

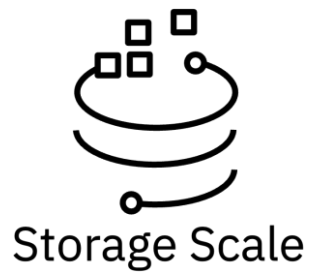
Perfekter Brückenschlag in die Cloud(s) mit IBM Storage Scale



Frank Kraemer
IBM Systems Architect
Storage Solutions for Data and AI
kraemerf@de.ibm.com



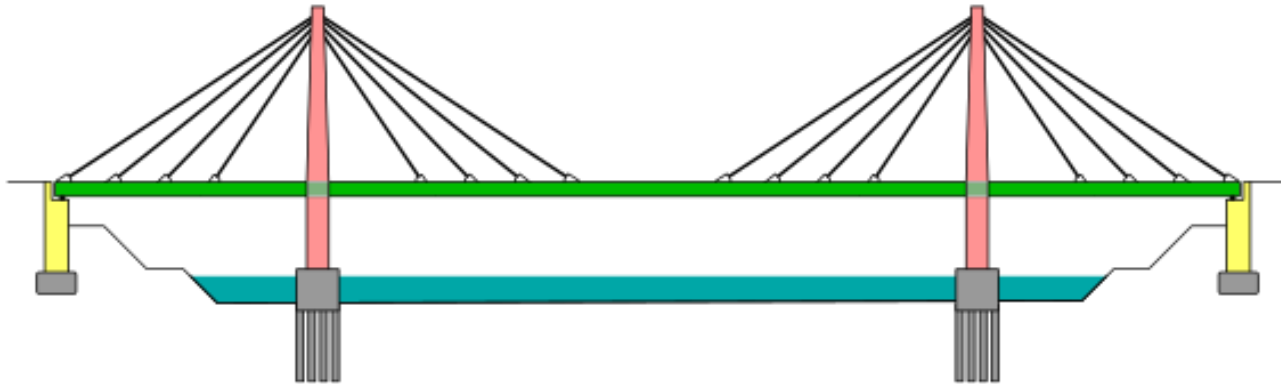
Disclaimer



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. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

IBM reserves the right to change product specifications and offerings at any time without notice. This publication could include technical inaccuracies or typographical errors. References herein to IBM products and services do not imply that IBM intends to make them available in all countries.

Was ist ein Brückenschlag?



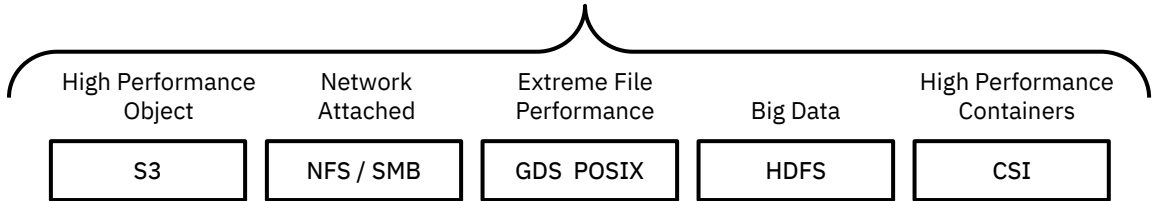
Brückenschlag bezeichnet den Bau einer Brücke. Eine Brücke ist ein Bauwerk, das Verkehrswege oder Versorgungseinrichtungen über natürliche Hindernisse oder andere Verkehrswege hinwegführt.
(Quelle: <https://de.wikipedia.org/wiki/Br%C3%BCckenschlag>)

A Global Data Platform for Unstructured Data

Unifying File and Object Storage to Provide Common Data Services



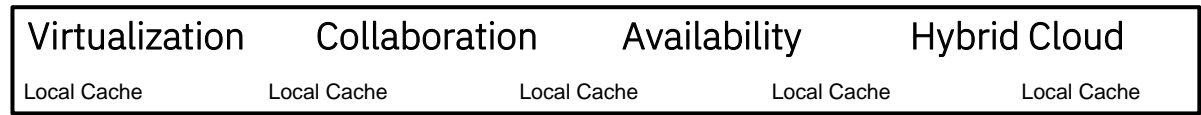
Data Access Services



Data Resiliency Services

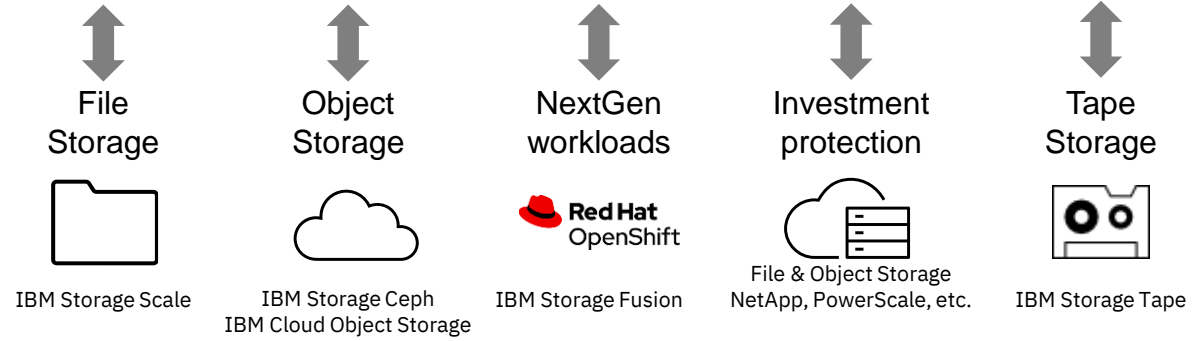
Identify

Data Caching & Core Services



Protect

Data Management Services



Detect

Respond

Recover

A Global Data Platform for Unstructured Data

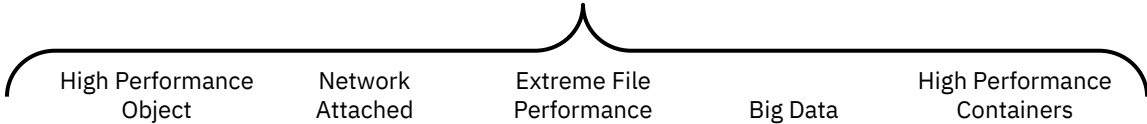
Unifying File and Object Storage to Provide Common Data Services



