Redbooks
ibm.com/redbooks

Monitoring Overview for IBM Spectrum Scale and IBM Elastic Storage Server

Kedar Karmarkar

Kaustubh Katruwar

Helene Wassmann

Cloud

Storage

IBM

Redpaper

International Technical Support Organization

Monitoring Overview for IBM Spectrum Scale and IBM Elastic Storage Server

July 2017

REDP-5418-00

Note: Before using this information and the product it supports, read the information in “Notices” on page v.

First Edition (July 2017)

This edition applies to Version 4, Release 2, Modification 2 of IBM Spectrum Scale.

This document was created or updated on July 28, 2017.

Contents

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.

developerWorks®
GPFS™
IBM®
IBM Elastic Storage™
IBM Spectrum™
IBM Spectrum Control™
IBM Spectrum Scale™
Power Systems™
Redbooks®
Redpaper™
Redbooks (logo)  ®
Storwize®
Tivoli®

The following terms are trademarks of other companies:

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

Preface

IBM® Spectrum Scale is software-defined storage for high-performance, large-scale workloads. IBM Spectrum™ Scale is a scalable data and file management solution that provides a global namespace for large data sets along with several enterprise features. IBM Spectrum Scale™ is used in clustered environments and provides file protocol (POSIX, NFS, and SMB) and object protocol (Swift and S3) access methods.

IBM Elastic Storage™ Server (ESS) is a software-defined storage system that is built upon proven IBM Power Systems™, IBM Spectrum Scale software, and storage enclosures. ESS allows for capacity scale up or scale out for performance in modular building blocks, which enables sharing for large data sets across workloads with unified storage for file, object, and Hadoop workloads. ESS uses erasure coding-based declustered RAID technology that was developed by IBM to rebuild failed disks in minutes instead of days.

IBM ESS and IBM Spectrum Scale are implemented in scalable environments that are running enterprise workloads. ESS and IBM Spectrum Scale are key components of the enterprise infrastructure. With growing expectations of availability on enterprise infrastructures, monitoring IBM Spectrum Scale, ESS health, and performance is an important function for any IT administrator.

This IBM Redpaper™ publication provides an overview of key parameters and methods of IBM Spectrum Scale and ESS monitoring.

The audience for this document is IT architects, IT administrators, storage administrators, and users who want to learn more about the administration of an IBM Spectrum Scale and ESS system.

This document can be used for the environments with IBM Spectrum Scale version 4.2.0 or later. The examples in the document are based on IBM Spectrum Scale 4.2.2.X and ESS 5.0.X.X versions.

Authors

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

Kedar Karmarkar is a Senior Engineer and Solution Architect with the IBM Spectrum Scale development team. Kedar is part of the IBM Spectrum Scale Client adoption team and was the IBM Storwize® V7000 Unified Level 3 support lead in his earlier role at IBM. Kedar has over 20 years of infrastructure software, storage development experience in management, and architect roles. He has led development of network-attached storage (NAS), Block level virtualization, replication, systems, and storage management products. Kedar has a Bachelor of Engineering (Computer Science) degree from University of Pune, India.

Kaustubh Katruwar is a Lead Engineer with the IBM Spectrum Scale development team. He leads FILE protocols (SMB/NFS) authentication efforts for this software-defined storage solution in strategizing, driving, designing, and developing integration with various authentication Directory Services. He also drives the authentication component for IBM Storwize V7000 Unified and IBM Scale Out NAS products, and participates in Openstack Object Store authentication service integration on IBM Spectrum Scale. He has over 9 years of experience and is a subject matter expert on FILE protocols, cross protocol access, authorization, and security aspects of scale-out parallel file system-based access storage. Kaustubh has a Bachelor of Engineering (Computer Science) degree from University of Pune, India.

Helene Wassmann is a Software Engineer from Germany who joined IBM in 1999. She has more than 10 years of experience with the IBM Spectrum Control™ product. During her time with IBM Spectrum Control, she held various roles across the software development process: Test, L3-Support, data layer, and back-end development. In March 2016, she joined the IBM Spectrum Scale development team. Her areas of expertise are System Health and Performance Data Management. Helene is Project Management Professional (PMP) certified since 2005.

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

Larry Coyne
International Technical Support Organization, Tucson Center

Norbert Schuld
Scott Fadden
Pallavi Galgali
Simon Lorenz
Sandeep R. Patil
IBM Systems

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

Comments welcome

Your comments are important to us!

We want our papers to be as helpful as possible. Send us your comments about this paper or other IBM Redbooks® publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

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 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

1

Health monitoring overview

This IBM Redpaper publication provides an overview of monitoring methods that are available for IBM Spectrum Scale and IBM Elastic Storage Server (ESS). The IBM Spectrum Scale software is one of the software components of ESS. In this paper (unless explicitly stated) all IBM Spectrum Scale-related monitoring methods can be used for monitoring ESS. Some of the sections in this publication are applicable to ESS only and the paper clearly calls out those sections.

For more information about IBM Spectrum Scale concepts and planning, see the IBM Spectrum Scale page of IBM Knowledge Center.

For more information about Elastic Storage Server, see the IBM Spectrum Scale page of IBM Knowledge Center.

IBM Spectrum Scale and IBM Elastic Storage Server (ESS) are deployed in enterprise workload environments with stringent availability, reliability, performance expectations from the infrastructure. Therefore, IBM Spectrum Scale infrastructure and its components must be monitored closely to help prevent any unplanned events that can result into storage and data access related issues.

This Redpaper publication broadly divides monitoring activities into two categories: Health monitoring and performance monitoring. This Redpaper publication focuses more on health monitoring, although performance monitoring also is described briefly.

Note: This Redpaper publication is focused on what should be monitored in an IBM Spectrum Scale environment and the various methods that are available to monitor the events, unplanned issues, performance, and so on. For more information about the actions to remediate any issues that are discovered through monitoring, see the Troubleshooting sections of the IBM Spectrum Scale page of IBM Knowledge Center.

IBM Spectrum Scale provides performance acceleration features and file system configuration options to increase availability and performance per your requirements. These features must be implemented and the options configured according to the best practices or recommendations that are provided in the IBM Spectrum Scale Knowledge Center page of IBM Knowledge Center or other documentation, such as IBM Redbooks publications and performance papers.

This chapter describes the various components that must be monitored to monitor the health of an IBM Spectrum Scale environment. How IBM Spectrum Scale GUI and CLI provide an easy way to set up the monitoring is described in Chapter 2, “Health monitoring methods” on page 7.

This chapter includes the following topics:

1.1 Cluster health

IBM Spectrum Scale cluster consists of various types of nodes in the cluster. Therefore, the health of the cluster depends on the health of nodes in the cluster. The following node types can affect the health of a cluster:

- Quorum node

  In the IBM Spectrum Scale cluster, quorum must be maintained to keep the cluster online. If quorum is not maintained because of node failure, IBM Spectrum Scale unmounts the file systems on the remaining nodes and attempts to reestablish the quorum. Therefore, each quorum node should be monitored for failures and appropriate actions should be taken if the quorum node fails to avoid the loss of quorum.

- NSD server node

  In a cluster that typically runs workloads on client nodes, NSD server nodes provide access to NSD disks. It is recommended that NSD disks be configured in a way where multiple NSD servers can access the NSD disks. If all NSD servers that provide access to NSD disks are down, it can result in loss of data access in certain configurations. Therefore, NSD server nodes must be monitored for failures and appropriate actions must be taken to get them back up quickly.

- Cluster manager node

  It is recommended that you configure primary and secondary cluster manager nodes within the cluster. If the primary cluster manager node fails when the secondary cluster manager is not configured, no administrative or configuration commands can be run in the cluster. Therefore, monitoring the cluster manager node is required to ensure that IBM Spectrum Scale management is not affected because of failed cluster manager nodes.

You must know which of the nodes in the cluster are configured as the quorum nodes, NSD server nodes, cluster manager nodes, or file system manager nodes. These nodes are monitored closely to maintain the health of the IBM Spectrum Scale cluster.

It is possible that a node cannot join the cluster or stay in the cluster because of some issue with the IBM Spectrum Scale daemon on the node. The state of the IBM Spectrum Scale daemon on the node should be monitored along with node health.

Another reason a node might not join the cluster can be that the IBM Spectrum Scale daemon on the node cannot communicate with IBM Spectrum Scale daemons that are running on other nodes. You can check for network issues or firewall configuration on the node or network as a possible cause for the communication failure.

1.2 File system health

Some of the file system state and configuration issues affect only the state of a specific file system and not all file systems that are configured in the IBM Spectrum Scale cluster. Such file system configuration parameters and state can be monitored from any of the IBM Spectrum Scale nodes that can run IBM Spectrum Scale administration commands, including the following examples:

- Configuration parameters

  Warnings or errors that are related to data and metadata space, the number of inodes in file system and filesets, or to fileset, user level quota usage.

- File system state

  File system unexpected unmounts, file system errors, inconsistent mount states, and file system descriptor quorum.

1.3 Protocol health

One or more Cluster Export Services (CES) nodes must be configured if you want to access the data by using the Network File System (NFS), Server Message Block (SMB), or Object protocols. You should monitor the health of such CES nodes to ensure that the protocol services are available in a IBM Spectrum Scale cluster.

Protocol services use the authentication server that is configured during protocol configuration. The authentication server, such as Active Directory, Lightweight Directory Access Protocol (LDAP), Network Information System (NIS), and Keystone, should be monitored along with their components.

The following protocol-specific components must be monitored to ensure that the protocol-specific services are available:

- NFS: NFS Ganesha daemon, Portmapper, statd (v3), and so on
- SMB: SMB daemon and port, ctdb status, and so on
- Object: Swift and keystone processes, postgreSQL, Ringfile checksum, and so on

In addition, protocol IPs should be monitored and checked to ensure that IPs are up and failing over if CES node failures or other protocol service failovers occur.

The protocol-related events are logged on the CES node and can be monitored by using the command line.

1.4 Network health

IBM Spectrum Scale cluster configurations contain a mix of SAN-attached and network attached cluster nodes. IBM Spectrum Scale nodes can share data within a location or across wide area network (WAN) connections. The CES or protocol nodes provide the protocol services by using CES IP addresses that are defined for protocol usage.

Network communication between nodes and with IBM Spectrum Scale services is a critical component of the IBM Spectrum Scale configuration. The following various network parameters must be monitored:

- Communication network between the nodes

  The nodes are configured to communicate with each other over cluster IP addresses. Network communication can encounter issues if the firewall is incorrectly configured between the nodes. The physical network and network components on the node must be monitored.

- Protocol CES IPs

  As described in 1.3, “Protocol health” on page 4, Protocol CES IP communication between protocol clients and CES nodes and between CES nodes must be monitored. In addition, the failover of Protocol CES IPs must be monitored.

The IBM Spectrum Scale client nodes send all I/O requests and receive read/write response over the network. Any issues with network performance directly affect I/O latency in the IBM Spectrum Scale cluster.

1.5 Disk health

In IBM Spectrum Scale configuration, the storage subsystem provides virtual disks (VDisks) or LUNs, which are used by IBM Spectrum Scale to create Network Service Disk (NSDs). IBM Spectrum Scale is independent of the storage configuration beneath the NSDs. The only way IBM Spectrum Scale can come to know about any issues with underlying storage is through I/O errors, which are warnings that are reported as a result of a read or write operation. You should monitor the storage subsystem to help prevent such I/O errors.

1.6 ESS system health

Elastic Storage Server consists of two Power-based I/O servers (or NSD servers) and storage enclosures that are managed by using IBM erasure-coding based declustered RAID system. The native RAID system uses Disk Hospital, which asynchronously detects and (if possible) fixes any issues with disks within the enclosures. Therefore, ESS provides a much more granular disk level view for storage subsystem-level monitoring. This monitoring should be used for overall monitoring of ESS storage subsystem health.

2

Health monitoring methods

This chapter describes the various methods and commands that are available in IBM Spectrum Scale and IBM Elastic Storage Server to set up monitoring. As described in Chapter 1., “Health monitoring overview” on page 1, monitoring should be set up for various components of IBM Spectrum Scale.

This chapter includes the following topics:

2.1 Monitoring health by using the GUI or CLI

IBM Spectrum Scale made significant improvements in the system health monitoring functionality, starting with release 4.2.1. A new infrastructure was introduced to monitor many components.

This system health monitoring infrastructure feeds the events data to IBM Spectrum Scale GUI and mmhealth command-line interface (CLI). The mmhealth CLI is now a single CLI for all health monitoring functionality to view IBM Spectrum Scale health status. The components that are monitored by the system health infrastructure are shown in Figure 2-1.

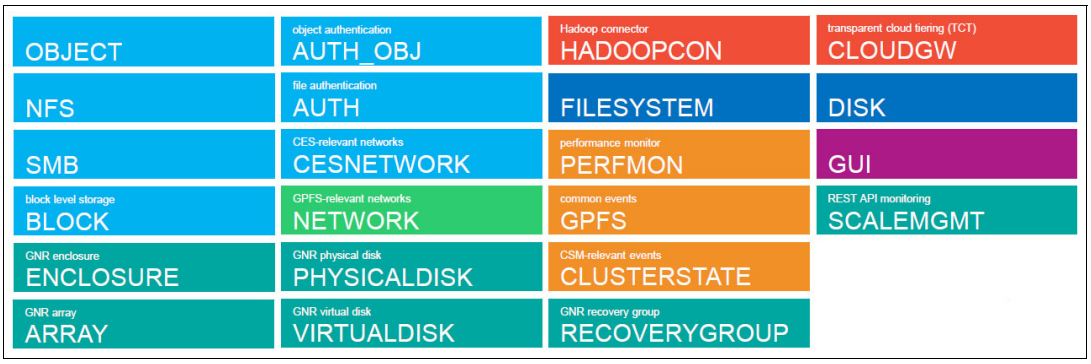

Figure 2-1 Components that are managed by System Health monitoring

2.1.1 Cluster health

The IBM Spectrum Scale GUI provides a monitoring menu for monitoring IBM Spectrum Scale cluster health and performance. The Home page and Dashboard page under Monitoring provides a consolidated view of system health and performance. The Dashboard can be customized in different layouts and to display widgets that display various health and performance data. You can use the Home page and Dashboard page as starting points to view the overall health of the IBM Spectrum Scale cluster.

The Home page of IBM Spectrum Scale GUI, which shows IBM Spectrum Scale health and configuration summary, is shown in Figure 2-2.

