

# Spectrum Scale Support for Container

---

Harald Seipp

STSM, CoE for Cloud Storage

IBM Systems EMEA Storage Competence Center

Credits: Ted Hoover, Sanjay Sudam

# EMEA Storage Competence Center

The EMEA Storage Competence Center (ESCC) provides **Pre- to Post-Sales Storage services to assist you across IBM's complete Storage portfolio (Hardware / Software / Solutions)** via:

## **Advanced Technical Skills** (ATS)

Skill Transfer, New Product Intro., Solution Enablement, Architectural Guidance

## **Lab Services** (LS)

Skill Enablement, Hands-On-Training, Implementation Assistance, Data Migration & Performance Tuning

## **Development & Product Test Systems Lab**

Proof of Concept (PoC), Hands-on Product Training ...

## **Product Field Engineering** (PFE)

Last level defect support

## **Client Care**

Pro-active Customer Care Management



### **ESCC@IBM Intranet**

[http://escc.mainz.de.ibm.com/e\\_dir/e4uweb.nsf/site.xsp](http://escc.mainz.de.ibm.com/e_dir/e4uweb.nsf/site.xsp)



### **ESCC@IBM Partnerworld**

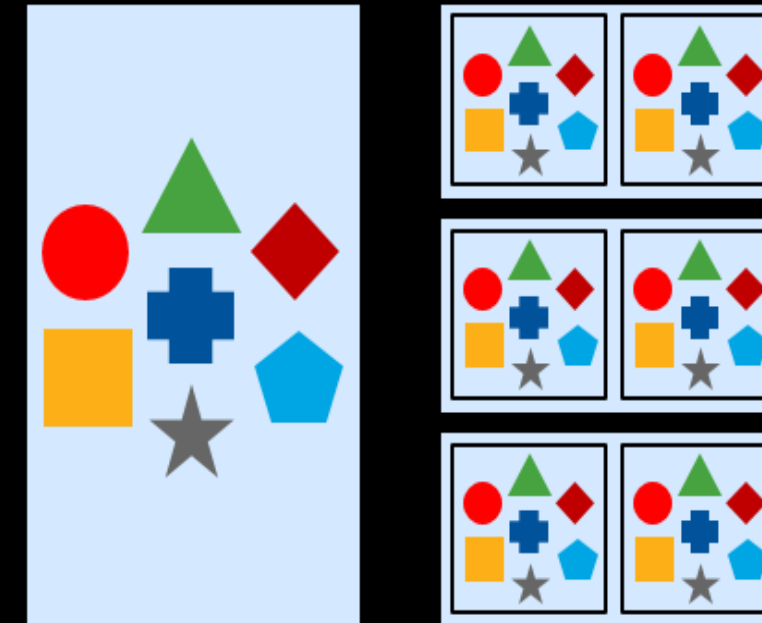
[https://www.ibm.com/partnerworld/page/pw\\_com\\_bp\\_ecampus](https://www.ibm.com/partnerworld/page/pw_com_bp_ecampus)

# Containers, Containers, Containers

- HPC and Scientific Computing
  - Portable and reproducible science
  - One-Click Laptop to Supercomputer
- On-premise and Public Clouds
  - Scheduling and Auto-Scaling
  - Improved resource utilization
  - Isolation and Multi-Tenancy
- Development, DevOps and continuous integration
  - Re-use of applications and services
  - Simplify and accelerate application deployment



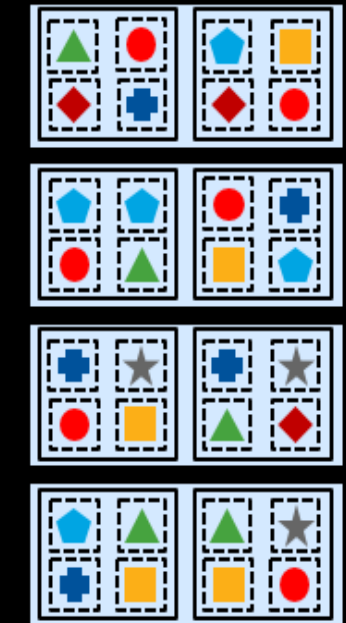
Traditional Monolithic Application



**Scales** by size... or monolithic replication.  
**Changes** monolithically.



Microservice Architecture



**Scales** by microservice replication.  
**Changes** by microservices.

# Open standards for cloud-native applications



Executable package  
of software that  
includes everything  
needed to run it  
locally or in the cloud

