High Availability – Cloud Pak for Integration

Mangesh Patankar

Developer Advocate - IBM

IBM Cloud Pak for Integration High Availability

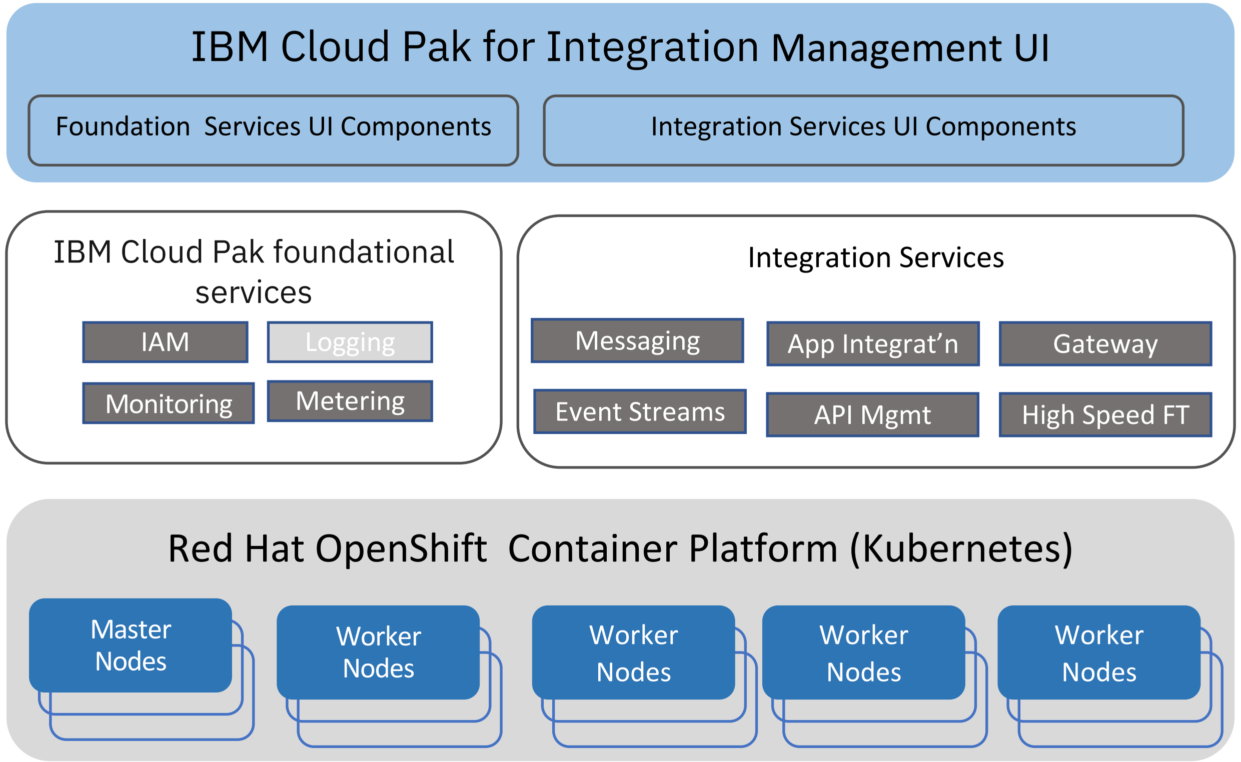
Understanding High Availability for the IBM Cloud Pak for Integration involves first understanding the overall architecture and the High availability considerations for the container platform. Based on this we then explore HA for each Integration service.