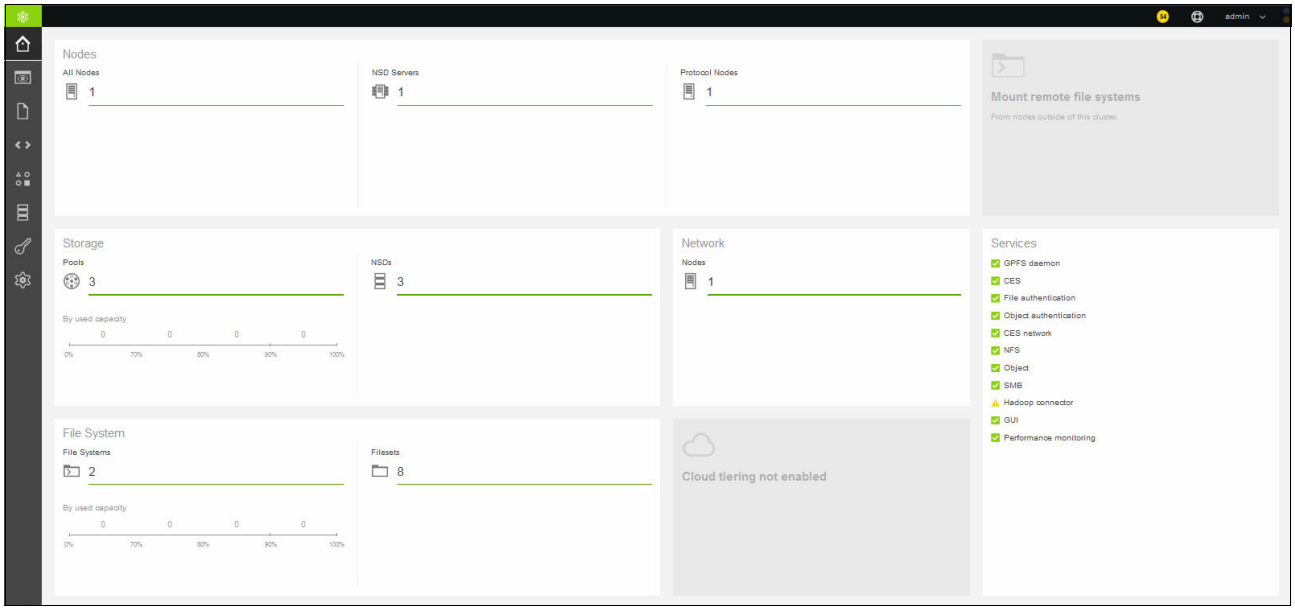

Figure 2-2 IBM Spectrum Scale Home page

The Spectrum Cluster health in the System Overview and System Health Events widgets is shown in Figure 2-3.

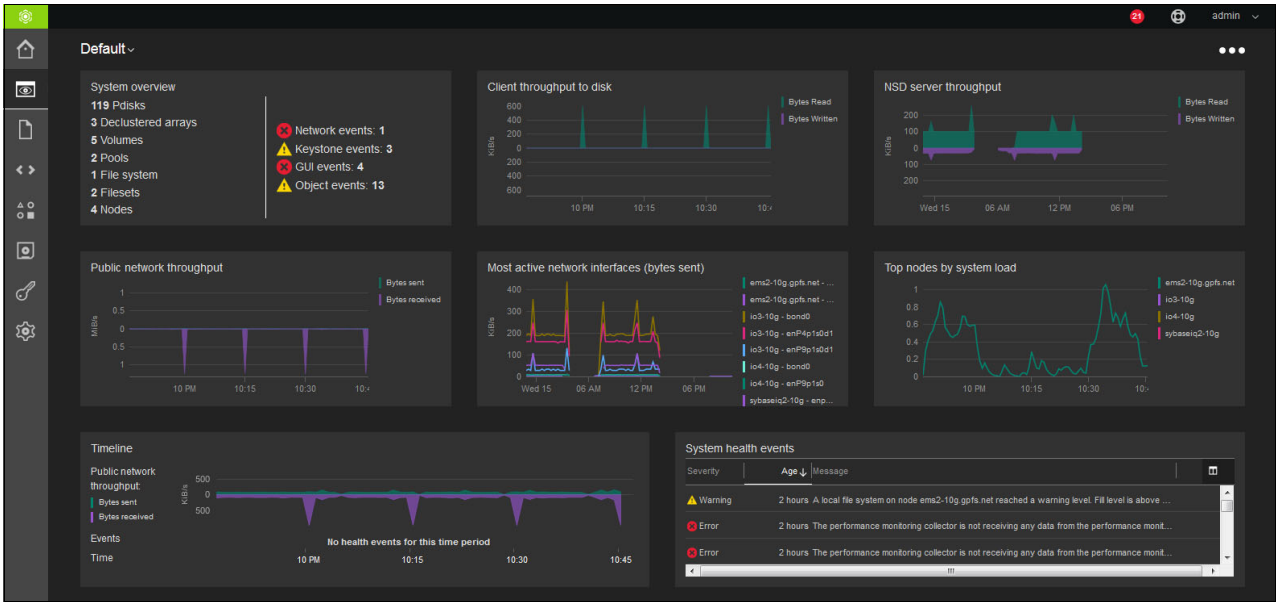

Figure 2-3 Default Dashboard view

The Events page under Monitoring displays all events that are reported in the system. The status bar on upper right corner of the GUI page also displays current Events.

You can use the Events page to get more information about current and past events that were reported in the IBM Spectrum Scale cluster. By default, the events page displays Current events. It also can display Unread Events and All events, as shown in Figure 2-4.

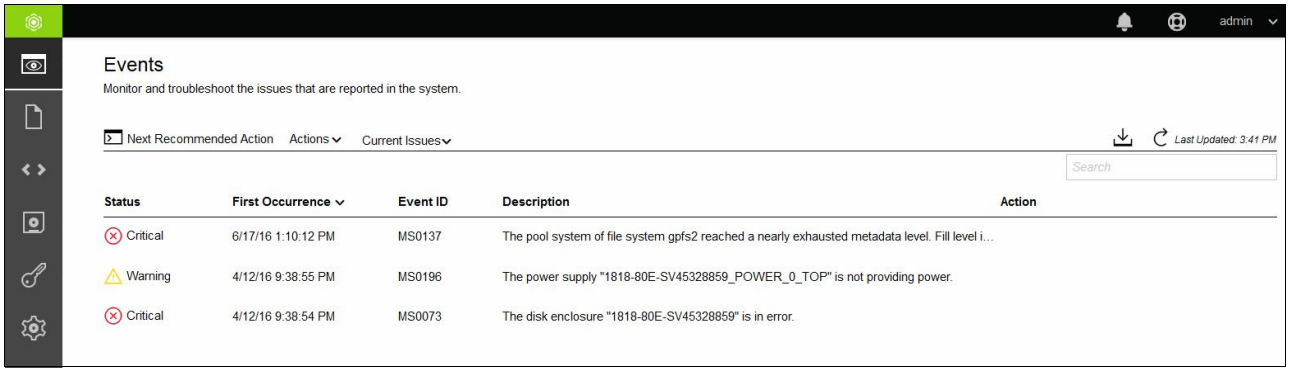

Figure 2-4 Events page details

The events are of Critical, Warning, and Information severity and can be sorted by the same status. More information is available about each of the events by clicking **Properties** under **Actions**, as shown in Figure 2-5.

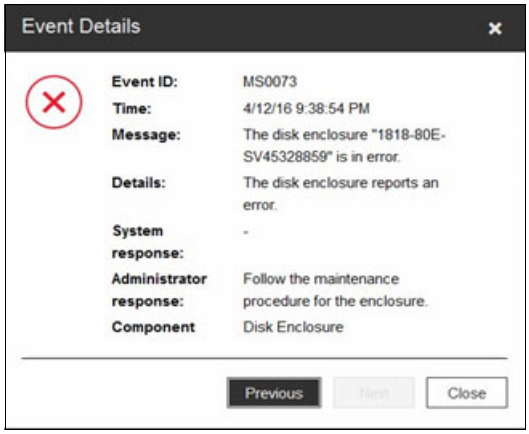

Figure 2-5 Event details

The health of the cluster depends on the health of the nodes in the cluster, especially quorum nodes and manager nodes in the cluster. The Nodes page under Monitoring shows the health of these nodes, as shown in Figure 2-6.

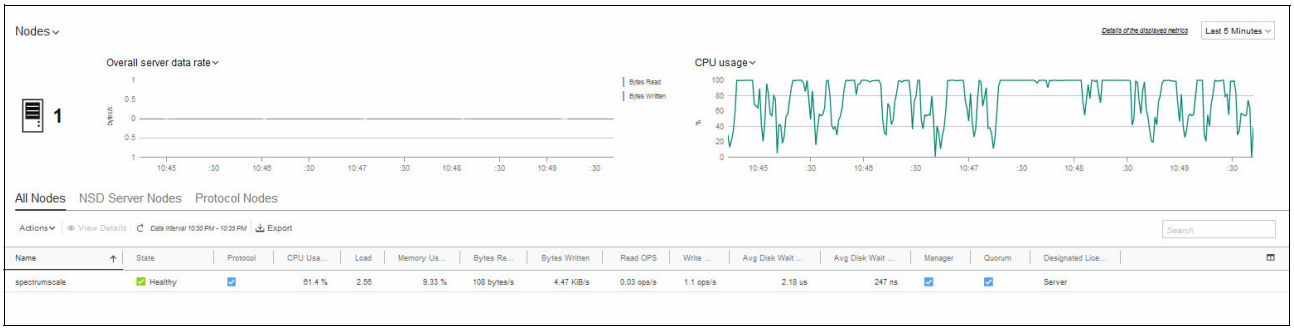

Figure 2-6 Nodes Health Monitoring page

Monitoring cluster health by using the CLI

The **mmhealth** command can be used for overall health monitoring of the cluster. By using the **mmhealth** command, the status of various components in the cluster is shown, including nodes, file system, protocols, network, NSDs, and physical disks. The use of the **mmhealth cluster show** command displays the overall state of the cluster, as shown in Example 2-1.

Example 2-1 Displaying overall state of the cluster

```
# mmhealth cluster show

Component       Total       Failed    Degraded     Healthy       Other
-------------------------------------------------------------------------
NODE              4           0         1            3            0
GPFS              3           0         0            3            0
NETWORK           4           1         0            3            0
FILESYSTEM        1           0         0            1            0
DISK              2           0         0            2            0
CES               1           0         0            1            0
GUI               1           0         1            0            0
```

```
NATIVE_RAID      2           0         0          2          0
PERFMON          4           0         0          4          0
```

Use the **mmhealth node show** command to view the health of a specific node with the status of various components on the node, as shown in Example 2-2.

Example 2-2 Viewing mmhealth node status

```
# mmhealth node show –N io3-10g
Node name:      io3-10g
Node status:    HEALTHY
Status Change:  20 hours ago

Component       Status        Status Change     Reasons
------------------------------------------------------------------------
GPFS            HEALTHY       21 hours ago      -
NETWORK         HEALTHY       13 hours ago      -
FILESYSTEM      HEALTHY       21 hours ago      -
DISK            HEALTHY       21 hours ago      -
NATIVE_RAID     HEALTHY       21 hours ago      -
PERFMON         HEALTHY       20 hours ago      -
```

You can use **options -verbose** command to view more information about each component status and **-N all** command to view more information about all of the nodes in the cluster.

You can use the **mmgetstate** command with various options, such as **-s**, **-aL**, and **-N nsdnodes**, to view state of cluster of various nodes. You can use the **mmlscluster** and **mmlsconfig** commands to check IBM Spectrum Scale configuration. Although these commands provide the necessary output, the **mmhealth** command is recommended as it provides consolidated output inclusive of states and configurations that are displayed by **mmgetstate**, **mmlscluster**, and **mmlsconfig** commands.

2.1.2 File system health

By clicking **Monitoring** → **Events** in the IBM Spectrum Scale GUI, all events (including the events that are related to File System health) are displayed. You can also get more information about file system, fileset, and quotas by clicking **Monitoring** → **Capacity** for any capacity-related events.

By clicking **Files** → **File Systems**, a page shows the file system health, including status, available capacity, and throughput data, as shown in Figure 2-7 on page 13.

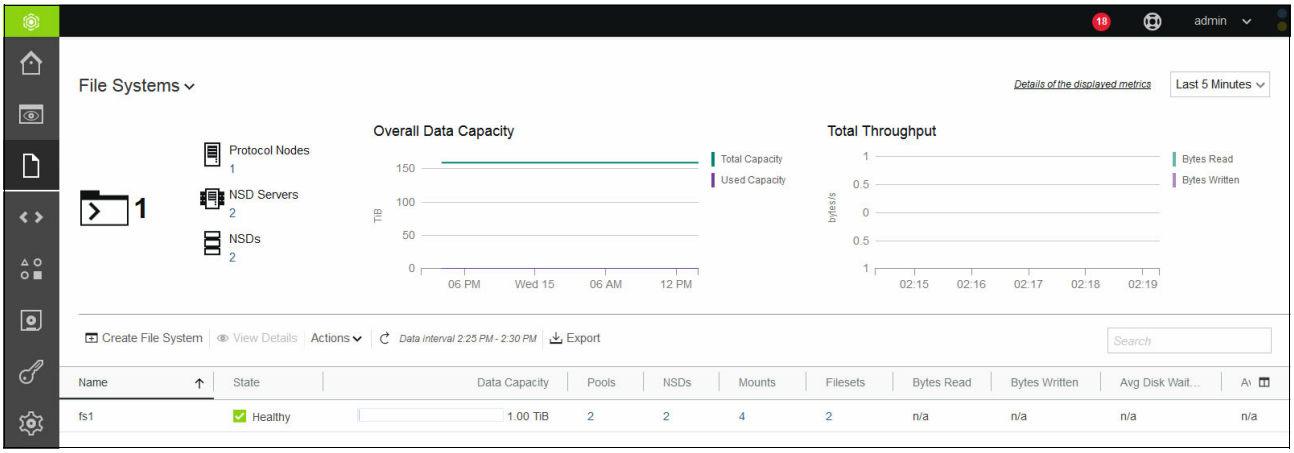

Figure 2-7 File System details page

The Filesets page provides more information about the fileset configuration, as shown in Figure 2-8.

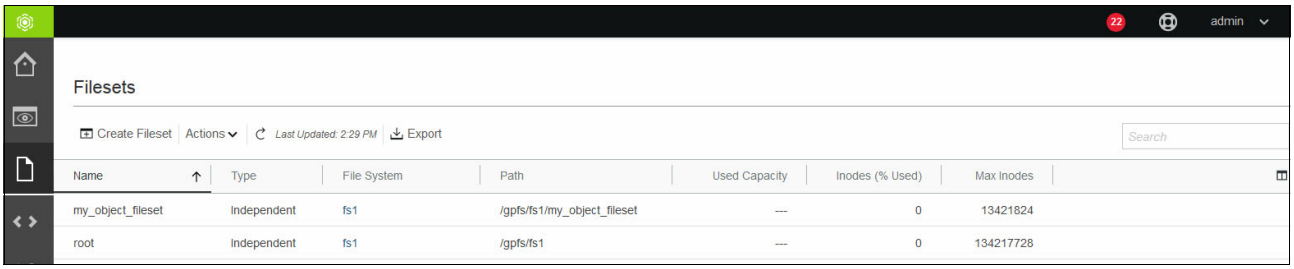

Figure 2-8 Fileset details page

Monitoring file system health by using the CLI

You can use the **mmhealth** command to monitor IBM Spectrum Scale daemon health on all nodes in the cluster (see Example 2-3).

Example 2-3 The mmhealth on all nodes in cluster

```
# mmhealth cluster show GPFS

Component                Node                     Status           Reasons
-------------------------------------------------------------------------------
GPFS                     io3-10g                  HEALTHY          -
GPFS                     ems2-10g.gpfs.net        HEALTHY          -
GPFS                     io4-10g                  HEALTHY          -
GPFS                     workload2-10g            HEALTHY          -
```

You can use the **mmhealth node show** command to view health of individual file systems on all nodes in the cluster (see Example 2-4).

