User Group
2017

# IBM Spectrum Scale 4.2.3

## *Security Overview*

Felipe Knop
STSM, File System Core

Sandeep Patil
STSM, IBM Master Inventor

# Please note

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 IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user’s job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.

# Acknowledgement – Spectrum Scale Development Team

Piyush Chaudhary, Michael Daubman, Cheng Ding, Dileep Dixith, Sasikanth Eda, Scott Fadden, Shuo Feng, Deepak Ghuge, Nils Haustein, Kaustubh I Katruwar, Ingo Meents, Kumaran Rajaram, John Olson, Aaron Palazzolo, Christof Schmitt, Gil Sharon, Gaurang Tapase, Carl Zetie, Yong Zheng, Wyane Sawdon, Wei Gong, Bill Owen, Sasikanth Eda.

# The world of data storage has indeed changed…

**IT is managing large amount of data**

- Doubling capacities every 6 months to 2 years depending on the industry

**IT is dealing with new applications/workloads…many of them didn’t exist 5 years ago**

- Think Hadoop, Spark, No-SQL or In-Memory databases

**IT is dealing with new types of deployments**

- Software Defined/Software Based
- Cloud
- Hyper-converged

**IT is dealing with new ways of storing data – it’s not file, it’s not block**

- Object Storage is becoming main stream for storing massive archives

**Storage is no more just a SAN or NAS box**

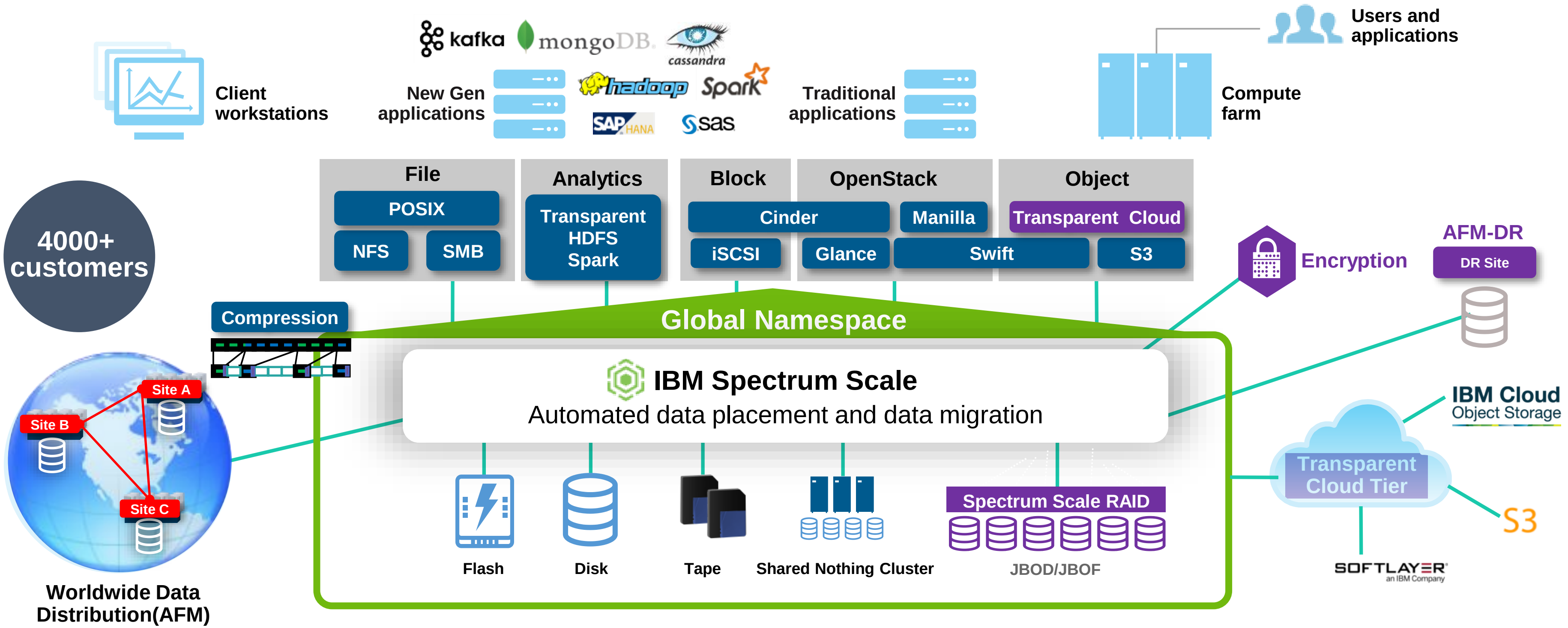
# ... But need for “Ubiquitous Security of Data” prevails !