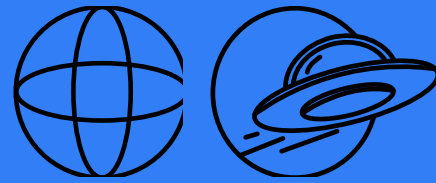
**Containers**



Automate  
deployment, scaling,  
and management of  
containerized  
applications

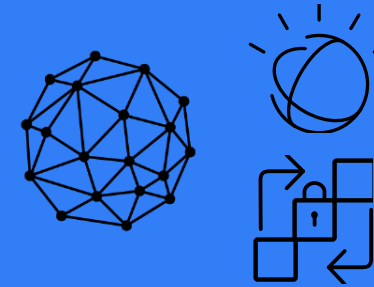
**Orchestration**

# IBM Cloud Private Architecture



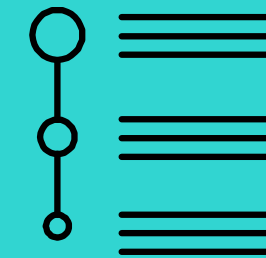
## Enterprise Content Catalog

Open Source and IBM Middleware, DevOps,  
Data, Analytics, and AI Software



## Core Operational Services

Logging, Monitoring, Metering, Security, Alerting



## Kubernetes Container Orchestration Platform

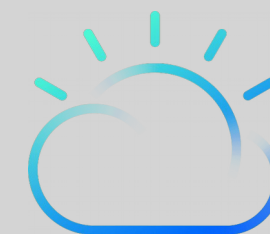


Choice of  
infrastructure



IBM Z

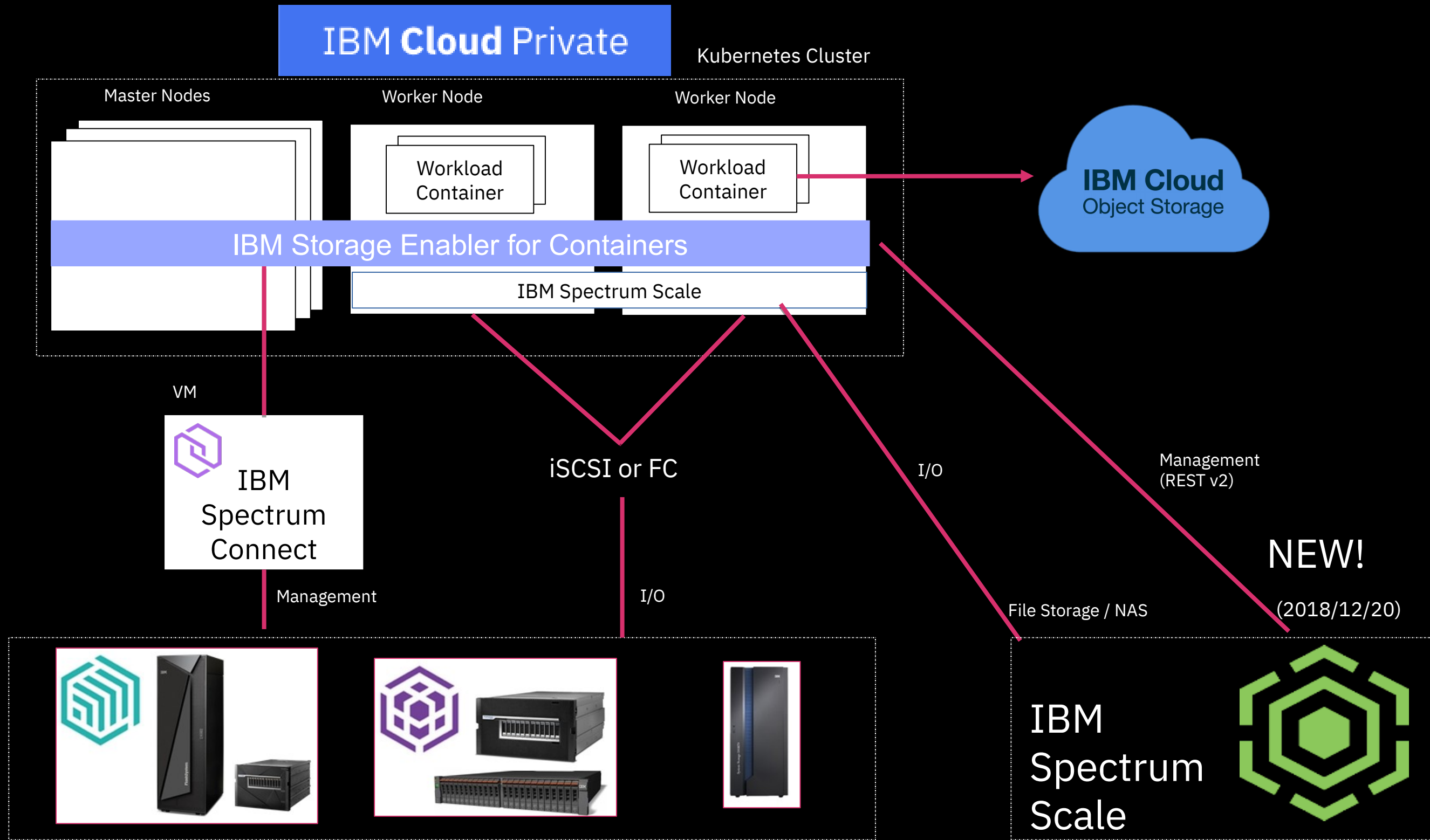
vmware®



- Industry's richest catalog of IBM and open source software containers
- Built-in management and security for simplified operations
- Runtimes and Orchestration capabilities for automated deployment with enterprise grade configurations
- Tools and expertise to help clients modernize with confidence
- Your choice of infrastructure