Example 2-4 Viewing health of individual file systems

```
# mmhealth node show FILESYSTEM -N all
Node name:      ems2-10g.gpfs.net

Component        Status        Status Change     Reasons
--------------------------------------------------------------------------------
FILESYSTEM       HEALTHY       29 min. ago       -
  fs1            HEALTHY       29 min. ago       -

There are no active error events for the component FILESYSTEM on this node (ems2-10g.gpfs.net).

Node name:      io3-10g

Component        Status        Status Change     Reasons
--------------------------------------------------------------------------------
FILESYSTEM       HEALTHY       1 day ago         -
  fs1            HEALTHY       1 day ago         -

There are no active error events for the component FILESYSTEM on this node (io3-10g).

Node name:      io4-10g

Component        Status        Status Change     Reasons
--------------------------------------------------------------------------------
FILESYSTEM       HEALTHY       14 days ago       -
  fs1            HEALTHY       14 days ago       -

There are no active error events for the component FILESYSTEM on this node (io4-10g).

Node name:      workload2-10g

Component        Status        Status Change     Reasons
--------------------------------------------------------------------------------
FILESYSTEM       HEALTHY       7 days ago        -
  fs1            HEALTHY       7 days ago        -

There are no active error events for the component FILESYSTEM on this node (workload2-10g).
```

To view more information about IBM GPFS™ health on a particular node in the cluster, add the **-N** parameter (see Example 2-5).

Example 2-5 Viewing more information about GPFS health

```
# mmhealth node show GPFS -N io3-10g --verbose

Node name:      io3-10g

Component     Status        Status Change            Reasons
-------------------------------------------------------------
GPFS          HEALTHY       2017-03-14 06:31:24      -


Event               Parameter Severity Active Since        Event Message
-------------------------------------------------------------------------------------------
gpfs_up               GPFS     INFO    2017-03-14 06:31:24 The Spectrum Scale service process is running
ccr_client_init_ok    GPFS     INFO    2017-03-14 14:21:54 GPFS CCR client initialization is ok CCR_CLIENT_INIT
quorum_up             GPFS     INFO    2017-03-14 14:21:54 Quorum achieved
```

```
ccr_auth_keys_ok       GPFS     INFO    2017-03-14 14:21:54 The security file used by GPFS CCR is ok FC_CCR_AUTH_KEYS
ccr_comm_dir_ok        GPFS     INFO    2017-03-14 14:21:54 The files committed to the GPFS CCR are complete and intact FC_COMMITTED_DIR
ccr_ip_lookup_ok       GPFS     INFO    2017-03-14 14:21:54 The IP address lookup for the GPFS CCR component is ok PC_IP_ADDR_LOOKUP
ccr_local_server_ok    GPFS     INFO    2017-03-14 14:21:54 The local GPFS CCR server is reachable PC_LOCAL_SERVER
ccr_paxos_12_ok        GPFS     INFO    2017-03-14 14:21:54 The stored GPFS CCR state is ok FC_CCR_PAXOS_12
ccr_paxos_cached_ok    GPFS     INFO    2017-03-14 14:21:54 The stored GPFS CCR state is ok FC_CCR_PAXOS_CACHED
ccr_quorum_nodes_ok    GPFS     INFO    2017-03-14 14:21:54 All quorum nodes are reachable PC_QUORUM_NODES
ccr_tiebreaker_dsk_ok  GPFS     INFO    2017-03-14 14:21:54 All tiebreaker disks used by the GPFS CCR are accessible TC_TIEBREAKER_DISKS
no_longwaiters_found   GPFS     INFO    2017-03-14 14:21:54 No Spectrum Scale long-waiters
gpfsport_up            GPFS     INFO    2017-03-14 14:21:54 Spectrum Scale port 1191 is active
```

The use of the **mmhealth** command (see Example 2-6) displays the status of file system on a particular node in the cluster.

Example 2-6 Viewing the status of file system on a node in cluster

```
# mmhealth node show FILESYSTEM -N io3-10g --verbose
Node name:      io3-10g

Component               Status        Status Change            Reasons
-------------------------------------------------------------------------------
FILESYSTEM              HEALTHY       2017-03-14 06:31:24      -
  fs1                   HEALTHY       2017-03-14 06:31:24      -

Event             Parameter  Severity    Active Since        Event Message
-------------------------------------------------------------------------------------------
mounted_fs_check    fs1        INFO     2017-03-14 14:21:54   The filesystem fs1 is mounted
```

Although the **mmlsmount**, **mmlsfileset**, **mmlsfs**, and **mmlsnsd** commands can be used to check file system status and configuration, the use of the **mmhealth** command is recommended.

2.1.3 Protocol/CES health

Nodes in IBM Spectrum Scale cluster can be assigned as Cluster Export Services (CES) nodes. CES nodes are also referred as *protocol nodes*. One or more protocols, including Network File System (NFS), Server Message Block (SMB), and Object can be enabled on the protocol nodes.

By clicking **Monitoring** → **Nodes** → **Protocol nodes**, a page shows the status of the protocol nodes in the cluster, as shown in Figure 2-9.

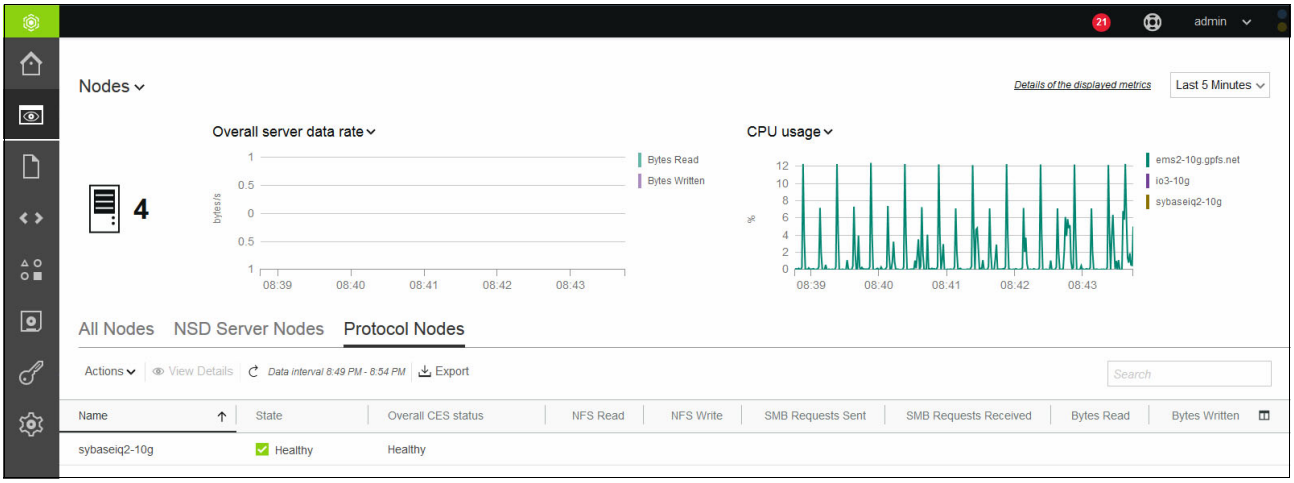

Figure 2-9 Protocol nodes status

By clicking **Settings** → **NFS Service** → **NFS Service status**, a page in the IBM Spectrum Scale GUI shows the status of NFS service on all CES nodes, as shown in Figure 2-10.

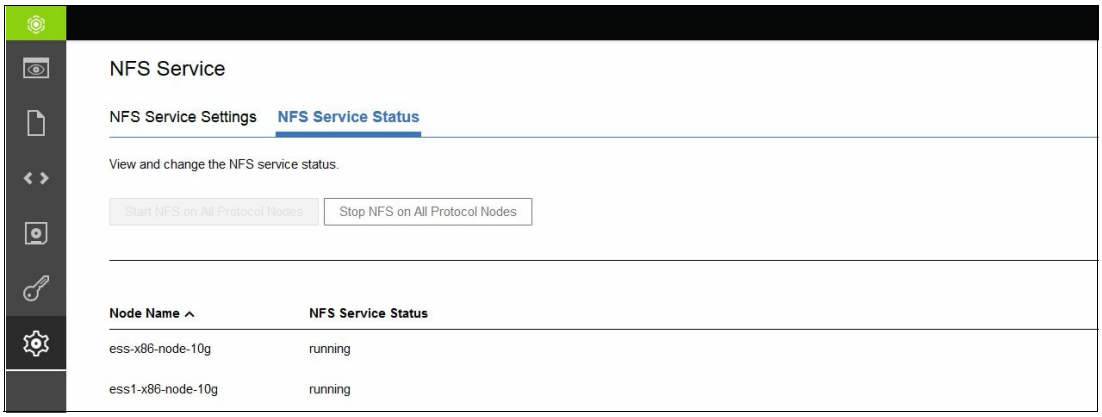

Figure 2-10 Protocol Service status

You can also view the status of other services, such as SMB and Object, by using the corresponding pages under the Settings menu.

Monitoring protocol health by using the CLI

You can view status of the protocols nodes by running the **mmhealth** command, as shown in Example 2-7.

Example 2-7 Monitoring the status of all protocol nodes

```
# mmhealth cluster show CES --verbose

Component             Node                Status           Reasons
-------------------------------------------------------------------------------
CES                   sybaseiq2-10g       HEALTHY          -
  OBJECT                                  HEALTHY          -
  CESNETWORK                              HEALTHY          -
  AUTH                                    DISABLED         -
  NFS                                     HEALTHY          -
  AUTH_OBJ                                HEALTHY          -
  SMB                                     HEALTHY          -
  BLOCK                                   DISABLED         -
```

To view the status of individual services that are related to protocols, you can use **mmhealth** command, as shown in Example 2-8.

Example 2-8 Viewing the status of individual services related to protocols

```
# mmhealth node show OBJECT --verbose -N sybaseiq2-10g

Node name:      sybaseiq2-10g

Component     Status        Status Change            Reasons
-----------------------------------------------------------
OBJECT        HEALTHY       2017-03-08 00:58:24      -

Event                          Parameter Severity  Active Since          Event Message
-------------------------------------------------------------------------------------------------------------------------
account-auditor_ok               OBJECT    INFO     2017-03-01 04:50:37   account-auditor process as expected, state is started
account-reaper_ok                OBJECT    INFO     2017-03-01 04:50:37   account-reaper process as expected, state is started
account-replicator_ok            OBJECT    INFO     2017-03-01 04:50:37   account-replicator process as expected, state is started
account-server_ok                OBJECT    INFO     2017-03-01 04:50:37   account process as expected, state is started
account_access_up                OBJECT    INFO     2017-03-01 04:50:37   Access to account service ip 0.0.0.0 port 6202 ok
```

```
container-auditor_ok              OBJECT    INFO     2017-03-01 04:50:37   container-auditor process as expected, state is started
container-replicator_ok           OBJECT    INFO     2017-03-01 04:50:37   container-replicator process as expected, state is started
container-server_ok               OBJECT    INFO     2017-03-01 04:50:37   container process as expected, state is started
container-updater_ok              OBJECT    INFO     2017-03-01 04:50:37   container-updater process as expected, state is started
container_access_up               OBJECT    INFO     2017-03-01 04:50:37   Access to container service ip 0.0.0.0 port 6201 ok
ibmobjectizer_ok                  OBJECT    INFO     2017-03-01 04:50:37   ibmobjectizer process as expected, state is stopped
memcached_ok                      OBJECT    INFO     2017-03-01 04:50:37   memcached process as expected, state is started
object-expirer_ok                 OBJECT    INFO     2017-03-08 00:57:15   object-expirer process as expected, state is stopped
object-replicator_ok              OBJECT    INFO     2017-03-01 04:50:37   object-replicator process as expected, state is started
object-server_ok                  OBJECT    INFO     2017-03-01 04:50:37   object process as expected, state is started
object-updater_ok                 OBJECT    INFO     2017-03-08 00:58:24   object-updater process as expected, state is stopped
object_access_up                  OBJECT    INFO     2017-03-01 04:50:37   Access to object store ip 0.0.0.0 port 6200 ok
openstack-object-sof_ok           OBJECT    INFO     2017-03-01 04:50:37   object-sof process as expected, state is stopped
openstack-swift-object-auditor_ok OBJECT    INFO     2017-03-01 04:50:37   object-auditor process as expected, state is stopped
proxy-server_ok                   OBJECT    INFO     2017-03-01 04:50:37   proxy process as expected, state is started
proxy_access_up                   OBJECT    INFO     2017-03-01 04:50:37   Access to proxy service ip 0.0.0.0 port 8080 ok
ring_checksum_ok                  OBJECT    INFO     2017-03-01 04:50:37   Checksum of ring file /etc/swift/object.ring.gz is OK
ring_checksum_ok                  OBJECT    INFO     2017-03-01 04:50:37   Checksum of ring file /etc/swift/account.ring.gz is OK
ring_checksum_ok                  OBJECT    INFO     2017-03-01 04:50:37   Checksum of ring file /etc/swift/container.ring.gz is OK
```

You can use the **mmces events list** command to list all of the events on CES nodes. The use of the **mmces** command provides various options to filter the events by components, severity, date, or nodes. You can use the **mmces events active** command to determine why components are showing up as unhealthy, as shown in Example 2-9.

Example 2-9 Event active to check reasons

```
# mmces events list --severity ERROR --time week -N sybaseiq2-10g
NODE            TIMESTAMP                       EVENT NAME        SEVERITY   DETAILS
sybaseiq2-10g   2017-03-14 06:14:28.594365 EST  pmsensors_down    ERROR      pmsensors service should be started and is stopped
sybaseiq2-10g   2017-03-14 06:46:43.679858 EST  pmsensors_down    ERROR      pmsensors service should be started and is stopped
sybaseiq2-10g   2017-03-14 06:46:58.678290 EST  pmcollector_down  ERROR      pmcollector service should be started and is stopped
```

2.1.4 Network health

The network is one of the most critical components of the IBM Spectrum Scale cluster. Network and networking components are used for communication between clustered nodes, providing protocol services, such as NFS, SMB, and Object protocols, and managing the CES IP addresses.

IBM Spectrum Scale can encounter some critical error conditions because of networking-related issues. For example, IBM Spectrum Scale cluster can encounter a quorum loss situation when communication between multiple quorum nodes is broken or the nodes cannot communicate with each other because of an incorrect firewall setting in the network.

IBM Spectrum Scale health monitoring infrastructure monitors the health of network interfaces and network configuration on all nodes in the cluster. Any issues related to the network are displayed in the Events page or Nodes page (see Figure 2-11).