Data Access Services

Data Caching & Core Services

Data Management Services



Source: Victorinox.com

Data Resiliency Services

Identify

Protect

Detect

Respond

Recover

Differentiators, Strengths, and Emerging Technologies



Unified Data Access

Performance Optimized,
Shared Multi Protocol
Access

Containerized and Non-
Containerized Workflows

High Fidelity S3 Object
Store

Workload Driven, Scaling to
Thousands of Clients

Caching & Core

Active File and Object
Management from Edge-
Core-Cloud

- File / Object Virtualization
- High Performance Tiering
- True Hybrid Multi Cloud

Enterprise Capabilities

- Replication, Snapshots,
Encryption, Tiering,
Compression

Performance Optimized

- Storage Scale System
- GPUDirect storage

Data Management

Integrated Data Cataloging

Automated Infrastructure
Management with Ansible
Playbooks

Policy Driven ILM

- Automatic, Intelligent, Cost
Efficient, Tiering of Data
- Seamless Archive to Tape

Data Resiliency

End to End Data Security
and Resiliency Capabilities

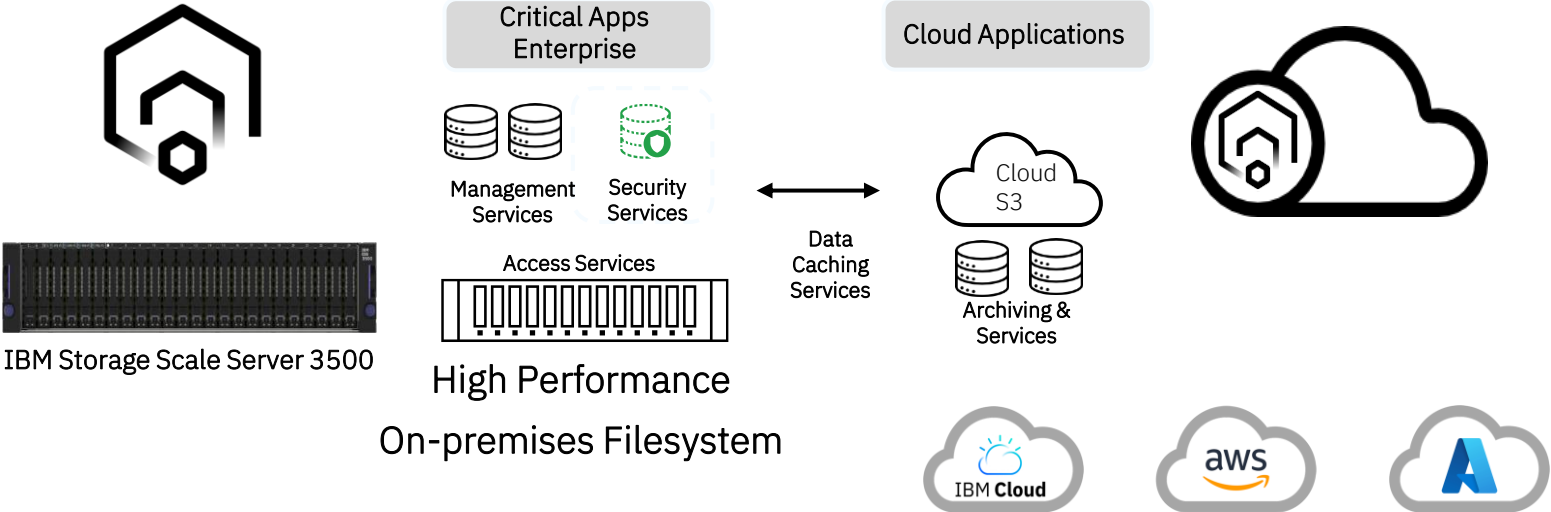
- SEIM Integration
- Data Protection through
Snapshots, Replication,
Backup, Disaster Recovery,
Immutability, Audit Logging