- **Industries understand that “Data is the new Oil”, hence**
  - Data Is Precious: Need to Secure it Accordingly

- **IT needs Data Repositories and Data serving platforms that are:**
  - Secure by Design
  - Caters to all aspects of Security
  - Enables Compliances

- **Spectrum Scale: secure-by-design principles for all new and existing major features**

# Spectrum Scale: Breath of new features / Catering to newer workloads

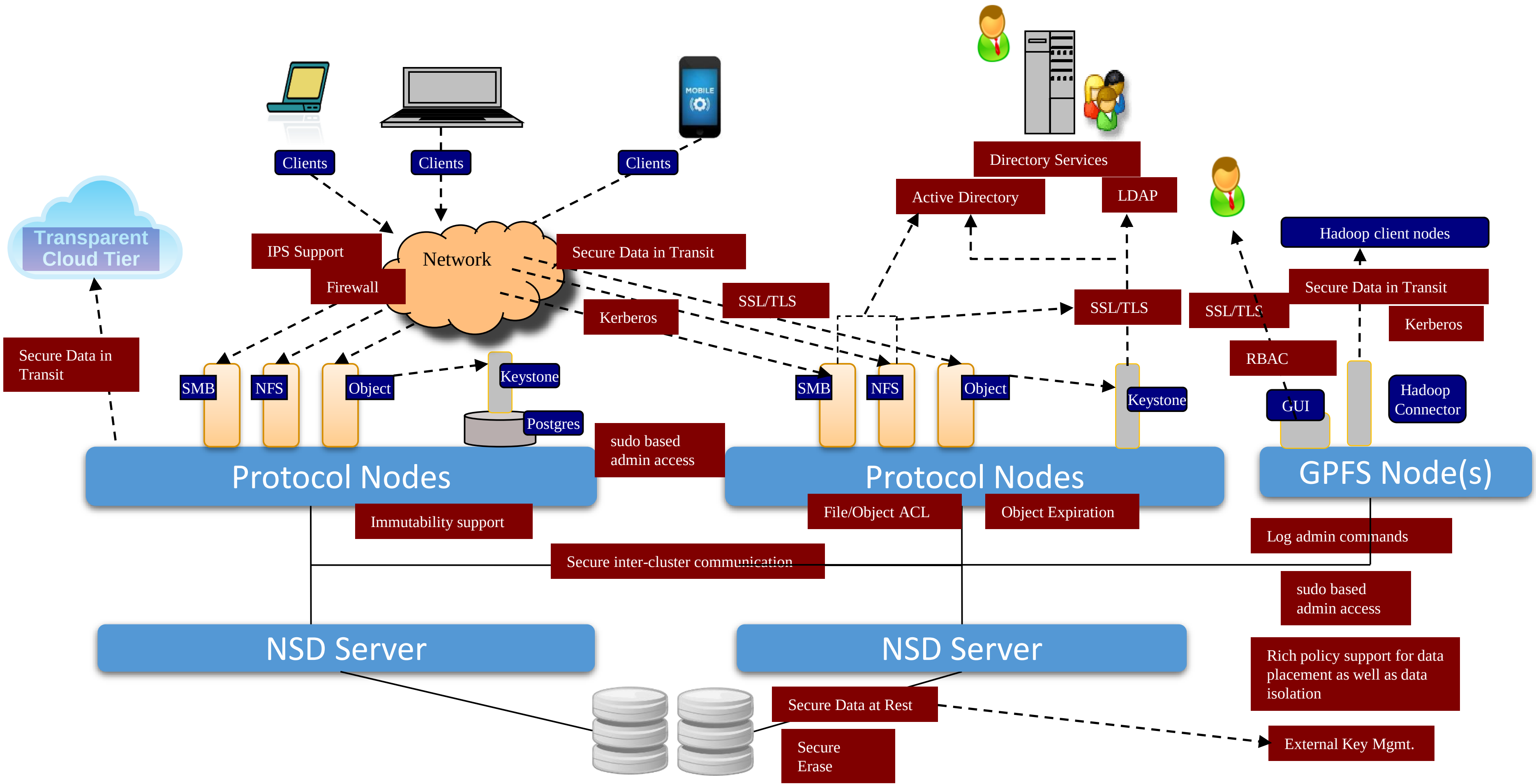


Consolidate all your unstructured data storage on Spectrum Scale with unlimited and painless scaling of capacity and performance – Ensuring data security

# Security Requirement Vs Spectrum Scale Security Capabilities

| Key Security requirement | Spectrum Scale Capability |
|---|---|
| Secure Data at Rest | ✓ |
| Secure Data in Transit | ✓ |
| Authentication | ✓ |
| Authorization | ✓ |
| Secure Administration | ✓ |
| Immutability | ✓ |
| Firewall | ✓ |
| Hadoop Security | ✓ |
| Cloud Tiering Security | ✓ |
| Audit Logging | ✓ Basic Covered (more coming) |
| Anti Virus | ✓ Basic Covered (Coming) |