Agenda:

- Overview of IBM Cloud Pak for Integration Architecture
- OpenShift High Availability Concerns
- Platform and Services Availability
- Deployment considerations
- IBM Cloud Pak for Integration High Availability Summary
- MQ High Availability
- Event Streams High Availability
- API Connect High Availability
- App Connect High Availability

IBM Cloud Pak for Integration Architecture

- IBM Cloud Pak for Integration deploys and manages instances of **Integration Services** running on Kubernetes infrastructure, namely Red Hat OpenShift Container Platform.
- Instances of Integration services are deployed individually as needed to satisfy each use-case.
- Services can be deployed in **Highly Available** topologies across multiple Kubernetes worker nodes or non HA on a single worker.
- Deployment and management is via the UI or CLI allowing integration with **CI/CD pipelines**.
- IBM Cloud Pak for Integration leverages **Cloud Pak Foundation Common Services.**
- The IBM Cloud Pak for Integration Platform Navigator UI unifies the management UIs of the Integration Services and the Common Services.


IBM Cloud Pak for Integration Management UI
Foundation Services UI Components
Integration Services UI Components
IBM Cloud Pak foundational services
IAM
Logging
Monitoring
Metering
Integration Services
Messaging
App Integrat'n
Gateway
Event Streams
API Mgmt
High Speed FT
Red Hat OpenShift Container Platform (Kubernetes)
Master Nodes
Worker Nodes
Worker Nodes
Worker Nodes
Worker Nodes

Deploying CP4I components in a HA configuration

- Each component of CP4I uses one of these approaches to high availability:

Active-active replication

- Active-active replication involves running multiple instances of a service that can process all workloads concurrently and use a load balancer that spreads the workload across instances. See Node configurations on Wikipedia.
- Active-active replication requires compute nodes in at least 2 failure domains.

Active-passive replication

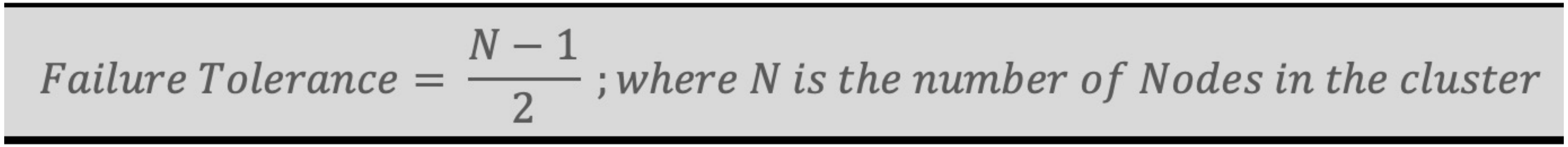
Active-passive replication involves running one or more instances of a service that process workloads (active), and one or more additional instances of the service that are on standby (passive) which are able to take over in the event of a failure in an active instance. See Node configurations on Wikipedia.

- Active-active replication requires compute nodes in at least 2 failure domains.

