IBM Spectrum Scale:
Use cases and (field) lessons learned with Kubernetes and OpenShift

—

Harald Seipp
Senior Technical Staff Member
Center of Excellence for Cloud Storage

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

Outline

- Baseline

- Use Cases
  - ➜ ADAS
  - ➜ CSP
  - ➜ Dev/Test
  - ➜ Multi-cloud

- Topics clients ask for
  - ➜ Unsupported platforms ;-)
  - ➜ Security / Multi-tenancy
 - ➜ Heterogenous environments

- Summary

Baseline: Containers Everywhere

- Multi-Cloud: On-premises and Public Clouds
  - Elastic scheduling and auto-scaling
  - Improved resource utilization
  - Secure isolation and Multi-Tenancy
  - Portable and reproducible service
  - One-click Laptop to Supercomputer

- Development, DevOps and continuous integration
  - Re-use of applications and services
  - Simplify and accelerate application deployment

- Microservices Architecture
  - Programming language and technology stack independence
  - Faster and easier development

Traditional Monolithic Application
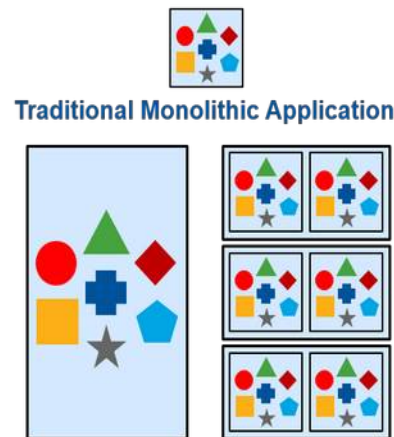
Scales by size … or monolithic replication.
Changes monolithically.

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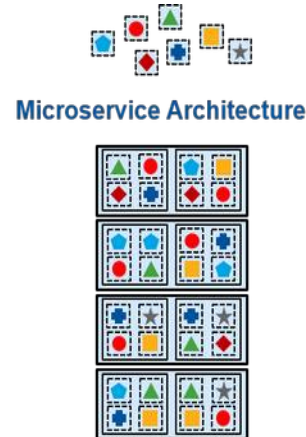
Scales by microservice replication.
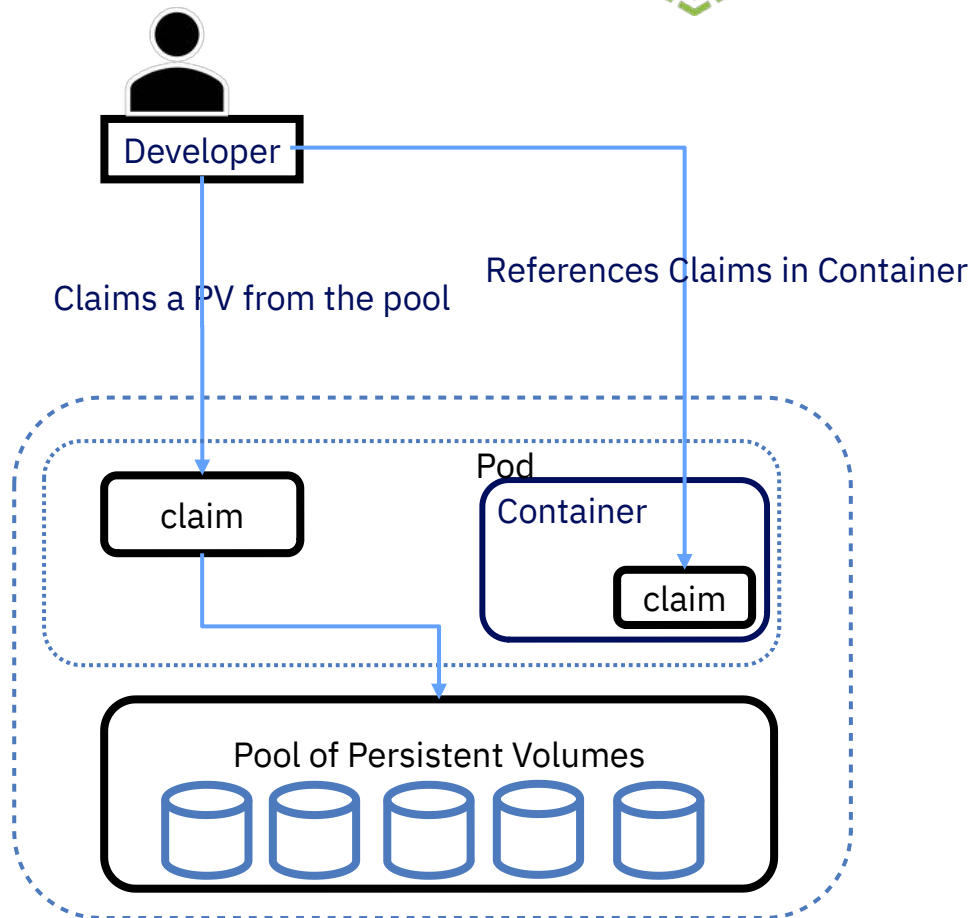*Changes* by microservices.

Baseline: 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/plug-in.

A PersistentVolumeClaim (PVC) is a request for storage by a user.

PVCs consume PV resources.

