IBM Spectrum Scale

File Audit Logging

May 16th, 2018

Boston User Group Event

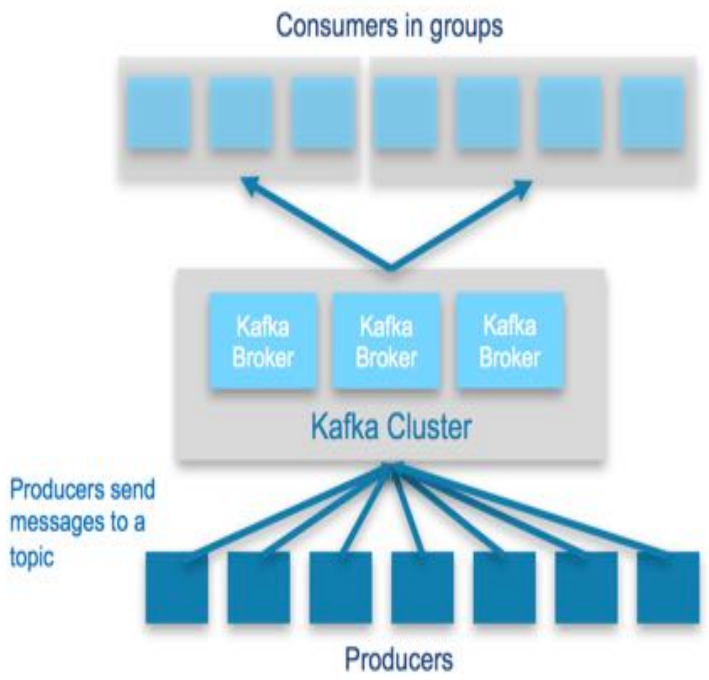
By

Subashini Balachandran

Motivation and Description

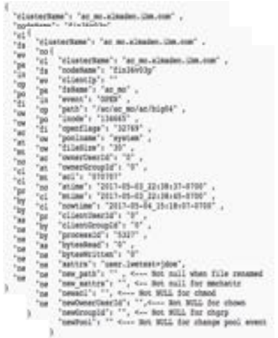
- Capture file operations on a given filesystem and log them for auditing purposes
- Display the stored events
- Capture most common types of file operation activity on the filesystem { create, open, close, destroy, rename, ACL changes, XATTR changes, rmdir, unlink }
- Protocol agnostic – Support Native GPFS, NFS, SMB
- Events are logged in a JSON formatted string
- Configurable options for log output include the device where it is mounted, name, retention period.
- Integrated into the system health infrastructure for easy monitoring of audit logging message queues and components

Kafka Publish-Subscribe model

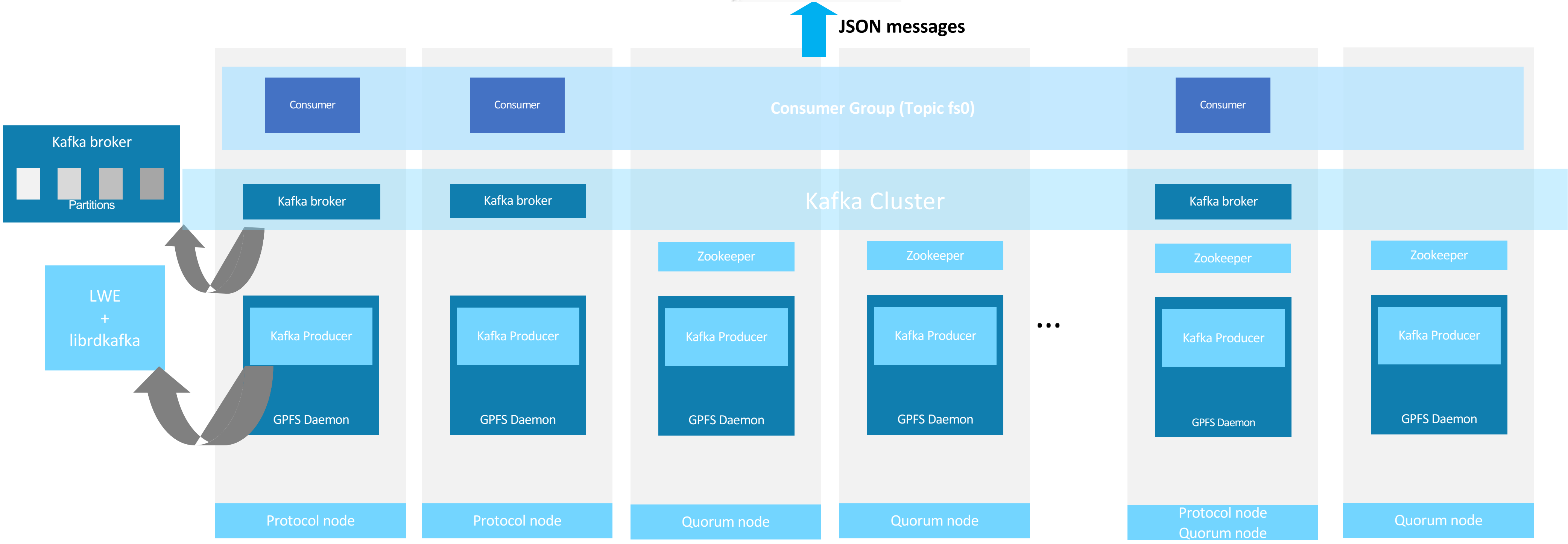


- Each audited filesystem will have an unique topic assigned to it in the MsgQueue
- Producers live inside the GPFS daemon publish events to the relevant topic
- Consumers subscribe to one topic
- Reliable architecture
  - Brokers are clustered
  - Consumer groups
  - Events replication across Brokers

Architecture Overview



```
{
  "clusterName": "ac_mo.almaden.ibm.com" ,
  "nodeName": "fin36v03p"
  "clientIp": ""
  "fsName": "ac_mo" ,
  "event": "OPEN" ,
  "path": "/ac/ac_mo/ac/big04" ,
  "inode": "136665" ,
  "openFlags": "32769" ,
  "poolName": "system" ,
  "fileSize": "30" ,
  "ownerUserId": "0" ,
  "ownerGroupId": "0" ,
  "acl": "070707"
  "atime": "2017-05-03_22:38:37-0700" ,
  "mtime": "2017-05-03_22:38:45-0700" ,
  "nowtime": "2017-05-04_15:18:07-0700" ,
  "clientUserId": "0" ,
  "clientGroupId": "0" ,
  "processId": "5327" ,
  "bytesRead": "0" ,
  "bytesWritten": "0"
  "xattrs": "user.lwetest=jdoe",
  "new_path": "" , <--- Not null when file renamed
  "new_xattrs": "", <-- Not null for mechattr
  "newacl": "", <--- Not NULL for chmod
  "newOwnerUserId": "",<--- Not NULL for chown
  "newGroupId": "", <-- Not NULL for chgrp
  "newPool": "" <--- Not NULL for change pool event
}
```

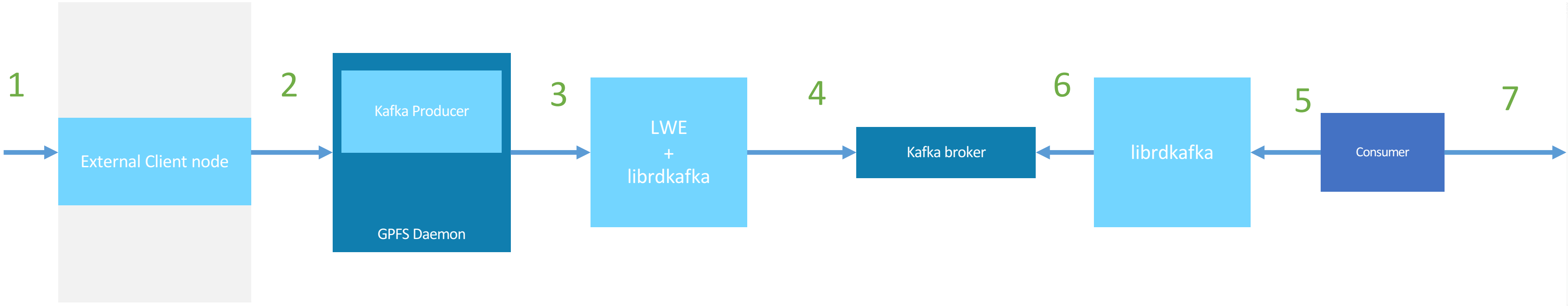


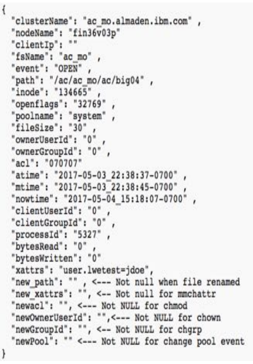
*Zookeeper resides on the quorum nodes

**Kafka Brokers can reside on any node (not confined to protocol nodes as depicted in this figure)

***Using the standardized JSON format, client facing API can be derived.

Flow of an event





```
{
  "clusterName": "ac_mo.almaden.ibm.com" ,
  "nodeName": "fin36v03p"
  "clientIp": ""
  "fsName": "ac_mo" ,
  "event": "OPEN" ,
  "path": "/ac/ac_mo/ac/big04" ,
  "inode": "134665" ,
  "openflags": "32769" ,
  "poolname": "system" ,
  "fileSize": "30" ,
  "ownerUserId": "0" ,
  "ownerGroupId": "0" ,
  "acl": "070707"
  "atime": "2017-05-03_22:38:37-0700" ,
  "mtime": "2017-05-03_22:38:45-0700" ,
  "nowtime": "2017-05-04_15:18:07-0700" ,
  "clientUserId": "0" ,
  "clientGroupId": "0" ,
  "processId": "5327" ,
  "bytesRead": "0" ,
  "bytesWritten": "0"
  "xattrs": "user.lwetest=jdoe",
  "new_path": "" , <--- Not null when file renamed
  "new_xattrs": "", <-- Not null for mmchattr
  "newacl": "", <--- Not NULL for chmod
  "newOwnerUserId": "",<--- Not NULL for chown
  "newGroupId": "", <-- Not NULL for chgrp
  "newPool": "" <--- Not NULL for change pool event
}
```

JSON messages

SeqNbr	Description
1	Client performs a file operation (read/ write/ remove, ..) on a file in an audited filesystem
2	External client node sends the client request to the relevant gpfs-node
3	Gpfs daemon using internal LWE (lightweight events) machinery sends the events to the msgQueue
4	Event messages are reliably delivered to msgQueue listening on this topic.

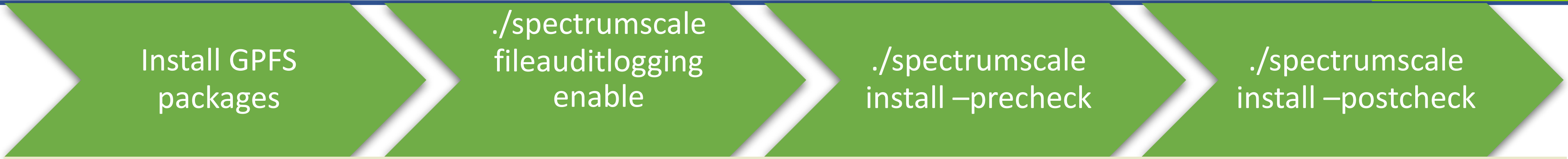
SeqNbr	Description
5, 6	Consumers belonging to a consumerGroup listening on this event topic, will periodically pull events from the msgQueue
7	Consumers will write the consumed events from the MsgQueue into the audited filesystem's ".audit_log" fileset.

Configuration and Setup

- Only Linux nodes (RHEL and Ubuntu)
- Linux Kernel version above > 3.10
- Minimum of 3 Linux quorum nodes
- Minimum of 3 nodes must be designated as Broker nodes
- Supported hardware platforms (x86 and PPCLE)
  - RHEL is supported on x86 and PPC LE
  - Ubuntu is only supported on x86
- Recommend that the ports 9092, 9093(not used currently, but will in future), 2181 and 2888-3888 are opened for TCP only.
- **Advanced License edition or the Data Management edition**

- During Installation, most configuration is automatically done and stored in /opt/kafka folder
- Free space requirements
  - min 5 GB local disk space per file system being audited
  - suggested 10 GB local disk space per file system being audited on all broker nodes
- 2 new rpms added to the package 5.0.0 release
  - gpfs.**kafka**-*
  - gpfs.librd**kafka**-*
- Java rpms installed on the Broker and Zookeeper nodes
  - **gpfs**.java-*

Installation - Linux Nodes Only


Install GPFS packages

./spectrumscale fileauditlogging enable

./spectrumscale install –precheck

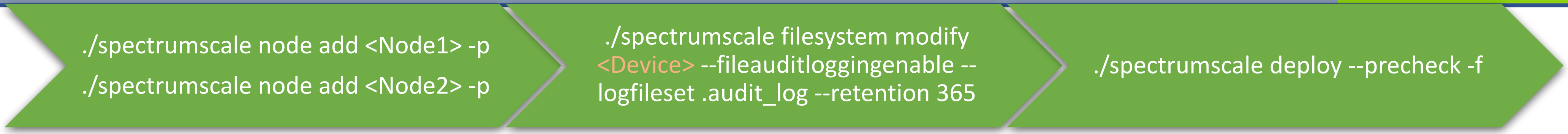
./spectrumscale install –postcheck

```
 # ./spectrumscale fileauditlogging enable
[ INFO  ] Enabling file audit logging in the cluster configuration file.
[ INFO  ] Tip :If all node designations and any required file audit logging configurations are complete,
proceed to assign filesystem to enable file audit logging configuration: ./spectrumscale filesystem
modify --fileauditloggingenable <filesystem name>.

 # ./spectrumscale node list
.
.
[ INFO  ] File Audit logging : Enabled

 # ./spectrumscale install –precheck
.
.
[ INFO  ] Performing FILE AUDIT LOGGING checks.
[ INFO  ] Running environment checks for file  Audit logging
[ INFO  ] File audit logging precheck OK≈
```