# Spectrum Scale High Level Security Outlook



# Spectrum Scale : Secure Data at Rest

**Encryption of Data at Rest**

- Files are encrypted before they are stored on disk
- Keys are never written to disk
- No “digital shredding”: secure delete is a cryptographic operator

**Secure Deletion**

- Ability to destroy files with no data remanence
- No “digital shredding” secure delete is cryptographic operator

**NIST & FIPS**

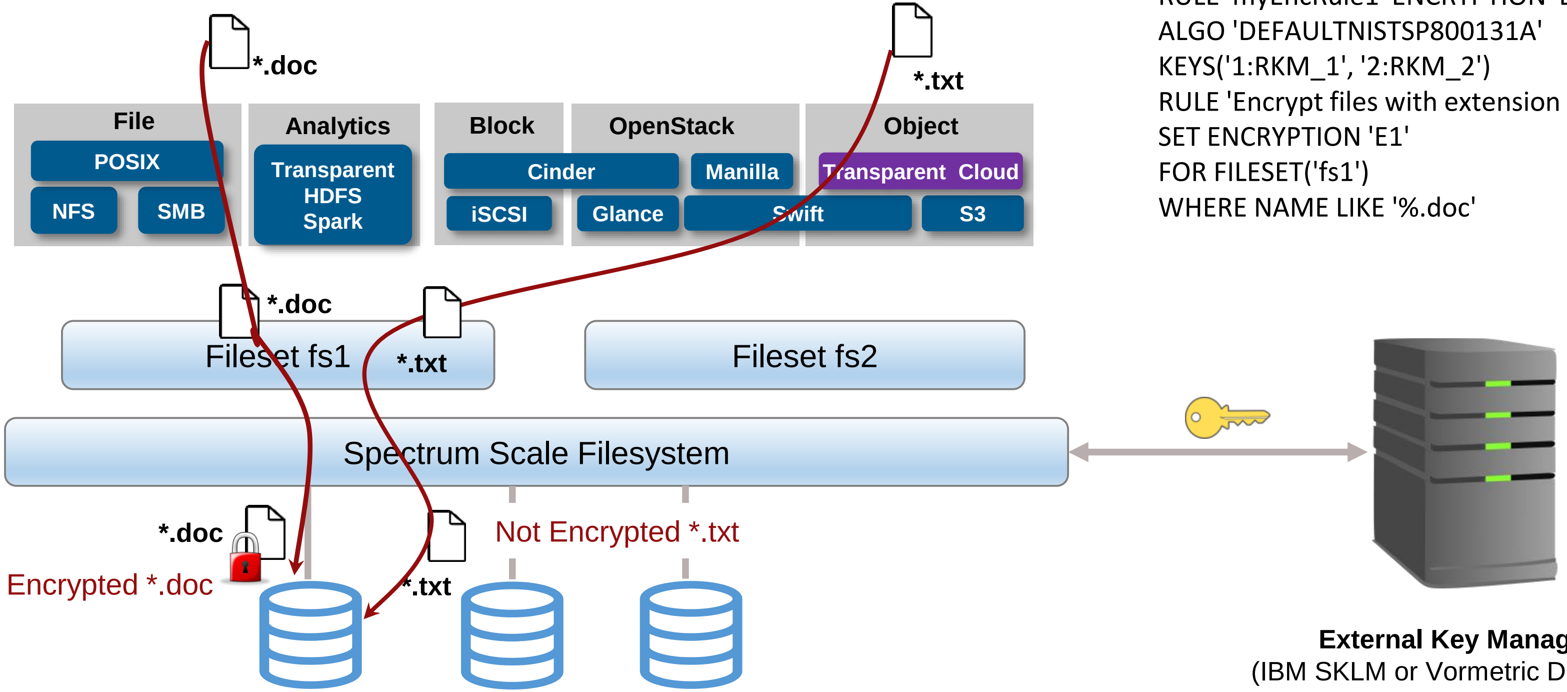
- Encryption algorithms used for file encryption are all compliant with NIST Special Publication 800-131A
- Allows cluster to be configured in FIPS mode

**Encryption policy rules:**

- which files are to be encrypted,
- with which algorithm,
- using which MEKs

***Example encryption policy rules***

```
RULE 'myEncRule1' ENCRYPTION 'E1' IS
ALGO 'DEFAULTNISTSP800131A'
KEYS('1:RKM_1', '2:RKM_2')
RULE 'Encrypt files with extension doc with rule E1'
SET ENCRYPTION 'E1'
FOR FILESET('fs1')
WHERE NAME LIKE '%.doc'
```