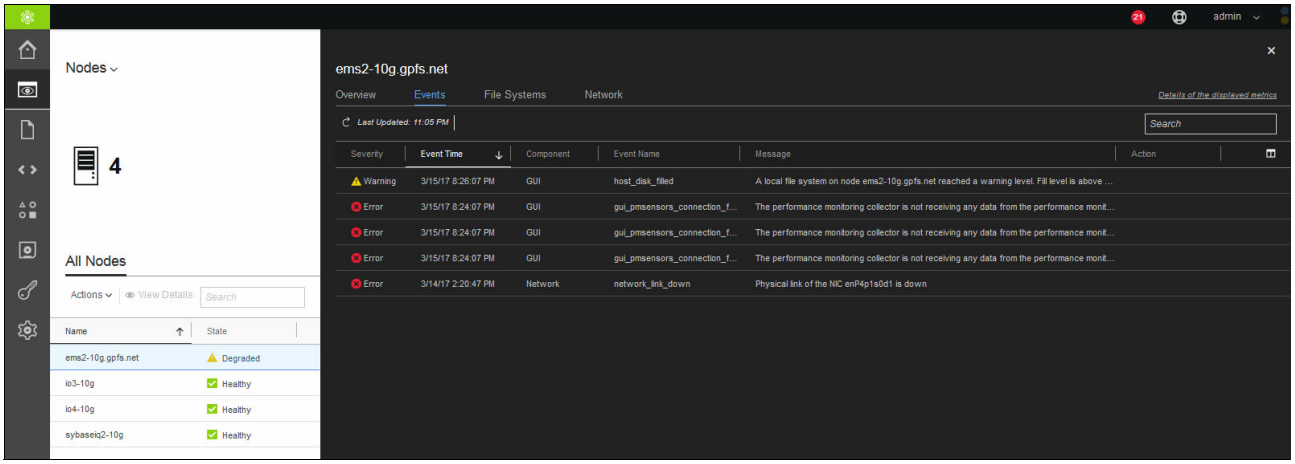

Figure 2-11 Network health events

Monitoring network health by using the CLI

The monitoring infrastructure monitors the various network configurations and network interfaces on the nodes in the cluster. The network in the cluster can be monitored by using the **mmhealth** command, as shown in Example 2-10.

Example 2-10 Using mmhealth to monitor a network in the cluster

```
# mmhealth cluster show NETWORK

Component        Node               Status       Reasons
-------------------------------------------------------------------------
NETWORK          io3-10g            HEALTHY       -
NETWORK          ems2-10g.gpfs.net  FAILED       network_link_down
NETWORK          io4-10g            HEALTHY       -
NETWORK          sybaseiq2-10g      HEALTHY       -
```

You can use the **mmhealth** command (see Example 2-11) to view all network components and their status on a particular node.

Example 2-11 Viewing all network components and their status

```
# mmhealth node show NETWORK --verbose -N io3-10g

Node name:      io3-10g

Component      Status        Status Change            Reasons
-----------------------------------------------------------
NETWORK        HEALTHY       2017-03-14 14:21:53      -
  bond0        HEALTHY       2017-03-14 14:21:53      -
  enP4p1s0     HEALTHY       2017-03-14 14:21:53      -
  enP4p1s0d1   HEALTHY       2017-03-14 14:21:53      -
  enP9p1s0     HEALTHY       2017-03-14 14:21:53      -
  enP9p1s0d1   HEALTHY       2017-03-14 14:21:53      -

Event               Parameter    Severity    Active Since                  Event Message
-------------------------------------------------------------------------------------------------------------------------
network_ips_up      NETWORK      INFO        2017-03-14 14:21:53     Relevant IPs are served by found NICs
bond_up             bond0        INFO        2017-03-14 14:21:53     All slaves of the bond bond0 are working as expected
network_link_up     enP9p1s0     INFO        2017-03-14 14:21:53     Physical link of the NIC enP9p1s0 is up
network_link_up     enP4p1s0d1   INFO        2017-03-14 14:21:53     Physical link of the NIC enP4p1s0d1 is up
network_link_up     bond0        INFO        2017-03-14 14:21:53     Physical link of the NIC bond0 is up
network_link_up     enP4p1s0     INFO        2017-03-14 14:21:53     Physical link of the NIC enP4p1s0 is up
network_link_up     enP9p1s0d1   INFO        2017-03-14 14:21:53     Physical link of the NIC enP9p1s0d1 is up
```

```
network_up          enP9p1s0      INFO        2017-03-14 14:21:53      NIC enP9p1s0 is up
network_up          enP4p1s0d1    INFO        2017-03-14 14:21:53      NIC enP4p1s0d1 is up
network_up          bond0         INFO        2017-03-14 14:21:53      NIC bond0 is up
network_up          enP4p1s0      INFO        2017-03-14 14:21:53      NIC enP4p1s0 is up
network_up          enP9p1s0d1    INFO        2017-03-14 14:21:53      NIC enP9p1s0d1 is up
no_tx_errors        enP9p1s0      INFO        2017-03-14 14:21:53      NIC enP9p1s0 had no or a tiny number of TX errors
no_tx_errors        enP4p1s0d1    INFO        2017-03-14 14:21:53      NIC enP4p1s0d1 had no or a tiny number of TX errors
no_tx_errors        bond0         INFO        2017-03-14 14:21:53      NIC bond0 had no or a tiny number of TX errors
no_tx_errors        enP4p1s0      INFO        2017-03-14 14:21:53      NIC enP4p1s0 had no or a tiny number of TX errors
no_tx_errors        enP9p1s0d1    INFO        2017-03-14 14:21:53      NIC enP9p1s0d1 had no or a tiny number of TX errors
```

In addition to the network interfaces on the nodes, other elements in the network (such as, network cables, switches, and routers) can cause the network to behave inconsistently. The **mmnetverify** command (available starting with IBM Spectrum Scale 4.2.2 release) can be used to check the network connectivity and network traffic behavior between nodes in the IBM Spectrum Scale cluster.

The following operations are available with the **mmnetverify** command:

- Network connectivity checks
- Port connectivity between IBM Spectrum Scale daemons
- Bandwidth that can be used between the nodes

The **mmnetverify** command performs various network checks between the current node and the `io3-10g` node, as shown in Example 2-12.

Example 2-12 The mmnetverify command

```
#  mmnetverify all --target-nodes io3-10g
Checking local configuration.
  Operation interface: Success.
Checking communication with node io3-10g.
  Operation resolution: Success.
  Operation ping: Success.
  Operation shell: Success.
  Operation copy: Success.
    Time on node io3-10g and local node differ by 1 second.
  Operation time: Success.
  Operation daemon-port: Success.
  Operation sdrserv-port: Success.
  Operation tsccmd-port: Success.
  Operation data-small: Success.
  Operation data-medium: Success.
  Operation data-large: Success.
    Bandwidth with io3-10g: 6.936G bits per second.
  Operation bandwidth-node: Success.
Checking cluster communications.
    Bandwidth with io3-10g: 7.179G bits per second.
  Operation bandwidth-cluster: Success.

No issues found.
```

2.1.5 IBM Spectrum Scale disk health

The IBM Spectrum Scale cluster includes Network Shared Disk (NSD) and NSD Server nodes to which the disks are attached. NSDs are created by using virtual disks (VDisks) that are provided by RAID controllers or Storage arrays. Therefore, the health of NSDs and NSD Server nodes must be monitored to avoid any I/O interruptions.

The health of all the NSD disks in the cluster can be viewed in the NSD page by clicking **Storage** → **NSD** (see Figure 2-12).

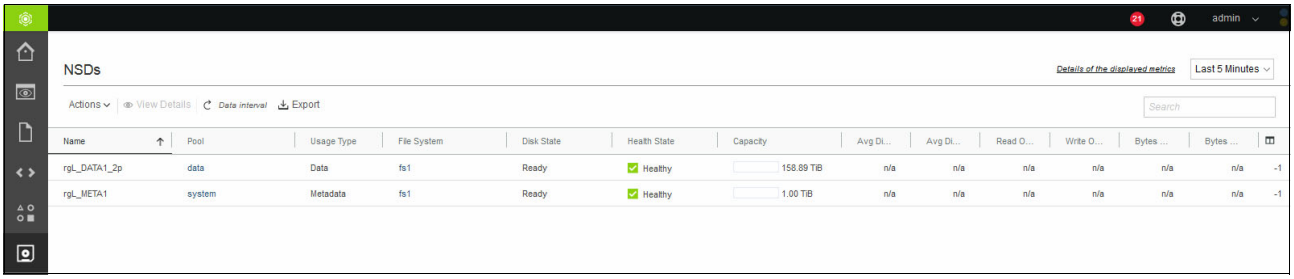

Figure 2-12 NSD Disks health in the cluster

You can view the health of the disks associated with a particular file system, as shown in Figure 2-13.

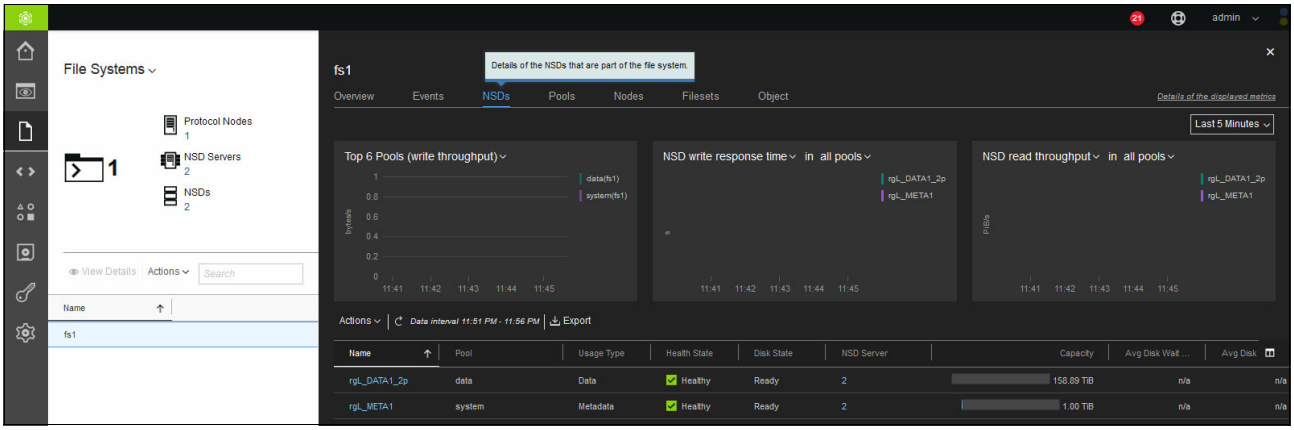

Figure 2-13 Health of NSD disks for a file system

If all NSDs server nodes that access an NSD are down, access to file system can be affected, depending on the file system configuration. Therefore, NSD Server node health also must be monitored along with NSDs. NSD nodes can be monitors by clicking **Monitoring** → **Nodes** in the IBM Spectrum Scale GUI (see Figure 2-14 on page 21).

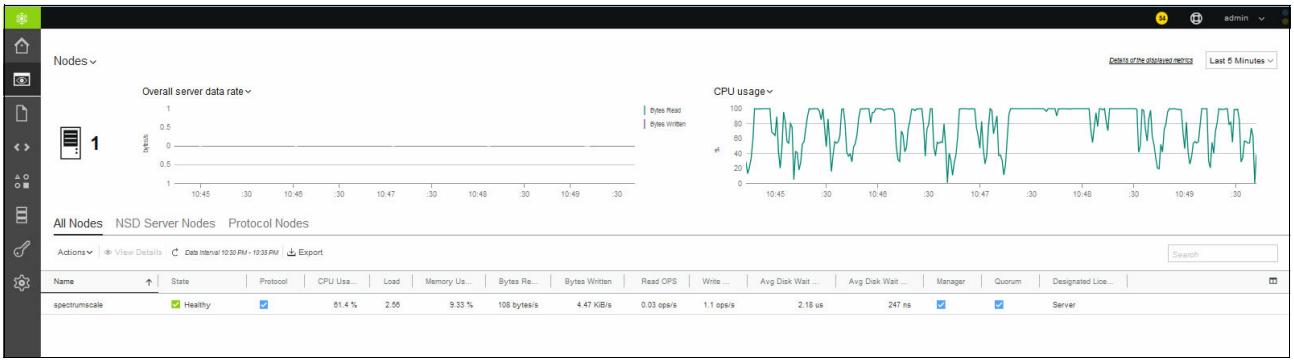

Figure 2-14 NSD Server nodes health

Any issues that are related to NSD disks are reported in the Events page of the GUI.

Monitoring Disk health by using the CLI

You can use the **DISK** option with the **mmhealth** command to view the state of the disks in the cluster, as shown in Example 2-13.

Example 2-13 The mmhealth DISK option

```
# mmhealth cluster show DISK --verbose

Component                Node           Status        Reasons
-------------------------------------------------------------------------
DISK                     io3-10g        HEALTHY          -
  rgL_DATA1_2p                          HEALTHY          -
  rgL_META1                             HEALTHY          -
DISK                     io4-10g        HEALTHY          -
  rgL_DATA1_2p                          HEALTHY          -
  rgL_META1                             HEALTHY           -
```

To view more information about the **DISK** component on a particular node, you can use the **mmhealth node show** command that is shown in Example 2-14.

Example 2-14 The mmhealth DISK option for a particular node

```
# mmhealth node show DISK --verbose -N io3-10g

Node name:      io3-10g

Component       Status        Status Change            Reasons
-------------------------------------------------------------------------------
--
DISK            HEALTHY       2017-03-14 06:31:24      -
  rgL_DATA1_2p  HEALTHY       2017-03-14 06:31:24      -
  rgL_META1     HEALTHY       2017-03-14 06:31:24      -

Event       Parameter     Severity  Active Since         Event Message
-------------------------------------------------------------------------------
-
disk_up     rgL_DATA1_2p  INFO      2017-03-14 14:21:55  Disk rgL_DATA1_2p is up
disk_up     rgL_META1     INFO      2017-03-14 14:21:55  Disk rgL_META1 is up
```

2.1.6 Elastic Storage Server system health

This section describes how you can monitor the health of various components within the Elastic Storage Server (ESS) system. This section is applicable to ESS environments only and is not applicable to IBM Spectrum Scale only environments.

The ESS GUI provides an overall status of the system on the Hardware page, as shown in Figure 2-15.

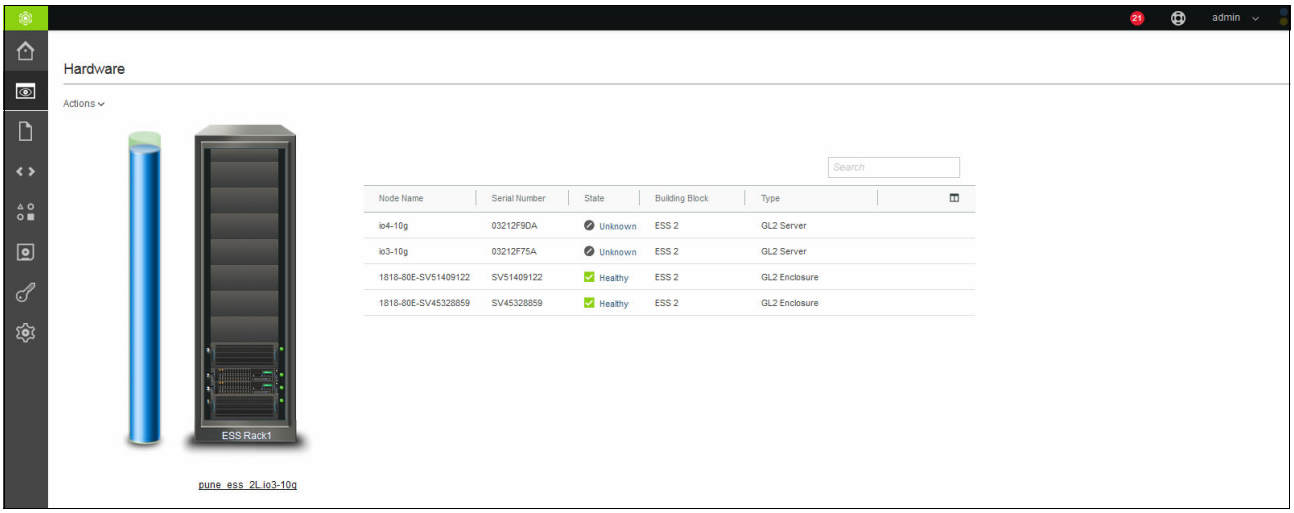

Figure 2-15 Monitoring Hardware page

Monitoring I/O nodes

The I/O nodes in ESS environment are the NSD Server nodes. As described in 1.5, “Disk health” on page 5, the I/O nodes must be monitored.

