system is created and those parameters that cannot. File system block size is the key parameter that must be determined at file system creation. After this parameter is set, the only way to change the block size is to re-create the file system. In a IBM Spectrum Scale file system, you can store the file metadata (inode information) on the same storage as data or on separate storage. Consider the following options:

- Store file system metadata and data on separate storage. For more information, see [IBM Knowledge Center](#).
- Here are the preferred block sizes for Hadoop workloads on an IBM Elastic Storage Server system:
 - 1 MiB for a metadata only pool
 - 8 MiB for a data only pool

► IBM Spectrum Scale Client-side settings (compute nodes)

Check that the appropriate `gssClientConfig` script was applied on clients that are added to the cluster. If the IBM Elastic Storage Server deployment infrastructure is used to add additional IBM Spectrum Scale client nodes, this application happens automatically. However, if the IBM Spectrum Scale/Ambari integration is used to add clients nodes to the IBM Elastic Storage Server cluster, the `gssClientConfig` scripts should be run manually on each compute node.

Shared Nothing Storage model

This section contains the details of deploying HDP with IBM FPO, which is a Shared Nothing Storage model.

IBM Spectrum Scale File Placement Optimizer

IBM Spectrum Scale FPO is a set of features that enables IBM Spectrum Scale to operate efficiently in a deployment that is based on a Shared Nothing Storage architecture. It is useful for big data applications that process massive amounts of data and rely on using data locality by scheduling tasks on nodes where the required data already is.

The IBM Spectrum Scale FPO enables file systems to be configured to support typical storage requirements that Hadoop applications expect:

- ▶ Large block size: IBM Spectrum Scale *chunks* enable a logical grouping of blocks to behave like one large block, which is useful for applications that need high sequential bandwidth.
- ▶ Topology-aware data replica placement: *Extended failure groups* enable the data block placement decisions about the disk selection to account for node locality. This technique enables users to specify which disks are within a node or rack and which are farther away.
- ▶ Data locality: *Write affinity depth* is a policy that enables the application to determine the layout of a file in the cluster to optimize for typical access patterns. Write affinity enables you to define the number of localized copies in the cluster (as opposed to wide striping, which is employed by default). *Write affinity failure group* is a policy that indicates the range of nodes (in a Shared Nothing architecture) where replicas of blocks in a file are to be written. The policy enables the application to determine the layout of a file in the cluster to optimize for typical access patterns.

IBM Spectrum Scale extended failure group and write affinity depth features keep the data closer to your compute workload. IBM Spectrum Scale auto recovery handles disk and node failure automatically. For more information, see [IBM Knowledge Center](#).

Scenarios for the Shared Nothing Storage model

A Shared Nothing Storage model architecture can be beneficial in the following situations:

- ▶ You start small with a minimum infrastructure investment.
- ▶ You have clusters with a few compute nodes (typically less than 20). Larger clusters of storage-rich servers are susceptible to multiple node failures, which result in data rebuilds that can decrease overall cluster performance.
- ▶ You have small storage clusters. The total storage requirements are projected to grow up to 250 TB of capacity.
- ▶ You have a data locality requirement for running analytics tasks. Hadoop workloads that are run on commodity hardware without access to high-performance network interconnects rely on data locality to achieve optimal performance. These types of clusters typically lend themselves to an FPO Shared Nothing Storage deployment.
- ▶ You have a dedicated cluster for locality aware analytics. If a cluster is expected to be dedicated for running analytics workloads alone, an FPO deployment is preferred. However, if Hadoop workloads are expected to run alongside and coexist with other traditional workloads or if in-place analytics is a requirement, a Shared Storage architecture is preferred.

Implementation guidelines

The following sections describe the architecture and implementation guidelines when implementing HDP with IBM Spectrum Scale FPO.

Cluster configuration

In an IBM Spectrum Scale FPO + HDP deployment model, IBM Spectrum Scale and IBM Spectrum Scale HDFS Transparency are both installed on each node, as shown in Figure 13. IBM Spectrum Scale creates a distributed file system with internal disks present in each node and provides I/O service to Hortonworks HDP.