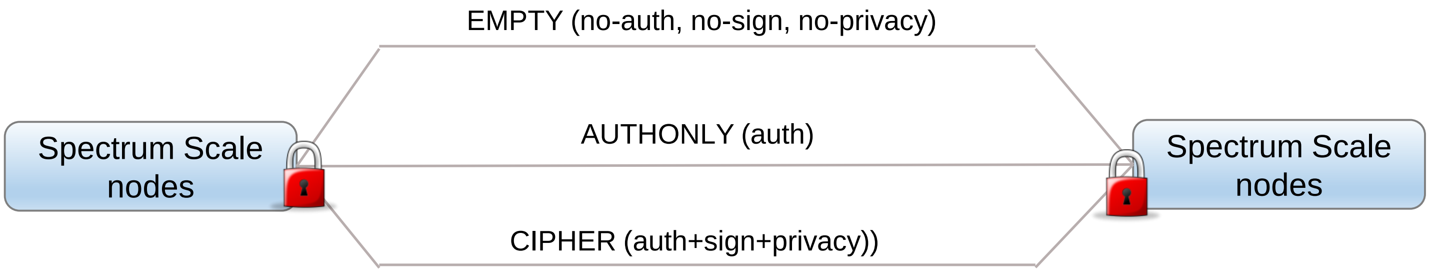


**External Key Manager Server**
(IBM SKLM or Vormetric DSM Key Server)

# Spectrum Scale : Secure Data at Motion

- Data in transit, also referred as Data in Motion or Data in Flight, is data that is being accessed over a network (internal or external) and can therefore be intercepted by malicious users on the network
- Based on your business needs or on the sensitivity of your data that is being accessed over the network, one needs to protect it by encryption over the wire

**Spectrum Scale Cluster Communication**


EMPTY (no-auth, no-sign, no-privacy)
AUTHONLY (auth)
CIPHER (auth+sign+privacy))
Spectrum Scale nodes
Spectrum Scale nodes

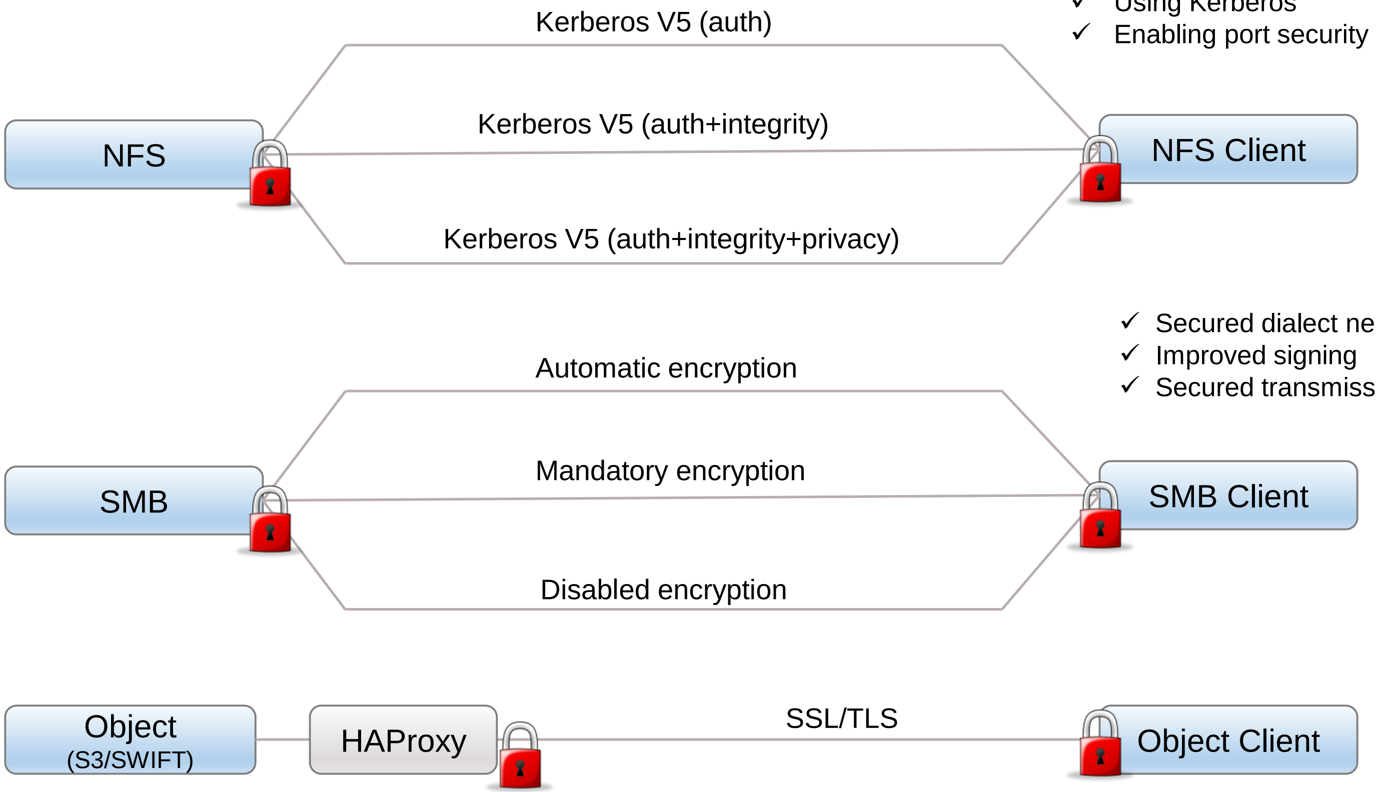
- When users are accessing a file system from another cluster, the cluster that owns a file system can designate a different security level for each connecting cluster.