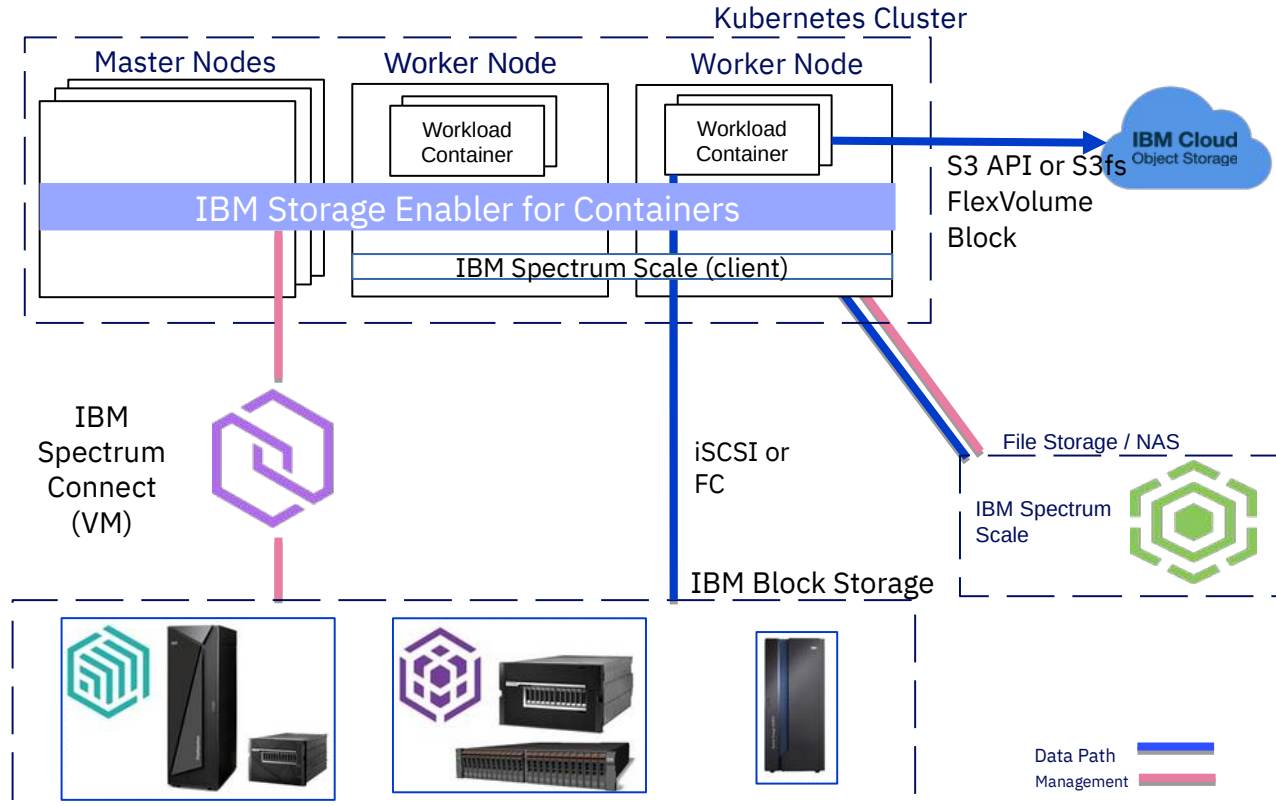A Storage Class is pre-defined by the admin to dynamically provision PVs.

Developer

Claims a PV from the pool

References Claims in Container

Pod

claim

Container

claim

Pool of Persistent Volumes

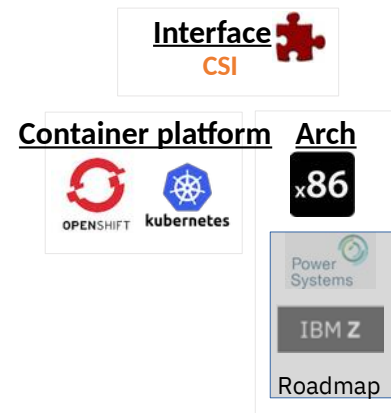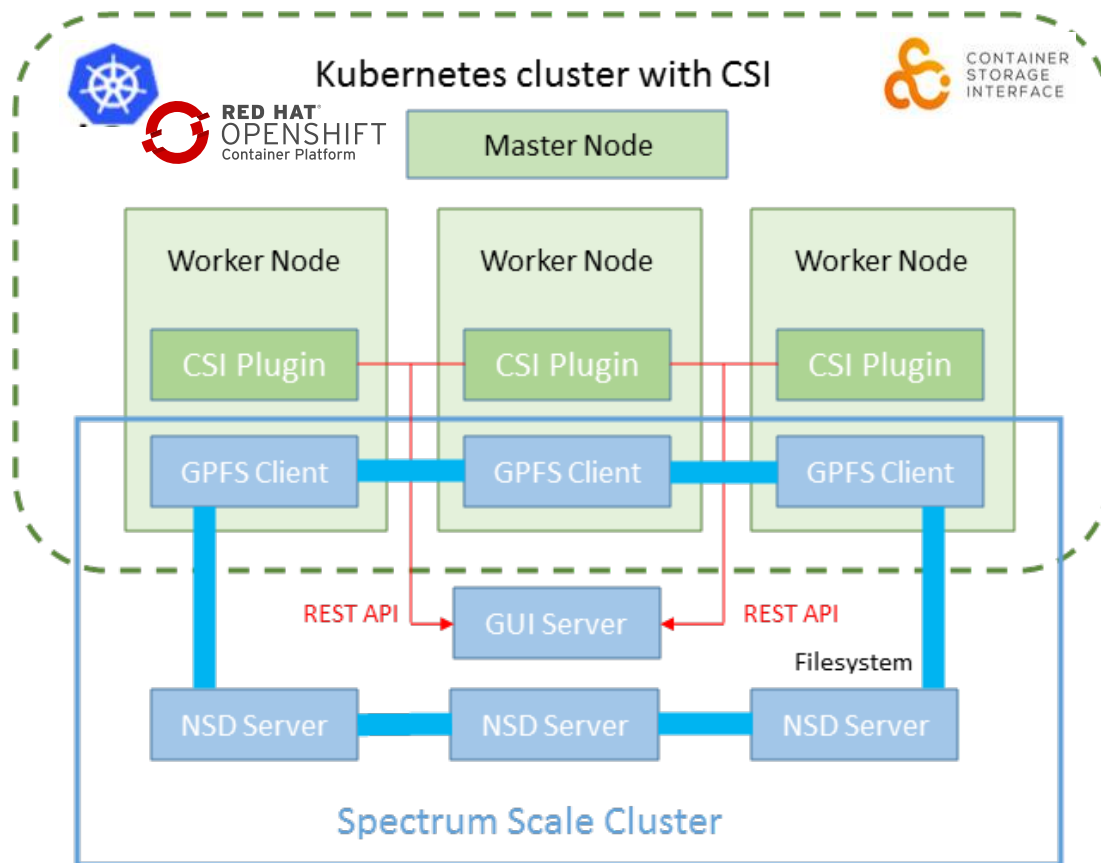# IBM persistent storage architecture 2017-2019