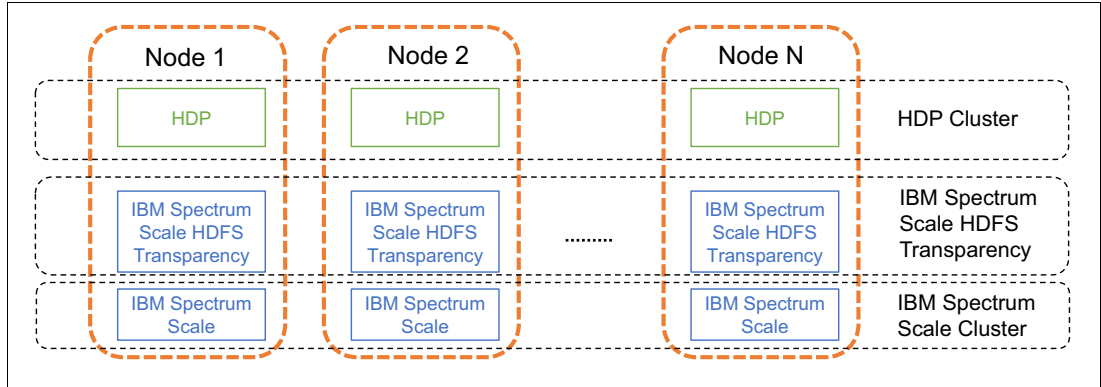


Figure 13 HDP + IBM Spectrum Scale FPO cluster configuration

System design

In the architecture shown in Figure 14, IBM Spectrum Scale is deployed in a Shared Nothing architecture running over storage-rich servers and a set of CES nodes (minimum of two) that are dedicated as protocol nodes. Protocol nodes are optional based on your requirement of having protocols, such as NFS, SMB, and Object.

Figure 14 shows the IBM Spectrum Scale configuration types that are needed on the nodes. For more information about IBM Spectrum Scale configurations, see [IBM Knowledge Center](#).

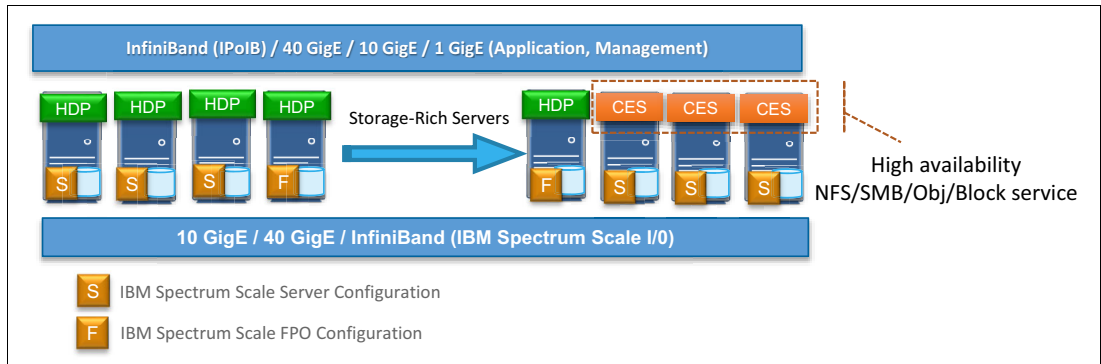


Figure 14 HDP + IBM Spectrum Scale with protocol nodes

Network

Use a high-speed network, such as 10 GigE, 40 GigE, or InfiniBand for the IBM Spectrum Scale I/O network. To maintain data copies, IBM Spectrum Scale FPO sends data to the other nodes for each write operation. A high-speed network helps IBM Spectrum Scale FPO provide better performance. In general, it is preferable (but not mandatory) to deploy a different network for IBM Spectrum Scale I/O and Hortonworks application and management.

Data protection

IBM Spectrum Scale in an FPO deployment uses replication to provide data redundancy. Three-way replication is preferable in an FPO model.

Scaling

You can add nodes to extend compute and storage capability at the same time, as shown in Figure 15. To simplify management and balance compute and storage performance, add nodes that have the same hardware configuration as existing ones.