# Spectrum Scale : Secure Data at Motion

## Secure access of Spectrum Scale File Interfaces

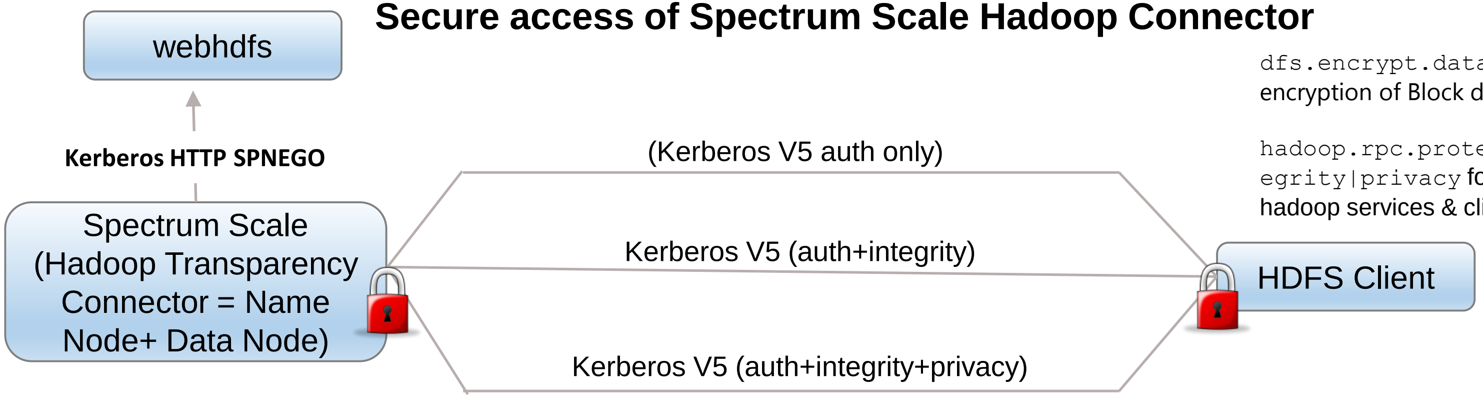
- ✓ Enabling squashing
- ✓ Using Kerberos
- ✓ Enabling port security

- ✓ Secured dialect negotiation
- ✓ Improved signing
- ✓ Secured transmission



# Spectrum Scale : Secure Data at Motion

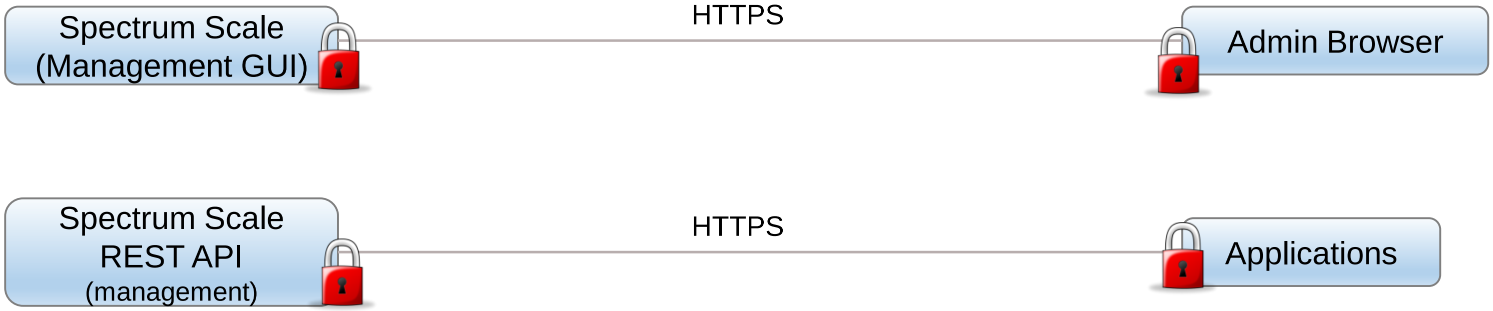
## Secure access of Spectrum Scale Hadoop Connector


webhdfs
Kerberos HTTP SPNEGO
Spectrum Scale (Hadoop Transparency Connector = Name Node+ Data Node)
(Kerberos V5 auth only)
Kerberos V5 (auth+integrity)
Kerberos V5 (auth+integrity+privacy)
HDFS Client