After install completes, verify that install installed the necessary GPFS rpms

```
# rpm -qa | egrep 'gpfs.java|kafka'
gpfs.java*
gpfs.kafka*
gpfs.librdkafka*

# ./spectrumscale install –postcheck
```

Installation – During deploy



./spectrumscale node add <Node1> -p
./spectrumscale node add <Node2> -p

./spectrumscale filesystem modify <Device> --fileauditloggingenable --logfileset .audit_log --retention 365

./spectrumscale deploy --precheck -f

1. Specify protocol nodes where Kafka Brokers will reside. Note: Shown below are 2 nodes for brevity, default configuration needs 3 protocol nodes.

```
 # ./spectrumscale node add my_protocol_node1 -p
[ INFO  ] Setting my_protocol_node1.xxx.com as a protocol node.
[ INFO  ] Configuration updated.
[ INFO  ] Tip : If all node designations are complete, configure the protocol environment as needed: ./spectrumscale
config protocols -f cesSharedRoot -m /ibm/cesSharedRoot
 # ./spectrumscale node add my_protocol_node2 -p
[ INFO  ] Setting my_protocol_node2.xxx.com as a protocol node.
[ INFO  ] Configuration updated.
[ INFO  ] Tip : If all node designations are complete, configure the protocol environment as needed: ./spectrumscale
config protocols -f cesSharedRoot -m /ibm/cesSharedRoot
```

2. Enable NFS and SMB during deploy

```
 # ./spectrumscale enable nfs
[ INFO  ] Enabling NFS on all protocol nodes.
[ INFO  ] Tip :If all node designations and any required protocol configurations are complete, proceed to check the
installation configuration:./spectrumscale deploy –precheck

 # ./spectrumscale enable smb
[ INFO  ] Enabling SMB on all protocol nodes.
[ INFO  ] Tip :If all node designations and any required protocol configurations are complete, proceed to check the
installation configuration:./spectrumscale deploy --precheck
```

3. During deploy configuration, modify filesystem(s) for audit logging

```
# ./spectrumscale filesystem modify fs0 --fileauditloggingenable --logfileset .audit_log --retention 2
[ INFO  ] The filesystem fs0 will be configured with file audit logging.
[ INFO  ] Tip : Now that you have modified this filesystem to use file audit logging, you need to enable it using the
 './spectrumscale fileauditlogging enable' command. please ignore if you have already enabled file audit logging.
[ INFO  ] The filesystem fs0 will be configured file audit logging with .audit_log log fileset.
[ INFO  ] The filesystem fs0 will be configured file audit logging with 2 retention days.
```

4. Deploy precheck will display precheck status of file audit logging

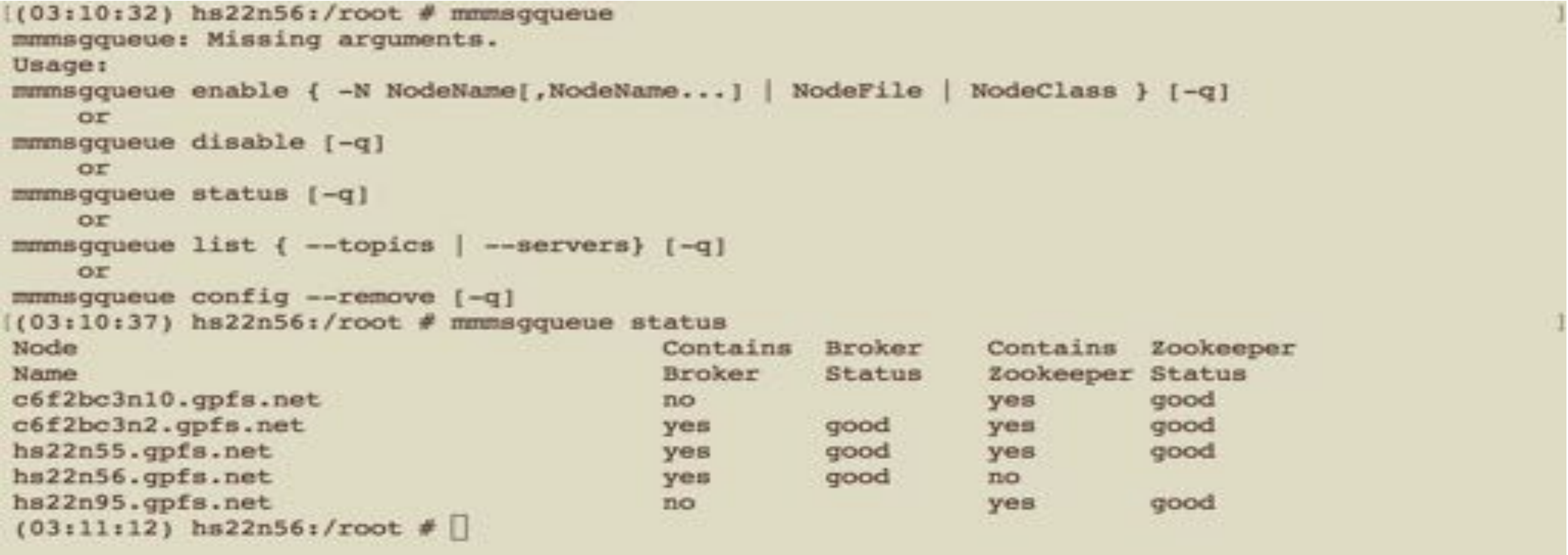
```
 # ./spectrumscale deploy --precheck -f
.
.
[ INFO  ] Performing FILE AUDIT LOGGING checks.
[ INFO  ] Running environment checks for file  Audit logging
[ INFO  ] File audit logging precheck OK
```

5. After running deploy, validate using mm-CLI commands to ensure file audit logging is

```
 # mmaudit all list
Audit     Cluster                           Fileset   Fileset          Retention
Device     ID                               Device     Name              (Days)
---------------------------------------------------------------------------------
fs0      4842233323150338002                 fs0      .audit_log        2
 # mmlsfs fs0 --file-audit-log
flag               value                description
------------------ ---------------------- ---------------------------------
 --file-audit-log   Yes                    File Audit Logging enabled?
```

Enablement - mmmsgqueue command

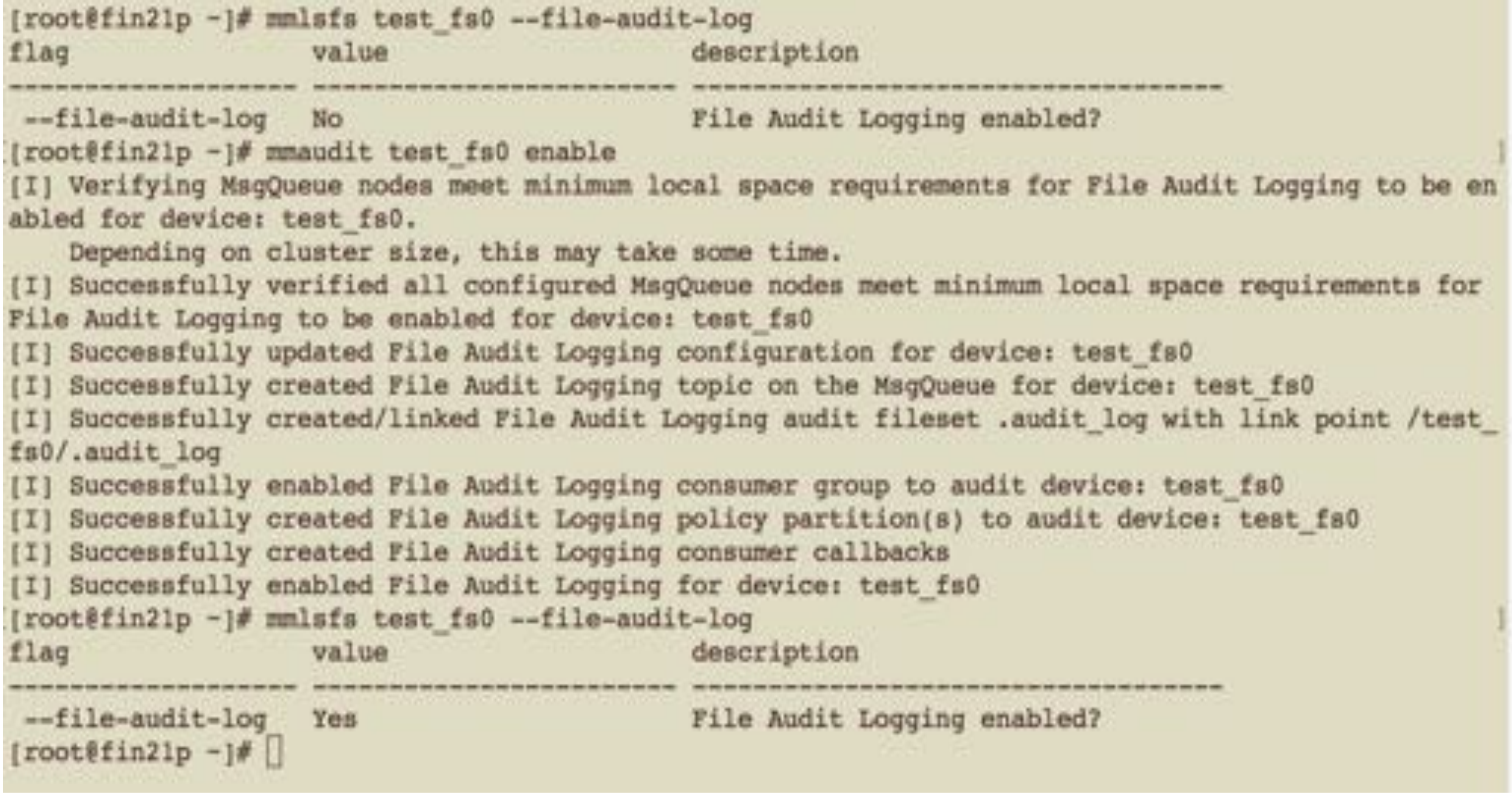
- Custom enablement of MsgQueue, to accommodate non-protocol nodes as Broker nodes



```
(03:10:32) hs22n56:/root # mmmsgqueue
mmmsgqueue: Missing arguments.
Usage:
mmmsgqueue enable { -N NodeName[,NodeName...] | NodeFile | NodeClass } [-q]
    or
mmmsgqueue disable [-q]
    or
mmmsgqueue status [-q]
    or
mmmsgqueue list { --topics | --servers} [-q]
    or
mmmsgqueue config --remove [-q]
(03:10:37) hs22n56:/root # mmmsgqueue status
Node                                    Contains  Broker    Contains  Zookeeper
Name                                    Broker    Status    Zookeeper Status
c6f2bc3n10.gpfs.net                     no                  yes       good
c6f2bc3n2.gpfs.net                      yes       good      yes       good
hs22n55.gpfs.net                        yes       good      yes       good
hs22n56.gpfs.net                        yes       good      no
hs22n95.gpfs.net                        no                  yes       good
(03:11:12) hs22n56:/root # 
```

Enablement - mmaudit command

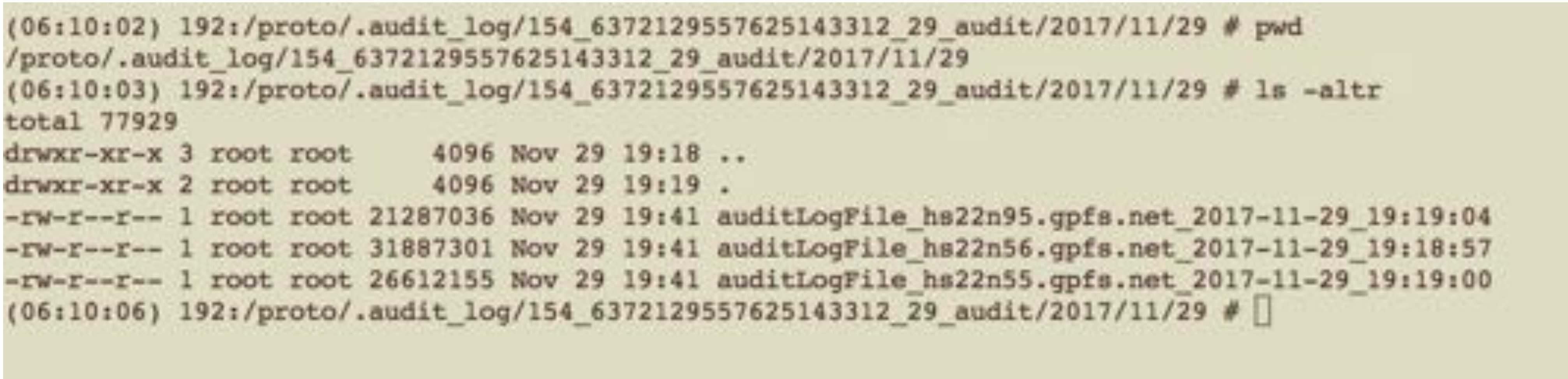
- Post Installation and deployment, File audit logging can be enabled using “mmaudit”



```
[root@fin21p ~]# mmlsfs test_fs0 --file-audit-log
flag                value                    description
------------------- ------------------------ -----------------------------------
 --file-audit-log   No                       File Audit Logging enabled?
[root@fin21p ~]# mmaudit test_fs0 enable
[I] Verifying MsgQueue nodes meet minimum local space requirements for File Audit Logging to be en
abled for device: test_fs0.
    Depending on cluster size, this may take some time.
[I] Successfully verified all configured MsgQueue nodes meet minimum local space requirements for
File Audit Logging to be enabled for device: test_fs0
[I] Successfully updated File Audit Logging configuration for device: test_fs0
[I] Successfully created File Audit Logging topic on the MsgQueue for device: test_fs0
[I] Successfully created/linked File Audit Logging audit fileset .audit_log with link point /test_
fs0/.audit_log
[I] Successfully enabled File Audit Logging consumer group to audit device: test_fs0
[I] Successfully created File Audit Logging policy partition(s) to audit device: test_fs0
[I] Successfully created File Audit Logging consumer callbacks
[I] Successfully enabled File Audit Logging for device: test_fs0
[root@fin21p ~]# mmlsfs test_fs0 --file-audit-log
flag                value                    description
------------------- ------------------------ -----------------------------------
 --file-audit-log   Yes                      File Audit Logging enabled?
[root@fin21p ~]# 
```