# Kubernetes Persistent Storage Architecture

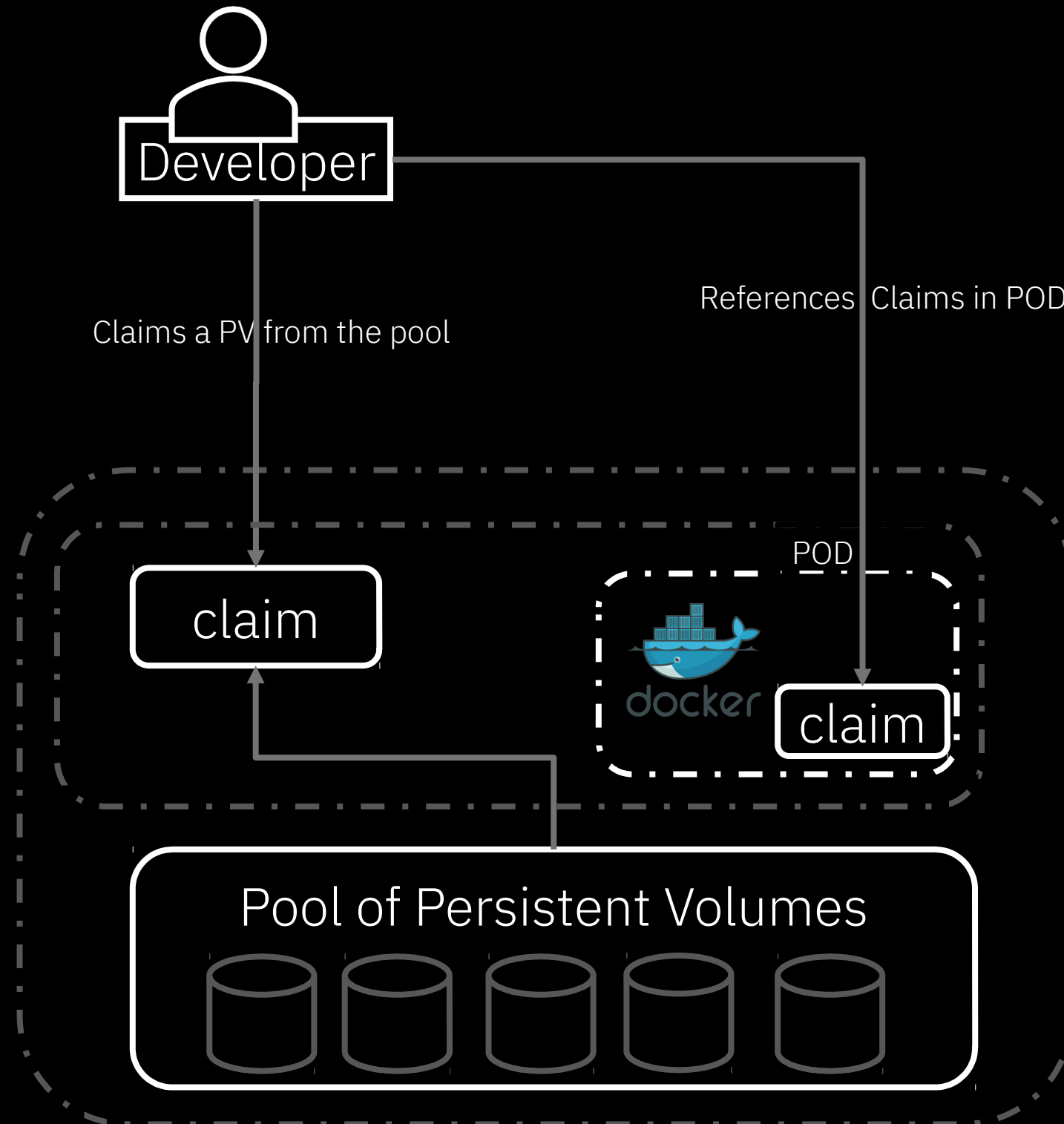


# Kubernetes Persistent Volumes

A Persistent Volume (PV) is a unit of storage that has been provisioned by an administrator or **dynamically** provisioned via a storage driver/plugin.

For Spectrum Scale PVs will be a fileset and the fileset's file system must have already been created  
A PersistentVolumeClaim (PVC) is a request for storage by a user.

PVCs consume PV resources.