IBM Spectrum Scale

IBM Storage Enabler for Containers provides Kubernetes FlexVolume Drivers

IBM Spectrum Connect provides the Storage management API and Storage classes **for IBM Block Storage**

Kubernetes Cluster

Master Nodes | Worker Node | Worker Node

Workload Container

Workload Container

IBM Storage Enabler for Containers

IBM Spectrum Scale (client)

S3 API or S3fs FlexVolume Block

IBM Cloud Object Storage

IBM Spectrum Connect (VM)

iSCSI or FC

File Storage / NAS

IBM Spectrum Scale

IBM Block Storage

Data Path
Management

IBM persistent file storage architecture 2019+
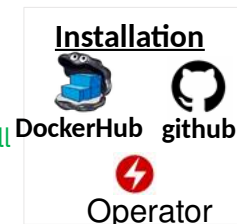
Multi-cluster support

Kubernetes Cluster with CSI

Master Node(s)

Infrastructure Node	Infrastructure Node	Worker Node	Worker Node	Worker Node
External-Provisioner	External-Attacher	Microservice A	Microservice B	
CSI - Plugin	CSI - Plugin	CSI - Plugin	CSI - Plugin	CSI - Plugin
Scale Client	Scale Client	Scale Client	Scale Client	Scale Client
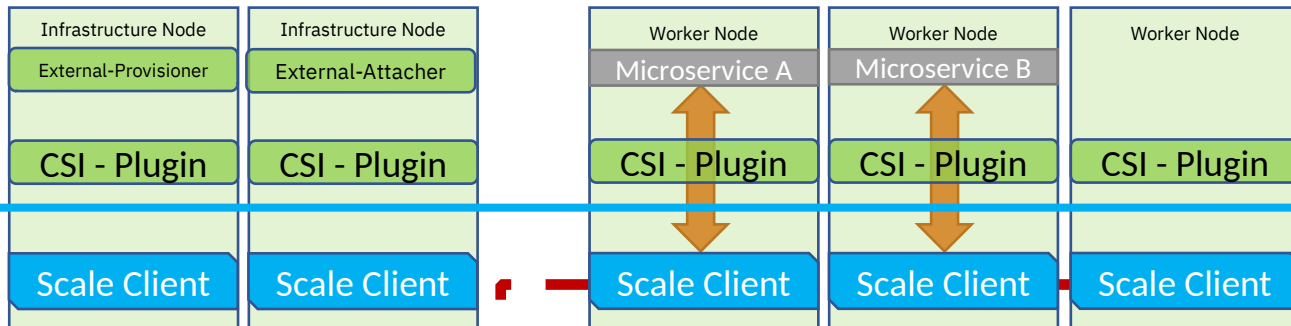
Scale GUI Server(s)

Spectrum Scale Client cluster

High speed network

Scale GUI Server(s)   Scale NSD Server   Scale NSD Server   Scale NSD Server

Spectrum Scale Storage cluster

SEC vs. CSI

Feature	Storage Enabler for Containers	CSI Driver 1.0.x
Static Provisioning	No K8s native static provisioning support, but there is way to use an existing fileset as a PV	Yes
Dynamic Provisioning	Fileset based (dependent and independent)	Lighweight (directory based) and fileset based (dependent and independent)
Multiple Filesystem Support	Yes	Yes
Remote cluster mounts	Limited support	Yes
Deployment	Helm Chart	Operator
Spectrum Scale Version	5.0.0+	5.0.4.1+
Platform support	RHEL 7.x on x86_64, ppc64le, s390x	RHEL 7.x on x86_64
Kubernetes Version	1.10, 1.11, and 1.12	1.13+
Openshift Support	3.11 via RPQ only	4.2
Migration path from SEC to CSI	N/A	Manual offline migration