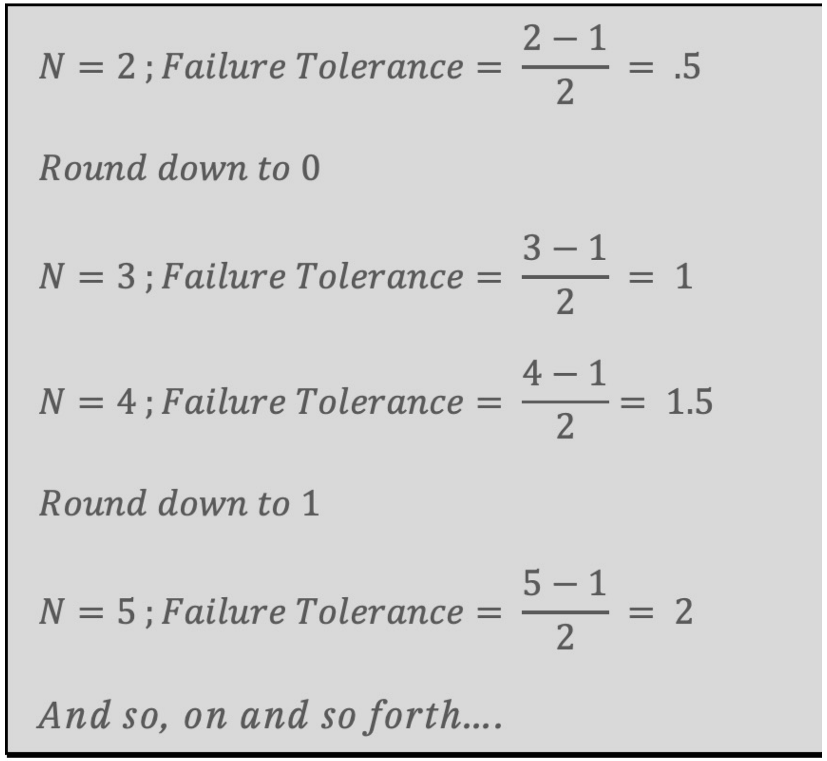
Quorum

- Quorum-based HA relies on a majority of nodes being available to vote. It requires a minimum of 3 nodes to be effective. See Quorum (distributed computing) on Wikipedia.
- Quorum requires compute nodes in at least 3 failure domains.

Quorum



$$Failure\ Tolerance = \frac{N-1}{2}\ ;where\ N\ is\ the\ number\ of\ Nodes\ in\ the\ cluster$$



$$N = 2\,; Failure\ Tolerance = \frac{2-1}{2} = .5$$

Round down to 0

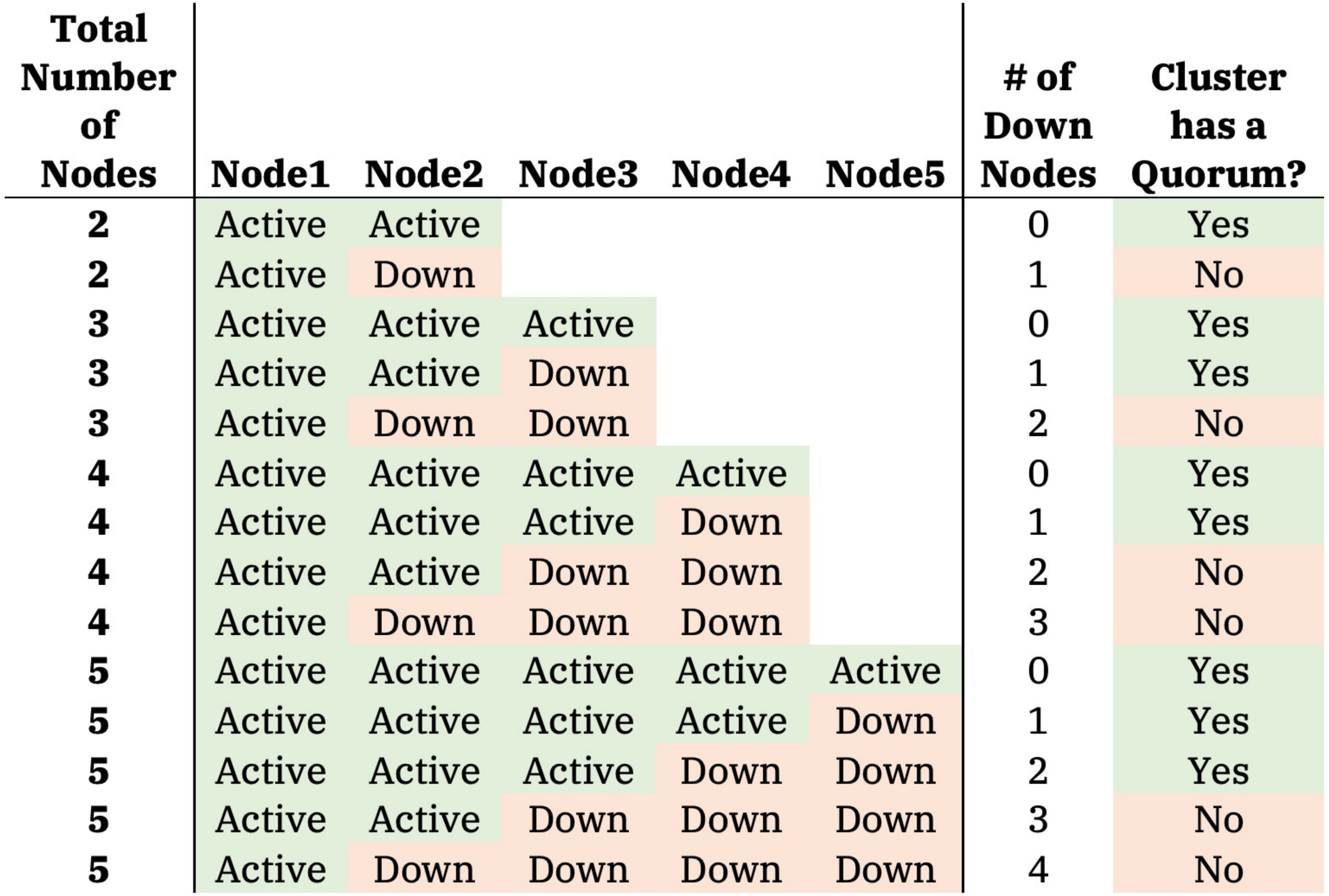
$$N = 3\,; Failure\ Tolerance = \frac{3-1}{2} = 1$$