Safeguarded Immutable
Snapshots with Cyber Vault
Services

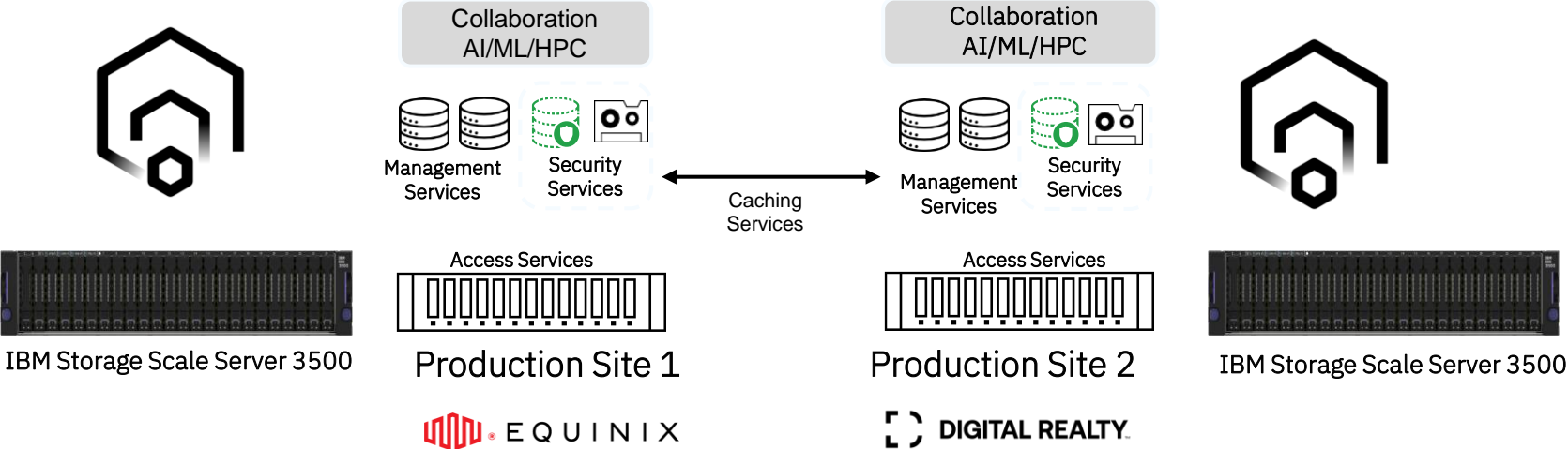
Near-Instant Data Access
Recovery Scenarios

Highly performant, enterprise ready, hybrid cloud-enabled global data platform from IBM that no competitor can match

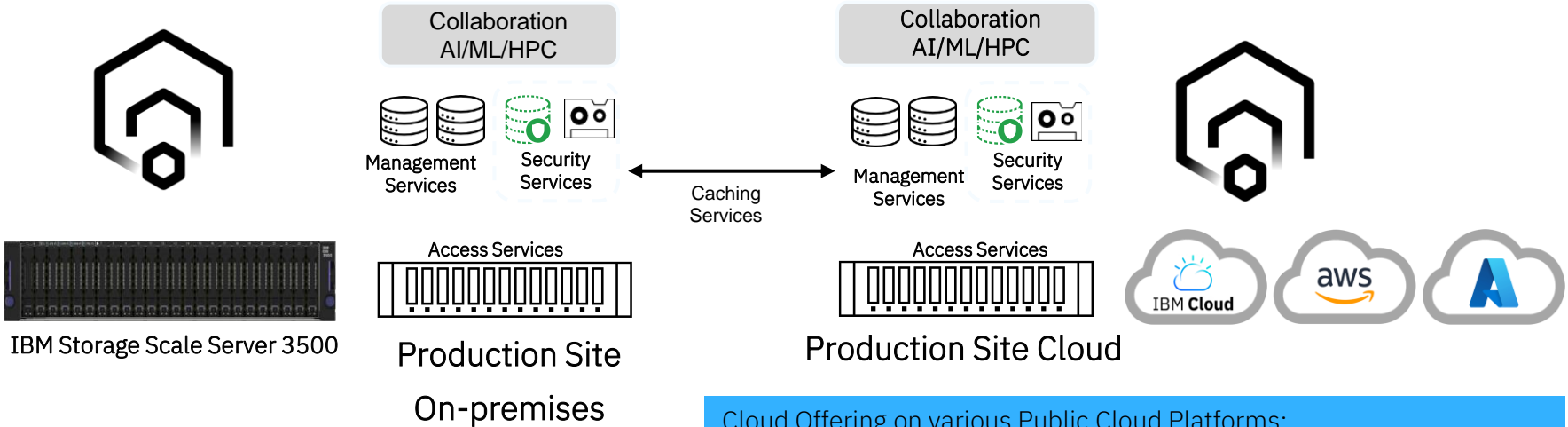
#1 Storage Scale to S3 (Cloud)



#2 Storage Scale to Storage Scale (Dual Site concept)

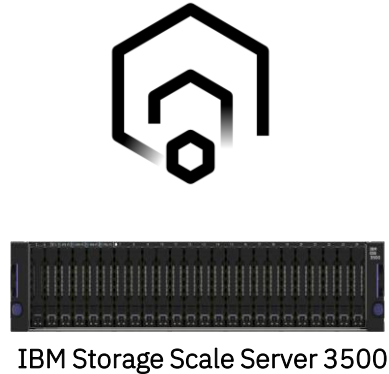


#3 Storage Scale to Storage Scale (Cloud Site concept)

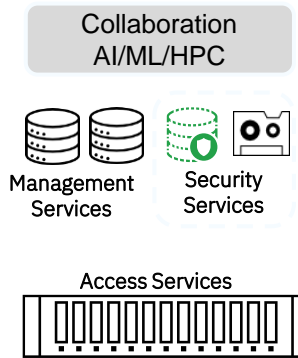


#3.1 Example: Oracle Cloud Infrastructure (OCI)

<https://blogs.oracle.com/cloud-infrastructure/post/accelerate-ai-ml-workloads-oci-nvidia-ibm>