The Events page in the ESS GUI reports any events that are related to I/O nodes.

The ESS GUI provides information about I/O nodes in the Hardware details page. This page displays status of I/O nodes, EMS, and disk enclosures in the ESS. As shown in Figure 2-16, the components marked with red warning icons need attention.

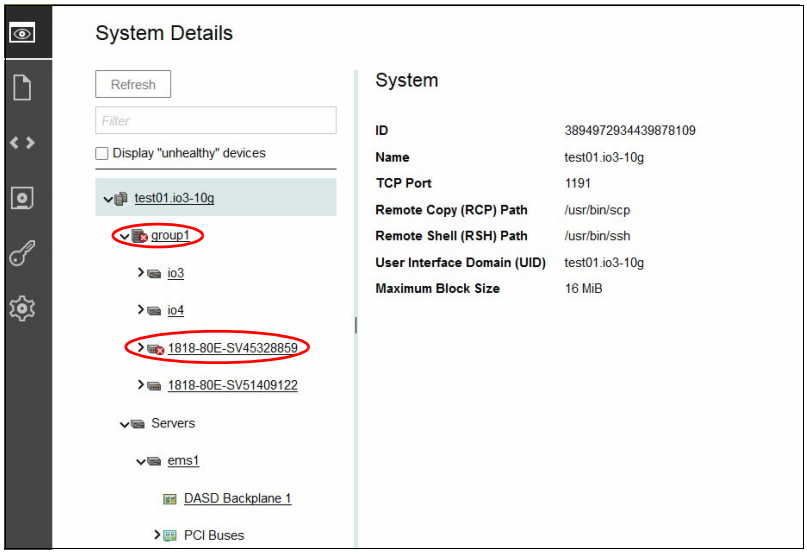

Figure 2-16 ESS Hardware details with error components

Monitoring disk enclosures

The ESS includes the Disk Hospital feature that automatically starts the disk rebuilding process without affecting I/O performance if a disk fails. The Events page in the ESS GUI displays any disk failures and rebuild-related events.

It is important that you subscribe to Hardware GUI notifications so that you receive notifications that are related to ESS I/O nodes and ESS disk enclosures, as described in section 2.2, “Monitoring with alerts and notifications” on page 27. The events that are related to physical disk that need further service are shown in Figure 2-17.

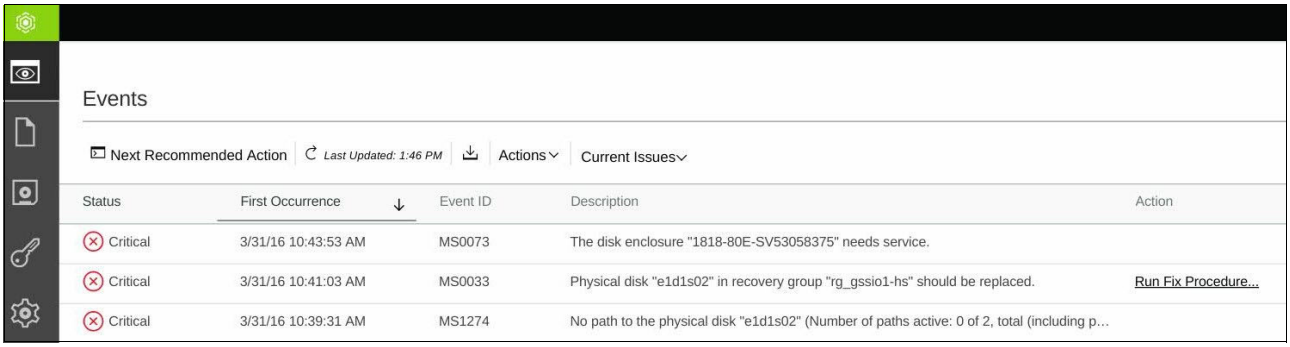

Figure 2-17 Disk errors reported under events page

The ESS GUI provides graphical representation for the disk enclosures in the Hardware page. As shown in Figure 2-18, the physical disk that requires service is marked in red.

Figure 2-18 Disk in error marked as red

You can also view more information about the disk and errors on the Hardware Details page and on the Physical page, as shown in Figure 2-19.

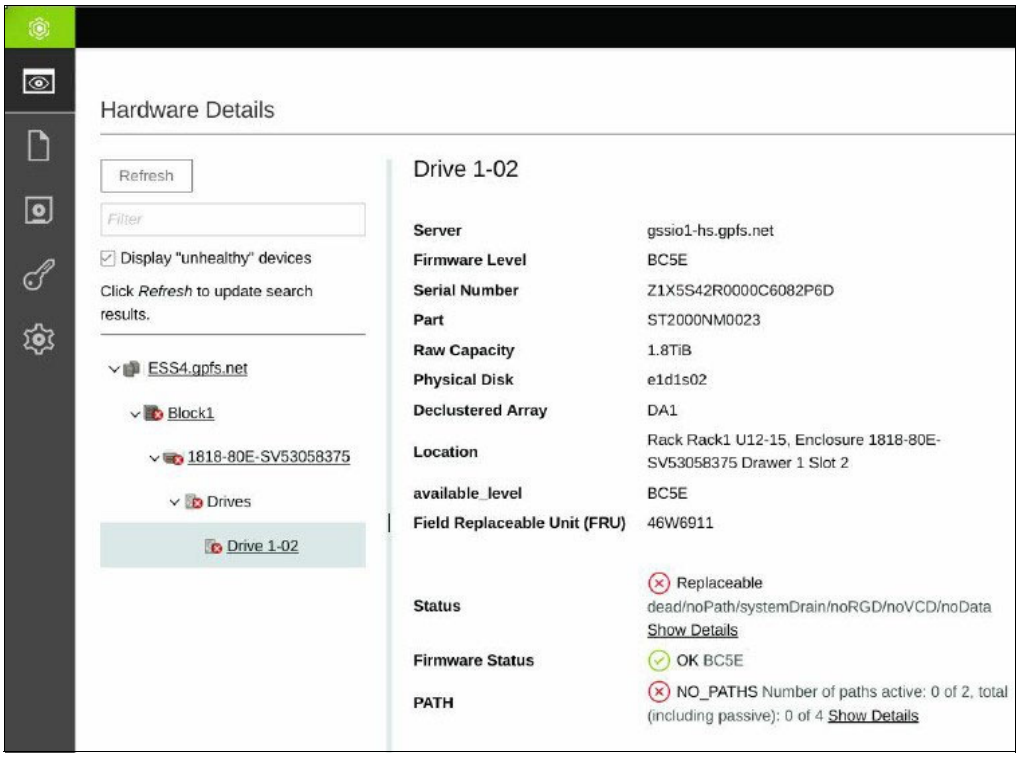

Figure 2-19 Hardware details page showing disk errors

The storage physical disk information is shown in Figure 2-20.

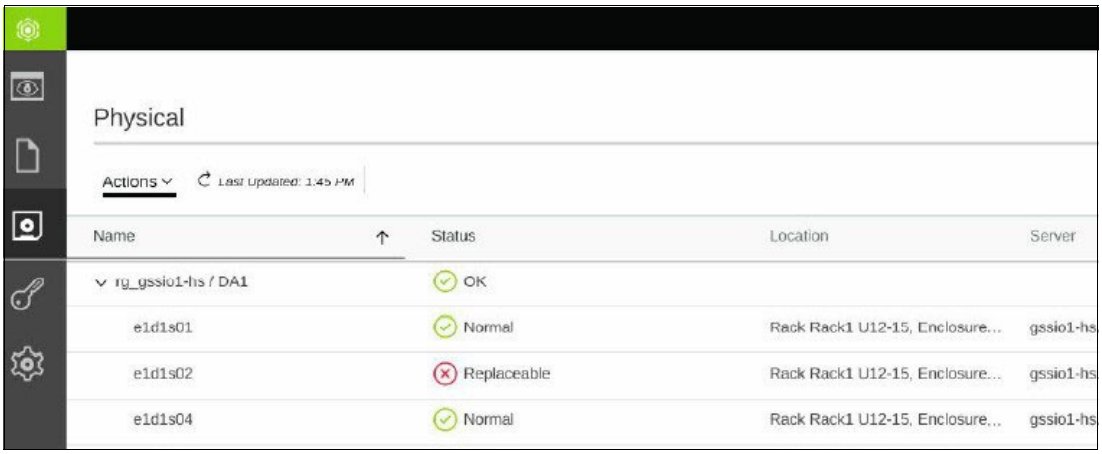

Figure 2-20 Storage Physical disk details

Monitoring ESS system health by using the CLI

The monitoring infrastructure in IBM Spectrum Scale also monitors the health of virtual disks, physical disks, arrays, and enclosures within ESS. You can view the status of the I/O nodes and physical disks, enclosures, and more in the cluster by using the **`mmhealth`** command, as shown in Example 2-15 on page 25. The information that is under the NATIVE_RAID component provides the overview and status of the virtual and physical disks.

Example 2-15 The mmhealth cluster showing verbose

```
# mmhealth cluster show --verbose

Component         Total     Failed    Degraded      Healthy          Other
-------------------------------------------------------------------------
NODE              4         0         1             3              0
GPFS              4         0         0             4              0
NETWORK           4         1         0             3              0
FILESYSTEM
  fs1             4         0         0             4              0
DISK
  rgL_DATA1_2p    2         0         0             2              0
  rgL_META1       2         0         0             2              0
CES               1         0         0             1              0
  AUTH            1         0         0             0              1
  AUTH_OBJ        1         0         0             1              0
  BLOCK           1         0         0             0              1
  CESNETWORK      1         0         0             1              0
  NFS             1         0         0             1              0
  OBJECT          1         0         0             1              0
  SMB             1         0         0             1              0
GUI               1         0         1             0              0
NATIVE_RAID       5         0         0             5              0
  ARRAY           2         0         0             2              0
  ENCLOSURE       2         0         0             2              0
  PHYSICALDISK    2         0         0             2              0
  RECOVERYGROUP   2         0         0             2              0
  VIRTUALDISK     2         0         0             2              0
PERFMON           4         0         0             4              0
```

You can view the overview health of physical disks on a particular node by using the **mmhealth node show** command, as shown in Example 2-16.

Example 2-16 The mmhealth show NATIVE_RAID

```
# mmhealth node show NATIVE_RAID -N io3-10g

Node name:      io3-10g

Component            Status        Status Change     Reasons
-------------------------------------------------------------------
NATIVE_RAID          HEALTHY       1 day ago         -
  ARRAY              HEALTHY       1 day ago         -
  ENCLOSURE          HEALTHY       1 day ago         -
  PHYSICALDISK       HEALTHY       1 day ago         -
  RECOVERYGROUP      HEALTHY       1 day ago         -
  VIRTUALDISK        HEALTHY       1 day ago         -

There are no active error events for the component NATIVE_RAID on this node
(io3-10g).
```

The use of the **mmlsrecoverygroup** command lists the recovery group that was created on the system. By using this command, information is displayed about physical disks, enclosure details, and the physical locations of the disks.

The use of the **mmlspdisk** command lists information about one or more IBM Spectrum Scale pdisks. This command includes filters that you can use to get information about only the pdisks that are not in use or are in an error (not ok) state. An example of information about a physical disk is shown in Example 2-17.

Example 2-17 The mmlspdisk command

```
# mmlspdisk BB01 --pdisk "n002v001"
pdisk:
   replacementPriority = 1000
   name = "n002v001"
   device = "//io4-10g/dev/sda10"
   recoveryGroup = "BB01"
   declusteredArray = "NVR"
   state = "ok"
   internalState = 0000.040
   capacity  = 2088763392
   freeSpace = 1904214016
   fru = "IPR-10  68C1D80"
   location = ""
   WWN = ""
   server = "io3-10g"
   reads = 0
   writes = 2
   bytesReadInGiB = 0.000
   bytesWrittenInGiB = 0.000
   IOErrors = 0
   IOTimeouts = 0
   mediaErrors = 0
   checksumErrors = 0
   pathErrors = 0
   relativePerformance = 1.000
   dataBadness = 0.000
   rgIndex = 1
   userLocation = ""
   hardware = "IBM IPR-10  68C1D800  "
   hardwareType = NVRAM
   nPaths = 1 active 1 total
   nsdFormatVersion = 2
   paxosAreaOffset = 2092826624
   paxosAreaSize = 4194304
   logicalBlockSize = 512
```

You can use the **mmlsrecoverygroupevents** command to display the events that are related to specific IBM RAID recovery group, as shown in Example 2-18.

Example 2-18 Use of the mmlsrecoverygroupevents command

```
# mmlsrecoverygroupevents BB01 --days 5
Tue Mar 14 01:58:47.398 2017 io3-10g ST [I] Start scrubbing tracks of
rgL_DATA1_2p.
Tue Mar 14 01:58:47.398 2017 io3-10g ST [I] End scrubbing tracks of rgL_META1.
Mon Mar 13 19:32:32.962 2017 io3-10g ST [I] Start scrubbing tracks of
rgL_LOGTIPBACKUP.
Mon Mar 13 19:32:32.960 2017 io3-10g ST [I] End scrubbing tracks of
rgL_LOGTIPBACKUP.
```

```
Mon Mar 13 19:30:07.528 2017 io3-10g ST [I] Start scrubbing tracks of rgL_LOGTIP.
Mon Mar 13 19:30:07.527 2017 io3-10g ST [I] End scrubbing tracks of rgL_LOGTIP.
Sun Mar 12 22:10:13.366 2017 io3-10g ST [I] Start scrubbing tracks of rgL_META1.
Sun Mar 12 22:10:13.366 2017 io3-10g ST [I] End scrubbing tracks of rgL_LOGHOME.
Sun Mar 12 20:48:42.613 2017 io3-10g ST [I] Start scrubbing tracks of rgL_LOGHOME.
Sun Mar 12 20:48:42.613 2017 io3-10g ST [I] End scrubbing tracks of rgL_DATA1_2p.
```

2.2 Monitoring with alerts and notifications

IBM Spectrum Scale is deployed as data and file management solution in enterprise workload environments. Because IBM Spectrum Scale environments are typically large scale, manually monitoring each component of IBM Spectrum Scale can be tedious. The recommended way to monitor the IBM Spectrum Scale environment is to subscribe to receive notifications when alerts or events are raised by IBM Spectrum Scale components and take appropriate actions on those received notifications.

The following sections describe the various ways to receive the notifications and how to configure IBM Spectrum Scale to receive these notifications.