$$N = 4\,; Failure\ Tolerance = \frac{4-1}{2} = 1.5$$

Round down to 1

$$N = 5\,; Failure\ Tolerance = \frac{5-1}{2} = 2$$

And so, on and so forth….



Total Number of Nodes	Node1	Node2	Node3	Node4	Node5	# of Down Nodes	Cluster has a Quorum?
2	Active	Active				0	Yes
2	Active	Down				1	No
3	Active	Active	Active			0	Yes
3	Active	Active	Down			1	Yes
3	Active	Down	Down			2	No
4	Active	Active	Active	Active		0	Yes
4	Active	Active	Active	Down		1	Yes
4	Active	Active	Down	Down		2	No
4	Active	Down	Down	Down		3	No
5	Active	Active	Active	Active	Active	0	Yes
5	Active	Active	Active	Active	Down	1	Yes
5	Active	Active	Active	Down	Down	2	Yes
5	Active	Active	Down	Down	Down	3	No
5	Active	Down	Down	Down	Down	4	No

- IBM Cloud Pak for Integration to have high availability, the associated common services instance (which includes logging, licensing, and related services) must also be highly available.

Configuring IBM Cloud Platform Common Services in HA

Node Types and Common Services

A kubernetes **Node** is a VM or bare metal machine which is part of a Kubernetes cluster.

OpenShift Nodes

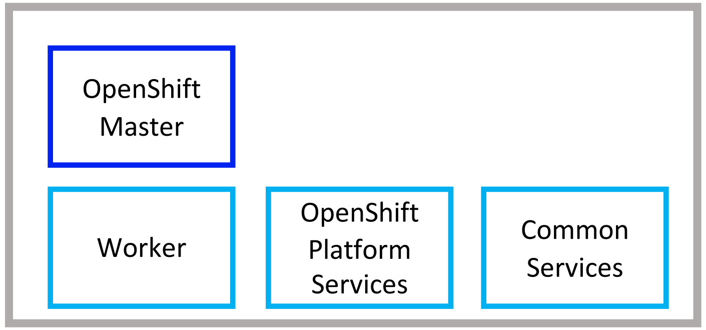
- **Master Nodes** run the services that control the cluster including the *etcd* database that stores the current state of the cluster.
- **Worker Nodes** run OpenShift Platform Services, Common Services, and Integration Services. If the IBM Cloud Pak for Integration is running in a cluster with other workloads then these also run on the worker nodes.

