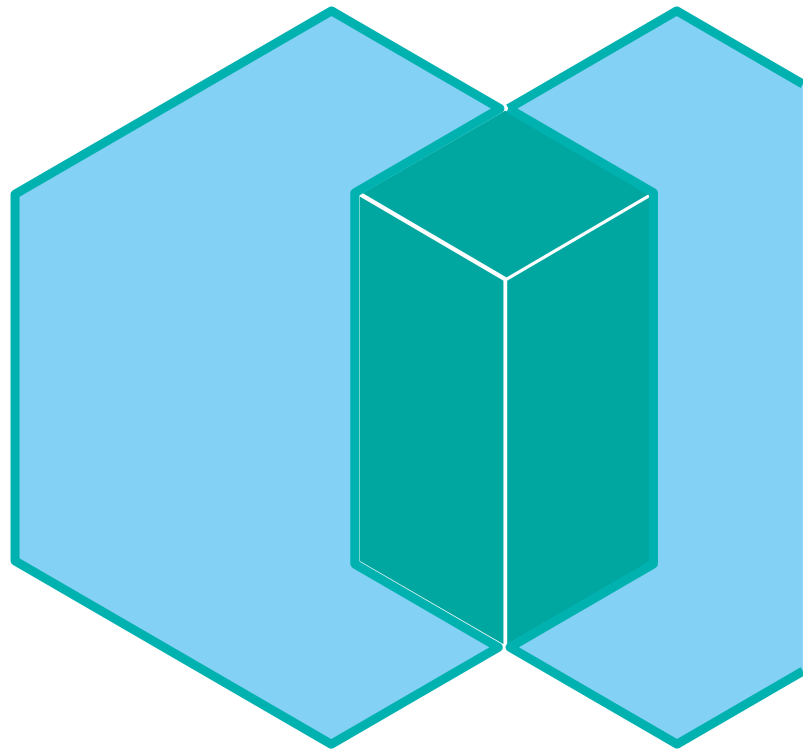# IBM Spectrum Scale
What's new in 5.0.0

Mathias Dietz
Spectrum Scale RAS Architect

IBM Research and Development
in Kelsterbach/Frankfurt, Germany
mdietz@de.ibm.com

Before we start ….

Which Spectrum Scale version is used ?

- Version 5.0
- Version 4.2.3
- Version 4.2.2
- Version < 4.2.2

How many use CES/Protocols ?

- SMB
- NFS
- Object

How many use the following features ?

- Spectrum Scale GUI
- REST API
- Install Toolkit for install or upgrade
- AFM
- Transparent Cloud Tiering (TCT)

Highlights in 5.0

Scaling and Performance enhancements

- Improved space efficiency and performance for small files
- Improved performance over RDMA transports
- Improved metadata performance (directory locking) and reduced CPU jitter
- Improved encryption performance (GSKit)

Currency & OS Support

- Ubuntu 16.04.03 (incl. Protocol support and zLinux)
- RHEL 7.4 & SLES 12 SP3
- Dropped support for RHEL6, SLES11, Ubuntu 14.04 and Debian

New License model

- Data Management License

File Audit Logging

Scaling and Performance enhancements

Improved space efficiency and performance for small files

- More than 32 subblocks per block
- Requires 5.0 Cluster release level
- Only for new filesystems !
- Default block size is now 4M

Blocksize	Subblocks	Subblock size
64K/128K/256K	32	2K/4K/8K
512K/1M/2M/4M	64/128/256/512	8KB
8M/16M/32M	512/1024/2048	16KB

RDMA enhancements – improved performance over RDMA transports

- Improve performance of VERBS RDMA send and recv and scale to 1000's of nodes
- Make RDMA primary code paths mutex free and improve cache line efficieny
- NUMA aware allocation of send and recv buffers

File Audit Logging

Logs filesystem operations to a retention-enabled fileset to track user access
- Events are monitored at a filesystem level
- Events that can be captured are: Open, Close, Destroy (Delete), Rename, Unlink, Remove Directory, Extended Attributed Change, ACL change
- Local events from each node are placed on a multi-node message queue (Kafka) for consumption, eliminating bottlenecks and providing redundancy
- Events are logged in a parseable, JSON formatted string

→ Join the Audit Logging Talk at 1:50 PM

JSON audit log

Consumer Group (Topic fs0)

| Consumer | Consumer | | Consumer |

Kafka Cluster

| Kafka broker | Kafka broker | | Kafka broker |

Zookeeper / Zookeeper / Zookeeper

LWE + librdkafka

| Kafka Producer | Kafka Producer | Kafka Producer | Kafka Producer | ... | Kafka Producer | Kafka Producer |
| GPFS Daemon | GPFS Daemon | GPFS Daemon | GPFS Daemon | | GPFS Daemon | GPFS Daemon |

Filesystem Updates

Usability

- Concurrent modification of preferred NSD servers
- Status/progress of a running filesystem check (mmfsck –status-report)
- Status of a running trace (mmtracectl --status)