`dfs.encrypt.data.transfer` = true for data encryption of Block data transfer

`hadoop.rpc.protection=authentication|integrity|privacy` for secure data transfer between hadoop services & clients

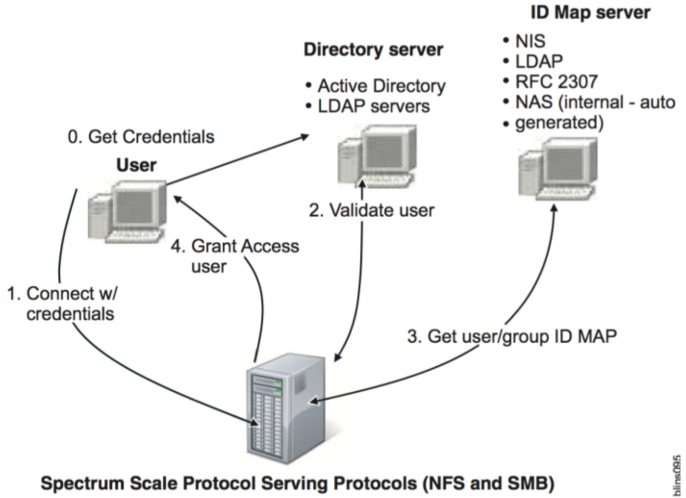
## Secure access of Spectrum Scale Management


Spectrum Scale (Management GUI)
HTTPS
Admin Browser
Spectrum Scale REST API (management)
HTTPS
Applications

# Spectrum Scale : Protocol Authentication

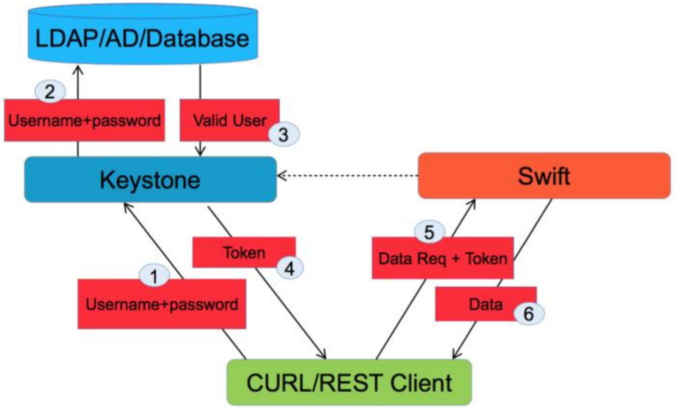
## Protocol Authentication (NFS/SMB)

- Directory Servers supported:
  - RFC2307 schema-compliant Lightweight Directory Access Protocol (LDAP) server
  - Microsoft Active Directory (AD) server
  - Network Information Service (NIS) server
- Kerberos authentication is supported by the AD and LDAP authentication schemes
- Securing NFS exports by using netgroup definitions that are stored on authentication servers is supported by LDAP and NIS-based authentication schemes



## Object Authentication

- IBM Spectrum Scale supports configuring Keystone with the following identity back ends:
  - Microsoft AD server
  - LDAP server
  - Postgres database (local)
- The Keystone service can be configured with https
- One can configure the communication between the Keystone service and identity back end (Microsoft AD/LDAP) to be over TLS



# Spectrum Scale : Authorization

- Spectrum Scale Client
  - Supports POSIX ACL
  - Supports NFS V4 ACL

- Authorizing NFS and SMB users
  - NFSv4 ACLs
  - ACL inheritance
  - SMB ACLs
  - Mapping between NFSv4 and SMB ACLs

- Authorizing Object (OpenStack Swift and S3) users
  - **Supports OpenStack Swift and S3 protocols** for object data access
  - Uses the **Keystone service for identity management**, and access by the object users to the object storage projects is controlled by these items:
    - **User roles** - Based on the roles that are defined for the user, object users can be administrative users or non-administrative users
    - **Container ACLs**
  - S3 ACLs are supported via the use of Swift3 Middleware for OpenStack Swift, which enables allowing access to IBM Spectrum Scale by using the Amazon Simple Storage Service (S3) API

# Spectrum Scale : Secure Administration

- Administration of Spectrum Scale requires Remote Shell and Remote Copy
  - SSH and scp are default and recommended