## What it is not

**NO** containerized  
Spectrum Scale

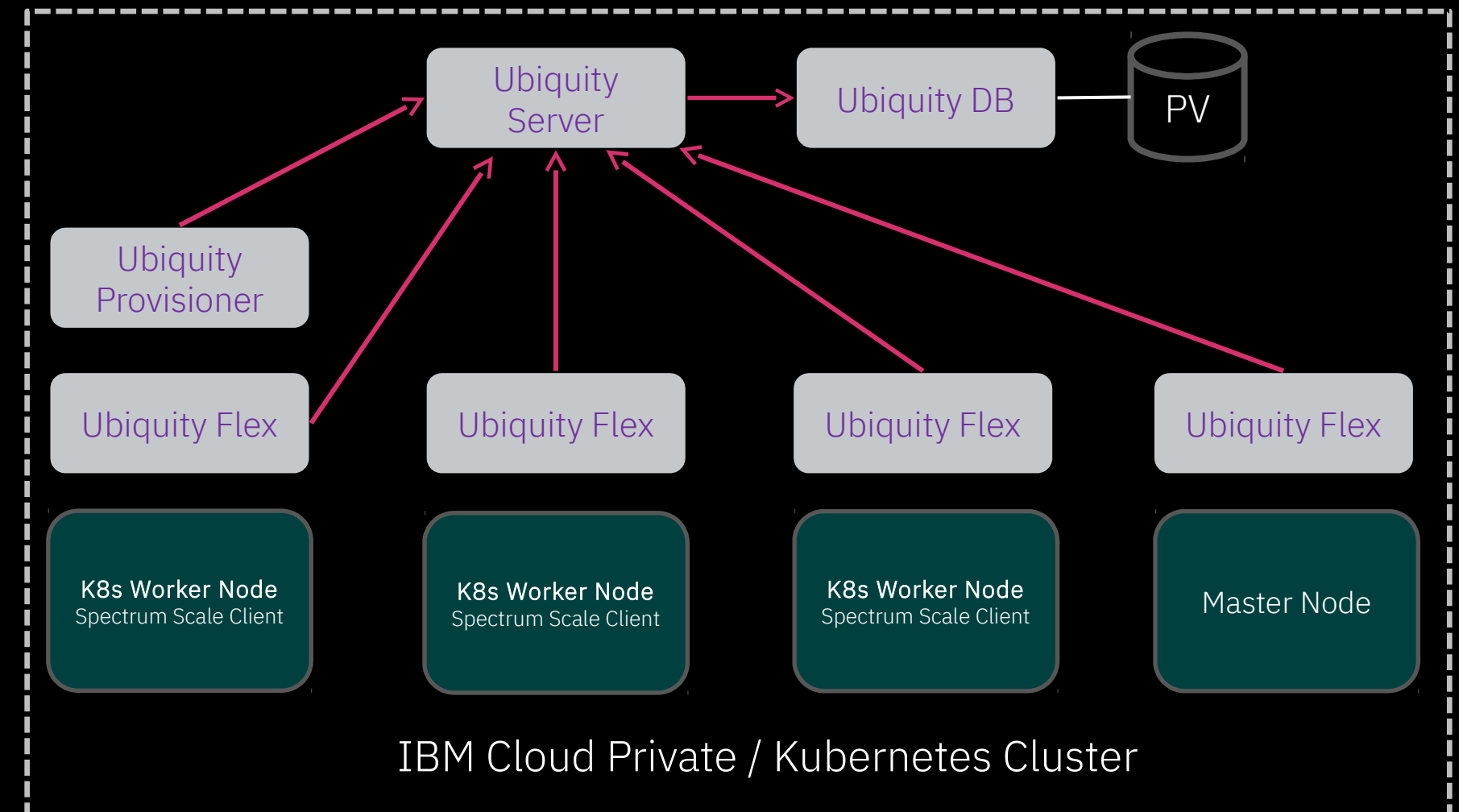
**NO** container deployment  
for the Spectrum Scale  
installation

**NO** storage for container  
images



# IBM Storage Enabler for Containers - Components

- Ubiquity Server
  - Upon request from Provisioner/Flex, Create/Delete/List fileset on Spectrum Scale and update the same on ubiquity-db
- Ubiquity DB
  - Housekeeping metadata – PostgreSQL database with own Persistent Volume
- Ubiquity Provisioner
  - Create/Provision Volume and Delete Volume
- Ubiquity Flex plugin
  - Attach, Detach, Mount, Unmount
- Helm Utils (only during install/uninstall/upgrade)



# IBM Storage Enabler for Containers - Capabilities

- PVC access modes
  - RWO (ReadWriteOnce)  
The volume can be mounted as read-write by a single node
  - RWX (ReadWriteMany)  
The volume can be mounted as read-write by many nodes
  - ROX (ReadOnlyMany)  
The volume can be mounted read-only by many nodes
- PVC name can be auto-generated (UUID) or provided through label
- Storage Class parameters
  - File system name
  - Independent (default) or dependent filesets
  - INode limit
  - Existing filesets
  - Custom UID or GID
  - Reclaim Policy Delete or Retain

# Lessons Learned

Strictly follow the steps outlined in the Storage Enabler for Containers KnowledgeCenter