2.2.1 Notifications by using email

You can configure IBM Spectrum Scale to receive notifications by using email. In the IBM Spectrum Scale GUI, click **Settings** → **Event Notifications** → **Email Server**. You must first configure the Email Server on this page, as shown in Figure 2-21.

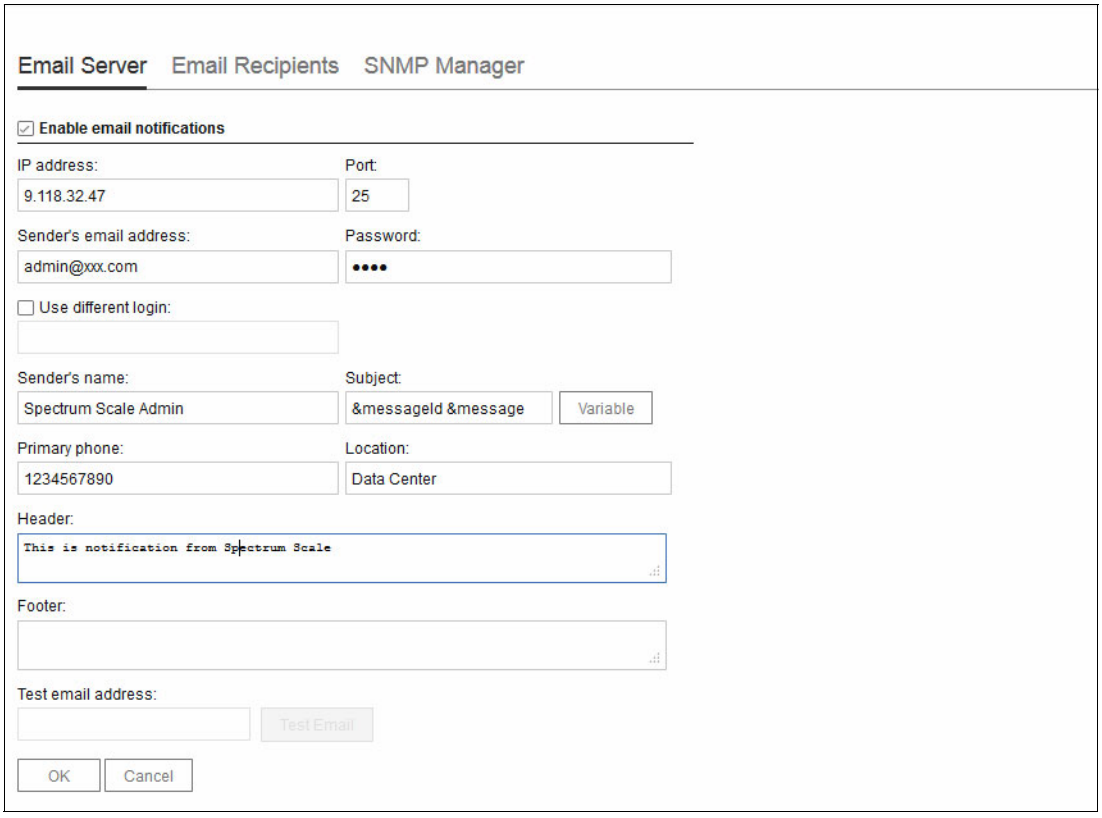

Figure 2-21 Configuring Email Server settings

You can customize the email subject by using following variables:

- &message
- &messageId
- &severity
- &dateAndTime
- &cluster
- &component

After you complete the Email Server configuration, click **Settings** → **Event Notifications** → **Email Recipients** page to add email recipients. You can choose to receive individual emails for each event or receive a combined report for all events, depending on the severity and type of event.

The severities of event notifications are listed in Table 2-1.

Table 2-1 Notification levels

Severity	Description
Error	Error notification is sent to indicate a problem that must be corrected immediately. This notification indicates a serious problem with the system. This type of notification can be result of an event that can lead to loss of access to data if not attended or corrected urgently. This notification typically is the result of a hardware failure, configuration errors, or fabric errors, which also are included in this notification level. **Note:** Hardware failures are reported on ESS only.
Warning	A warning notification is sent to indicate a problem or unexpected condition with the system. Always attend and remediate the notifications to avoid any future errors.
Information	An informational notification is sent to indicate that an expected event is occurred. For example, an SMB service restarted.

Configuring email recipient is shown in Figure 2-22.

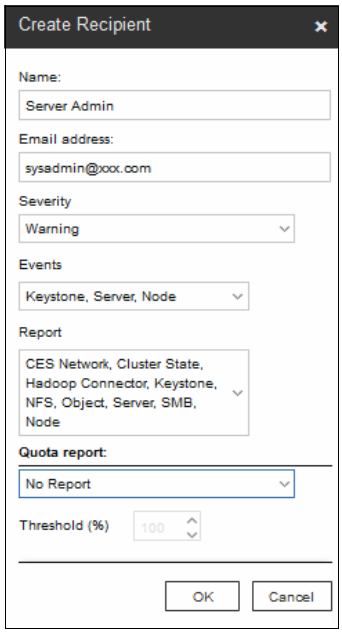

Figure 2-22 Configuring email recipient

When a recipient for Email notification is made, consider the following points:

- Choose events in the “Event” drop-down menu for which you want to receive immediate notification.
- Choose events in the “Report” drop-down menu for which you want to receive daily summary report.
- Choose event in the “Quota report” for which you want to receive quota notifications.

You can use the following practices to avoid receiving a voluminous number of emails from the notifications:

- Do not create any recipient for notification for messages with “Information” as severity. Otherwise, choose to receive Daily summary reports for “Information” messages.
- Create a separate recipient for all messages with “Error” as severity so that these events can be acted upon immediately.
- Create a separate recipient who might be interested in receiving daily reports for all events.
- Certain warning messages might need to be acted upon immediately whereas other warning messages can be acted upon with the next day or so. Choose the appropriate category under Events and Report for Warning messages.
- It is recommended that any “Hard quota exceeded messages” be handled immediately as it can affect the users’ or groups’ ability to from write any other data on the file system. The rest of the quota message can be notified in a daily report and be acted upon the next day.

For more information about the various functions for which notifications are sent, see to IBM Knowledge Center.

2.2.2 Notifications by using Simple Network Management Protocol

Simple Network Management Protocol (SNMP) is a standard protocol for managing networks and exchanging messages. The IBM Spectrum Scale SNMP agent can send SNMP messages that can notify an SNMP manager about an event. You can use an SNMP manager, such as IBM Tivoli®, to view the SNMP messages that are sent by IBM Spectrum Scale.

IBM Spectrum Scale supports the following types of SNMP mechanisms:

- SNMP notifications sent by using IBM Spectrum Scale GUI

  When new event occurs, IBM Spectrum Scale can send SNMP notifications (known as *traps*) to the SNMP manager. In the IBM Spectrum Scale GUI, click **Settings** → **Event Notifications** → **SNMP** manager to configure SNMP managers for event notifications.

  For more information about SNMP configuration and notification by using IBM Spectrum Scale GUI, see the Configuring SNMP manager section of IBM Knowledge Center.

- SNMP notification sent by using IBM Spectrum Scale core (NET-SNMP)

  You can configure an SNMP agent that collects information from subagents in the cluster and sends various cluster, file system, storage, and disk-related event notifications as traps.

  For more information about SNMP configuration and notification by using IBM Spectrum Scale GUI, see the GPFS SNMP Support section of IBM Knowledge Center.

2.2.3 Setting up alerts with IBM Spectrum Scale CLI (mmcallback)

IBM Spectrum Scale provides callbacks as a method for you to take notice or action when an important event occurs. IBM Spectrum Scale enables you to register a script to be run as an action on receiving the event. For example, administrator can use the pdReplacePdisk callback on ESS to send an email to notify system operators that the replacement threshold for a declustered array was reached and that pdisks must be replaced.

The event can be cluster-wide or local to a node. Based on the nature of the event, the registered script is started when the event occurs on all of the nodes of the cluster or only a node of the cluster. This behavior is the default behavior.

While registering the callback, you can specify the nodes in the cluster on which the callback must be started. Based on the nature of event, the callback and the script are started on the registered nodes only.

For global events, it is recommended to register the callback to be started on a single node only, typically the administrative node of the cluster. In the case of local events, the callback is started on a node only if the node is among the set of registered nodes to receive callback. The script that is run for the registered callback is placed on every cluster node wherever the callback is received. Certain events (for example, lowDiskSpace) are triggered more than once until the resolution condition are met.

It is not recommended to start complex or long-running IBM Spectrum Scale commands in the registered callback scripts. Restriping a file system, deleting a snapshot, applying policy rules, and evaluating quota are, but not limited to, examples of long-running operations on IBM Spectrum Scale.

Callback can be registered to be started asynchronously or synchronously when the event occurs. It is generally guided to register callback with asynchronous execution. Consider the following examples:

- Example 1: Email notification to user on exceeding soft quota

  In scenarios where quota is enabled on IBM Spectrum Scale, a quite common ask by a user is intimating usage to avoid exhausting the quota limit. With the registration of a callback that informs when quota usage limits were exceeded, you can enable sending email notification to the accessing user and storage administrator. Accordingly, the user and administrator can take action.

- Example 2: Notifying applications about file system mount state

  Your applications that are running on IBM Spectrum Scale expect the file system to be available. Otherwise, it can lead to unexpected behavior. You can register a callback for notifying applications of file system pre-mount, mount, pre-unmount, and unmount events to avoid application issues.

- Example 3: Increasing quota of critical users

  Another use case can be automatically increasing quota limits for critical users when they meet soft quota limits.

On the ESS system, several callbacks are pre-created, which sends notification to the ESS GUI. These callbacks are registered only when you designate and install the GUI on cluster nodes. The scripts of these registered callbacks are strictly disallowed for edits and updates.

You can add callbacks by using the **mmaddcallback** CLI command. Pre-created callback and newly created callbacks can be listed by using the **mmlscallback** CLI command. For more information, see the **mmaddcallback**, **mmlscallback** command **man** pages.

The list of events (global and local) for IBM Spectrum Scale are captured on the **man** page of **mmaddcallback** command. On ESS, the **mmaddcallback** man page includes the list of events for IBM Spectrum Scale RAID.

The callback mechanism is an obvious choice when you want to automate certain actions on receiving particular events. The actions can be notifying daily status of processes and statistics, monitoring, and notifying disruptive events for auditing. The following actions also are available:

- Notifying daily status of processes
- Notifying disruptive events for auditing

Sample callback scripts are available on IBM Spectrum Scale and ESS, sample scripts are available in the `/usr/lpp/mmfs/samples` directory.

3

Performance monitoring

The IBM Spectrum Scale GUI provides a graphical representation of current and historical performance indicators. The Dashboard and Statistics monitoring pages in the GUI can be used to monitor the performance of the system based on various aspects.

You can use the Statistics page to select the resources based on what you want to monitor and compare the performance of the system. By using the Performance page, you can create customized charts for various resources and metrics to be included in those charts. You can mark the frequently used or favorite charts and then add them as widgets to the Dashboard page. The Dashboard page can contain one or more current, historical performance charts and you can view these performance charts in a single page.

You can configure the GUI to monitor the performance of the following functional areas (called as resources in the Performance page) in the system (for IBM Spectrum Scale and ESS GUI):

- Network
- System resources
- NSD server
- IBM Spectrum Scale client
- NFS

The following functional areas are available for IBM Spectrum Scale GUI only:

- SMB
- Object
- CTDB

The Native RAID functional area is available for ESS GUI only.

This chapter includes the following topics:

3.1 Performance data collection for GUI

The IBM Spectrum Scale GUI collects the performance data based on the following components:

- Sensor

  The sensors are available on all the nodes in IBM Spectrum Scale cluster. The sensors collect data based on the required metrics and send the data to the collector. Sensors are started by default on the protocol nodes only.

- Collector

  The Collector component runs on each of the GUI node. The collector gathers metrics from all the nodes that are running the required sensors. The metrics are stored in a database on the collector node. The collector aggregates older data.

3.2 Configuring performance metrics

Click **Monitoring** → **Statistics** to monitor the performance of various functional areas or resources. The IBM Spectrum Scale GUI provides several predefined charts. You can select the required charts and monitor the performance based on the filter criteria. The predefined performance charts and metrics help in investigating every node or any particular node that is collecting the metrics.

The list of the available predefined performance charts are shown in Figure 3-1 on page 35.

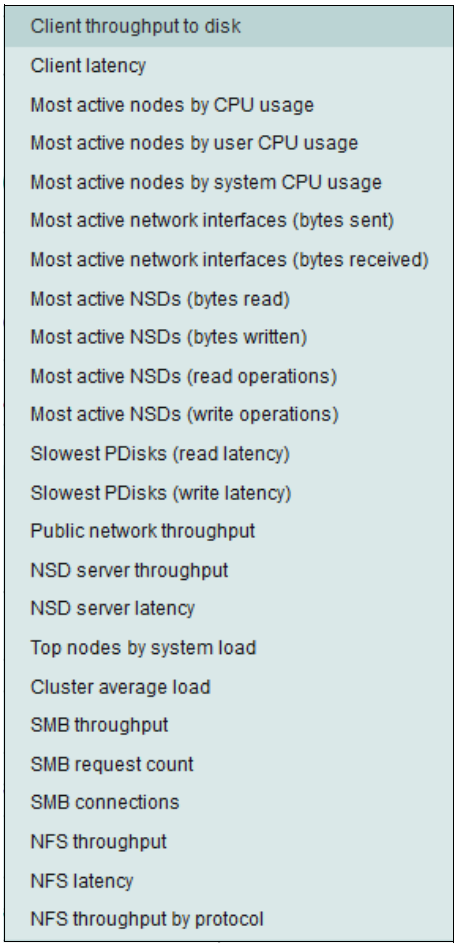

Figure 3-1 List of predefined performance charts

3.2.1 Selecting performance and capacity metrics

You can create customized charts by selecting the appropriate metrics to be displayed in the performance charts. The Metrics are grouped under the combination of resource types and aggregation levels. The resource type corresponds to the functional area from where data is collected to create performance analysis. The aggregation level determines the level at which the data is aggregated.

Sensors are configured against each resource type. Table 3-1 provides a mapping between resource types and sensors under the Performance category.

Table 3-1 Sensors available for each resource type

Resource type	Sensor name	Candidate nodes
Network	Network	All
System Resources	CPU	All
	Load	
	Memory	
NSD Server	GPFSNSDDisk	NSD Server nodes

Resource type	Sensor name	Candidate nodes
IBM Spectrum Scale Client	GPFSFilesystem	IBM Spectrum Scale Client nodes
	GPFSVFS	
	GPFSFilesystemAPI	
NFS	NFSIO	Protocol nodes running NFS service
SMB	SMBStats	Protocol nodes running SMB service
	SMBGlobalStats	
Waiters	GPFSWaiters	All nodes
CTDB	CTDBStats	Protocol nodes running SMB service
Object	SwiftAccount	Protocol nodes running Object service
	SwiftContainer	
	SwiftObject	
	SwiftProxy	

You can edit a chart by clicking the icon that is available in the upper right corner of the performance chart and select **Edit** to modify the metrics selections. Complete the following steps to drill down to the metric you are interested in:

1. Select the Resource type. This area is the area from which the data is taken to create the performance analysis.
2. Select the Aggregation level. The aggregation level determines the level at which the data is aggregated.
3. Select the entities that must be graphed. The table lists all entities that are available for the chosen resource type and aggregation level. When a metric is selected, you can also see the selected metrics in the same grid and use methods, such as sorting, filtering, or adjusting the time frame to select the entities that you want to select.
4. Select Metrics. Metrics is the type of data that must be included in the performance chart. The aggregation levels that are available for selection varies based on the resource type.
5. Use the filter option to further narrow down to the objects and metrics selection. Depending on the selected object category and aggregation level, the Filter section can be displayed underneath the aggregation level, which allows one or more filters to be set.