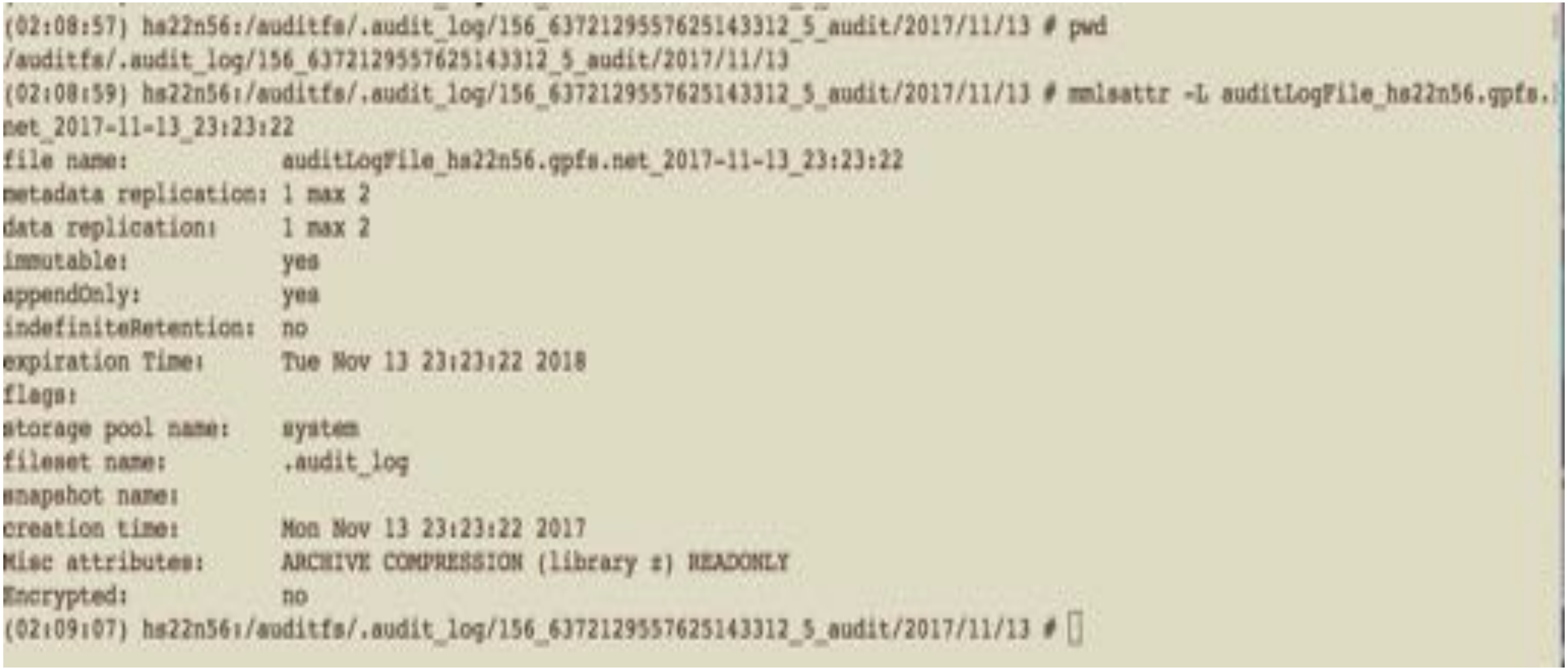
Logging details - Where is it logged

- Each file system enabled for file audit logging, has a dedicated fileset where the audit logs will go. Default option is .audit_log
- .audit_log fileset is created as IAM mode noncompliant.
  - Files cannot be deleted if retention time is not expired.
  - But retention times can be reset and files can be deleted but not changed, by root user only.
- AuditLog files are nested within /FS/.audit_log/topic/year/month/date/*
- Easy to search and consume



```
(06:10:02) 192:/proto/.audit_log/154_6372129557625143312_29_audit/2017/11/29 # pwd
/proto/.audit_log/154_6372129557625143312_29_audit/2017/11/29
(06:10:03) 192:/proto/.audit_log/154_6372129557625143312_29_audit/2017/11/29 # ls -altr
total 77929
drwxr-xr-x 3 root root     4096 Nov 29 19:18 ..
drwxr-xr-x 2 root root     4096 Nov 29 19:19 .
-rw-r--r-- 1 root root 21287036 Nov 29 19:41 auditLogFile_hs22n95.gpfs.net_2017-11-29_19:19:04
-rw-r--r-- 1 root root 31887301 Nov 29 19:41 auditLogFile_hs22n56.gpfs.net_2017-11-29_19:18:57
-rw-r--r-- 1 root root 26612155 Nov 29 19:41 auditLogFile_hs22n55.gpfs.net_2017-11-29_19:19:00
(06:10:06) 192:/proto/.audit_log/154_6372129557625143312_29_audit/2017/11/29 # ▯
```

- Live events can be monitored by tailing the current auditLogFile<…>
- Log file is written to an append only mode
- Rotation to a new log file ,upon reaching a threshold(500,000 events), is compressed and marked immutable for the retention period.
- Default retention period is 365 days



```
(02:08:57) hs22n56:/auditfs/.audit_log/156_637212955762514331 2_5_audit/2017/11/13 # pwd
/auditfs/.audit_log/156_6372129557625143312_5_audit/2017/11/13
(02:08:59) hs22n56:/auditfs/.audit_log/156_6372129557625143312_5_audit/2017/11/13 # mmlsattr -L auditLogFile_hs22n56.gpfs.
net_2017-11-13_23:23:22
file name:              auditLogFile_hs22n56.gpfs.net_2017-11-13_23:23:22
metadata replication: 1 max 2
data replication:     1 max 2
immutable:            yes
appendOnly:           yes
indefiniteRetention:  no
expiration Time:      Tue Nov 13 23:23:22 2018
flags:
storage pool name:    system
fileset name:         .audit_log
snapshot name:
creation time:        Mon Nov 13 23:23:22 2017
Misc attributes:      ARCHIVE COMPRESSION (library z) READONLY
Encrypted:            no
(02:09:07) hs22n56:/auditfs/.audit_log/156_6372129557625143312_5_audit/2017/11/13 # 
```

Logging details-What is logged (JSON)

```
{"LWE_JSON": "0.0.1", "path":  "/newfs/1Kfile2.restore", "oldPath":  null,  "clusterName":  "pardie.cluster",  "nodeName": "c6f2bc3n10", "nfsClientIp":  "",  "fsName":  "newfs",  "event": "OPEN", "inode":  "26626",  "openFlags":  "32962",  "poolName": "sp1", "fileSize":  "0",  "ownerUserId":  "0",  "ownerGroupId":  "0",  "atime":  "2017-10-25_12:36:22-0400",  "ctime":  "2017-10-25_12:36:22-0400",  "eventTime":  "2017-10-25_12:36:22-0400",  "clientUserId":  "0",  "clientGroupId":  "0", "processId":  "10437",  "permissions":  "200100644",  "acls":  "u::rwc,  g::r,  o::r,  ",  "xattrs":  null  }
```

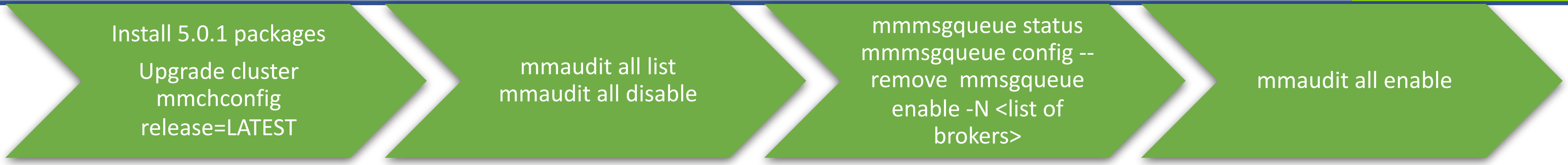
Attribute Name	Description
LWE_JSON	Version of the record
Path	Path name of the file involved in the event
oldPath	Previous path name of the file during RENAME event. For all other events indicated as null.
clusterName	Name of the cluster where the event took place
nodeName	Name of the node where the event took place
nfsClientIp	IP address of the remote client involved in the event
fsName	name of the file system involved in the event
event	event type. One of the following events {OPEN, CREATE, CLOSE,RENAME, XATTRCHANGE, ACLCHANGE, UNLINK, DESTROY, RMDIR}
inode	inode number of the file involved in the event

Attribute Name	Description
openFlags	open flags specified during the event (O_RDONLY, O_WRONLY,O_RDWR, O_CREAT, ...) as defined in fcntl.h
poolName	pool name where the file resides
fileSize	current size of the file in bytes
ownerUserId	owner id of the file involved in the event
ownerGroupId	group id of the file involved in the event
atime	The time in UTC format of the last access of the file involved in the event
ctime	The time in UTC format of the last status change of the file involved in the event
eventTime	The time in UTC format of the event
clientUserId	user id of process involved in the event
clientGroupId	group id of the process involved in the event
processId	process id involved in the event
permissions	permissions on the file involved in the event
acls	the access control lists involved in the event (Only in case of acl change event)
xattrs	the extended attributes involved in the event (Only in case of an Xattr change event)

Authentication

- Protection for non-GPFS producer / consumers from connecting to the MsgQueue
- Brokers (MsgQueue) is started with auth mode
  - SASL_PLAINTEXT (msgQ-gen=0) – for release 5.0.0
  - SASL_SCRAM (SHA-512) -- starting 5.0.1 release
- SASL_SCRAM the default authentication mode going forward.
- username and password are stored in the CCR
- Producer and Consumers will fetch {username:password} from CCR at FAL-enable / mount of the filesystem
- Whenever MsgQueue is disabled and re-enabled, MsgQueue generation number is incremented generating new {username:password}
- Additional level of validation with Producer and Consumers registering with the CCR using the MsgQueue-genNbr when fetching {username:password}

Upgrade from 5.0.0 to 5.0.1



Install 5.0.1 packages

Upgrade cluster mmchconfig release=LATEST

mmaudit all list
mmaudit all disable

mmmsgqueue status
mmmsgqueue config --remove mmsgqueue enable -N <list of brokers>

mmaudit all enable

- Change in authentication mode from PLAINTEXT to SCRAM
- One time re-configuration of the MsgQueue with SCRAM configuration
- Additional openssl and libssl-dev Linux libraries needed for the new authentication mode
    - For RHEL, openssl-devel and cyrus-sasl-devel packages
    - For Ubuntu, libssl-dev and libsasl2-dev packages

Manually upgrading FAL from 5.0.0 to 5.0.1

1. Upgrade cluster to latest release (5.0.0 to 5.0.1)

```
root@windwalker-vm1:~# mmchconfig release=LATEST
Verifying that all nodes in the cluster are up-to-date ...
mmchconfig: Command successfully completed
mmchconfig: Propagating the cluster configuration data to all
  affected nodes.  This is an asynchronous process.
```

2. List the existing file systems that are file audit logging enabled

```
root@windwalker-vm1:~# mmaudit all list
Audit       Cluster                          Fileset          Fileset         Retention
Device   ID                       Device     Name                             (Days)
---------------------------------------------------------------------------------------
fs0         6391413883505451835   fs0        .audit_log_wind_fs0              25
fs1         6391413883505451835   fs1        .audit_log_wind_fs1              365
```

3. Disabling all the file audit logging enabled file systems, in this example

```
root@windwalker-vm1:~# mmaudit fs0 disable
[I] Successfully deleted File Audit Logging policy partition(s) for device: fs0
[I] Successfully disabled File Audit Logging consumer group for device: fs0
[I] Successfully disabled ACL access to the File Audit Logging topic of the MsgQueue for device: fs0
[I] Successfully deleted File Audit Logging topic from the MsgQueue for device: fs0
[I] Successfully updated File Audit Logging configuration for device: fs0
[I] Successfully disabled File Audit Logging for device: fs0
```

```
root@windwalker-vm1:~# mmaudit fs1 disable
[I] Successfully deleted File Audit Logging policy partition(s) for device: fs1
[I] Successfully disabled File Audit Logging consumer group for device: fs1
[I] Successfully disabled ACL access to the File Audit Logging topic of the MsgQueue for device: fs1
[I] Successfully deleted File Audit Logging topic from the MsgQueue for device: fs1
[I] Successfully updated File Audit Logging configuration for device: fs1
[I] Successfully removed File Audit Logging consumer callbacks
[I] Successfully removed File Audit Logging consumer node class kafkaAuditConsumerServers
[I] Successfully disabled File Audit Logging for device: fs1
```

4. Checking the message queue status, recording which nodes are broker nodes, and removing the message queue

```
root@windwalker-vm1:~# mmmsgqueue status
Node                                Contains   Broker    Contains     Zookeeper
Name                                Broker     Status    Zookeeper    Status
windwalker-vm1.tuc.stglabs.ibm.com  yes        good      yes          good
windwalker-vm2.tuc.stglabs.ibm.com  yes        good      yes          good
windwalker-vm3.tuc.stglabs.ibm.com  yes        good      yes          good
windwalker-vm4.tuc.stglabs.ibm.com  yes        good      no
windwalker-vm5.tuc.stglabs.ibm.com  no                   yes          good
windwalker-vm6.tuc.stglabs.ibm.com  no                   yes          good
root@windwalker-vm1:~# mmmsgqueue config --remove
[I] Attempting to disable the MsgQueue.  This may take some time.
[I] Disabling MsgQueue daemons.
[I] Removing callbacks that control starting and stopping the MsgQueue daemons.
[I] MsgQueue successfully disabled.
[I] Removing MsgQueue callbacks, node classes and configuration information if present.
[I] MsgQueue successfully disabled and configuration removed.
```