Encryption and Compression

- Optimized compression and support for LZ4 (result in better read performance)
- LROC for encrypted files (lrocEnableStoringClearText)
- Improved encryption configuration (simplified IBM SKLM setup)

More enhancements

- Faster file system rebalancing (lenient vs strict round robin)
- InfiniBand addresses can be specified for RDMA transfers
- mmnetverify support for GPFS subnets and verify different MTU sizes

Protocol Support

SMB

- Update to Samba 4.6
- Detect and warn about version mismatch
- Option to use "unix primary group" in AD

NFS

- Update to Ganesha 2.5
- Modify NFS exports without impacting connected clients (Dynamic exports)

Object

- Improved upgrade support for Object (Removes intermediate levels in upgrade paths)

General

- Ubuntu support as a CES protocol node (16.04.x)

Monitoring & RAS Improvements

Health & Performance Monitoring Improvements

- Warn customers about filled /var/mmfs
- Detect and display when the GPFS daemon is unresponsive
- Kafka & audit consumer monitor
- Event script callout (undocumented feature)
- User defined thresholds – priorities and additional measurements
- Significant reduction of CPU Jitter and GSKit initialization

Callhome aka Proactive Service

- Simplified setup and configuration (through Install toolkit and GUI)
- Upload of diagnostic data directly from the system (CLI/GUI)
- Added Callhome support for zLinux and Ubuntu
- Many improvements under the hood to provide the basis for proactive service

→ Join the System Health Talk tomorrow 9AM for more details

GUI & REST API Updates

Filesystem creation wizard for non-ESS environments

Enhanced management of AFM relationships

- Ability to access remote GUIs and view AFM status
- Enhanced monitoring: View events from multiple cache sites on a home cluster

Health & Performance Monitoring Improvements

- Performance and capacity threshold management
- Network monitoring for both IP and RDMA transports
- Additional directed maintenance procedures to assist with failure recovery

REST API has been expanded to allow additional tasks

- Query Performance data and manage thresholds
- Add/remove nodes from a cluster and manage node classes
- The ability to cancel jobs

Transparent Cloud Tiering support

Install Toolkit Enhancements

- Improved integration between the Installation Toolkit and ESS
  - Installation toolkit will automatically detect ESS nodes for easier deployment
- Enhanced pre-checking to find common network issues prior to deployment
- Installation and configuration of file audit logging
- Simplified callhome deployment and configuration
- Support for Ubuntu 16

→ Join the Install & Upgrade Talk tomorrow 10 AM

AFM Updates

- File compression is now supported for files residing in AFM and AFM DR filesets
- Load balancing improvements for AFM gateways
  - Updated hash algorithm (afmHashVersion) tunable to modify how filesets are distributed across gateway nodes
- New limit on the number of parallel recoveries being run on gateway nodes
  - Can be tuned to reduce the resource requirements on a gateway node
- ILM for snapshots is now supported for AFM and AFM DR filesets
- GUI improvements to assist with AFM management

Transparent Cloud Tiering

- Remote mounted filesystem support
  - Clients can access tiered files on a remotely mounted filesystem
- Ability to tier different filesets to different cloud containers
- Enhanced support for multiple cloud accounts and containers
- Usability improvements, updated CLI and enhanced GUI support
  - CLI has been enhanced for ease of deployment with multiple cloud accounts

→ Join the Transparent Cloud Tiering Talk tomorrow 1 PM

Big Data and Analytics

- Certification with HortonWorks Data Platorm 2.6
  - Certified on both Power8 and x86 platforms
  - Certified with Ambari 2.5 for rapid deployment
- Support short circuit write for better performance
- Support 2 or more file systems in one HDFS Transparency cluster
- Support local read/write bytes for Spectrum Scale nodes with internal disks
- Integration with Spectrum Conductor for Spark

→ Join the Big Data & Analytics Talk tomorrow 11AM

5.0.1 just released …..

Filesystem Enhancements

Support for data checksums in non-ESS environments with unreliable network

Improvement in the area of Systemd integration, adding disks to full file systems

Health Monitoring Updates

Simplified performance monitoring configuration

Several improvements around CES, network monitoring and Call Home

NFS Enhancements

Automatic restart of nfs-ganesha for better error recovery

Automatic backtrace generation for better problem determination

GUI Updates

New Services Panel to manage and monitor services like SMB, NFS, GPFS, Zimon

Display performance data from remote clusters

...and more….

QUESTIONS ?

THANK YOU

LEARN MORE

Client Enablement Material

New videos, blogs and presentations:

➢ https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/General%20Parallel%20File%20System%20(GPFS)/page/White%20Papers%20%26%20Media

Watch for new blogs at:

➢ https://developer.ibm.com/storage/blog/

Turn-key Spectrum Scale VM available for download
- Try the latest Spectrum Scale enhancements
- Full functionality on laptop, desktop or server
- Incorporate external storage

Spectrum Scale Blueprints for Genomic Medicine Workloads
- Enablement and Redpapers available for these workloads