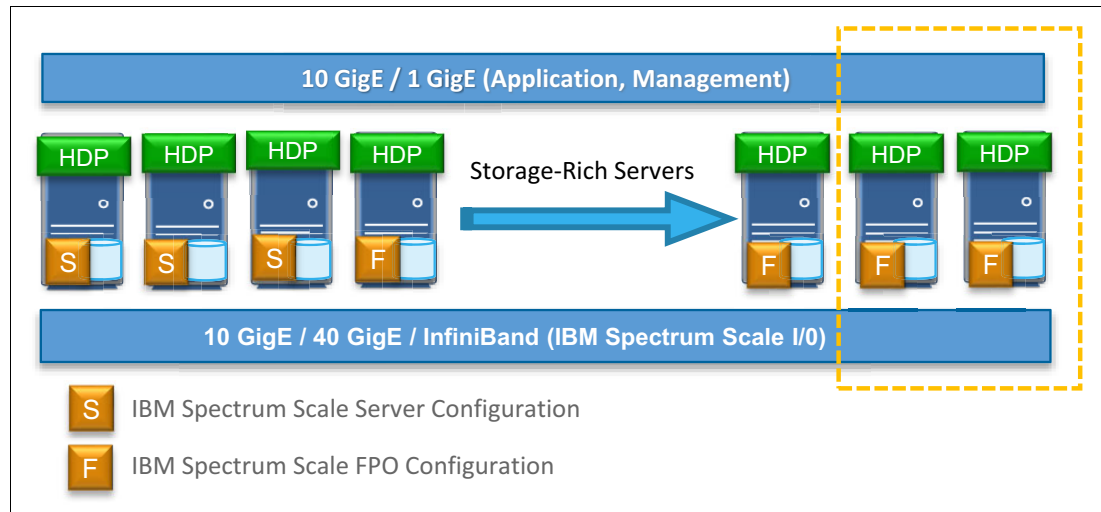


Figure 15 FPO scaling

After a new node joins the cluster, new data is striped to the whole cluster immediately. The decision to balance existing data across the whole cluster depends on data lifecycle and performance requirements. If the data will be removed some day and disk space is not a problem, it is not necessary to balance the old data. Otherwise, you can follow the guide to balance the data.

For more information, see [Big Data Best practices](#).

Other preferred practices

The following are some preferred practices to consider while planning the deployment of HDP with IBM Spectrum Scale FPO:

► Tuning

For more information about IBM Spectrum Scale FPO preferred practices, see [Big Data Best practices](#).

For IBM Spectrum Scale tuning preferred practices for a Shared-Nothing Cluster (SNC) environment, see [IBM Spectrum Scale Tuning Recommendations for Shared Nothing Environments](#).

Aligning the IBM Spectrum Scale data block size and chunk size with your workload I/O size can provide better performance.

► SSD consideration for metadata or tiering

Generally use SSD for IBM Spectrum Scale metadata in an FPO deployment model. It improves IBM Spectrum Scale performance with fast metadata operations, and reduces the time that it takes for maintenance operations such as replacing a broken disk drive and recovery from disk or node failure.

An SSD has lower latency compared to SAS or SATA disks. You can create a storage pool based on SSD for metadata and data usage. When your application requires a lower latency I/O response for a file or set of files, place data into the storage pool with SSD by using the policy engine.

Multiple IBM Spectrum Scale file systems support

Multiple IBM Spectrum Scale file system support is designed to make multiple IBM Spectrum Scale file systems service the same Hadoop cluster to save on IT costs.

As shown in Figure 16, an IBM Spectrum Scale FPO file system and an IBM Spectrum Scale ESS file system service the same Hadoop cluster through HDFS Transparency.