5. Re-enabling the message queue using the same broker nodes from before

```
root@windwalker-vm1:~# mmmsgqueue enable -N windwalker-vm1.tuc.stglabs.ibm.com,windwalker-vm2.tuc.stglabs.ibm.com,
windwalker-vm3.tuc.stglabs.ibm.com,windwalker-vm4.tuc.stglabs.ibm.com
[I] The kafkaZookeeperServers node class was successfully created with 5 member nodes.
[I] The kafkaBrokerServers node class was successfully created with 4 member nodes.
[I] Successfully created Kafka broker configuration file and added to CCR.
[I] Successfully created Kafka Zookeeper configuration file and added to CCR.
[I] Enabling MsgQueue daemons.
[I] Creating callbacks to control starting and stopping the MsgQueue daemons.
[I] Pushing producer authentication information to eligible cluster nodes.
    Depending on cluster size, this may take some time.
[I] MsgQueue successfully enabled.
```

6. Enable FAL for fs0 and fs1

```
root@windwalker-vm1:~# mmaudit fs0 enable
[I] Successfully created File Audit Logging consumer node class kafkaAuditConsumerServers
[I] Verifying MsgQueue nodes meet minimum local space requirements for File Audit Logging to be enabled for device: fs0.
    Depending on cluster size, this may take some time.
[I] Successfully verified all configured MsgQueue nodes meet minimum local space requirements for File Audit Logging to be enabled
 for device: fs0
[I] Successfully updated File Audit Logging configuration for device: fs0
[I] Successfully created File Audit Logging topic on the MsgQueue for device: fs0
[I] Successfully enabled ACL access to the topic for producers and consumers for device: fs0
[I] Successfully created/linked File Audit Logging audit fileset .audit_log with link point /fs0/.audit_log
[I] Successfully enabled File Audit Logging consumer group to audit device: fs0
[I] Successfully created File Audit Logging policy partition(s) to audit device: fs0
[I] Successfully created File Audit Logging consumer callbacks
[I] Successfully enabled File Audit Logging for device: fs0
```

```
root@windwalker-vm1 [root@fin21p ~]# mmaudit fs1 enable --log-fileset .audit_log_SCRAM_fs1 --retention 10
[I] Successfully created File Audit Logging consumer node class kafkaAuditConsumerServers
[I] Verifying MsgQueue nodes meet minimum local space requirements for File Audit Logging to be enabled for device: fs1.
   Depending on cluster size, this may take some time.
[I] Successfully verified all configured MsgQueue nodes meet minimum local space requirements for File Audit Logging to
be enabled for device: fs1
[I] Successfully updated File Audit Logging configuration for device: fs1
[I] Successfully created File Audit Logging topic on the MsgQueue for device: fs1
[I] Successfully enabled ACL access to the topic for producers and consumers for device: fs1
[I] Successfully created/linked File Audit Logging audit fileset .audit_log_SCRAM_lroc_fs with link point /fs1/.audit_log_SCRAM_fs1
[I] Successfully enabled File Audit Logging consumer group to audit device: fs1
[I] Successfully created File Audit Logging policy partition(s) to audit device: fs1
[I] Successfully created File Audit Logging consumer callbacks
[I] Successfully enabled File Audit Logging for device: fs1
```

6. Finally, view the new file audit logging configuration

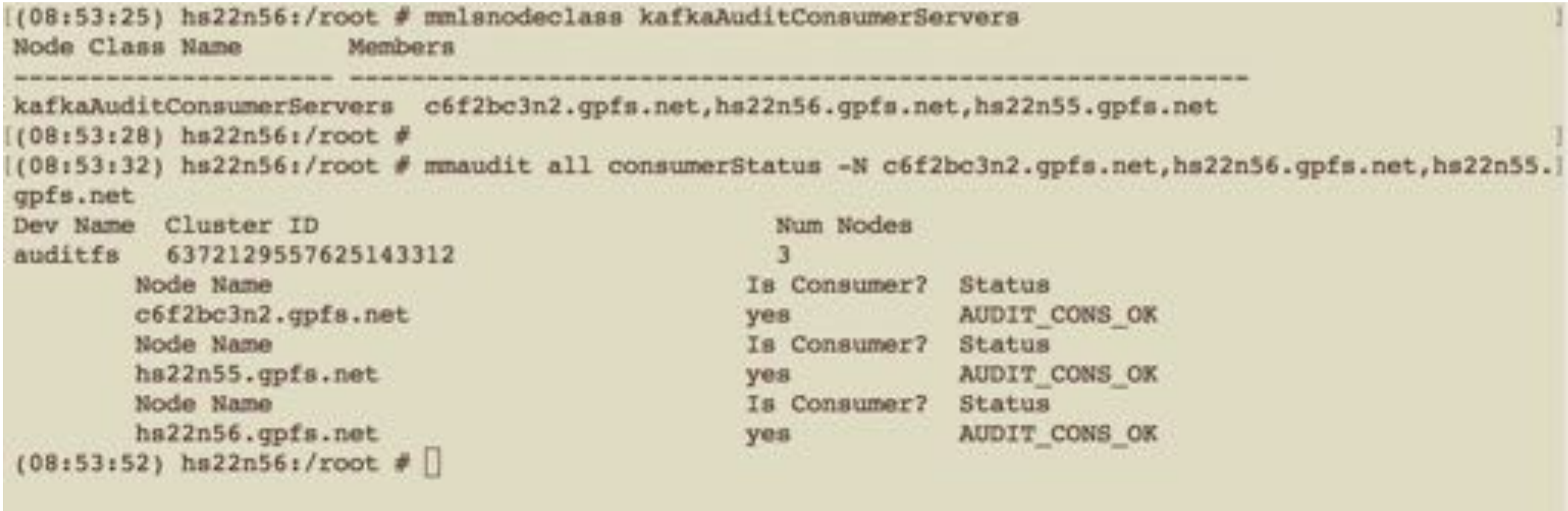
```
root@windwalker-vm1:~# mmaudit all list
Audit     Cluster                        Fileset    Fileset                  Retention
Device    ID                             Device     Name                     (Days)
------------------------------------------------------------------------------------------------
fs0       6391413883505451835             fs0       .audit_log               365
fs1       6391413883505451835             fs1       .audit_log_SCRAM_fs1      10
```

Health monitoring for FAL

- Monitoring using CLI commands
  - mmaudit
  - mmmsgqueue
  - mmpmon
- Monitoring using mmhealth
  - Cluster wide
  - Node view
- Monitoring of FILEAUDITLOG component
  - auditc_xxx events raised for various error and warning scenarios
- Monitoring of MSGQUEUE component
  - Kafka_xxx | zookeeper_xxx events raised for various msgQueue error and warning scenarios
- Monitoring using GUI
  - Via the Service and Events panel

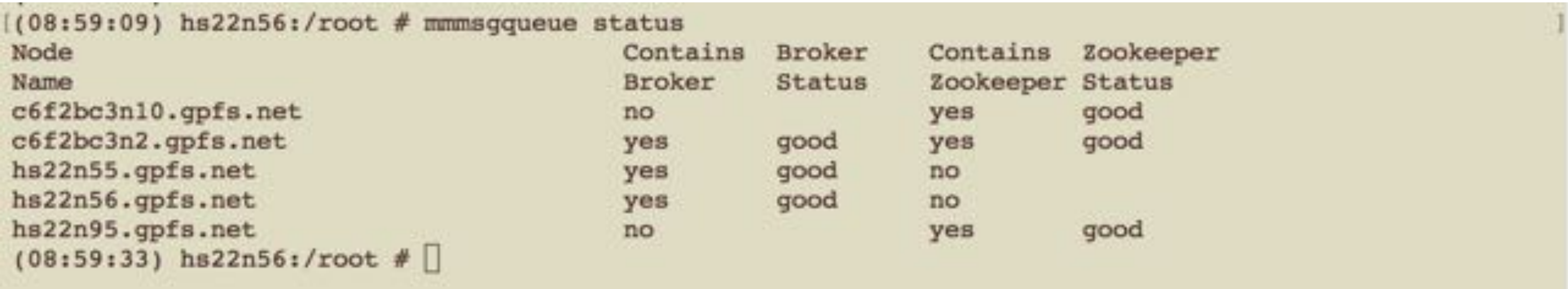
FAL monitoring using CLI-cmds

- mmaudit all consumerStatus –N …



```
(08:53:25) hs22n56:/root # mmlsnodeclass kafkaAuditConsumerServers
Node Class Name       Members
--------------------- -----------------------------------------------------------
kafkaAuditConsumerServers  c6f2bc3n2.gpfs.net,hs22n56.gpfs.net,hs22n55.gpfs.net
(08:53:28) hs22n56:/root #
(08:53:32) hs22n56:/root # mmaudit all consumerStatus -N c6f2bc3n2.gpfs.net,hs22n56.gpfs.net,hs22n55.
gpfs.net
Dev Name  Cluster ID                              Num Nodes
auditfs   6372129557625143312                     3
        Node Name                               Is Consumer?  Status
        c6f2bc3n2.gpfs.net                      yes           AUDIT_CONS_OK
        Node Name                               Is Consumer?  Status
        hs22n55.gpfs.net                        yes           AUDIT_CONS_OK
        Node Name                               Is Consumer?  Status
        hs22n56.gpfs.net                        yes           AUDIT_CONS_OK
(08:53:52) hs22n56:/root # 
```

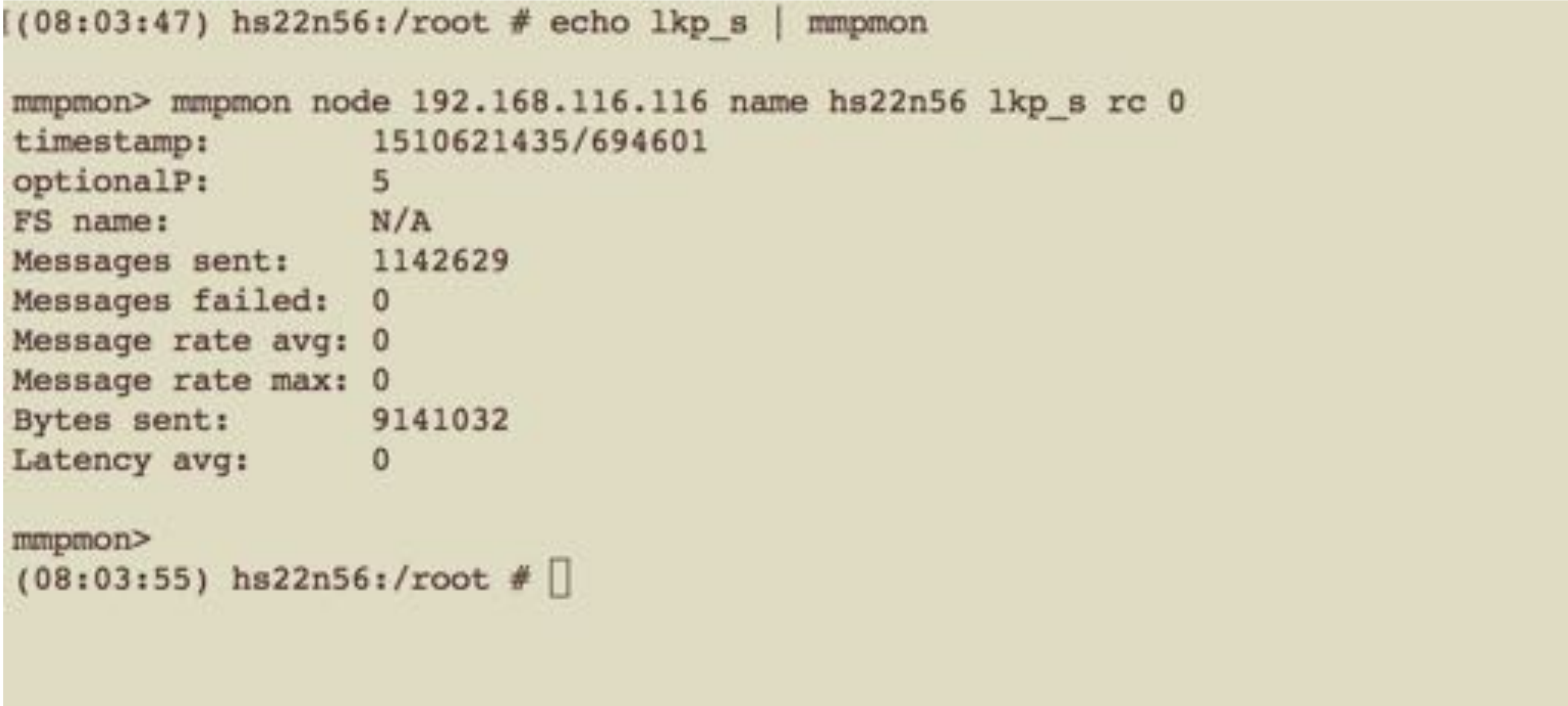
- mmmsgqueue status



```
(08:59:09) hs22n56:/root # mmmsgqueue status
Node                                    Contains  Broker    Contains  Zookeeper
Name                                    Broker    Status    Zookeeper Status
c6f2bc3n10.gpfs.net                     no                  yes       good
c6f2bc3n2.gpfs.net                      yes       good      yes       good
hs22n55.gpfs.net                        yes       good      no
hs22n56.gpfs.net                        yes       good      no
hs22n95.gpfs.net                        no                  yes       good
(08:59:33) hs22n56:/root # 
```