Use Cases

Spectrum Scale use case at automotive client

- Containerized platform to train and test the AI for an ADAS (Advanced Driver Assistance Systems) project

- High-bandwidth data ingest (double-digit TB per day) through Spectrum Scale/ESS

- Sophisticated cloud architecture (see next slide)

- Skilled admins and developers

Spectrum Scale use case
at automotive client – cluster architecture

RedHat OpenShift

RHEL

OpenStack

VMware ESX

x86 Servers + Nvidia GPUs

IBM **Spectrum Scale** / IBM **Elastic Storage Server** (ESS)

Spectrum Scale use case
at automotive client – lessons learned

- To get the Spectrum Scale client up & running

  - Assigned an additional OpenStack network with a dedicated NIC to the VM

    - With OpenStack floating IP the Scale Cluster IP was not visible within VM

  - Adjusted OpenStack security groups to allow inbound traffic to the Spectrum Scale ports

- To ensure that persistent Pods are placed on the Spectrum Scale node(s)

  - Labeled the node and added a nodeSelector to the persistent Pod deployment configs
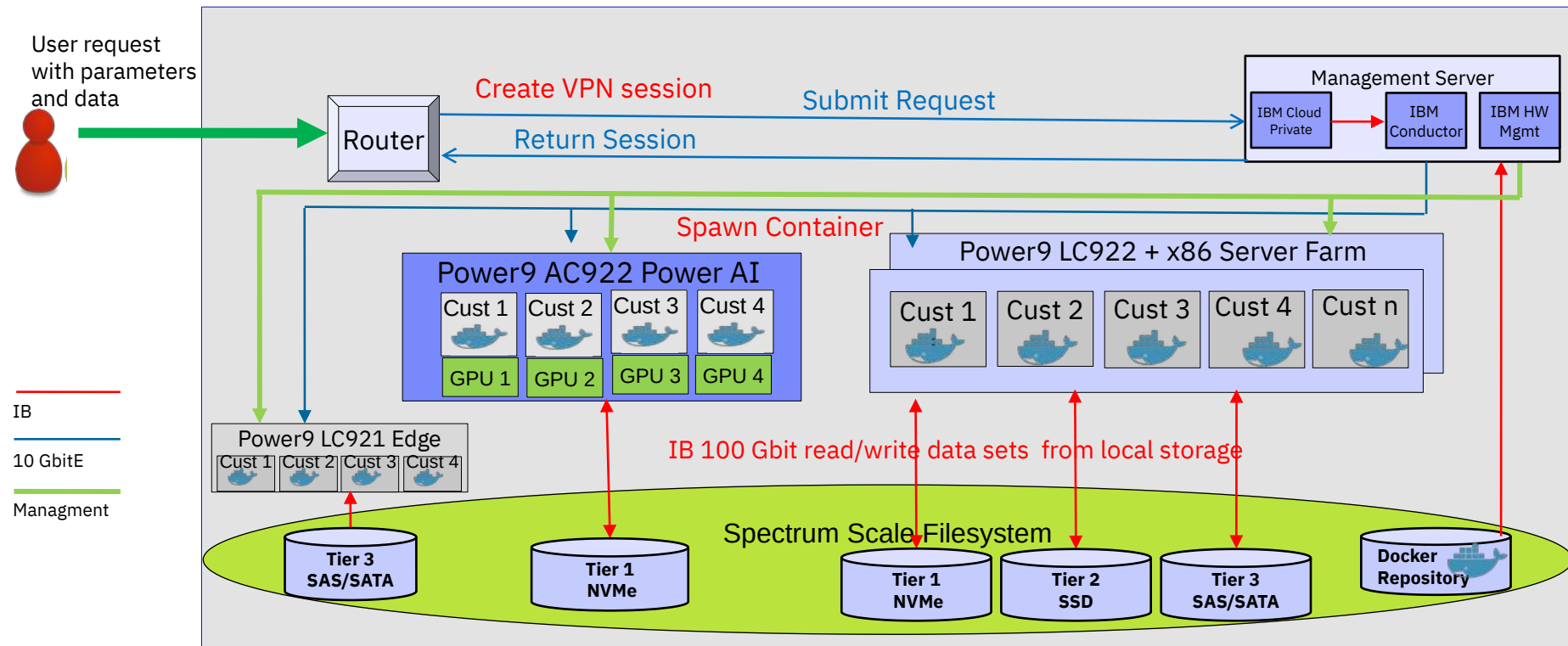
Spectrum Scale use case at automotive client – lessons learned (cont.)

- Made Storage Enabler for Containers 2.0.0 work
  - Steps are now documented as solution blueprint

- Re-installed with SEC 2.1.0
  - SEC helm chart 1.0.1 requires container privilege adjustments for the deployment to succeed

- Existing data can be integrated though additional storage class for existing filesets
  - Created a PVC to re-use an existing Spectrum Scale fileset (ingest directory)
 - Ensure sufficient access rights for the container process

Spectrum Scale use case at Cloud Service Provider: AIaaS*

User request with parameters and data

Router

Create VPN session

Submit Request

Return Session

Management Server

IBM Cloud Private → IBM Conductor | IBM HW Mgmt

Spawn Container

Power9 AC922 Power AI
Cust 1 | Cust 2 | Cust 3 | Cust 4
GPU 1 | GPU 2 | GPU 3 | GPU 4