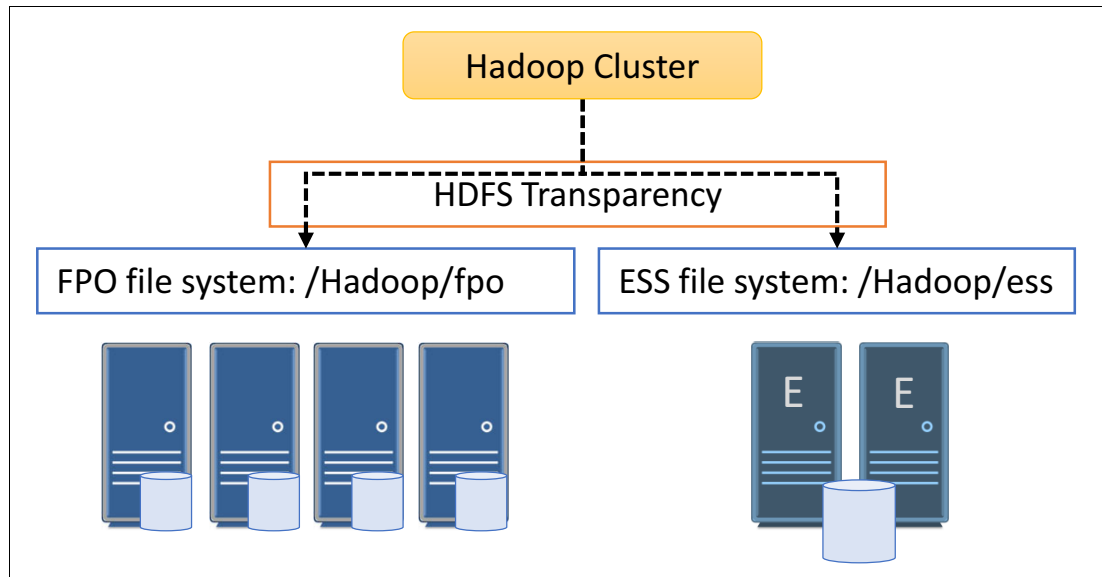


Figure 16 IBM Spectrum Scale FPO and ESS file systems service the same Hadoop cluster

System configuration

This section describes configuration preferred practices when running HDP on IBM Spectrum Scale.

Preferred software version levels

Table 1 shows the preferred software version levels.

Table 1 Preferred software version levels

Component	Preferred release level	More information
HDP on IBM Spectrum Scale Integration of HDP, Ambari, Ambari Management Pack (Mpack), HDFS Transparency, and IBM Spectrum Scale file system	See the wiki support matrix for the latest information	IBM Knowledge Center for HDP 2.6 with IBM Spectrum Scale
IBM Spectrum Scale Hadoop Transparency Connector	2.7.3	IBM Knowledge Center for IBM Spectrum Scale support for Hadoop
HDP	2.6	HDP Documentation
IBM Spectrum Scale Client	4.2.3 or later	IBM Knowledge Center for IBM Spectrum Scale
IBM Spectrum Scale Protocol Nodes (Optional)	4.2.3 or later	IBM Knowledge Center for IBM Spectrum Scale
IBM Elastic Storage Server	5.3 or later	Introducing IBM Spectrum Scale RAID

Other system configuration details

Consider the following system configuration details:

▶ IBM Elastic Storage Server nodes

IBM Elastic Storage Server I/O nodes and the EMS run Red Hat Enterprise Linux 7 in an IBM Power Systems architecture. For IBM Elastic Storage Server specifications, see [IBM Elastic Storage Server: Specifications](#).

▶ HDP Compute Nodes running IBM Spectrum Scale Client software + IBM Spectrum Scale Transparency Connector (Shared Storage deployment model)

HDP compute nodes can be of any Linux version that is supported by IBM Spectrum Scale and Hortonworks HDP.

For HDP Compute nodes hardware requirements, see [Hardware Recommendations for Apache Hadoop](#).

The following are the preferred operating systems:

- 64-bit Red Hat Enterprise Linux (RHEL) 7 and later
- 64-bit SUSE Linux Enterprise Server 12, SP1

Note: When using shared storage, you are not required to have storage-rich servers for your compute nodes because data is centrally stored in IBM Elastic Storage Server.

For preferred performance, reserve 20% physical memory or up to 20 GB memory when you configure more than a 100 GB pagepool for IBM Spectrum Scales:

- ▶ HDP Compute Nodes with IBM Spectrum Scale Server

Table 2 contains references that are related to the supported operating systems.

Table 2 References to supported operating systems

OS support	Supported OS recommendations
Preferred operating systems	64-bit Red Hat Enterprise Linux (RHEL) 7 and later 64-bit SUSE Linux Enterprise Server 12, SP1
HDFS Transparency supported Linux version	IBM Spectrum Scale Wiki - 2nd generation HDFS Transparency Protocol Support
Hortonworks supported operating systems	Hortonworks - Meet Minimum System Requirements

- ▶ IBM Spectrum Scale protocol nodes (CES nodes)

Supported on RHEL 7.x and later.

The protocol function (NFS/SMB/Object) is software-only delivery, so the capability and performance is based on the configuration that you choose. If you are going to enable only one of either NFS or Object, have a minimum of 1 CPU socket server of the latest POWER or Intel variety with at least 64 GB of memory. If you are going to enable multiple protocols or if you enable SMB, have a minimum two CPU socket server of the latest POWER or Intel variety with at least 128 GB of memory. Network configuration is important, so use at least a 10 Gb Ethernet connection for protocol access.

For more information about hardware requirements for IBM Spectrum Scale Protocol services, see [IBM Spectrum Scale Frequently Asked Questions and Answers](#).