Common Services

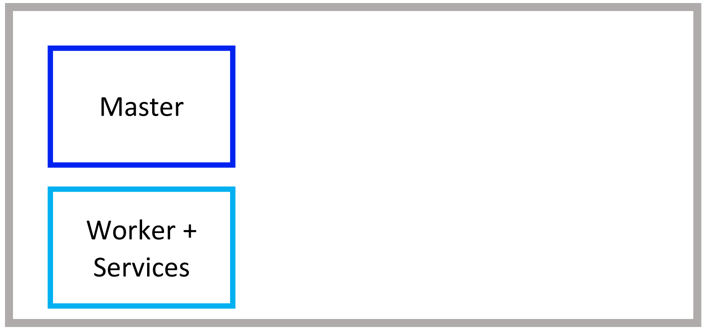
The common services run on OpenShift worker nodes. They are grouped into Master, Proxy, and Management services, and can all be run on the same node, run on separate nodes, and run on multiple nodes for HA.

- **Master services** IAM, Catalog, helm/tiller
- **Proxy services** ingress controller for app compatibility with OpenShift
- **Management services** Metering, Logging, Monitoring

The minimum *theoretical* configuration is one Master node and one Worker node. This would not be Highly available and is unlikely to have enough CPU to be usable for more than limited demos.



IBM Cloud Pak for Integration
Cluster Node types



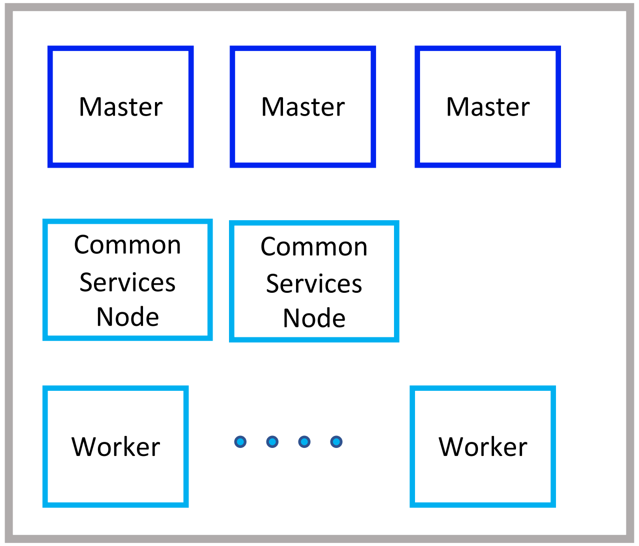
IBM Cloud Pak for Integration
Theoretical minimum cluster

Note: There are other more obscure types such as dedicated etcd nodes and vulnerability advisor nodes services but these are beyond the scope of this presentation

IBM CONFIDENTIAL – INTERNAL AND BUSINESS PARTNER USE ONLY

Platform and Services Availability

- To make the solution fully Highly Available, each component must be deployed in HA topology.
- Master Nodes contain software that uses a **quorum*** paradigm for high availability and so these must be deployed as an odd number of nodes. Typically either **3** or **5** masters are used in a HA cluster depending on the size of the cluster and the type of load.
- Common services do not require a quorum so each group of common services needs to be assigned to 2 or more worker nodes for HA.
- Worker Nodes run Integration Services. Depending on the integration services required 2 or more worker nodes may be needed for HA. (More detail in subsequent slides)
- A topology that is often used is:
    - 3 Master nodes
    - 2 Worker nodes for Common Services
    - 3(+) Worker nodes for Integration Services



Master

Master

Master

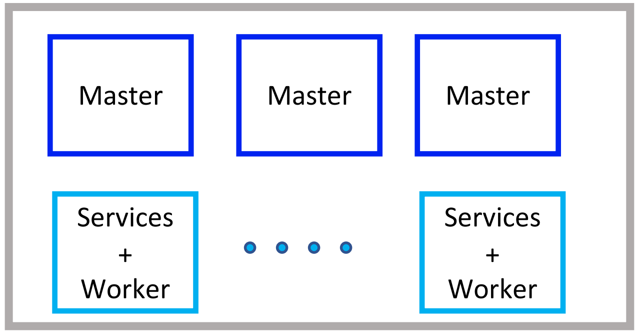
Common Services Node

Common Services Node

Worker

Worker

IBM Cloud Pak for Integration HA Cluster



Master

Master

Master

Services + Worker

Services + Worker

IBM Cloud Pak for Integration minimal HA Cluster

High Availability for Integration services.