3.2.2 Creating favorite charts

You can mark customized or predefined charts as your Favorite charts. You can add a favorite chart by using the single chart configuration as a widget in the Dashboards page. Favorite charts along with the predefined charts are available when you add widgets in the dashboard.

To create favorite charts, click the star symbol that is placed next to the chart title, enter the label, and select whether you need to display the favorite chart in single chart or dual chart mode.

3.3 Performance monitoring by using CLI

You can use the **mmpmon** or **mmperfmon** command to monitor the IBM Spectrum Scale performance by using CLI. These commands require configuration and allow data to be collected and then aggregated on certain nodes in the cluster.

The **mmpmon** command is based on various performance counters from GPFS daemon. The performance monitoring request is sent to the GPFS daemon on the same node that is running the **mmpmon** command. For more information about how to monitor GPFS I/O performance by using the **mmpmon** command, see IBM Spectrum Scale Knowledge Center.

The **mmperfmon** command is used to configure zimon sensors on the node and to report the data that is collected by zimon sensors for various IBM Spectrum Scale components. In addition to the data that is collected by GPFS daemon, the **mmperfmon** command reports data that is related to network, protocols, and so on.

The **mmperfmon** command can also report the historical data that is collected by zimon sensors. Therefore, the use of the **mmperfmon** command is recommended whenever you need historical data and data that is related to network, protocols, and so on.

For more information about how to monitor IBM Spectrum Scale performance by using the **mmpeform** command, see IBM Spectrum Scale Knowledge Center.

3.4 Performance data graphical visualization

A good start to IBM Spectrum Scale performance monitoring is to use the IBM Spectrum Scale GUI dashboard. The GUI provides an easy-to-read and real-time user interface that displays a graphical representation of the status and historical trends of key performance indicators. This information helps the users to make decisions easily without wasting time evaluating CLI data queries output or installing and configuring more software tools. However, the IBM Spectrum Scale Performance Monitoring GUI user has limited flexibility in customizing the dashboard and adapting the individual operational needs.

Starting with release 4.2.1, IBM Spectrum Scale users can allow third-party performance visualization tools to access the internal IBM Spectrum Scale performance data. This functionality is provided by the IBM Spectrum Scale Monitoring Bridge. The user can create performance graphs by using Grafana dashboard with the IBM Spectrum Scale Monitoring Bridge.

3.4.1 Grafana overview

Grafana is an open source tool for visualizing time series infrastructure and application metrics. It provides a powerful and elegant way to create, explore, and share dashboards and data with your team and the world. The following Data Source plug-ins are officially supported by Grafana:

- Graphite
- Elasticsearch
- CloudWatch
- InfluxDB
- OpenTSDB
- KairosDB
- Prometheus

It is up to the plug-in to establish the communication with a database over HTTP and transform the data into a time series so that any Grafana panel can display it.

3.4.2 IBM Spectrum Scale monitoring bridge

The IBM Spectrum Scale performance monitoring bridge is a stand-alone Python application that uses Grafana to display performance data. The monitoring bridge emulates an openTSDB API, which is used by Grafana to set up and populate the graphs.

The metadata that is received from IBM Spectrum Scale is used to create the Grafana graphs, and the data from IBM Spectrum Scale is used to populate these graphs. The integration framework of IBM Spectrum Scale for Grafana is shown in Figure 3-2.

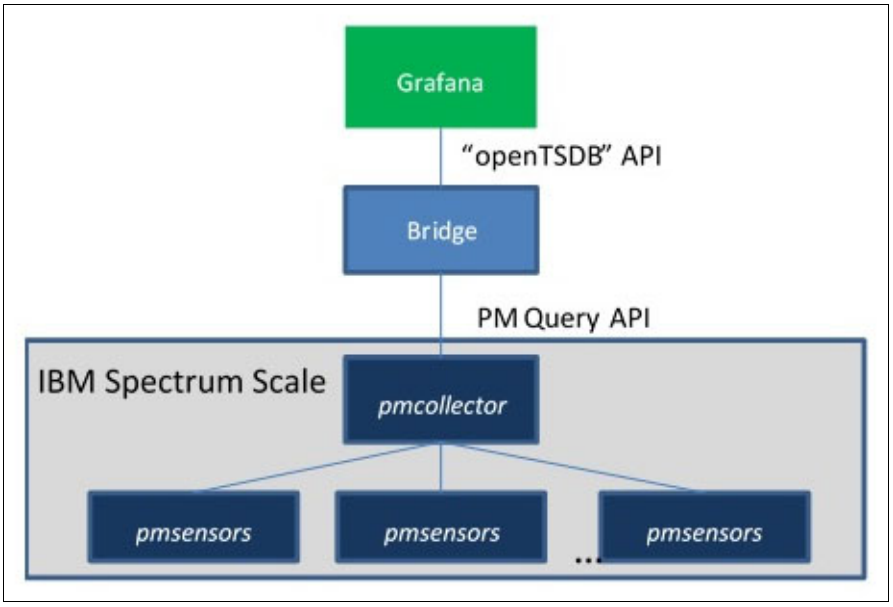

Figure 3-2 IBM Spectrum Scale integration framework for Grafana

The IBM Spectrum Scale Performance monitoring bridge is easy to set up and requires few resources overhead. It acts as the data source for Grafana dashboards and provides the full set of IBM Spectrum Scale performance metrics.

Note: Grafana is a separate component and not a part of the IBM Spectrum Scale 4.2.2 package. Grafana can be downloaded from IBM developerWorks Wiki.

For more information about Grafana, see the Grafana web page.

For more information about how to set up the IBM Spectrum Scale performance monitoring bridge for Grafana, see the Setting up IBM Spectrum Scale performance monitoring bridge for Grafana page of IBM Knowledge Center.

3.4.3 Creating a Grafana dashboard

From the Grafana icon in the upper left corner, click **Dashboards** → **+new**. Click the **Manage dashboard** icon and select **settings** from the drop box for modifying the dashboard name and other attributes. Save the dashboard when changes are made or if you want to browse away from the dashboard.

Dashboards can be thought of as a set of panels that are organized and arranged in rows. Grafana version 3 and later supports the following panel types:

- Dashboard list
- Graph
- Plugin list
- Singlestat
- Table
- Text

The most frequently used panel is Graph. It provides a rich set of graphing options, as shown in Figure 3-3.

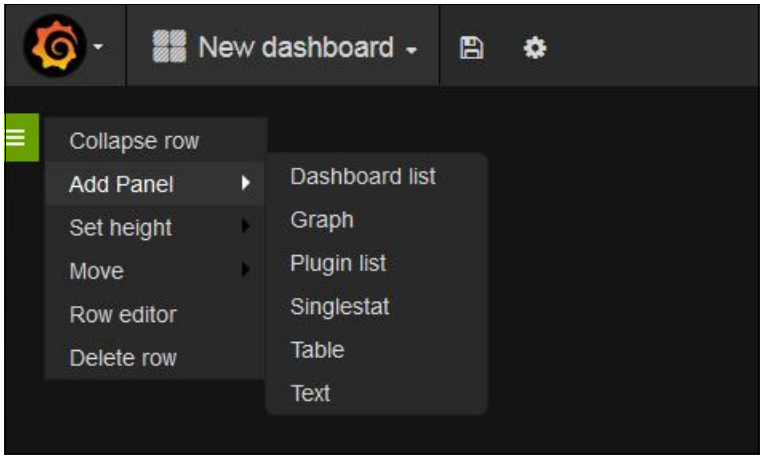

Figure 3-3 Grafana Panel overview

Complete the following steps to create your own performance graphs (see Figure 3-4):

1. Select **Graph** from the Panel list.
2. Click the **graph title** → **Edit**.
3. Click the **General** tab to modify the graph title.

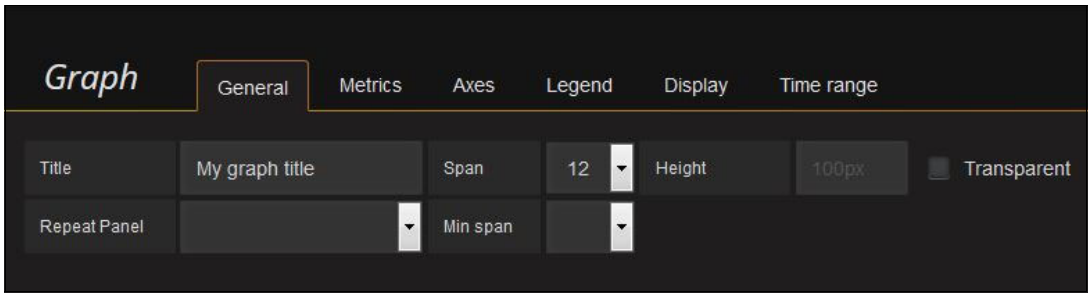

Figure 3-4 Specifying General settings for the Graph Panel

4. Under the **Metrics** tab, specify the metric name, sample interval, and filters, as shown in Figure 3-5.

   Filter is a tuple of query attributes `{ key= Tag_name, type="pm_filter", filter=Tag_value, groupBy=Yes|No }`.

   It is useful to limit the data query results to a specific item view (for example: a node, host, or disk), or an items group (for example, all nodes with name `gpfsgui*`).

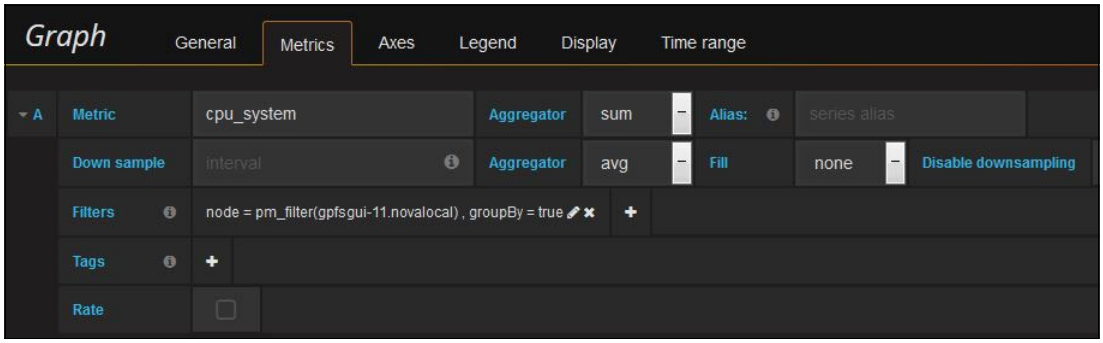

Figure 3-5 Configuring the Metrics section on the Graph Panel

> **Note:** IBM Spectrum Scale Performance monitoring bridge supports only the custom pm_filter that allows values to be constants or regular expressions.
>
> Also, Alias, Downsampling, and Rate options are not supported.

5. Format legend, time range, and other graph settings until a working graph is created.

   The time range can be managed at the dashboard-level. Also, the auto-refresh function can be enabled and the reload frequency for all Panels in the dashboard can be specified.

   Panels can be made more dynamic by using Template variables instead of hardcoded parameters in the data queries. An example of a variable that makes the node name selection flexible (`$nodeName`) is shown in Figure 3-6.

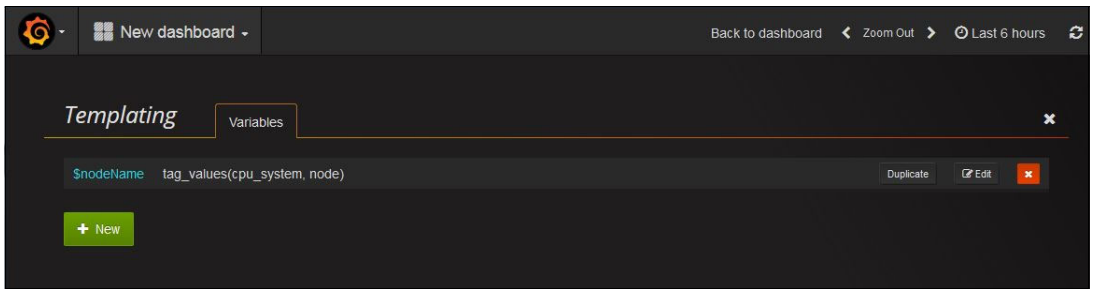

Figure 3-6 Example of creating Dashboard Template variables

After replacing the node name `gpfsgui-11.novalocal` with the variable `$nodeName` in the query filter, the graph displays another drop-down menu that includes all of the node names on the selection list, as shown in Figure 3-7.

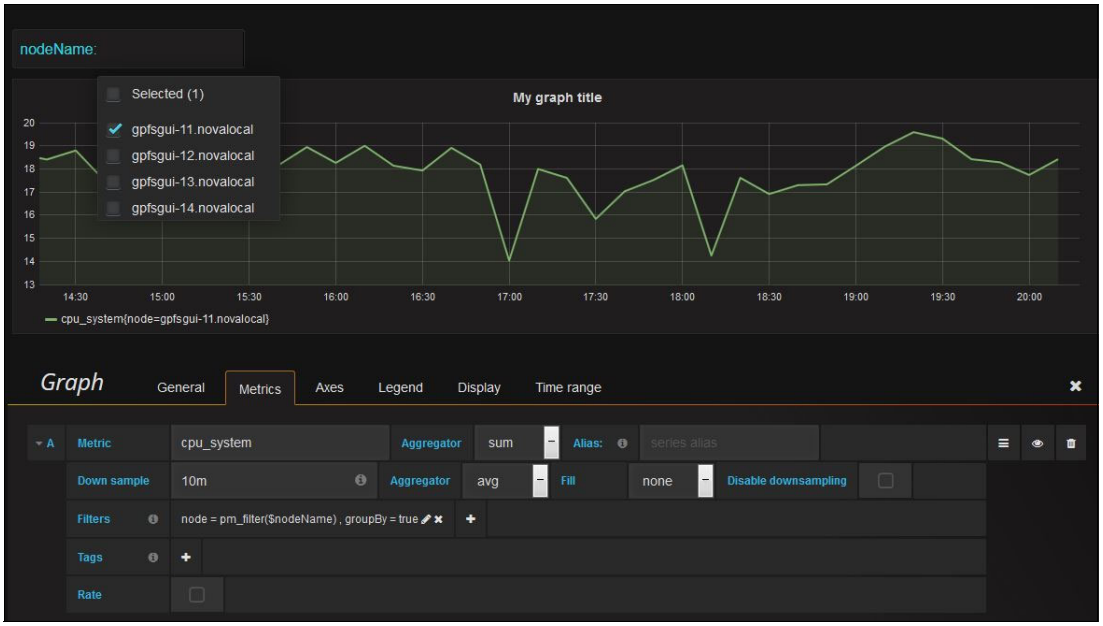

Figure 3-7 Example of Dashboard Template variables usage

Dashboard Templating is one of the most powerful and most used features of Grafana. Generic dashboards can be distributed and used in different environments.