Power9 LC922 + x86 Server Farm
Cust 1 | Cust 2 | Cust 3 | Cust 4 | Cust n

Power9 LC921 Edge
Cust 1 | Cust 2 | Cust 3 | Cust 4

IB 100 Gbit read/write data sets from local storage

Spectrum Scale Filesystem

Tier 3 SAS/SATA | Tier 1 NVMe | Tier 1 NVMe | Tier 2 SSD | Tier 3 SAS/SATA | Docker Repository

Legend:
- IB
- 10 GbitE
- Managment

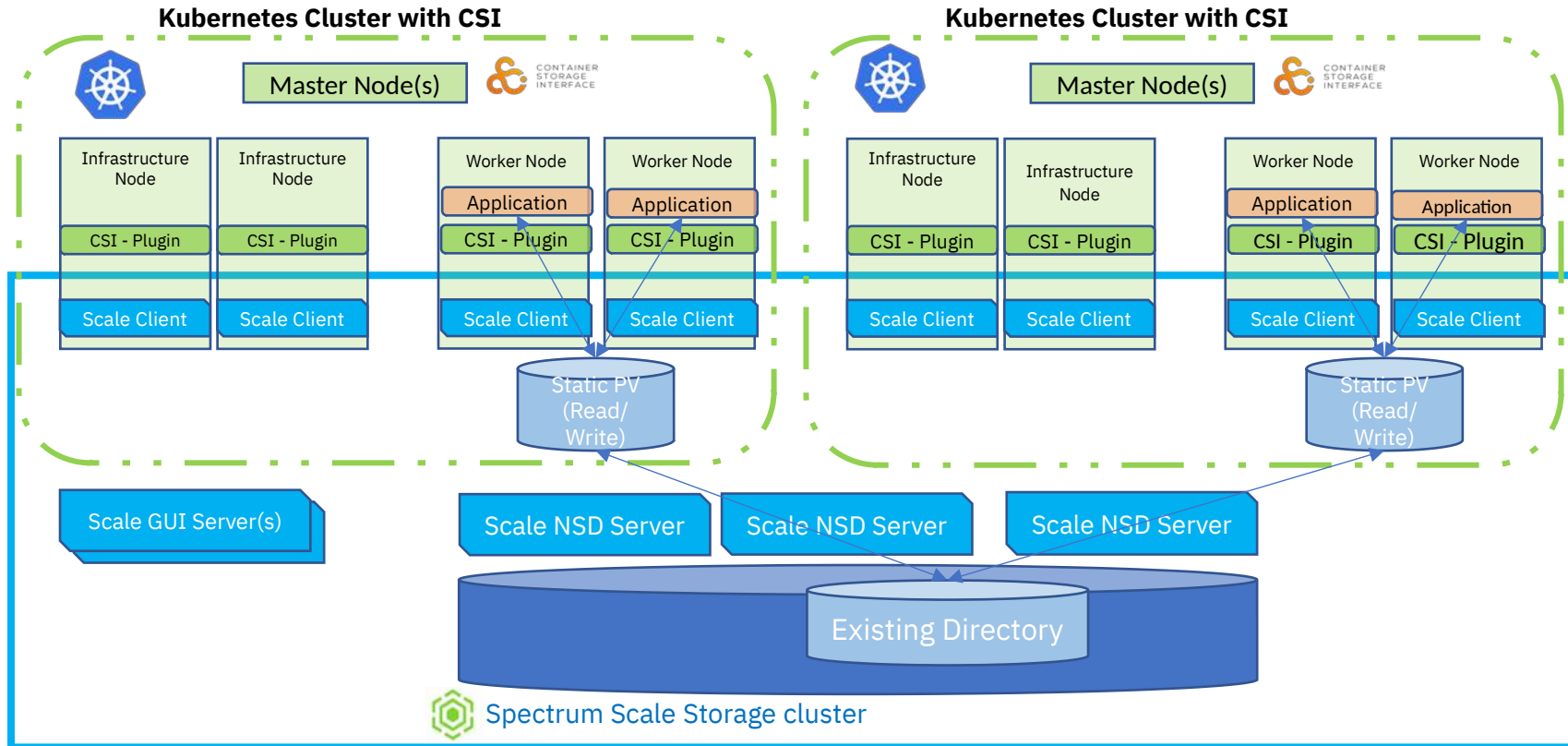# Spectrum Scale use case at CSP: AIaas – lessons learned

- For the CSP, multi-tenant isolation and data management is more important than dynamic provisioning
  - Used (semi-)static provisioning to prevent Fileset sprawl
  - CSP pre-creates PVs/Filesets with own naming conventions
  - Better control of tenant-related services (Snapshots, Backup)

- Integration with 3-Tier concept
  - Through Fileset placement policies
  - One Kubernetes Storage Class per Tier
  - Leveraging CSI driver lightweight volumes