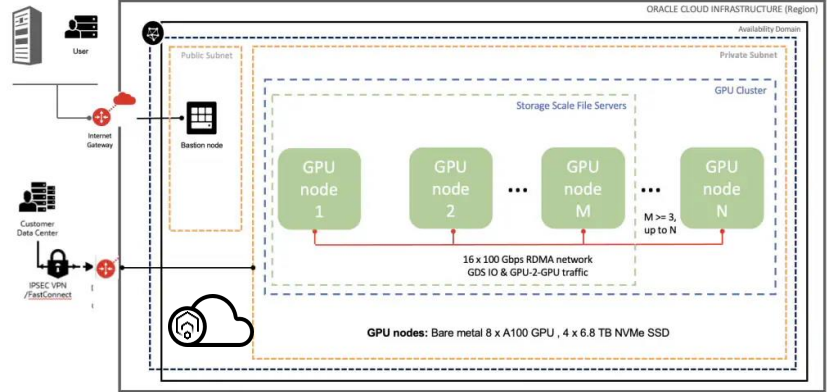
IBM Storage Scale Server 3500



Production Site
On-premises



Caching Services

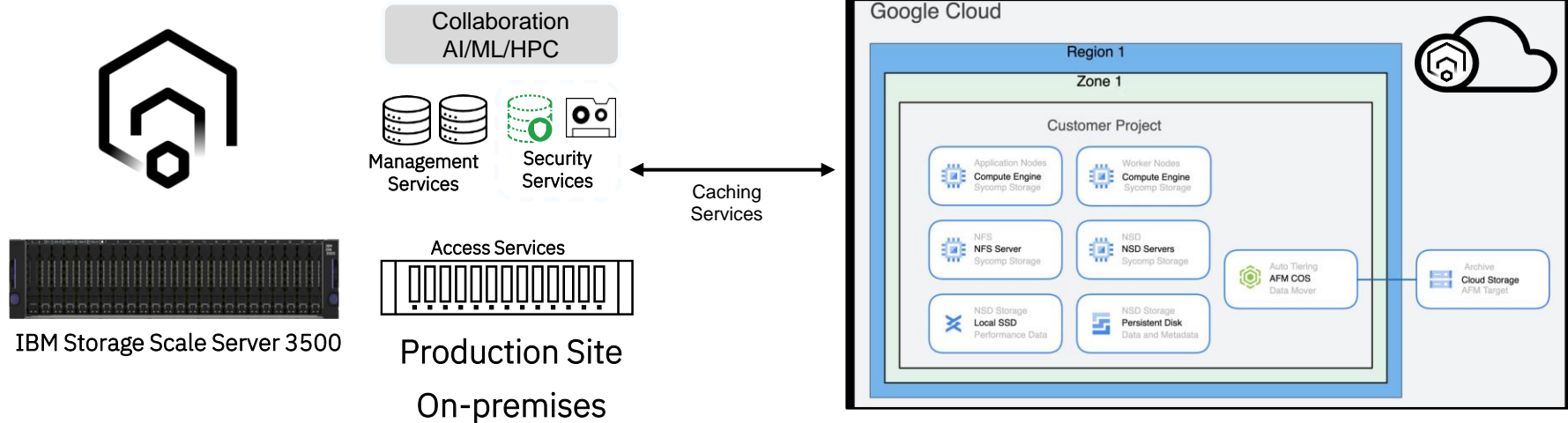


In collaboration with NVIDIA and IBM, Oracle Cloud Infrastructure (OCI) validated NVIDIA Magnum IO Storage on OCI using a hyperconverged cluster of bare metal NVIDIA A100 Tensor Core GPU compute nodes.

#3.2 Example: Google Marketplace Solution

(by IBM partner Sycomp)

<https://cloud.google.com/blog/products/storage-data-transfer/announcing-ibm-spectrum-scale-is-now-available-in-google-cloud/>