Integration Services run on the **worker nodes**.

The **Cloud Pak for Integration** is composed from the **Certified Container Software** of the component products.

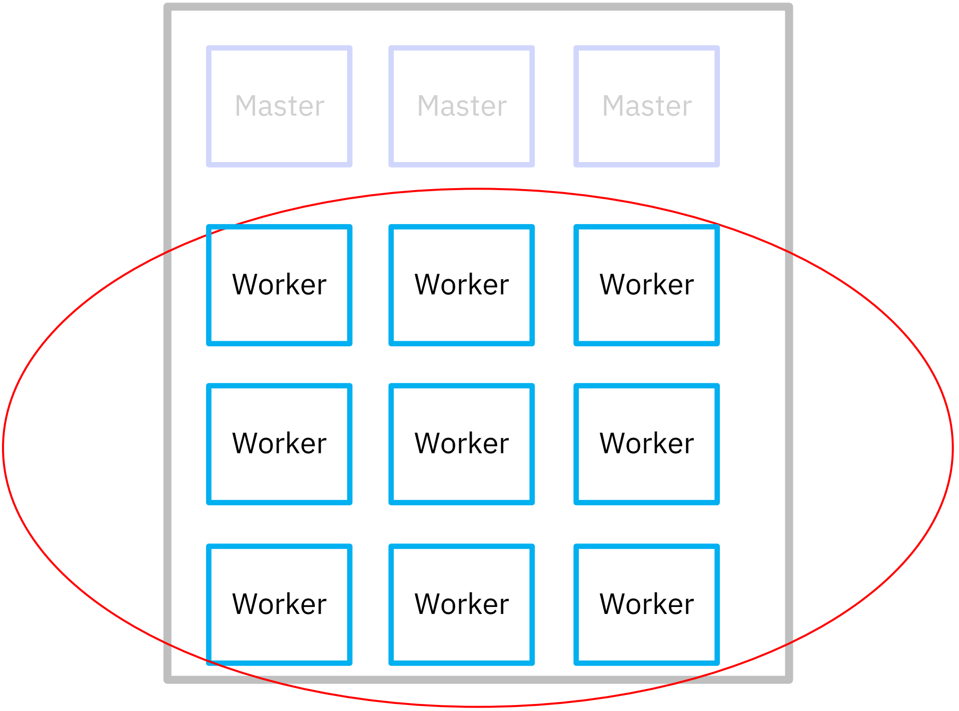
IBM Certified Container Software is developed by the product development teams to embody best practice for deploying the product onto Kubernetes as a secure, scalable **Highly Available** deployment.

Thus, at a high level, all that has to be ensured to provide a highly available deployment is:

- There are sufficient nodes for the solution *
- The nodes are deployed into separate failure domains so that a single failure will not take out multiple nodes. **

Kubernetes takes care of:

- Running appropriate numbers of instances of the Cloud Pak, spread across the available workers.
- Mixing together the workloads on the available worker nodes taking account the resources they need.
- Restarting workloads that fail and recovering from failed nodes by scheduling workloads onto alternative nodes.



IBM Cloud Pak for Integration HA Cluster

* Some of the component products require 2 worker nodes for active/standby style HA and others require 3 or more workers for quorum style deployment.