HDP and IBM Spectrum Scale frequently asked questions

The following are some HDP and IBM Spectrum Scale frequently asked questions:

- ▶ Is HDF supported by HDP in IBM Spectrum Scale environments?
Yes. Hortonworks is paper-certified HDF 3.0 with IBM Spectrum Scale. The paper certification means that currently Hortonworks does not believe that additional testing with HDF is required.
- ▶ Can a IBM Spectrum Scale pool that is used for HDP be tiered to IBM Cloud Object Storage by using the Transparent Cloud Tiering feature?
Yes. If it is just for using the Cloud tier for archival, then you should be okay. You cannot use any of the data in the Cloud tier actively in HDP.
- ▶ Can Hortonworks SmartSense work with HDP + IBM Spectrum Scale solution?
SmartSense works as normal with all the HDP components in the HDP + IBM Spectrum Scale solution. There are no special hooks in SmartSense to get IBM Spectrum Scale related diagnostic information. Therefore, IBM Spectrum Scale diagnosis must be performed outside of SmartSense.
- ▶ Can the HDP Atlas and Ranger components work as normal in an HDP + IBM Spectrum Scale solution?
Yes.

- ▶ Can IBM Spectrum Scale be used with HDP running on Power Systems and x86 servers?
Yes. IBM Spectrum Scale/IBM Elastic Storage Server is certified and supported to work with HDP on Power Systems and x86 server.
- ▶ Is there any additional sizing guidance that is available for building an HDP + IBM Spectrum Scale solution by using IBM hardware?
For information and assistance about sizing and configuring the HDP on a Power Systems + IBM Spectrum Scale/IBM Elastic Storage Server solution, contact the Cognitive Systems Solution Center (cssc@us.ibm.com).

Additional references

- ▶ IBM Knowledge Center for IBM Spectrum Scale
https://www.ibm.com/support/knowledgecenter/en/STXKQY_4.2.3/ibmspectrumscale423_welcome.html
- ▶ IBM Knowledge Center for IBM Elastic Storage Server
https://www.ibm.com/support/knowledgecenter/en/SSYSP8/sts_welcome.html
- ▶ *IBM Spectrum Scale Security*, REDP-5426
<http://www.redbooks.ibm.com/abstracts/redp5426.html>
- ▶ Download Hadoop Transparency Connector
<https://ibm.biz/BdiUME>
- ▶ IBM Spectrum Scale Protocol quick overview
[https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/General%20Parallel%20File%20System%20\(GPFS\)/page/Protocols%20Quick%20overview%20for%20IBM%20Spectrum%20Scale](https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/General%20Parallel%20File%20System%20(GPFS)/page/Protocols%20Quick%20overview%20for%20IBM%20Spectrum%20Scale)
- ▶ *Enterprise Data Warehouse Optimization with Hadoop on IBM Power Systems Servers*, REDP-5476
<http://www.redbooks.ibm.com/abstracts/redp5476.html>
- ▶ Hortonworks documentation
<https://hortonworks.com/partner/ibm>

Contributors

This paper was produced by a team of specialists from around the world working with the International Technical Support Organization, Tucson Center.

Thanks to the following people for their contributions to this project:

Larry Coyne, Wei G. Gong, Sandeep R. Patil, Piyush Chaudhary, Muthu Muthiah, Yong ZY Zheng, Pallavi Galgali, Theodore Hoover Jr.

IBM Systems

Doug Reid, Prasad Menon, Mike Hovermale
Hortonworks

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Stay connected to IBM Redbooks

- ▶ Find us on Facebook:
<http://www.facebook.com/IBMRedbooks>
- ▶ Follow us on Twitter:
<http://twitter.com/ibmredbooks>
- ▶ Look for us on LinkedIn:
<http://www.linkedin.com/groups?home=&gid=2130806>
- ▶ Explore new IBM Redbooks® publications, residencies, and workshops with the IBM Redbooks weekly newsletter:
<https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm>
- ▶ Stay current on recent Redbooks publications with RSS Feeds:
<http://www.redbooks.ibm.com/rss.html>

Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.


COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at <http://www.ibm.com/legal/copytrade.shtml>

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

Redbooks (logo) ®
IBM®
IBM Cloud™
IBM Elastic Storage™

IBM Spectrum™
IBM Spectrum Scale™
POWER®
Power Systems™

Redbooks®
Redpaper™

The following terms are trademarks of other companies:

Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.



REDP-5448-01

ISBN 0738456276

Printed in U.S.A.

Get connected