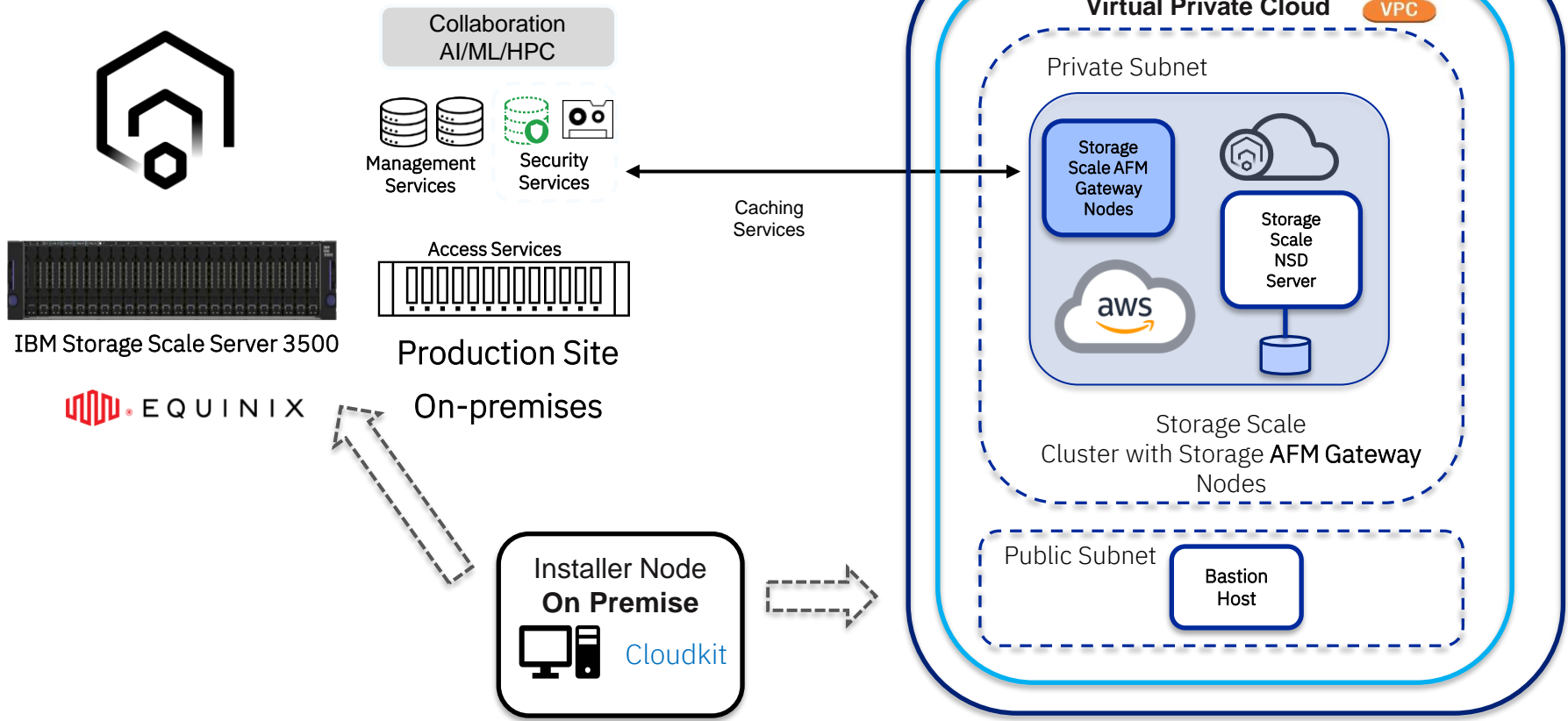
IBM Storage Scale Server 3500

Production Site
On-premises



Deploy an instance of Sycomp Storage Fueled by IBM Spectrum Scale using Terraform provided by Sycomp through Google marketplace.

#3.3 Example: AWS by Scale Cloudkit (Self Managed)



IBM Storage Scale Cloudkit: Introduction



What is Storage Scale Cloudkit?

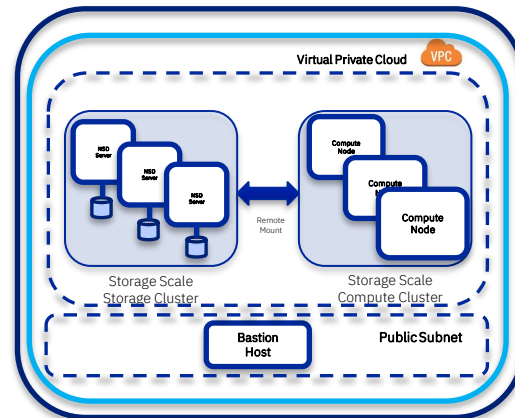
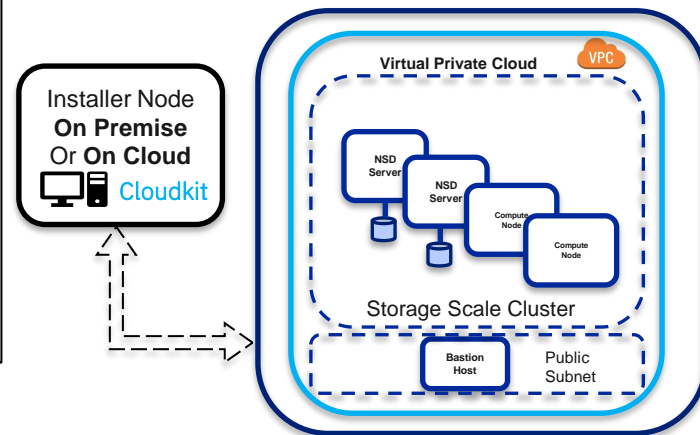
- Command Line Interface to create Storage Scale clusters on the cloud
- Provides end to end automation to deploy and configure Storage Scale cluster(s) on public clouds:
 - Automates infrastructure provisioning on the cloud
 - Automates the deployment of Spectrum Scale on the cloud
 - Applies Spectrum Scale best practises for deploying on the cloud

Advantages of Storage Scale Cloudkit

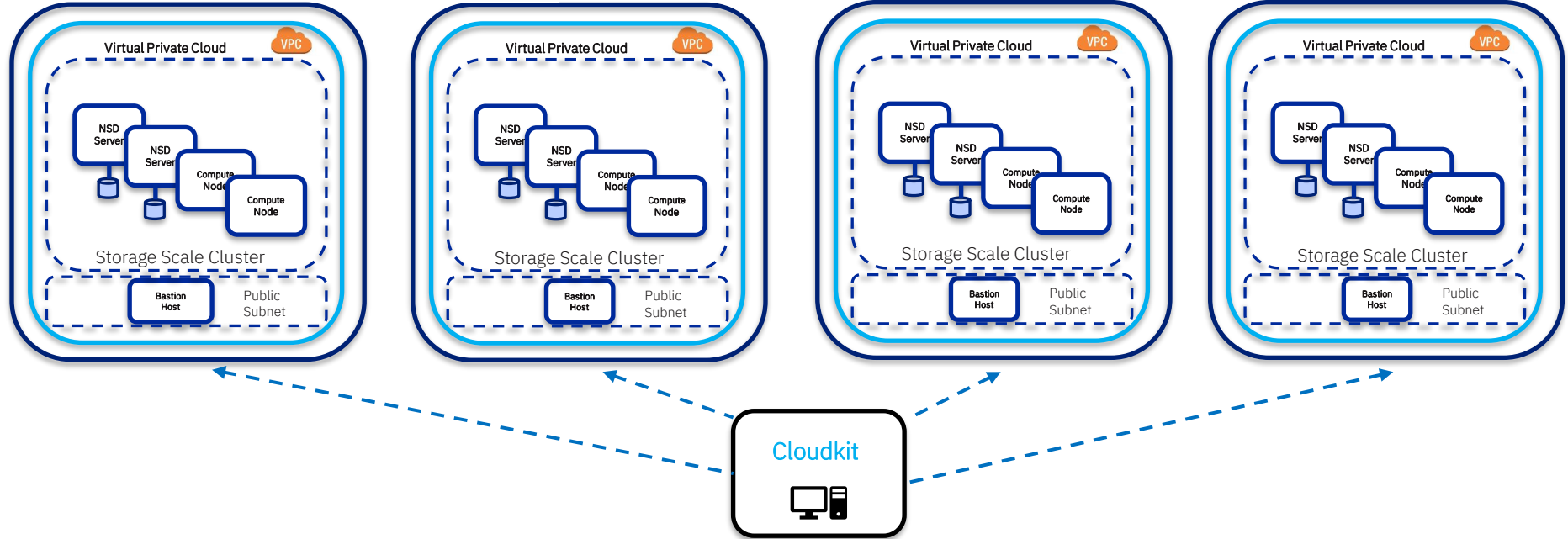
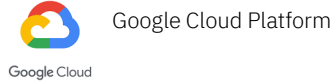
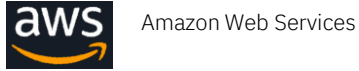
- Support for all major public clouds (more details see Roadmap)
- Rapid deployment and configuration of Storage Scale on the Cloud (in minutes)
- Out of the box “best practice” configuration
- Easy to use, guided interface
- Ability to provision multiple clusters from a single cloudkit instance
- Flexible deployment (supports multiple deployment models to suit customer needs)