** In larger clusters the nodes should be spread between 3 ore more availability zones.

IBM CONFIDENTIAL – INTERNAL AND BUSINESS PARTNER USE ONLY

System requirements for high availability environments

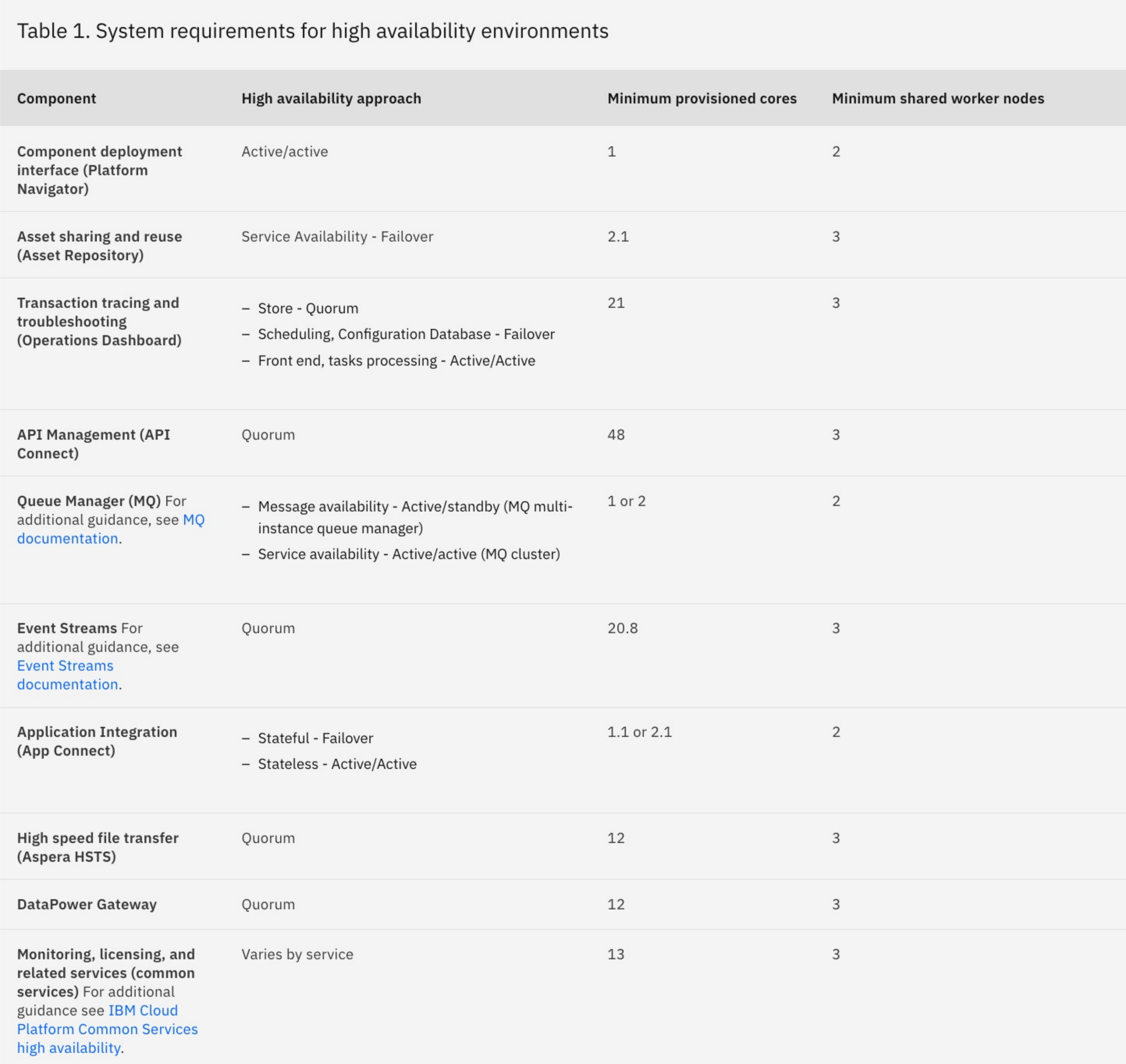


Table 1. System requirements for high availability environments

Component	High availability approach	Minimum provisioned cores	Minimum shared worker nodes
Component deployment interface (Platform Navigator)	Active/active	1	2
Asset sharing and reuse (Asset Repository)	Service Availability - Failover	2.1	3
Transaction tracing and troubleshooting (Operations Dashboard)	– Store - Quorum – Scheduling, Configuration Database - Failover – Front end, tasks processing - Active/Active	21	3
API Management (API Connect)	Quorum	48	3
Queue Manager (MQ) For additional guidance, see MQ documentation.	– Message availability - Active/standby (MQ multi-instance queue manager) – Service availability - Active/active (MQ cluster)	1 or 2	2