[https://www.ibm.com/support/knowledgecenter/SSC\\_KLT\\_2.1.0/UG/sec\\_ug\\_requirements\\_file.html](https://www.ibm.com/support/knowledgecenter/SSC_KLT_2.1.0/UG/sec_ug_requirements_file.html)

To check if enable-controller-attach-detach is activated on Kubernetes, use

```
$ kubectl cluster-info dump | grep attach-detach
```

```
"volumes.kubernetes.io/controller-managed-attach-detach": "true"
```

In addition to activating Quotas `mmchfs <fs> -Q yes` on the Scale file system, you need to:

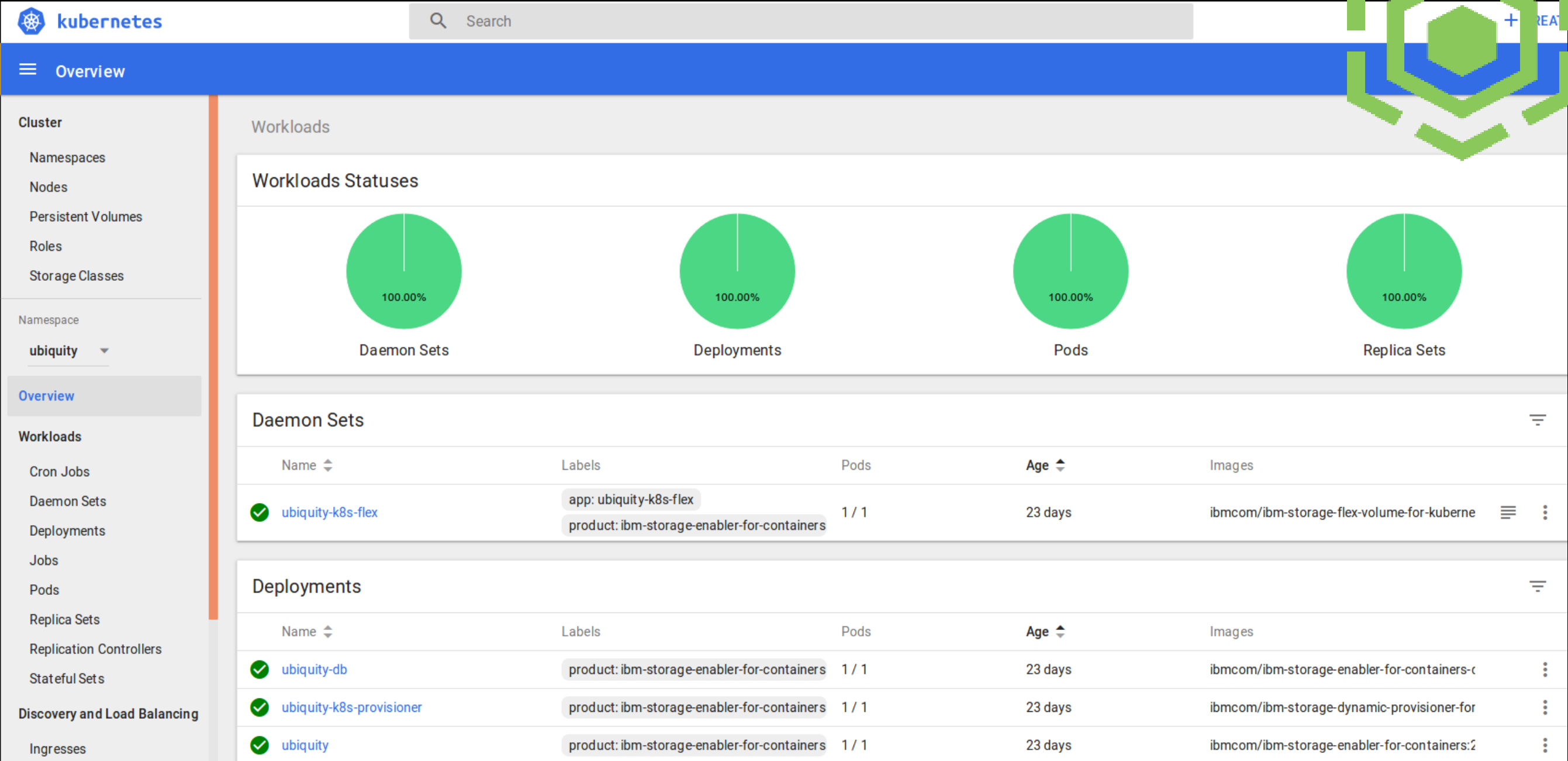
```
$ mmdefquotaon <fs>
```

....