- Deploy into existing customer environment (existing VPC)
- Create and deploy into new VPC
- Support Direct Access Network connection between On-prem and Cloud clusters

- Deploy separate Storage Scale Storage and Compute clusters
- Deploy a single Storage Scale cluster with Storage and Compute nodes



Cloudkit: Multi Cloud Support



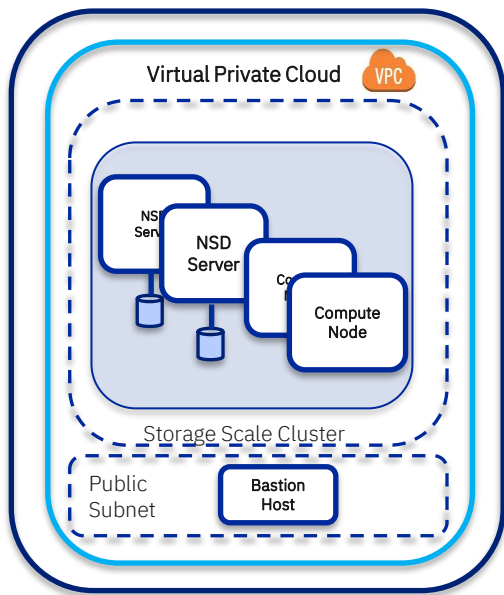
A single instance of Cloudkit can deploy Storage Scale clusters on different public clouds:

- Unified Deployment Experience
- No longer necessary to use different Cloud specific deployment technologies to provision on different clouds
- Cloud Agnostic Best Practise Deployment
- Enables Global Data Platform

Cloudkit: Flexible Deployment Architecture

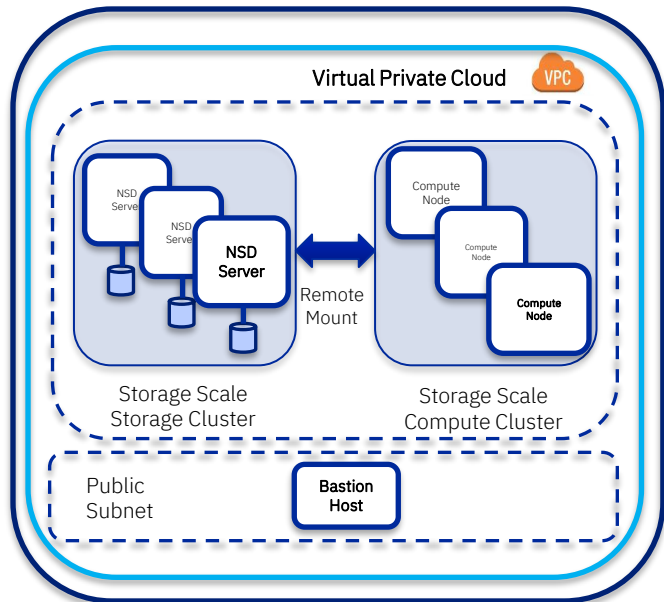


Single Unified Cluster



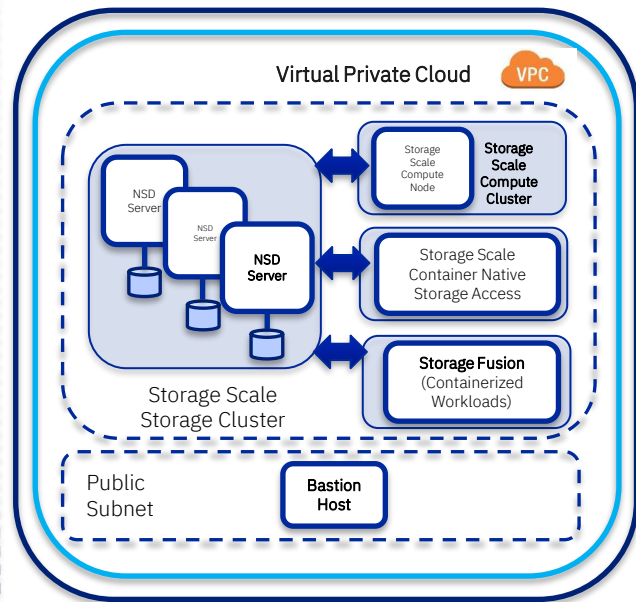
- Single administrative domain
- Storage and Compute nodes in the same cluster
- Ideal for small organizations

Separate Storage Cluster remote mounted Compute Cluster (s)