- CSI Driver works with IBM Cloud Private 3.2.1 on Power
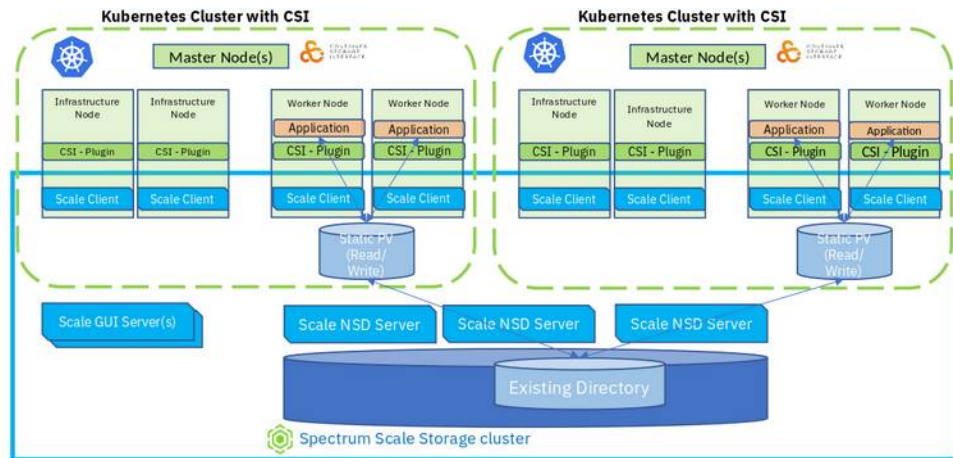 - No IBM CSI Driver Support here, CSP is supporting
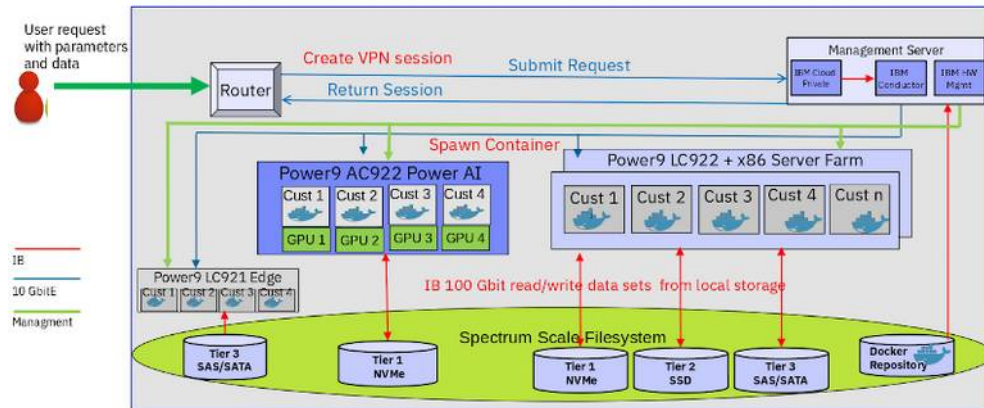
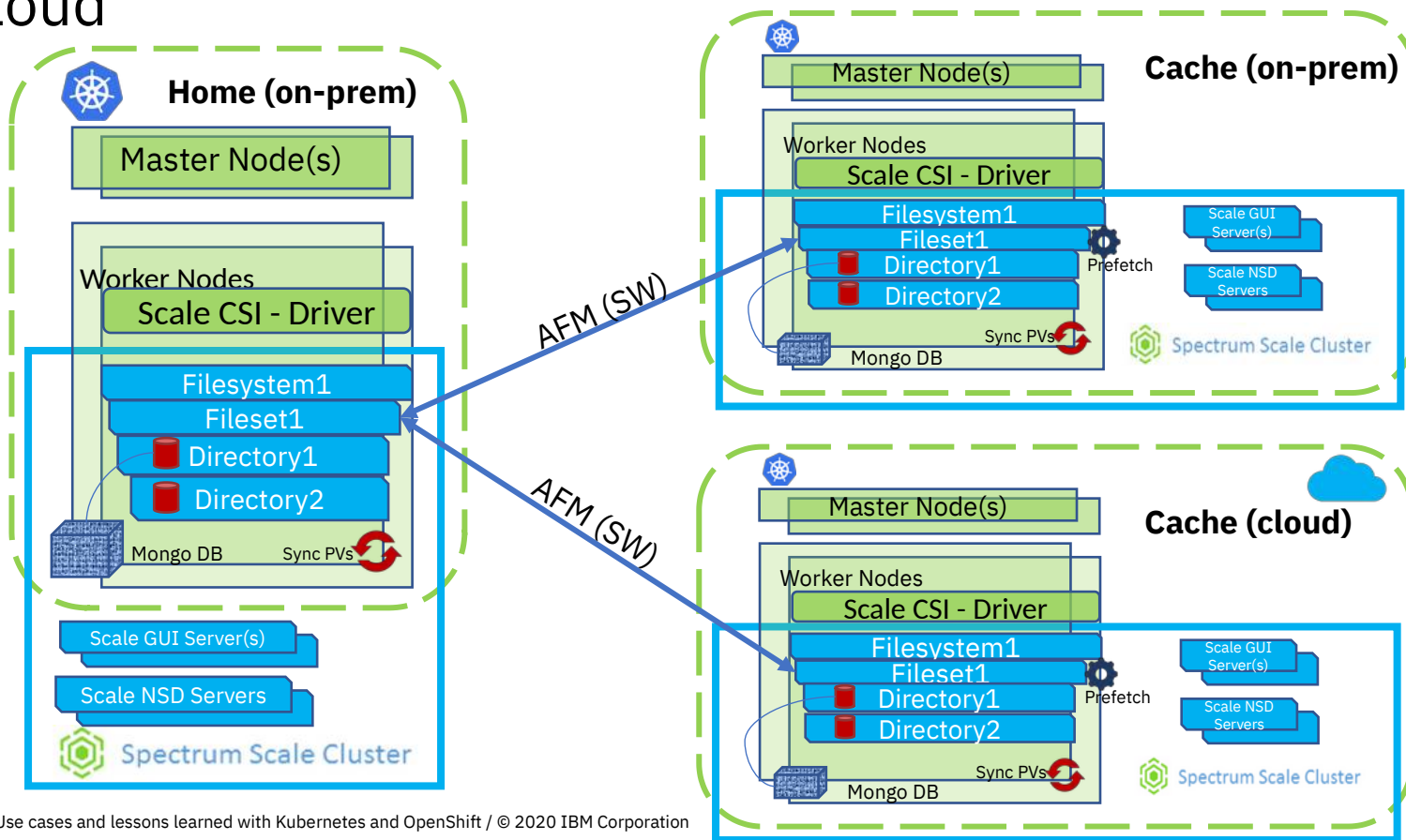# Spectrum Scale use case Dev/Test