- Limited Admin Nodes
  - The **adminMode** configuration attribute specifies whether all nodes in the cluster can be used for issuing IBM Spectrum Scale administration commands or just a subset of the nodes
    - **allToAll** - indicates that all nodes in the cluster can be used for running IBM Spectrum Scale administration commands
    - **central** - indicates that only a subset of the nodes can be used for running IBM Spectrum Scale commands
  - The **major advantage of the central mode of administration is that the number of nodes that must have root level access to the rest of the nodes is limited**, and can be as low as one

# Spectrum Scale : Secure Administration

- Running IBM Spectrum Scale without remote root login

  - In several environments, corporate IT policies **require that the ssh PermitRootLogin parameter be disabled to prevent remote login as root**

  - By **using sudo and the IBM Spectrum Scale sudo wrappers**, IBM Spectrum Scale administration can be **performed securely by using a non-root ID**.

  - The IBM Spectrum Scale sudo wrappers enable :
    - IBM Spectrum Scale administrative operations to be securely performed by using a non-root user. One needs to rely on ssh wrappers to start remote commands with a non-root user ID.  Sudo is then used on the remote node to run the necessary commands

# Spectrum Scale : Secure Administration

- Secure administration by using the GUI
- Role-based access control for administration by using the GUI
  - The IBM Spectrum Scale GUI **supports different administrative roles.** These predefined roles are associated with user groups that define the working scope within the GUI
  - This feature **enables division of responsibilities** among multiple administrators based on roles
- The GUI allows users to be local users or even ones from central directory services such as **Microsoft Active Directory (AD) or LDAP**
- Support for **sudo wrappers**
- IBM Spectrum Scale supports **secure access to the GUI by using https** with the support for self-signed or trusted certificate authority (CA)

- REST APIs provide secure administration
  - by mandating **authenticated requests,**
  - supporting role-based access control (**RBAC**),
  - ensuring secure administration over the wire by **leveraging SSL/TLS**,
  - supporting **IBM Spectrum Scale sudo wrappers** for deployments for secure administration with non-root remote credentials

# Spectrum Scale : Immutability

- **Tamper-proof data is ensured by the immutability feature**
- Spectrum Scale can be **used for archiving use cases where regulatory requirements demand that the implementation prevent modification and deletion of files.**
- IBM Spectrum Scale immutability is based on immutable filesets
- Immutable filesets can be exported by using the Network File System (NFS) protocol and Server Message Block (SMB) protocol.
- In an immutable fileset, **files can be immutable or append-only** for a configurable retention time by using standard file system commands.
- IBM Spectrum Scale supports one of the following immutability ("IAM") modes for an immutable fileset:
  - **None**: No immutability mode is set (default). The fileset is a regular fileset
  - **Advisory**: Allows setting retention times and immutability, but files can be deleted with the proper file permission
  - **Noncompliant**: Advisory mode plus files cannot be deleted if retention time has not yet expired. However, retention times can be reset, and files can be deleted but not changed
  - **Compliant**: Noncompliant mode plus retention time cannot be reset. When retention time expires, files can be deleted but not changed
- The immutability function in IBM Spectrum Scale Version 4.2 **was assessed for compliance in accordance to the US SEC17a-4f, German, and Swiss laws** and regulations by a recognized auditor

# Spectrum Scale : Hadoop Security

## Spectrum Scale Benefits for Hadoop

- In an Hadoop deployment, Spectrum Scale replaces HDFS with following benefits:
  - In-place Analytics – No data movement
  - Enterprise grade data security
  - Ability to scale compute and storage separately
  - Unified file and object access to data
  - Enterprise data management with ILM capabilities
  - Enterprise Backup capabilities
  - Scalability
  - Federation

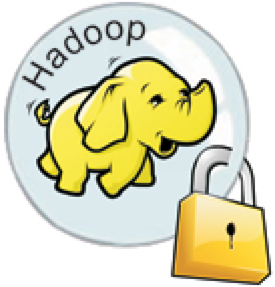
## Hadoop Ecosystem Security via Hadoop Distro

Hadoop Distro provides the following security for Hadoop based deployments:

- Authentication : Kerberos
- Authorization: POSIX based for Data, Service-level authorization support
- Centralized administration of Security Policies: Sentry & Ranger provide a central location for managing all security-related tasks (role-based or attribute-based access control, fine-grained authorization, authentication, auditing, and data protection)
- Secure REST Access: Apache Knox Gateway provides a single access point for all REST interactions with the Hadoop cluster. It integrates with popular enterprise identity management services, and provides a single point of control, management, monitoring, and auditing of REST access to the Hadoop cluster.