- Separate administrative domains
- Rapid Provision/Deprovision of Compute Clusters as required
- Storage Cluster persistent and isolated

Single Storage Cluster remote mounted to multiple Compute Clusters

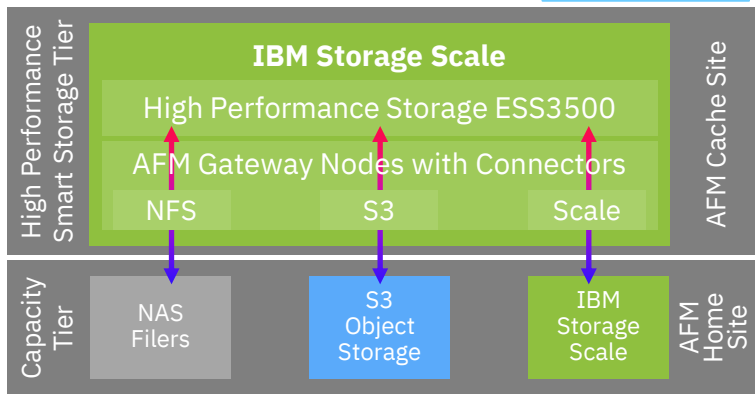


- Single Storage Data Lake for entire organization (Storage Cluster)
- Separate Compute for each Department administrated independently
- Provides Persistence, High Performance Storage for IBM Storage Fusion (Containerized Workloads)

Scale Data Caching Services (AFM) - Use case details #1

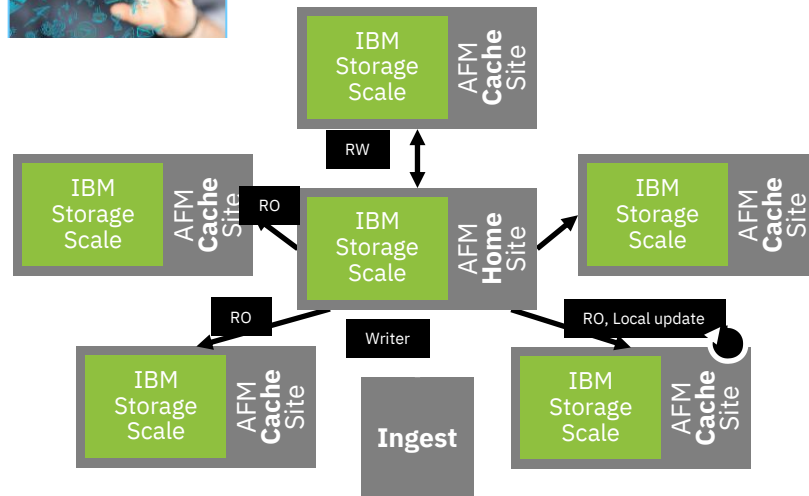


Data Virtualization



- Vertical caching
- Common namespace across isolated data silos in legacy 3rd party data stores
- Transparent access to all data regardless of silos
- Scale-out Posix performance
- Data export via NFS, SMB, HDFS, Object
- Can be used to seamlessly migrate data to new storage

Data Collaboration

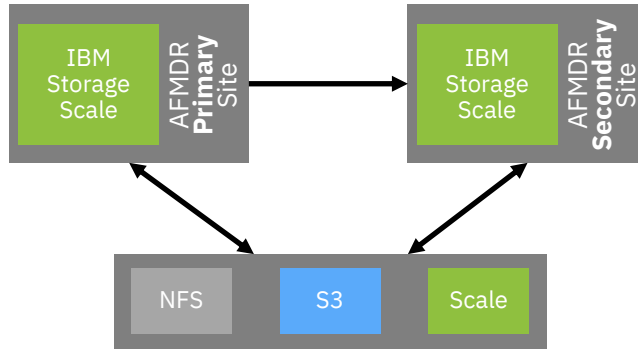


- Consistent cache provides a single source of truth with no stale data copies
- Horizontal caching
- Bi-direction traffic from Edge to Center
- Eventually Consistent data cache
- Transparent on-demand data access and transfer
- Policy driven data prefetch and eviction

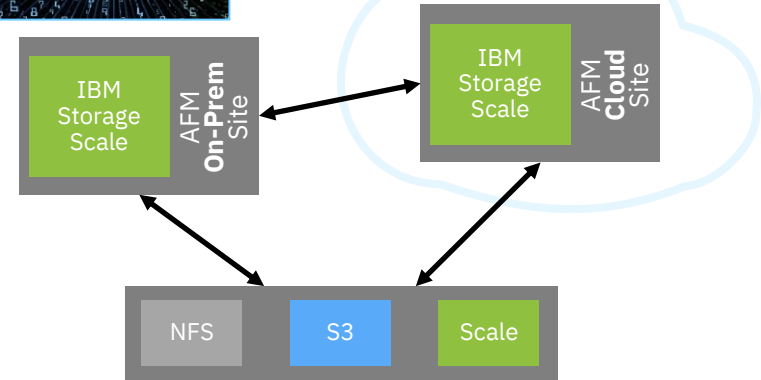
Scale Data Caching Services (AFM) - Use case details #2



Data Resilience



Hybrid cloud / Bursting



- Active-Passive DR over WAN or Cloud
- Designed for high latency and asynchronous DR
- Hot standby failover to DR site
- Automatic fallback data reconciliation
- Read-only access / analytics to all data at passive site






- Rapidly expand compute resources to cloud or data centers
- Common file system creates a single namespace across all locations
- Transparent access to data
- Cost effective way to increase compute on existing data
- Analytic results automatically pushed to home site