Spectrum Scale use case Dev/Test – lessons learned

- Stronger isolation in Dev/Test clusters beyond Kubernetes namespace separation

  - Separate Kubernetes clusters with different cluster admins
  - Still single data plane wanted
    - With or without access to same data

Spectrum Scale use case Multicloud

Spectrum Scale use case Multicloud – lessons learned

- To service workload on cloud
  - Single-writer (home site) only
  - Processed data should be pushed to separate file system

- For DR purposes
  - Only one workload container should be up at given time
  - Independent writer can be used, but monitor cloud data outgoing traffic (EGRESS)

- Can be used with Spectrum Scale on AWS automation
  - Requires (not-yet-supported) Spectrum Scale GUI
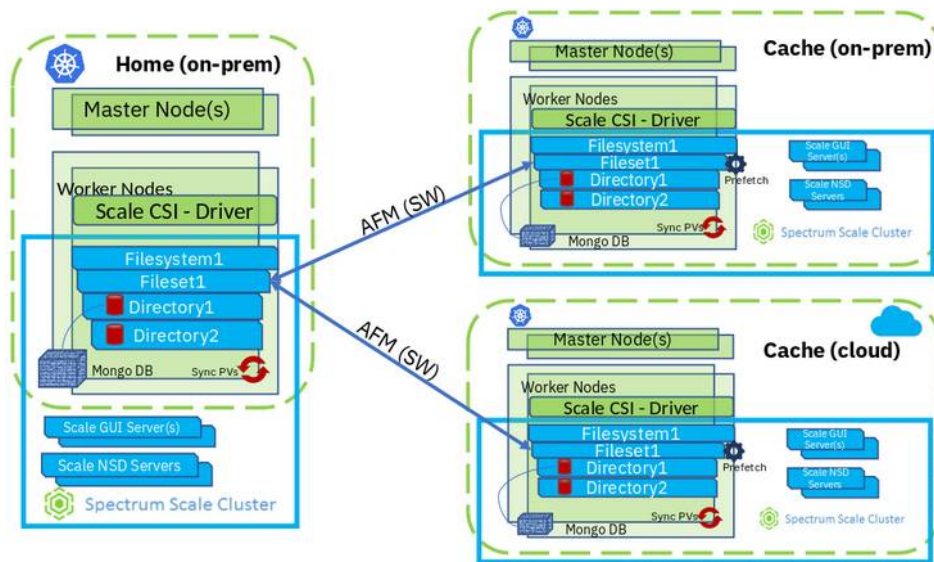
Topics clients ask for...

CSI driver unsupported platforms

In Research

- Operating Systems
  - Ubuntu 18.04.x
    - No issues probe testing with plain k8s

- Orchestration Platforms
  - IBM Cloud Private
    - See Cloud Service Provider use case
    - Cross-architecture (x86, Power, Z) support in a single k8s cluster
    - To get the CSI driver running
      - add `ibm-anyuid-hostaccess-psp` Pod Security Policy to ClusterRole CRs
  - Rancher
    - Sold as "plain k8s", easy deployment & nice GUI
    - To get the CSI driver running
      - Remove the `type: Directory` specifier from the `gpfs-classic` hostPath mount of the ibm-spectrum-scale-csi daemonSet
      - Add a PodSecurityPolicy to the `ibm-spectrum-scale-csi*` ClusterRole definitions when running Rancher with PSP enforcement

```
- apiGroups: ['policy']
  resources: ['podsecuritypolicies']
  resourceNames: ["ibm-anyuid-hostaccess-psp"]
  verbs:      ['use']
- apiGroups: ["extensions"]
  resources: ["podsecuritypolicies"]
  resourceNames: ["ibm-anyuid-hostaccess-psp"]
  verbs: ["use"]
```
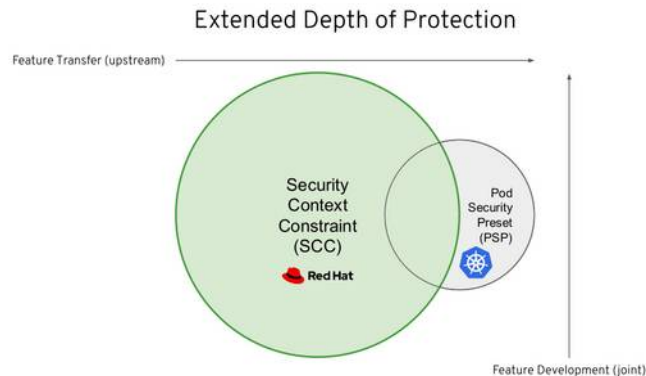
```
- hostPath:
    path: /gpfs/fs1
    type: Directory
  name: gpfs-classic
```

```
- apiGroups:
  - policy
  resourceNames:
  - csi-psp
  resources:
  - podsecuritypolicies
  verbs:
  - use
```

```
---
apiVersion: policy/v1beta1
kind: PodSecurityPolicy
metadata:
  name: csi-psp
spec:
  requiredDropCapabilities:
  - NET_RAW
  privileged: false
  allowPrivilegeEscalation: false
  defaultAllowPrivilegeEscalation:
false
  hostNetwork: true
  hostPorts:
  - min: 0
    max: 65535
  fsGroup:
    rule: RunAsAny
  runAsUser:
    rule: RunAsAny
  seLinux:
    rule: RunAsAny
  supplementalGroups:
    rule: RunAsAny
  volumes:
  - emptyDir
  - secret
  - persistentVolumeClaim
  - downwardAPI
  - configMap
  - projected
  - hostPath
```