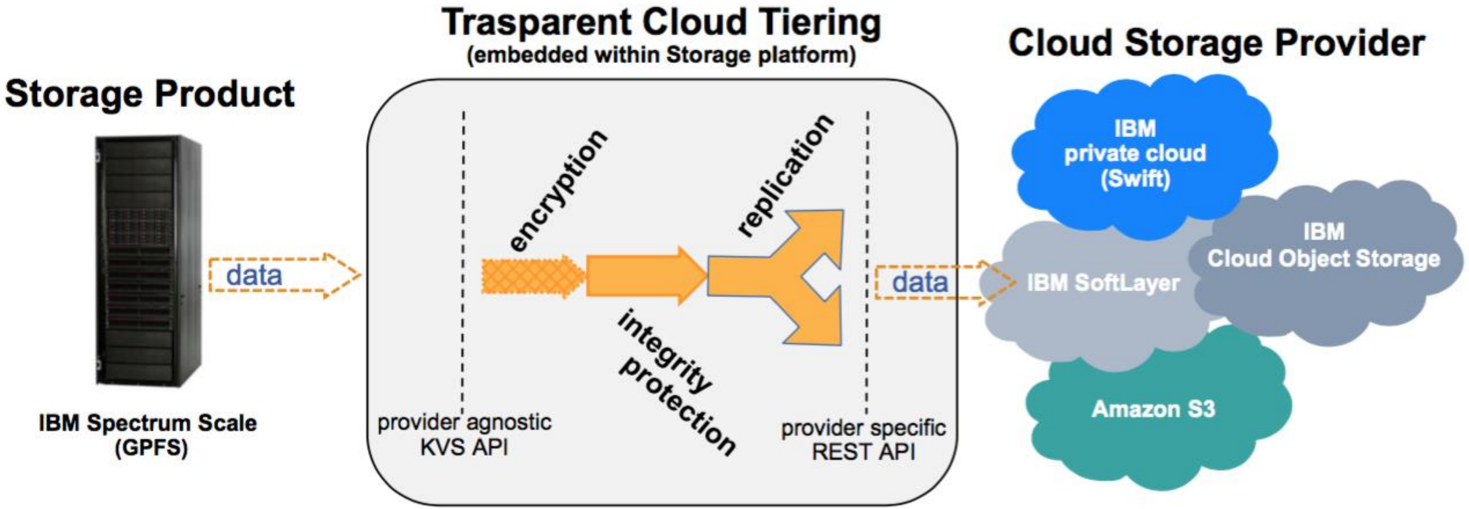
## Spectrum Scale Security for Hadoop Deployment

- Spectrum Scale provides following security in an Hadoop Deployment
  - Secure Data at Rest via Filesystem encryption (FIPS compliant)
  - Enterprise Key lifecycle management
  - Secure data in transit with Spectrum Scale Hadoop Transparency connector (Data nodes + Name Nodes)
  - Secure data in transit across all other access interface (NFS/SMB/Object)
  - Secure Delete of Data
  - Immutability  of classified data for compliance
  - Secure Backup of Data



# Spectrum Scale : Cloud Tiering Security

- Data is encrypted (AES 256) before it is pushed to Cloud Object Storage (on-premises or off-premises)
- Supports two types of Encryption Key Management Providers to store the encryption key
  - IBM Security Key Lifecycle Manager and Java Key Store
- TLS protocol is used when communicating with the cloud.



# Spectrum Scale : Audit Logging

- Auditing file system activities is an important security aspect in a number of deployments

- **File Access Audit logging with Varonis DatAdvantage**
  - IBM Spectrum Scale is integrated with Varonis DatAdvantage to log file activity within IBM Spectrum Scale protocol shares.
    - Major file operations (file creation, deletion, and directory creation and deletion) can be detected in Ganesha, unified file and object, and SMB shares.
    - Varonis agent software is installed on protocol nodes that interface with one or more Probes, running on nodes that are external to the IBM Spectrum Scale cluster. The DatAdvantage software and console run on an external Windows server.

## Audit logging for cluster configuration changes

- To help with problem determination and in auditing changes to the cluster configuration, audit messages can be sent to syslog or to the GPFS log whenever an IBM Spectrum Scale command changes the configuration of the cluster.

# IBM Spectrum Scale : Security Redpaper & Blogs

- Redpaper – Released Jan 2017
  http://www.redbooks.ibm.com/abstracts/redp5426.html?Open
- Security Blogs by Developers:
  https://developer.ibm.com/storage
- Enhanced Knowledge Center with all details.
- Assessment report Spectrum Scale 4.2 immutability:

http://www.kpmg.de/bescheinigungen/RequestReport.aspx?41742

- Spectrum Scale Immutability whitepaper:

http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102620

Redbooks
ibm.com/redbooks

IBM Spectrum Scale Security

Felipe Knop
Sandeep R. Patil
Larry Coyne

Cloud

Storage

In partnership with
IBM Academy of Technology

Thank You !

User Group
2017

IBM