FAL monitoring using CLI-cmds

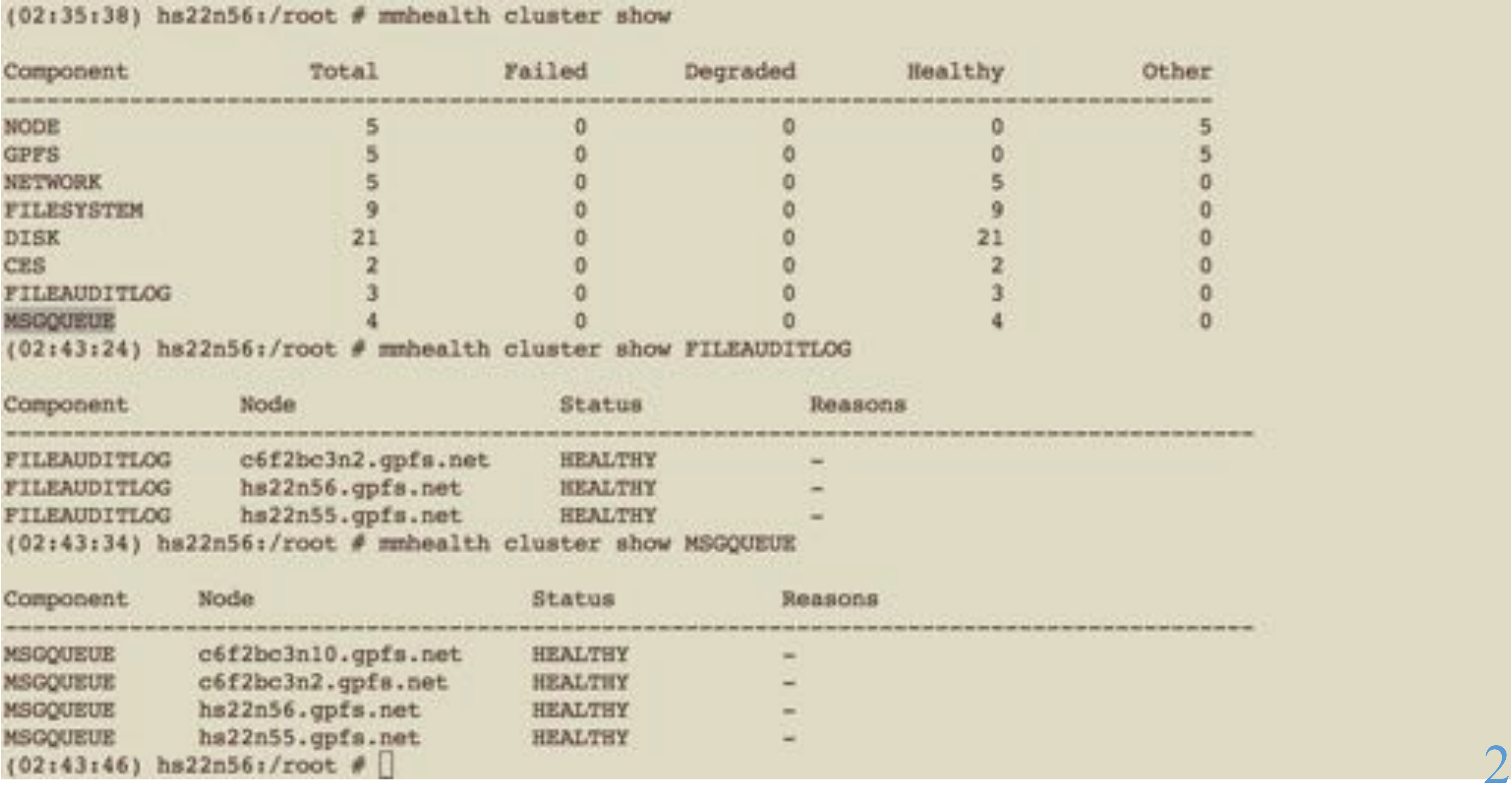
- mmpmon lkp_s



```
(08:03:47) hs22n56:/root # echo lkp_s | mmpmon

mmpmon> mmpmon node 192.168.116.116 name hs22n56 lkp_s rc 0
timestamp:          1510621435/694601
optionalP:          5
FS name:            N/A
Messages sent:      1142629
Messages failed:    0
Message rate avg:   0
Message rate max:   0
Bytes sent:         9141032
Latency avg:        0

mmpmon>
(08:03:55) hs22n56:/root #
```

Cluster wide: mmhealth cluster show

- Periodic polling and event callback registration mechanism is used.
- Possible lag in determining the health due to polling constraints.

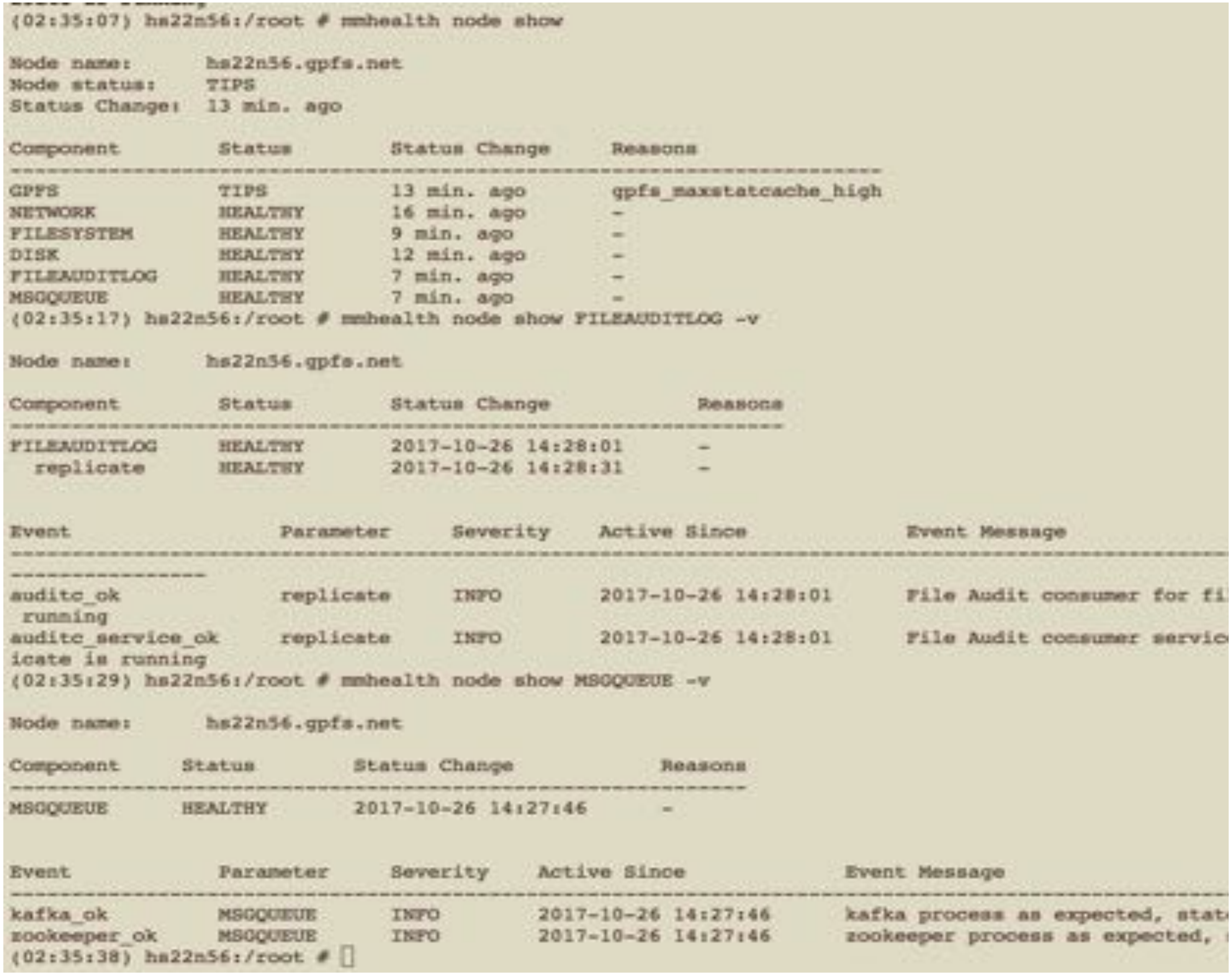


```
(02:35:38) hs22n56:/root # mmhealth cluster show

Component              Total          Failed        Degraded         Healthy          Other
-------------------------------------------------------------------------------------------
NODE                       5               0               0               0              5
GPFS                       5               0               0               0              5
NETWORK                    5               0               0               5              0
FILESYSTEM                 9               0               0               9              0
DISK                      21               0               0              21              0
CES                        2               0               0               2              0
FILEAUDITLOG               3               0               0               3              0
MSGQUEUE                   4               0               0               4              0
(02:43:24) hs22n56:/root # mmhealth cluster show FILEAUDITLOG

Component        Node                     Status           Reasons
-------------------------------------------------------------------------------------------
FILEAUDITLOG     c6f2bc3n2.gpfs.net       HEALTHY          -
FILEAUDITLOG     hs22n56.gpfs.net         HEALTHY          -
FILEAUDITLOG     hs22n55.gpfs.net         HEALTHY          -
(02:43:34) hs22n56:/root # mmhealth cluster show MSGQUEUE

Component     Node                     Status           Reasons
-------------------------------------------------------------------------------------------
MSGQUEUE      c6f2bc3n10.gpfs.net      HEALTHY          -
MSGQUEUE      c6f2bc3n2.gpfs.net       HEALTHY          -
MSGQUEUE      hs22n56.gpfs.net         HEALTHY          -
MSGQUEUE      hs22n55.gpfs.net         HEALTHY          -
(02:43:46) hs22n56:/root #
```

Node view: mmhealth node show

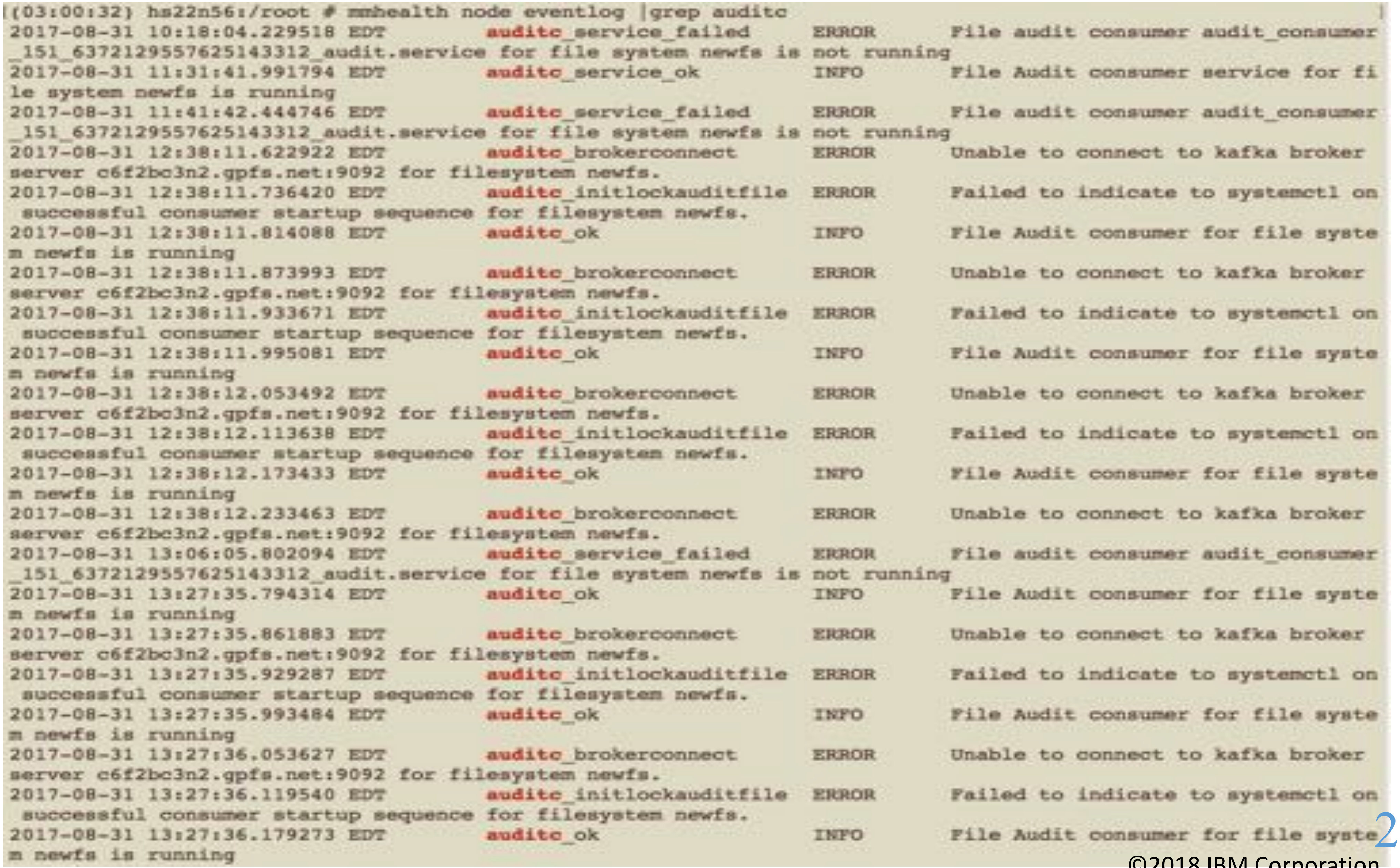
Two separate components monitored

- FILEAUDITLOG
- MSGQUEUE



```
(02:35:07) hs22n56:/root # mmhealth node show

Node name:      hs22n56.gpfs.net
Node status:    TIPS
Status Change:  13 min. ago

Component        Status        Status Change     Reasons
----------------------------------------------------------------------
GPFS             TIPS          13 min. ago       gpfs_maxstatcache_high
NETWORK          HEALTHY       16 min. ago       -
FILESYSTEM       HEALTHY       9 min. ago        -
DISK             HEALTHY       12 min. ago       -
FILEAUDITLOG     HEALTHY       7 min. ago        -
MSGQUEUE         HEALTHY       7 min. ago        -
(02:35:17) hs22n56:/root # mmhealth node show FILEAUDITLOG -v

Node name:      hs22n56.gpfs.net

Component        Status        Status Change            Reasons
---------------------------------------------------------------
FILEAUDITLOG     HEALTHY       2017-10-26 14:28:01      -
  replicate      HEALTHY       2017-10-26 14:28:31      -


Event                 Parameter     Severity    Active Since             Event Message
------------------------------------------------------------------------------------------------
----------------
auditc_ok             replicate     INFO        2017-10-26 14:28:01      File Audit consumer for fi
 running
auditc_service_ok     replicate     INFO        2017-10-26 14:28:01      File Audit consumer servic
icate is running
(02:35:29) hs22n56:/root # mmhealth node show MSGQUEUE -v

Node name:      hs22n56.gpfs.net

Component     Status        Status Change            Reasons
------------------------------------------------------------
MSGQUEUE      HEALTHY       2017-10-26 14:27:46      -


Event            Parameter     Severity    Active Since             Event Message
------------------------------------------------------------------------------------------------
kafka_ok         MSGQUEUE      INFO        2017-10-26 14:27:46      kafka process as expected, stat
zookeeper_ok     MSGQUEUE      INFO        2017-10-26 14:27:46      zookeeper process as expected,
(02:35:38) hs22n56:/root # ▯
```