For more information about configuration options, see the Screencasts page of the Grafana Documentation website.

3.4.4 Importing predefined Grafana dashboards

Predefined Grafana dashboards, as shown in Figure 3-8, can be downloaded from the default dashboards set. Import these dashboards into your running Grafana environment. These dashboards can be used directly without any changes.

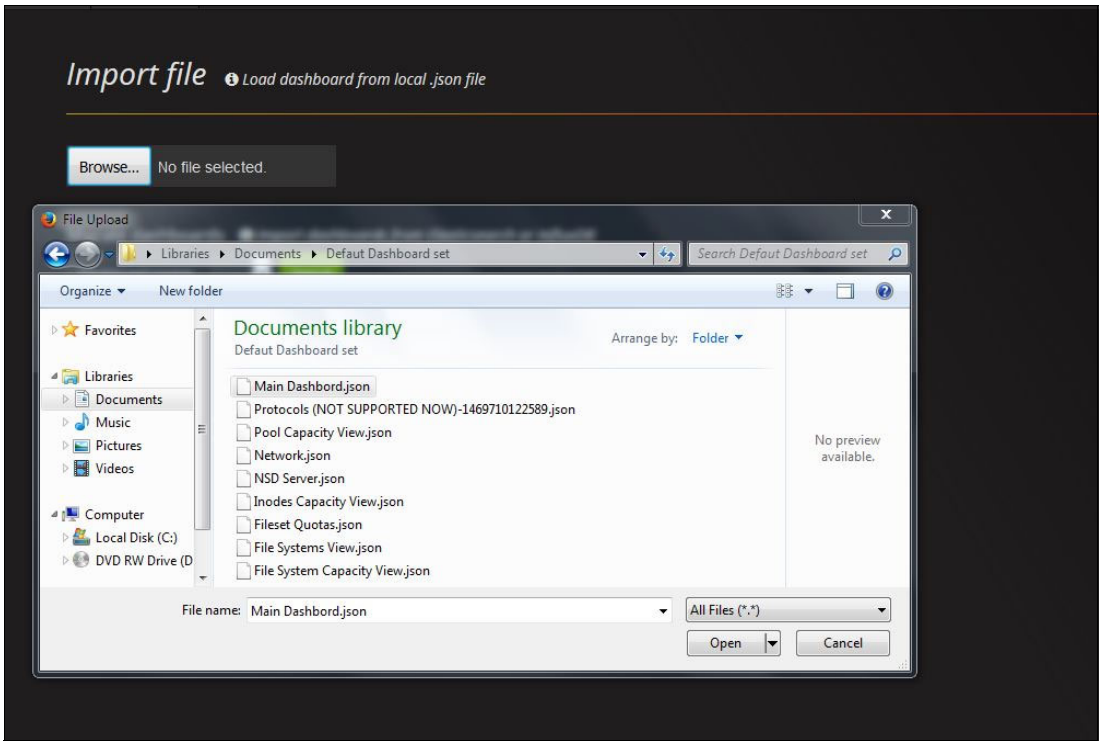

Figure 3-8 Importing predefined dashboards into Grafana

Each dashboard that is included in the download package can be imported and used as a stand-alone dashboard that is independent from all other dashboards in the package. Only the Main Dashboard is linked with other dashboards. Therefore, the Main Dashboard requires the import of all dashboards in the default dashboards set package for its full functionality.

Note: If you are running a Grafana version older than version 3, you must perform some manual steps before or after the import process so that the dashboards function correctly.

For more information, see the Importing a dashboard section of the Grafana Export and Import page.

Main Dashboard overview

The Main Dashboard includes two dashboards list panels: Node View and Capacity View. This splitting refers to the distinction of performance metrics in node-wide and cluster-wide performance measurement views. Each row in both lists is linked with one of the dashboards, which is included in the download package.

In addition to listing panels, the Main Dashboard shows the following high-level performance view graphs:

- CPU usage
- Memory usage
- Network

- NSD Server Nodes
- Protocol Nodes

These graphs include the most commonly used metrics, which describe the overall device and system performance, as shown in Figure 3-9.

Figure 3-9 Main Dashboard view

The Main Dashboard is an example of how the performance views can be organized. In a productive environment, other measurement categories, such as splitting data in application and operating system view, can be more relevant. Also, the type of server usage, such as IBM Spectrum Scale NSD servers, user login nodes, and compute nodes, are important criteria for defining workload measurement. That is, always clearly define a performance target before you start to create performance monitoring views.

For more information about defining measurable targets, see the Performance Tuning Checklist page of the IBM developerWorks® website.

Detail Dashboard examples

Other dashboards that are included in the download package include a small text field that describes the performance data details that can be viewed on it. The following predefined dashboards can be modified and expanded based on the user individual needs:

- File Systems View

  As shown in Figure 3-10 on page 44, this dashboard gives an overview of the file system performance aspects for individual nodes. This dashboard can be used for observing the file system read and write throughput, average read and write transactions size, and file system read and write latency.

Figure 3-10 File System view dashboard

- Network

As shown in Figure 3-11, this dashboard can be used to measure the network throughput in a cluster and test the network behavior that is under heavy load.

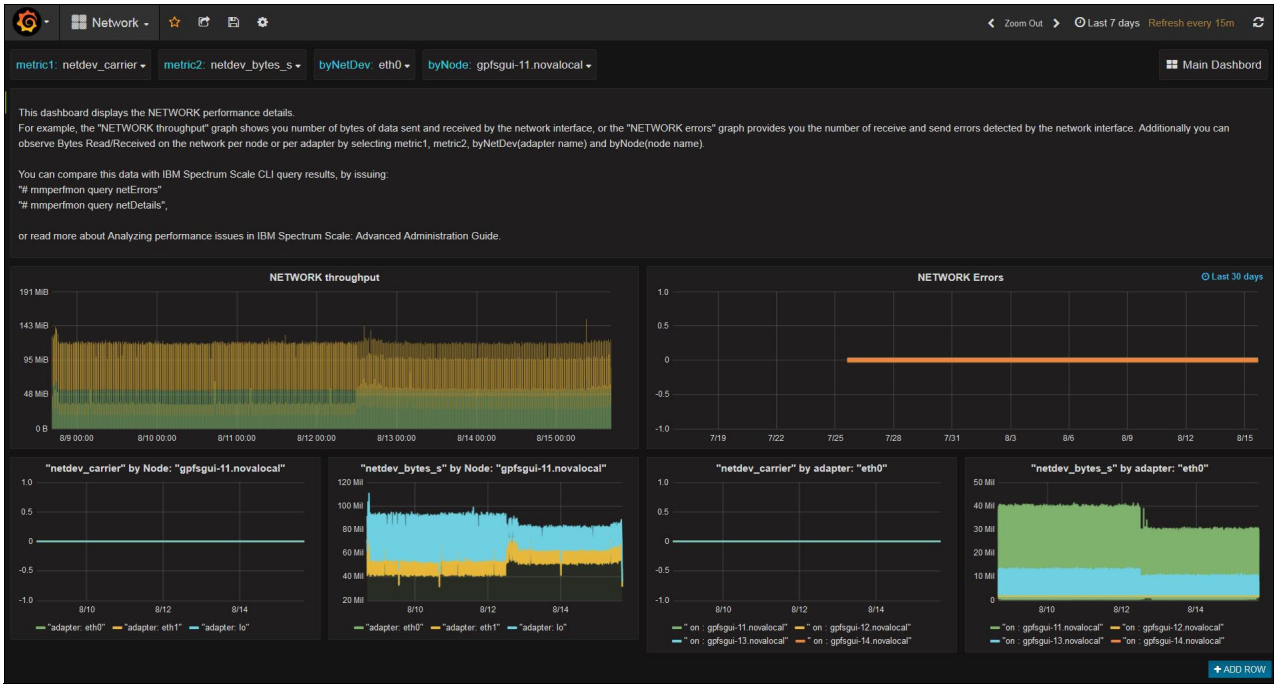

Figure 3-11 Network view dashboard

- File System Capacity

The historical capacity data collection for file systems depend on correctly configured data collection sensors for pool and disk capacity. When the IBM Spectrum Scale system is installed by using the installation toolkit, the collection is configured by default. Otherwise, use the **mmperfmon** command to enable data collection for the GPFSDiskCap sensor.

As shown in Figure 3-12, the File System Capacity View graphs display the capacity metrics for each file system that is aggregated on the disk and pool level. This dashboard indicates the cause of file system free space overflow. For example, a file system having enough free space on the disk level can cause problems if the disk pool reaches a nearly exhausted level. More information about the capacity of the listed pools is provided in the Pool Capacity View dashboard graph.

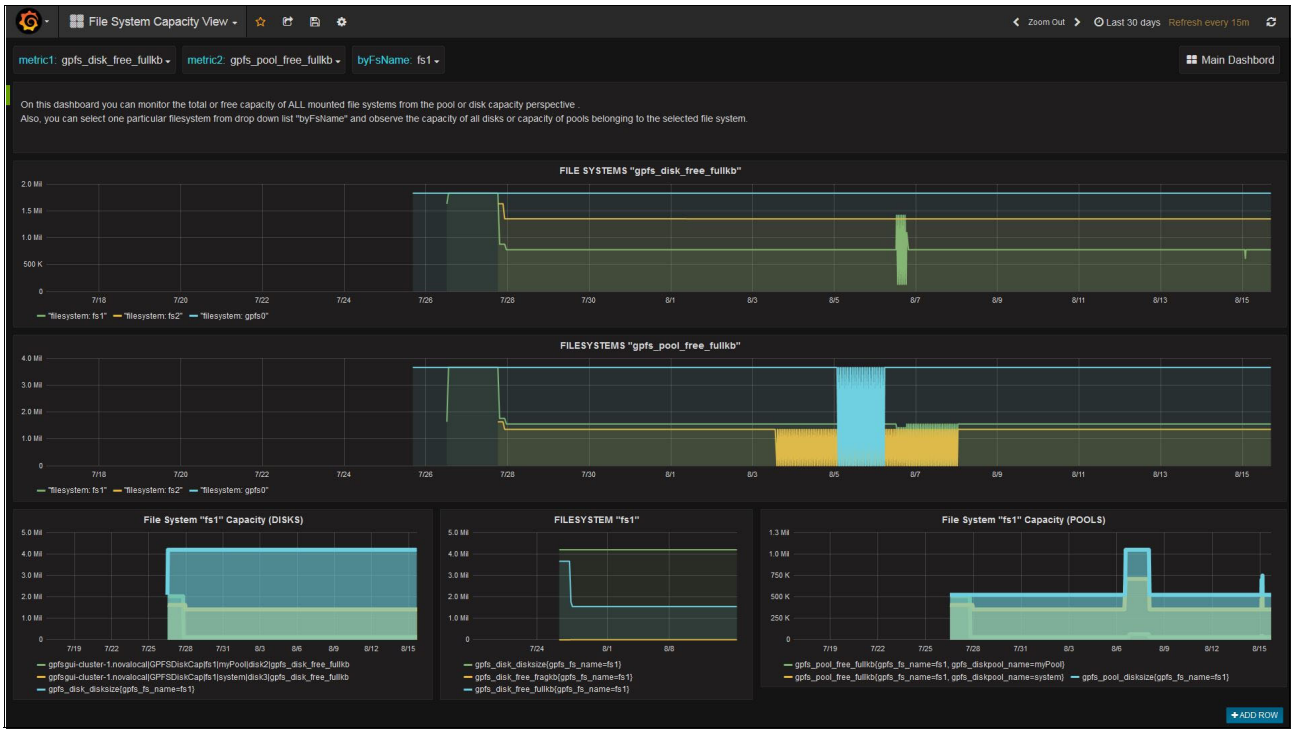

Figure 3-12 File System Capacity View dashboard

3.4.5 Basic troubleshooting tips

In this section, we provide some basic troubleshooting tips.

Connection issues

Consider the following troubleshooting tips for connection issues:

- If the bridge establishes the connection to the specified pmcollector and the metadata was initialized successfully, the message “server starting” appears printed at the end of the line. Otherwise, check the `zserver.log` that is stored in the `/zimonGrafanaIntf` directory.
- Check the pmcollector service that is running properly by using the following command:

  ```
  # systemctl status pmcollector
  ```

- Check the port. IBM Spectrum Scale bridge listens on port 4242.

Metadata issue

The bridge code reads the metadata from the pmcollector on start and does not automatically check for updates. If the metadata changed (for example, new sensors or nodes were added), the metadata can be updated by sending `http://<grafana_bridge_host>:4242/api/update` to the Grafana bridge.

This http request returns a response of `{“msg”: “metadata successfully updated.”, “rc”: 0}` if the command was successful.

3.5 Conclusion

IBM Spectrum Scale and Elastic Storage Server solutions can be monitored for health and performance by using the various methods that are described in this IBM Redpaper publication.

You need to monitor the various components to monitor the overall storage solution. It is important that you set up the monitoring method per your requirements immediately following the installation and configuration of IBM Spectrum Scale. Monitoring helps you to receive notifications about any issues in your system and enable you take timely actions to resolve those issues.

Related publications

The publications that are listed in this section are considered particularly suitable for a more detailed discussion of the topics that are covered in this paper.

IBM Redbooks

The following IBM Redbooks publications provide more information about the topic in this document. Note that some publications that are referenced in this list might be available in softcopy only:

- *IBM Spectrum Scale (formerly GPFS),* SG24-8254
- *Introduction Guide to the IBM Elastic Storage Server,* REDP-5253
- *Implementing IBM Spectrum Scale*, REDP-5254

You can search for, view, download, or order these documents and other Redbooks publications, Redpaper publications, Web Docs, draft, and other materials at the following website:

ibm.com/redbooks

Other publications

The following publications are also relevant as further information sources:

- *IBM Spectrum Scale V4.2: Concepts, Planning, and Installation Guide*, GA76-0441
- *IBM Spectrum Scale V4.2: Administration and Programming Reference*, SA23-1452
- *IBM Spectrum Scale V4.2: Advanced Administration Guide*, SC23-7032
- *IBM Spectrum Scale V4.2: Data Management API Guide*, GA76-0442
- *IBM Spectrum Scale V4.2: Problem Determination Guide,* GA76-0443

Online resources

The following websites are also relevant as further information sources:

- IBM Elastic Storage Server:
  http://www.ibm.com/systems/storage/spectrum/ess/
- IBM Spectrum Scale:
  http://www.ibm.com/systems/storage/spectrum/scale/index.html
- IBM Spectrum Scale resources:
  http://www.ibm.com/systems/storage/spectrum/scale/resources.html
- IBM Spectrum Scale (GPFS) in IBM Knowledge Center:
  https://www.ibm.com/support/knowledgecenter/SSFKCN/gpfs_welcome.html
- IBM Spectrum Scale Frequently Asked Questions and Answers:
  http://ibm.co/1IKO6PN
- IBM Spectrum Scale Wiki:
  https://ibm.biz/BdXVxv

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

REDP-5418-00

ISBN 0738456306

Printed in U.S.A.