... otherwise the PV can't be provisioned and you'll see an error like this:

```
ubiquity/flex ubiquity-k8s-provisioner-59c4995f56-g52f84aa9d611-1b33-11e9-a9fc-0242ac110007 Failed to provision volume with StorageClass "scalefs1": error creating volume: Quota not enabled for Filesystem [fs1]
```

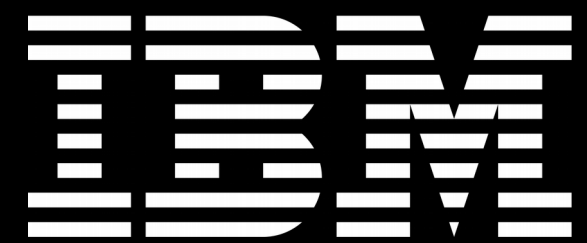
# Live Demo: IBM Storage Enabler for Containers 2.1.0 and Spectrum Scale



# Thank you!



Harald Seipp  
Senior Technical Staff Member, Center of Excellence for Cloud Storage  
IBM Systems EMEA Storage Competence Center  
[seipp@de.ibm.com](mailto:seipp@de.ibm.com)





# Legal notices

Copyright © 2019 by International Business Machines Corporation. All rights reserved.

No part of this document may be reproduced or transmitted in any form without written permission from IBM Corporation.

Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This document could include technical inaccuracies or typographical errors. IBM may make improvements and/or changes in the product(s) and/or program(s) described herein at any time without notice. Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does not infringe IBM's intellectually property rights, may be used instead.

THE INFORMATION PROVIDED IN THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER OR IMPLIED. IBM LY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted, if at all, according to the terms and conditions of the agreements (e.g., IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. 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 in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. IBM makes no representations or warranties, ed or implied, regarding non-IBM products and services.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing  
IBM Corporation  
North Castle Drive  
Armonk, NY 10504- 785  
U.S.A.

# Information and Trademarks

IBM, the IBM logo, ibm.com, IBM System Storage, IBM Spectrum Storage, IBM Spectrum Control, IBM Spectrum Protect, IBM Spectrum Archive, IBM Spectrum Virtualize, IBM Spectrum Scale, IBM Spectrum Accelerate, Softlayer, and XIV are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. 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 are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

IT Infrastructure Library is a Registered Trade Mark of AXELOS Limited.

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and other countries.

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

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

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

ITIL is a Registered Trade Mark of AXELOS Limited.

UNIX is a registered trademark of The Open Group in the United States and other countries.

\* All other products may be trademarks or registered trademarks of their respective companies.

## Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon 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 throughput improvements equivalent to the performance ratios stated here.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

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

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the 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.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

This presentation and the claims outlined in it were reviewed for compliance with US law. Adaptations of these claims for use in other geographies must be reviewed by the local country counsel for compliance with local laws.

# Special notices

This document was developed for IBM offerings in the United States as of the date of publication. IBM may not make these offerings available in other countries, and the information is subject to change without notice. Consult your local IBM business contact for information on the IBM offerings available in your area.

Information in this document concerning non-IBM products was obtained from the suppliers of these products or other public sources. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. Send license inquires, in writing, to IBM Director of Licensing, IBM Corporation, New Castle Drive, Armonk, NY 10504-1785 USA.

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

The information contained in this document has not been submitted to any formal IBM test and is provided "AS IS" with no warranties or guarantees either expressed or implied.