Events view: mmhealth eventlog show

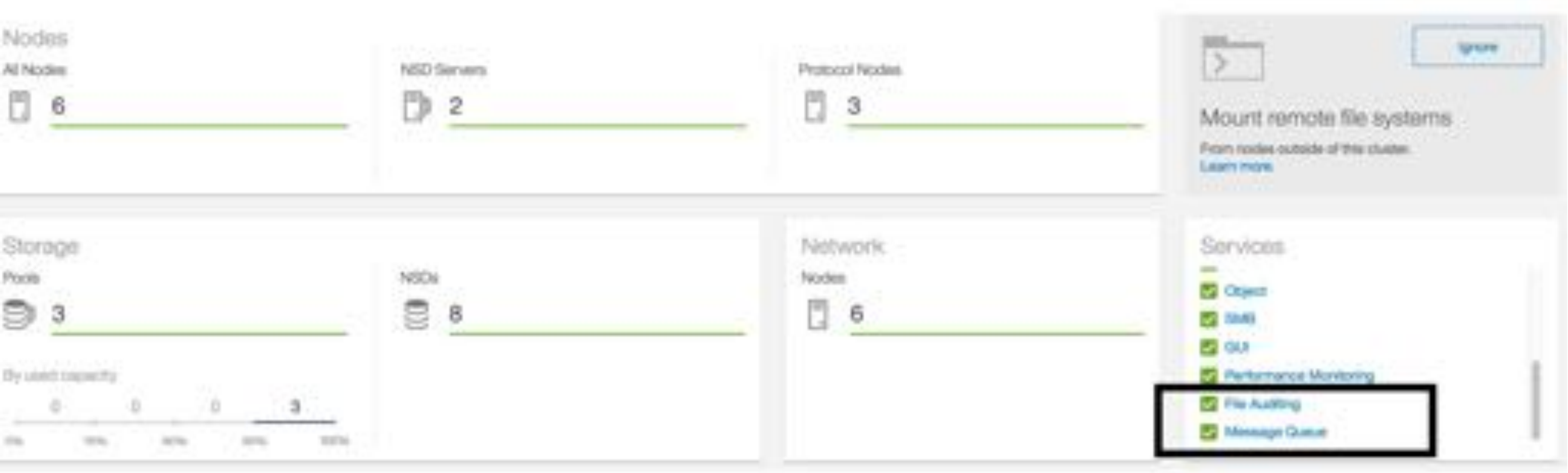


```
(03:00:32) hs22n56:/root # mmhealth node eventlog |grep auditc
2017-08-31 10:18:04.229518 EDT         auditc_service_failed     ERROR      File audit consumer audit_consumer
_151_6372129557625143312_audit.service for file system newfs is not running
2017-08-31 11:31:41.991794 EDT         auditc_service_ok         INFO       File Audit consumer service for fi
le system newfs is running
2017-08-31 11:41:42.444746 EDT         auditc_service_failed     ERROR      File audit consumer audit_consumer
_151_6372129557625143312_audit.service for file system newfs is not running
2017-08-31 12:38:11.622922 EDT         auditc_brokerconnect      ERROR      Unable to connect to kafka broker
server c6f2bc3n2.gpfs.net:9092 for filesystem newfs.
2017-08-31 12:38:11.736420 EDT         auditc_initlockauditfile  ERROR      Failed to indicate to systemctl on
 successful consumer startup sequence for filesystem newfs.
2017-08-31 12:38:11.814088 EDT         auditc_ok                 INFO       File Audit consumer for file syste
m newfs is running
2017-08-31 12:38:11.873993 EDT         auditc_brokerconnect      ERROR      Unable to connect to kafka broker
server c6f2bc3n2.gpfs.net:9092 for filesystem newfs.
2017-08-31 12:38:11.933671 EDT         auditc_initlockauditfile  ERROR      Failed to indicate to systemctl on
 successful consumer startup sequence for filesystem newfs.
2017-08-31 12:38:11.995081 EDT         auditc_ok                 INFO       File Audit consumer for file syste
m newfs is running
2017-08-31 12:38:12.053492 EDT         auditc_brokerconnect      ERROR      Unable to connect to kafka broker
server c6f2bc3n2.gpfs.net:9092 for filesystem newfs.
2017-08-31 12:38:12.113638 EDT         auditc_initlockauditfile  ERROR      Failed to indicate to systemctl on
 successful consumer startup sequence for filesystem newfs.
2017-08-31 12:38:12.173433 EDT         auditc_ok                 INFO       File Audit consumer for file syste
m newfs is running
2017-08-31 12:38:12.233463 EDT         auditc_brokerconnect      ERROR      Unable to connect to kafka broker
server c6f2bc3n2.gpfs.net:9092 for filesystem newfs.
2017-08-31 13:06:05.802094 EDT         auditc_service_failed     ERROR      File audit consumer audit_consumer
_151_6372129557625143312_audit.service for file system newfs is not running
2017-08-31 13:27:35.794314 EDT         auditc_ok                 INFO       File Audit consumer for file syste
m newfs is running
2017-08-31 13:27:35.861883 EDT         auditc_brokerconnect      ERROR      Unable to connect to kafka broker
server c6f2bc3n2.gpfs.net:9092 for filesystem newfs.
2017-08-31 13:27:35.929287 EDT         auditc_initlockauditfile  ERROR      Failed to indicate to systemctl on
 successful consumer startup sequence for filesystem newfs.
2017-08-31 13:27:35.993484 EDT         auditc_ok                 INFO       File Audit consumer for file syste
m newfs is running
2017-08-31 13:27:36.053627 EDT         auditc_brokerconnect      ERROR      Unable to connect to kafka broker
server c6f2bc3n2.gpfs.net:9092 for filesystem newfs.
2017-08-31 13:27:36.119540 EDT         auditc_initlockauditfile  ERROR      Failed to indicate to systemctl on
 successful consumer startup sequence for filesystem newfs.
2017-08-31 13:27:36.179273 EDT         auditc_ok                 INFO       File Audit consumer for file syste
m newfs is running
```

FAL monitoring from the GUI

Home screen

- On the right-hand you can see the overall File Auditing and Message Queue status



Nodes

All Nodes 6

NSD Servers 2

Protocol Nodes 3

Ignore

Mount remote file systems

From nodes outside of this cluster.
Learn more

Storage

Pools 3

By used capacity

NSDs 8

Network

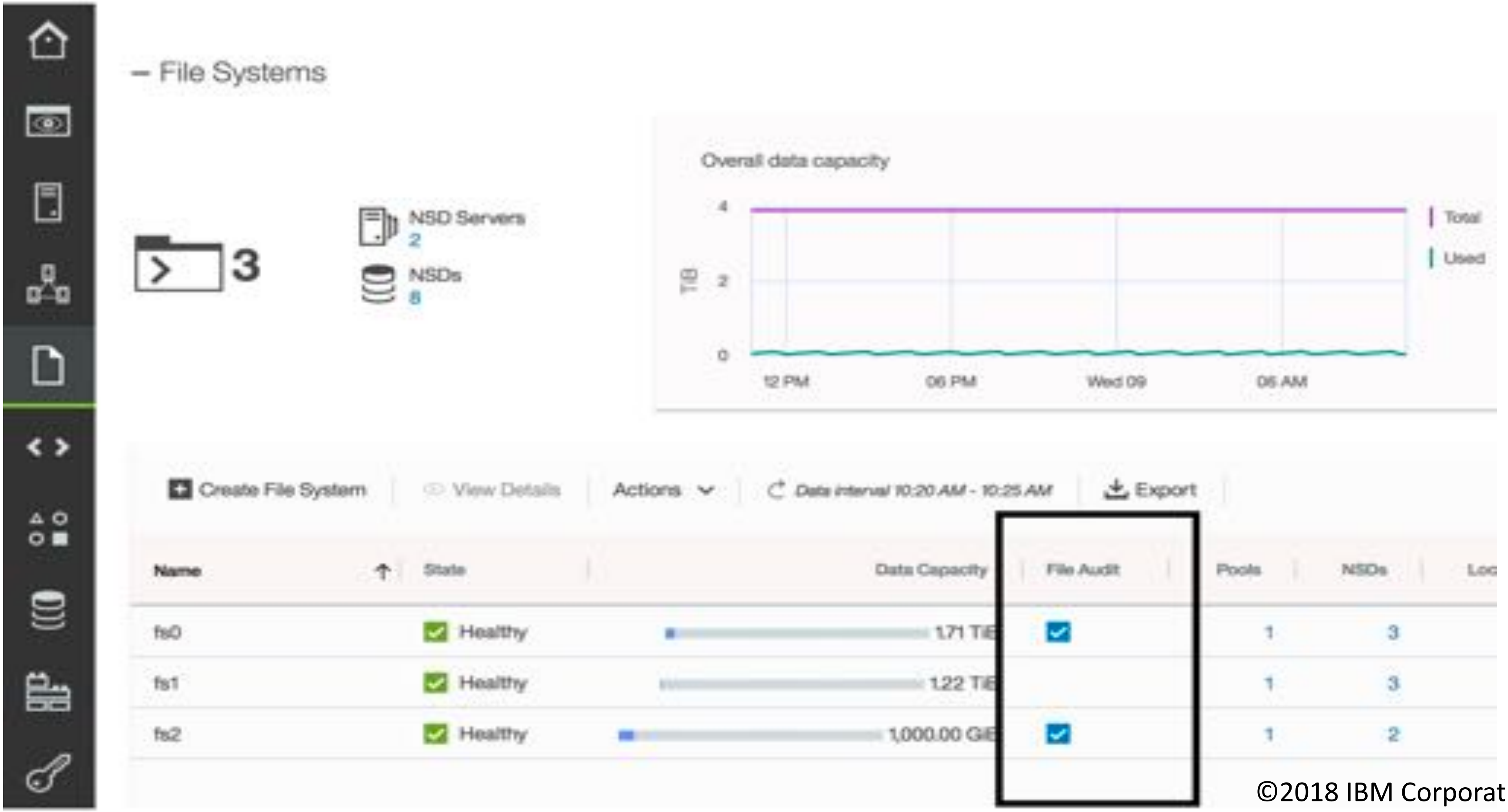
Nodes 6

Services

☑ Object
☑ SMB
☑ GUI
☑ Performance Monitoring
☑ File Auditing
☑ Message Queue

GUI – File Systems Panel

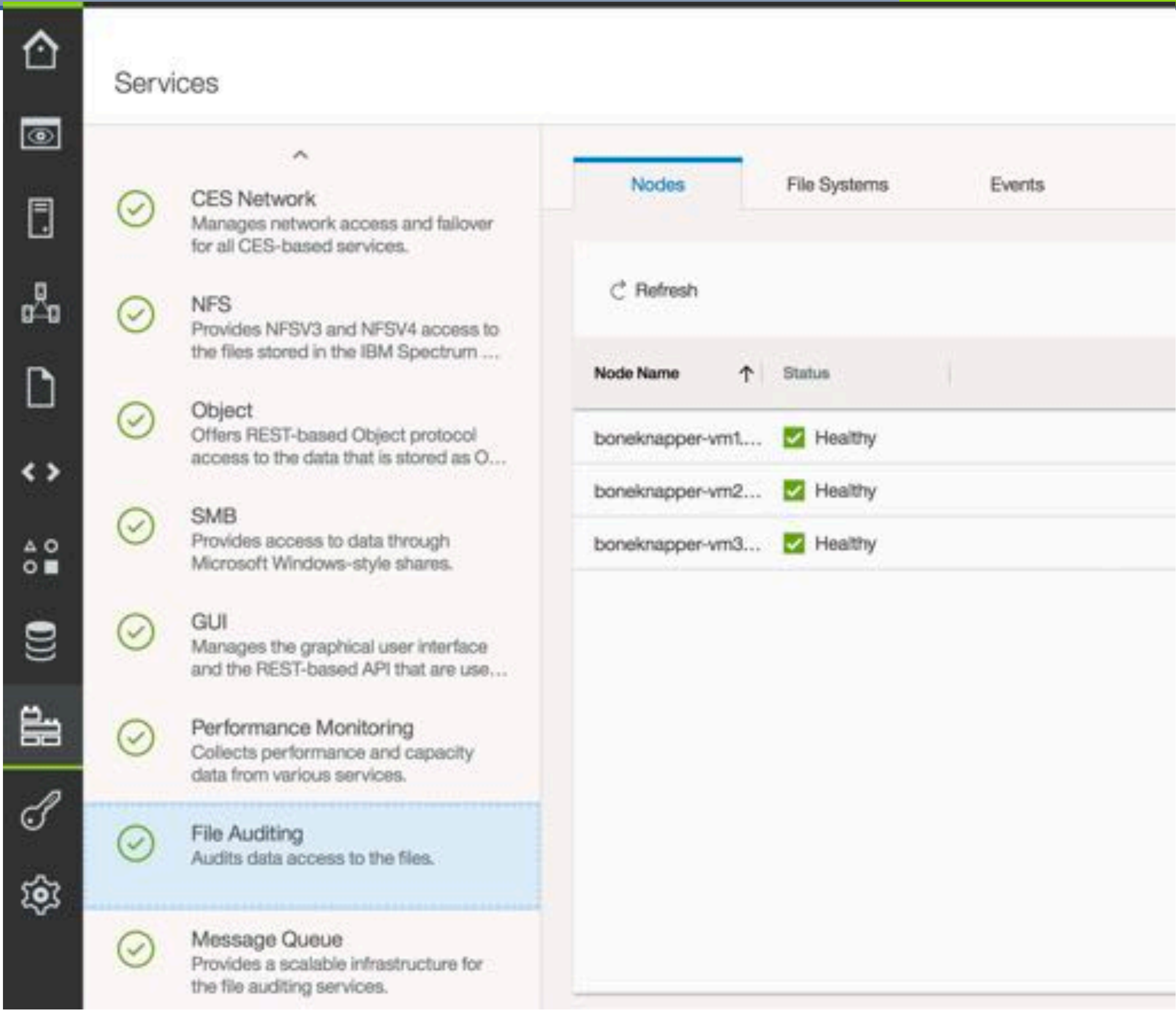
- Which file systems are enabled for FAL.
- Request this by using the Actions pull-down that is shown and then customize the columns to view the file audited file systems.