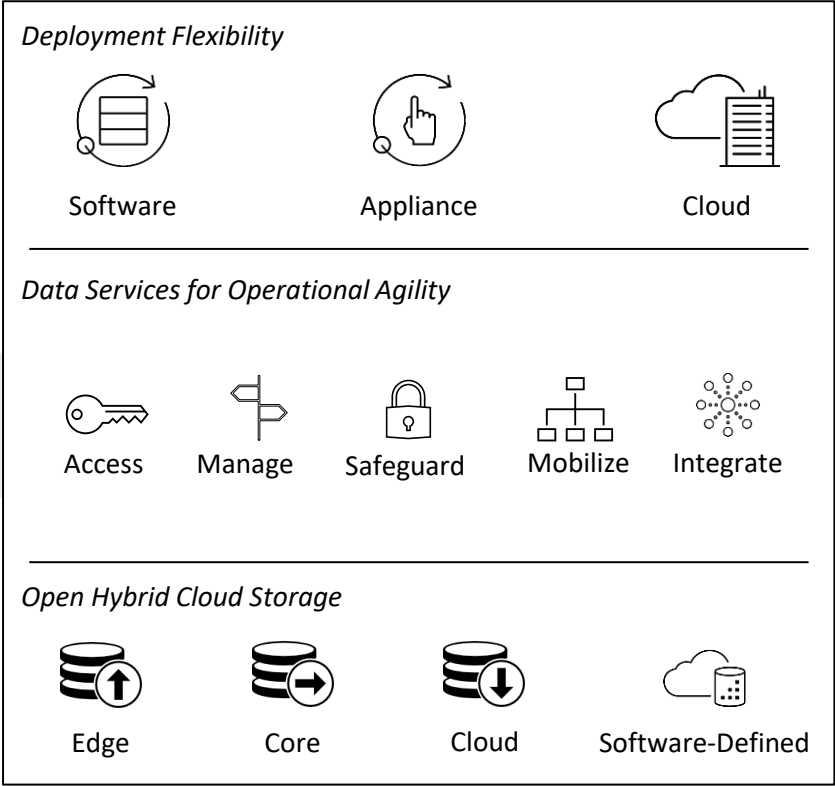
Summary: IBM's Global Data Platform



An Open Hybrid Cloud Experience

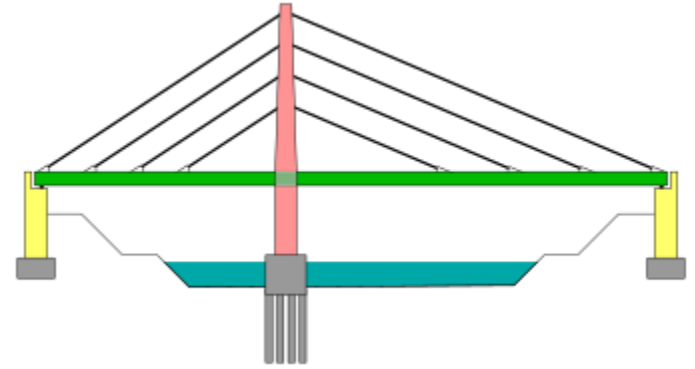
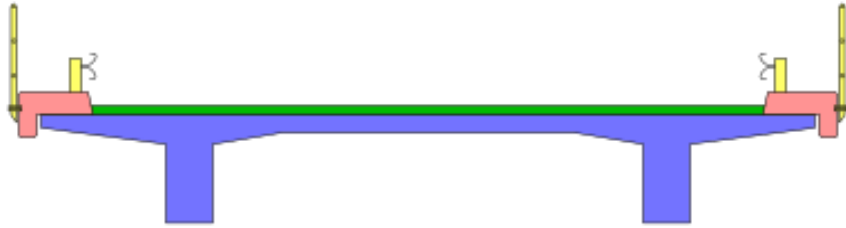
Data Sources

- 
Transactional
- 
Application
- 
Semi-Structured
- 
Mixed Media
- 
IoT

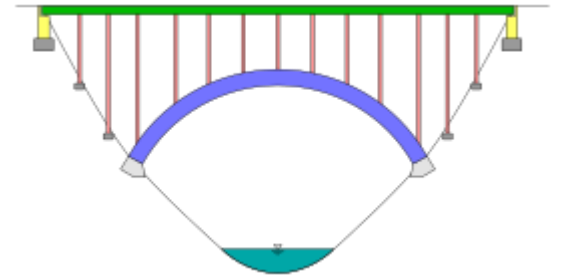
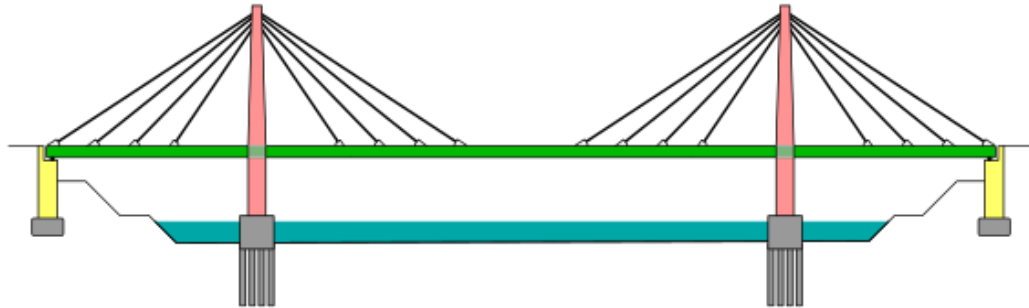


Outcomes

- 
Data Monetization
- 
Faster Time to Market
- 
Org Growth
- 
Digital Business



IBM



(Quelle: <https://de.wikipedia.org/wiki/Br%C3%BCckenschlag>)