All examples cited or described in this document are presented as illustrations of the manner in which some IBM products can be used and the results that may be achieved. Actual environmental costs and performance characteristics will vary depending on individual client configurations and conditions.

IBM Global Financing offerings are provided through IBM Credit Corporation in the United States and other IBM subsidiaries and divisions worldwide to qualified commercial and government clients. Rates are based on a client's credit rating, financing terms, offering type, equipment type and options, and may vary by country. Other restrictions may apply. Rates and offerings are subject to change, extension or withdrawal without notice.

IBM is not responsible for printing errors in this document that result in pricing or information inaccuracies.

All prices shown are IBM's United States suggested list prices and are subject to change without notice; reseller prices may vary.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

Any performance data contained in this document was determined in a controlled environment. Actual results may vary significantly and are dependent on many factors including system hardware configuration and software design and configuration. Some measurements quoted in this document may have been made on development-level systems. There is no guarantee these measurements will be the same on generally-available systems. Some measurements quoted in this document may have been estimated through extrapolation. Users of this document should verify the applicable data for their specific environment.

Backup

# SEC 2.0.0 Limitations

Only one backend ie. either Block or Spectrum Scale can be enabled on a Kubernetes Cluster/ICP cluster.

Spectrum Scale 4.2.3 and above must be preinstalled along with GUI.

At least one filesystem must exist and mounted on all worker nodes.

Quota must be enabled for all the filesystems being used for creating persistent volumes.

(All) Kubernetes worker nodes must have Spectrum Scale client installed on them.

RWX support – Same PVC cannot be attached/mounted on more than one pod on same host.

If single PVC is used by multiple Pods then it is application's responsibility to maintain the data consistency.

## SEC 2.0.0 Limitations (cont.)

It is recommended to create PVC one after another. Create new PVC once all earlier PVC's created using SEC are in bound state.

Creation of large number of PVC in a single batch and deleting all PVC in one go is not recommended. Such action might result in overloading of Spectrum Scale GUI node which in turn lead to failure of creation/deletion of fileset on Spectrum Scale.

uid/gid/fileset-type/inode-limit storage class parameters are only allowed for new fileset creation.

Volume for ubiquity-db is created as a dependent fileset and changing it to independent fileset is not supported.

For each uid-gid combination a new storage class need to be defined. User with uid/gid must exist on Spectrum Scale.

Advanced Spectrum Scale functionality like AFM, Remote Mount, Encryption, Compression etc are not supported via SEC.



## SEC 2.0.0 Limitations (cont.)

The persistent volumes created using SEC with Spectrum Scale backend use Spectrum Scale quota to make sure one cannot overuse the storage space than specified in the PVC but it does not guarantee that storage specified in the PVC is actually available. It is up to the storage admin to make sure required storage is available on the Spectrum Scale filesystem.

SEC does not check the storage space available on the Spectrum Scale filesystem before creating the PVC. One has use Kubernetes storage resource quota to limit the number of PVC's or storage space.

Fileset created by SEC should not be unlinked or deleted from any other interface. Filesystem used by SEC for persistent volumes must be mounted on all worker nodes all the time.

ICP and Spectrum Scale GUI use port 443 so plan setup accordingly.

SEC does not support allowVolumeExpansion for storageclass.

# In a nutshell: What is CSI and why is it important?



- Standardized storage interface for Container Orchestration Systems (Kubernetes, Mesos, Docker, and CloudFoundry)
  - In addition, makes installing new volume plugins as easy as deploying a pod
- CSI spec V1.0 GA level of support in Kubernetes 1.13 (Dec 2018)
  - ICP 3.2 plans to support Kubernetes 1.13 in May 2019
  - OpenShift 4.1 plans to support Kubernetes 1.13 in May/June 2019
- Flex Volume plugins can coexist with CSI plugins. SIG Storage will continue to maintain the Flex API so that existing plugins will continue to work.
- New volume features (e.g. snapshots, resize) will be added only to CSI
- The new standard requested by customers and adopted by the industry.

# IBM Cloud Private and RedHat OpenShift