Name	State	Data Capacity	File Audit	Pools	NSDs	Loc
fs0	☑ Healthy	1.71 TiB	☑	1	3	
fs1	☑ Healthy	1.22 TiB		1	3	
fs2	☑ Healthy	1,000.00 GiB	☑	1	2	

GUI – Services → File Auditing Panel

- View the overall File Auditing status for each node.
- This is a healthy system, so there is nothing in the Events section.



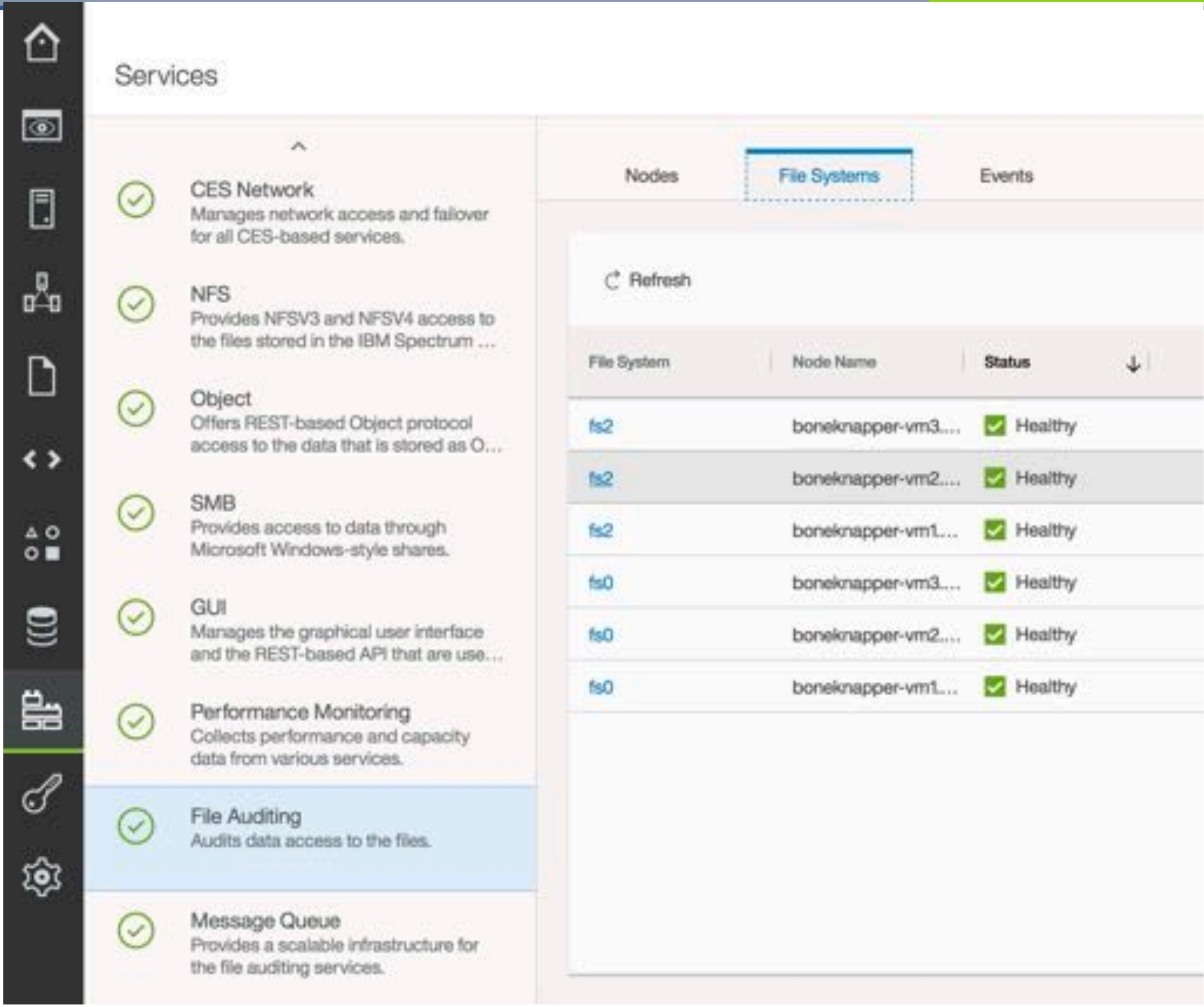
Services

CES Network
Manages network access and failover for all CES-based services.
NFS
Provides NFSV3 and NFSV4 access to the files stored in the IBM Spectrum ...
Object
Offers REST-based Object protocol access to the data that is stored as O...
SMB
Provides access to data through Microsoft Windows-style shares.
GUI
Manages the graphical user interface and the REST-based API that are use...
Performance Monitoring
Collects performance and capacity data from various services.
File Auditing
Audits data access to the files.
Message Queue
Provides a scalable infrastructure for the file auditing services.
Nodes
File Systems
Events
Refresh

Node Name	Status
boneknapper-vm1....	☑ Healthy
boneknapper-vm2...	☑ Healthy
boneknapper-vm3...	☑ Healthy

GUI – Services → File Auditing Panel

- View the Auditing status at the File System level.



Services

CES Network
Manages network access and failover for all CES-based services.

NFS
Provides NFSV3 and NFSV4 access to the files stored in the IBM Spectrum ...

Object
Offers REST-based Object protocol access to the data that is stored as O...

SMB
Provides access to data through Microsoft Windows-style shares.

GUI
Manages the graphical user interface and the REST-based API that are use...

Performance Monitoring
Collects performance and capacity data from various services.

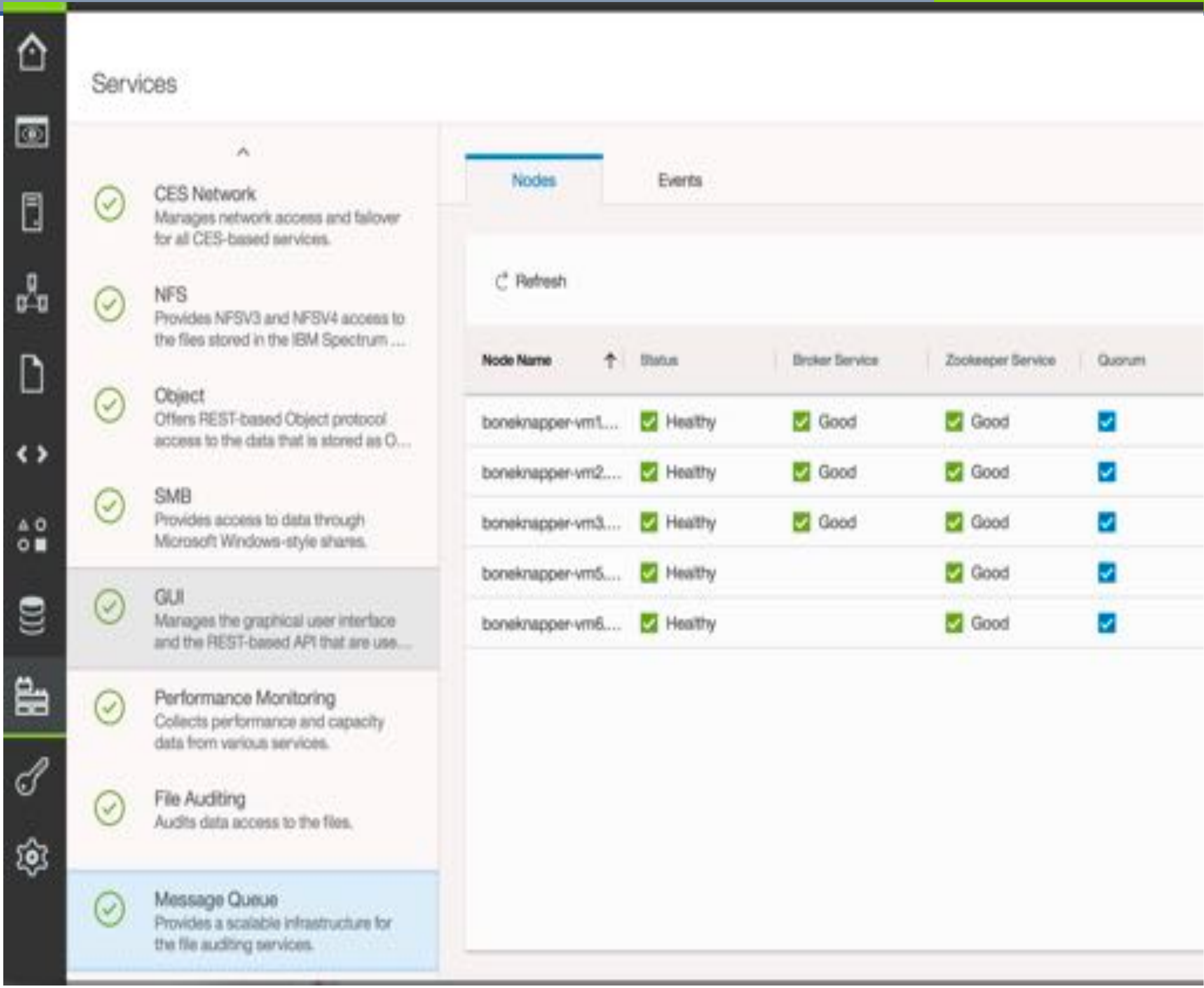
File Auditing
Audits data access to the files.

Message Queue
Provides a scalable infrastructure for the file auditing services.

File System	Node Name	Status
fs2	boneknapper-vm3....	✓ Healthy
fs2	boneknapper-vm2....	✓ Healthy
fs2	boneknapper-vm1....	✓ Healthy
fs0	boneknapper-vm3....	✓ Healthy
fs0	boneknapper-vm2....	✓ Healthy
fs0	boneknapper-vm1....	✓ Healthy

GUI – Services → Message Queue Panel

- view the members of the message queue.
- aligns with the "mmmsgqueue status" CLI command.
- This is a healthy system, so there is nothing in the Events section.



Services

CES Network
Manages network access and failover for all CES-based services.

NFS
Provides NFSV3 and NFSV4 access to the files stored in the IBM Spectrum ...

Object
Offers REST-based Object protocol access to the data that is stored as O...

SMB
Provides access to data through Microsoft Windows-style shares.

GUI
Manages the graphical user interface and the REST-based API that are use...

Performance Monitoring
Collects performance and capacity data from various services.

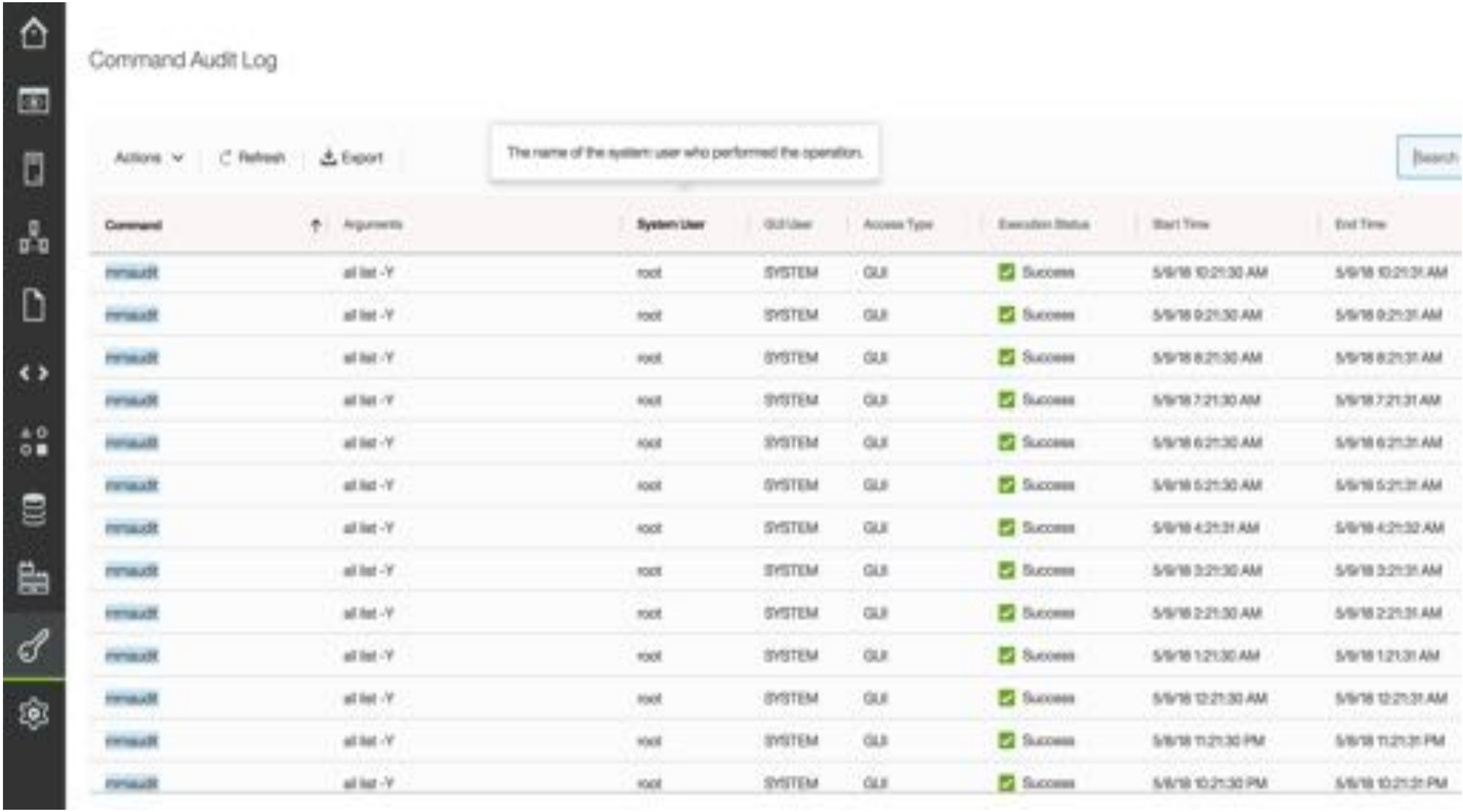
File Auditing
Audits data access to the files.