Security / Multi-tenancy

- For maximum container security, consider OpenShift

    – Uses K8s RBAC plus Security Context Constraints (plus SELinux) to restrict container process security impact

- Kubernetes Administrator (kubeadmin) user can access all PVs

    – One could use RBAC /PSP or SCCs to create "less privileged admin" role

    – For strict isolation use separate Kubernetes clusters

        - Optionally use Kubernetes federation (v2!)

Extended Depth of Protection

Feature Transfer (upstream)

Security Context Constraint (SCC)

Red Hat

Pod Security Preset (PSP)

Feature Development (joint)

Heterogenous environments

- Multi-platform
  - X86, IBM Power, IBM Z

- IBM Cloud Private supports multi-architecture clusters
  - IBM Storage Enabler for Containers supports all three platforms
  - Spectrum Scale supports all three platforms (Linux)

- OpenShift support for IBM Power and IBM Z is coming
  - Initial releases might have limitations regarding RHEL worker node and multi-architecture support

IBM **Power**
Systems

x86

IBM Z

Summary and call to action

- Spectrum Scale provides tight Kubernetes and OpenShift integration through

  - IBM Storage Enabler for Containers (FlexVolume)
    for Kubernetes < 1.13 and OpenShift 3.x

  - IBM CSI drivers
    for Kubernetes 1.13 and later and OpenShift 4.x

- Use cases (ADAS, CSP, Dev/Test and more) are evolving

- Involve us (ATS / Lab Services)

  - To discuss your Spectrum Scale Container Platform use case

  - If you need help with IBM Storage Orchestration for Containers

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

Thank you!

Please help us to improve Spectrum Scale with your feedback

- If you get a survey in your email or a popup from the GUI, please respond

- We read every single reply

Teilen Sie sich mit!

Nutzen Sie den Hashtag #IBMStorageDays auf LinkedIn, Twitter und Co. und berichten von unseren Vorträgen!

Let's connect!

Bei Fragen einfach melden!

 @haraldseipp

 linkedin.com/in/harald-seipp-43458356

Notices and Disclaimers

© 2020 International Business Machines Corporation. No part of this document may be reproduced or transmitted in any form without written permission from IBM.

U.S. Government Users Restricted Rights — use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.

Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. **This document is distributed "as is" without any warranty, either express or implied. In no event, shall IBM be liable for any damage arising from the use of this information, including but not limited to, loss of data, business interruption, loss of profit or loss of opportunity.** IBM products and services are warranted per the terms and conditions of the agreements under which they are provided.

IBM products are manufactured from new parts or new and used parts.
In some cases, a product may not be new and may have been previously installed. Regardless, our warranty terms apply."

Any statements regarding IBM's future direction, intent or product plans are subject to change or withdrawal without notice.

Performance data contained herein was generally obtained in a controlled, isolated environments. Customer examples are presented as illustrations of how thosecustomers have used IBM products and the results they may have achieved. Actual performance, cost, savings or other results in other operating environments may vary.

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.

Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall constitute legal or other guidance or advice to any individual participant or their specific situation.

It is the customer's responsibility to insure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer follows any law.

Notices and Disclaimers (cont.)

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 about this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. IBM does not warrant the quality of any third-party products, or the ability of any such third-party products to interoperate with IBM's products. **IBM expressly disclaims all warranties, expressed or implied, including but not limited to, the implied warranties of merchantability and fitness for a purpose.**

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents, copyrights, trademarks or other intellectual property right.

IBM, the IBM logo, ibm.com and [names of other referenced IBM products and services used in the presentation] are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at: www.ibm.com/legal/copytrade.shtml

Bookmarks

- IBM Spectrum Scale CSI Driver documentation:
  https://www.ibm.com/support/knowledgecenter/en/STXKQY_5.0.4/com.ibm.spectrum.scale.csi.v5r04.doc/bl1csi_kc_landing.html

- Storage Enabler for Containers documentation:
  https://www.ibm.com/support/knowledgecenter/SSCKLT_2.1.0/sec_kc_welcome.html

- IBM Spectrum Scale CSI Driver github:
  https://github.com/IBM/ibm-spectrum-scale-csi

- IBM Spectrum Scale CSI Driver video blogs:
  https://developer.ibm.com/storage/2019/12/26/ibm-spectrum-scale-csi-driver-video-blogs/

- IBM Spectrum Scale CSI Driver at OperatorHub:
  https://operatorhub.io/operator/ibm-spectrum-scale-csi-operator

- Kubernetes CSI documentation:
  https://kubernetes-csi.github.io/docs/