Message Queue
Provides a scalable infrastructure for the file auditing services.

Node Name	Status	Broker Service	Zookeeper Service	Quorum
boneknapper-vm1....	☑ Healthy	☑ Good	☑ Good	☑
boneknapper-vm2....	☑ Healthy	☑ Good	☑ Good	☑
boneknapper-vm3....	☑ Healthy	☑ Good	☑ Good	☑
boneknapper-vm5....	☑ Healthy		☑ Good	☑
boneknapper-vm6....	☑ Healthy		☑ Good	☑

GUI – Access → Command Audit Log Panel

- Every time a command related to FAL is ran (mmaudit, mmmsgqueue, mmcrnodeclass, etc.), it is logged in this panel.



The name of the system user who performed the operation.

Command	Arguments	System User	GUI User	Access Type	Execution Status	Start Time	End Time
mmaudit	all list -Y	root	SYSTEM	GUI	✓ Success	5/9/18 10:21:30 AM	5/9/18 10:21:31 AM
mmaudit	all list -Y	root	SYSTEM	GUI	✓ Success	5/9/18 9:21:30 AM	5/9/18 9:21:31 AM
mmaudit	all list -Y	root	SYSTEM	GUI	✓ Success	5/9/18 8:21:30 AM	5/9/18 8:21:31 AM
mmaudit	all list -Y	root	SYSTEM	GUI	✓ Success	5/9/18 7:21:30 AM	5/9/18 7:21:31 AM
mmaudit	all list -Y	root	SYSTEM	GUI	✓ Success	5/9/18 6:21:30 AM	5/9/18 6:21:31 AM
mmaudit	all list -Y	root	SYSTEM	GUI	✓ Success	5/9/18 5:21:30 AM	5/9/18 5:21:31 AM
mmaudit	all list -Y	root	SYSTEM	GUI	✓ Success	5/9/18 4:21:31 AM	5/9/18 4:21:32 AM
mmaudit	all list -Y	root	SYSTEM	GUI	✓ Success	5/9/18 3:21:30 AM	5/9/18 3:21:31 AM
mmaudit	all list -Y	root	SYSTEM	GUI	✓ Success	5/9/18 2:21:30 AM	5/9/18 2:21:31 AM
mmaudit	all list -Y	root	SYSTEM	GUI	✓ Success	5/9/18 1:21:30 AM	5/9/18 1:21:31 AM
mmaudit	all list -Y	root	SYSTEM	GUI	✓ Success	5/9/18 12:21:30 AM	5/9/18 12:21:31 AM
mmaudit	all list -Y	root	SYSTEM	GUI	✓ Success	5/8/18 11:21:30 PM	5/8/18 11:21:31 PM
mmaudit	all list -Y	root	SYSTEM	GUI	✓ Success	5/8/18 10:21:30 PM	5/8/18 10:21:31 PM

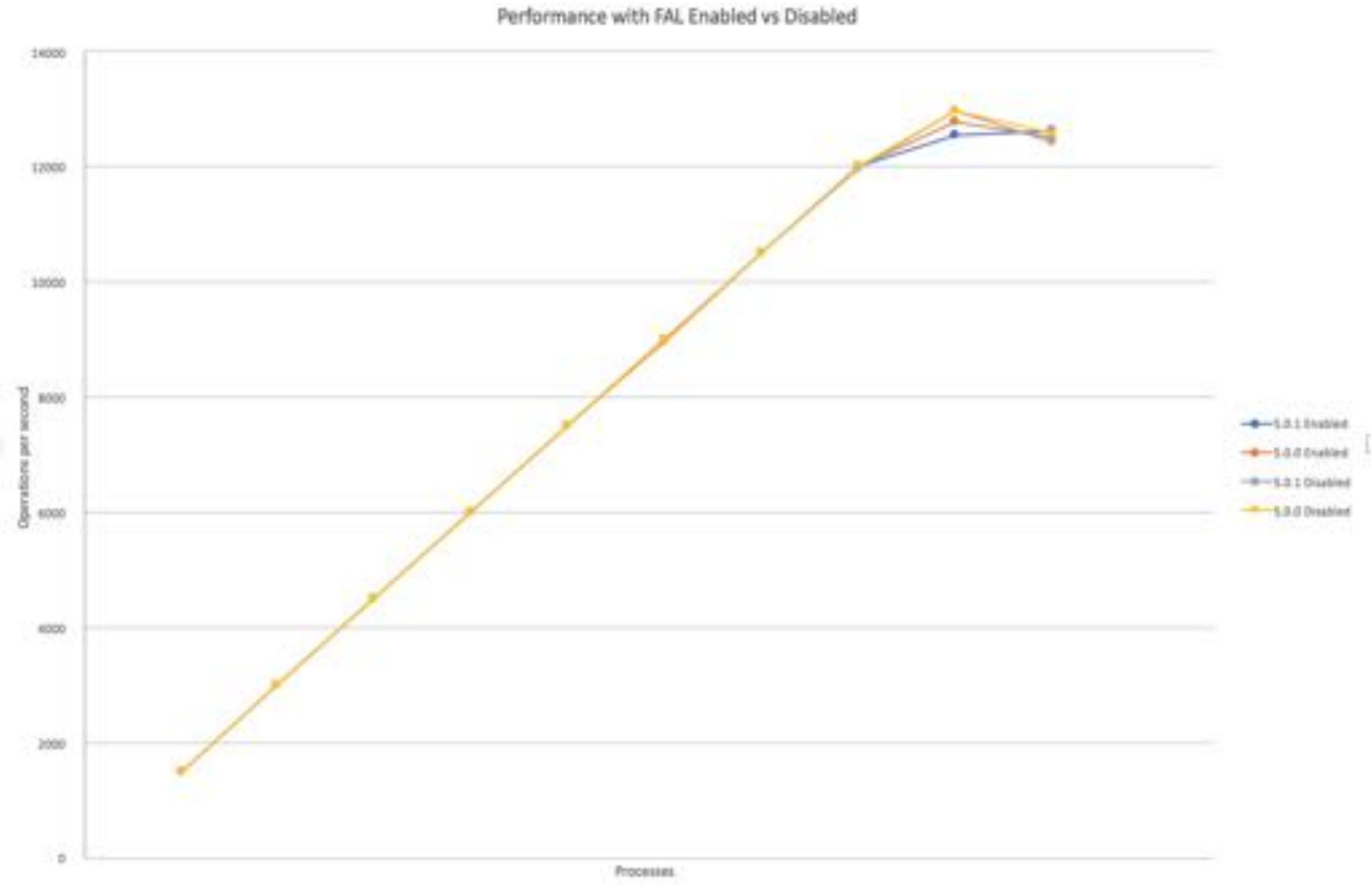
Performance

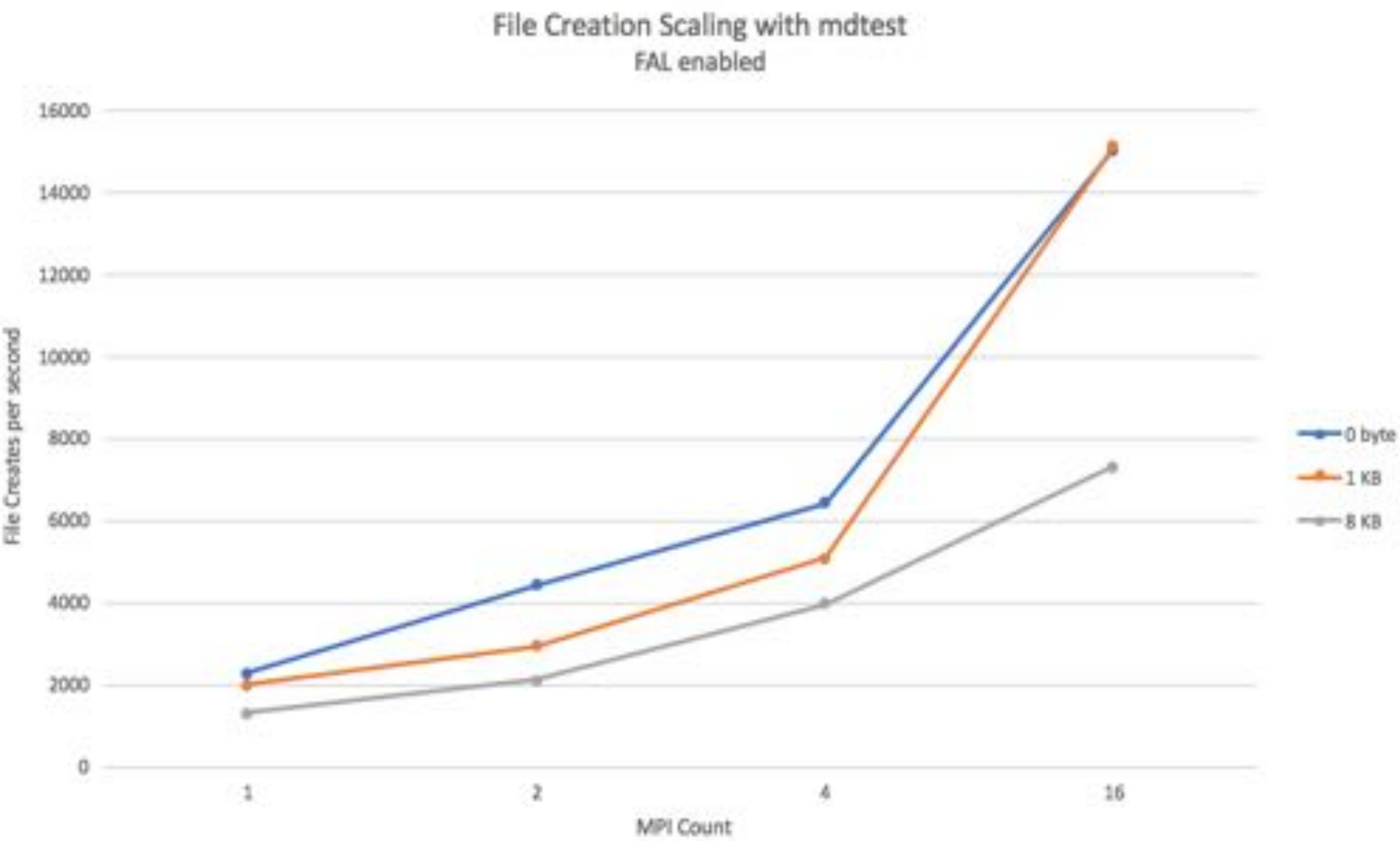
- Concerns
  - Does enabling FAL impact IO-performance on my filesystem?
  - How performant is FAL?
- Run perf tests to evaluate the above concerns
- Setup
  - Kafka cluster: 4 Broker nodes, 3 zookeeper nodes, 4 consumer nodes
  - Gpfs Cluster: 4 protocol nodes, 2 NSD server nodes (Linux 3.10.0-229.el7.x86_64)
  - Network: 10 GE
  - Storage: IBM DCS3700
- Tests run
  - Metadata intensive workload benchmark
    - With and without FAL
  - mdtest
    - With FAL enabled
    - File create with MPI-count

Disclaimer

- IBM’s statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM’s sole discretion. Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision. The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

- Performance is based on measurements and projections using standard benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors such as the I/O configuration, the storage configuration, and the workload characteristics. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.


Performance with FAL Enabled vs Disabled
14000
12000
10000
8000
6000
4000
2000
0
Operations per second
Processes
5.0.1 Enabled
5.0.0 Enabled
5.0.1 Disabled
5.0.0 Disabled



File Creation Scaling with mdtest

FAL enabled

Troubleshooting

- /var/adm/ras/mmmsgqueue.log
  - Contains information regarding the set up and configuration operations that take place that affect the message queue
  - Valid on any node containing a broker and/or zookeeper
- /var/adm/ras/mmaudit.log
  - Contains information regarding the set up and configuration operations that take place that affect the File Audit Logging
  - Valid on any node running the File Audit Logging command or location where the subcommand may be run (such as a consumer)
- /var/adm/ras/mmfs.log.latest
  - Daemon log, and contains entries when major message queue or File Audit Logging activity occurs.
- /var/log/messages (Redhat)
- /var/log/syslog (Ubuntu)
  - Contains messages from Kafka components as well as the producer and consumers that are running on a node.
- Logs collected via gpfs.snap

References

- https://www.ibm.com/support/knowledgecenter/en/STXKQY_5.0.0/com.ibm.spectrum.scale.v5r00.doc/bl1ins_quickrefadlg.htm

धन्यवाद

Hindi

Спасибо

Russian

شكراً

Arabic

Grazie

Italian

நன்றி

Tamil

谢 谢

Simplified Chinese

Thank You

English

감사합니다

Korean

謝 謝

Traditional Chinese

תודה רבה

Hebrew

Gracias

Spanish

Obrigado

Brazilian Portuguese

Danke

German

Merci

French

ขอบคุณ

Thai