Event Streams For additional guidance, see Event Streams documentation.	Quorum	20.8	3
Application Integration (App Connect)	– Stateful - Failover – Stateless - Active/Active	1.1 or 2.1	2
High speed file transfer (Aspera HSTS)	Quorum	12	3
DataPower Gateway	Quorum	12	3
Monitoring, licensing, and related services (common services) For additional guidance see IBM Cloud Platform Common Services high availability.	Varies by service	13	3

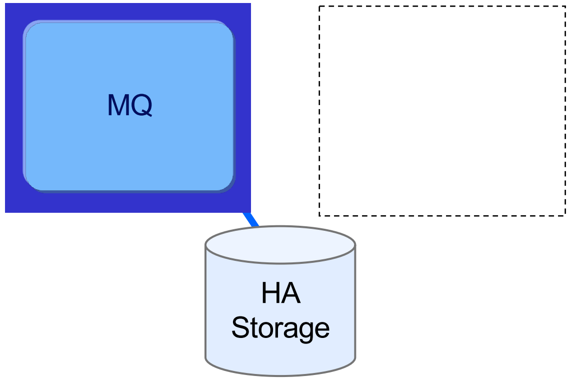
CP4I Components - HA

- MQ High Availability
- Event Streams High Availability
- API Connect High Availability
- App Connect High Availability

Foundational: Containerization

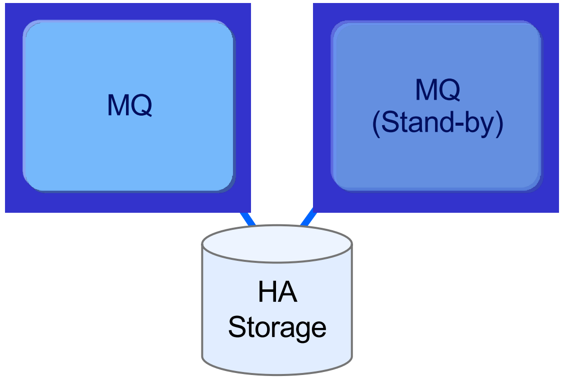
IBM MQ High Availability on a container platform

Single resilient queue manager



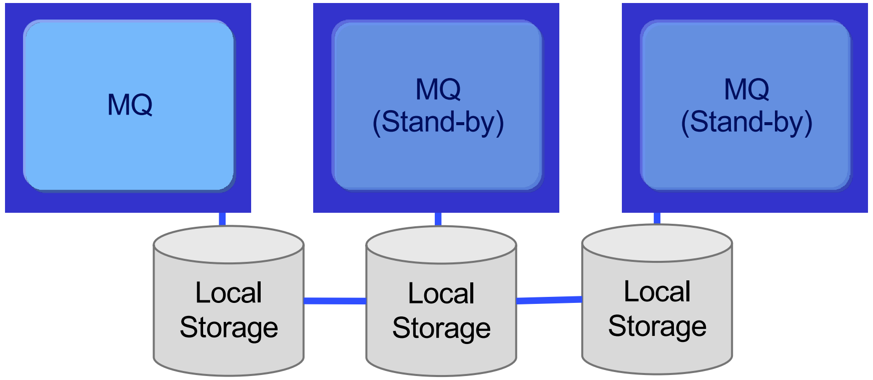
- Platform controlled resilience
- Resilience monitoring required outside of the platform

Multi-instance queue manager



- Removes edge conditions around a node failure.
- Requires shared filesystem with lease locking

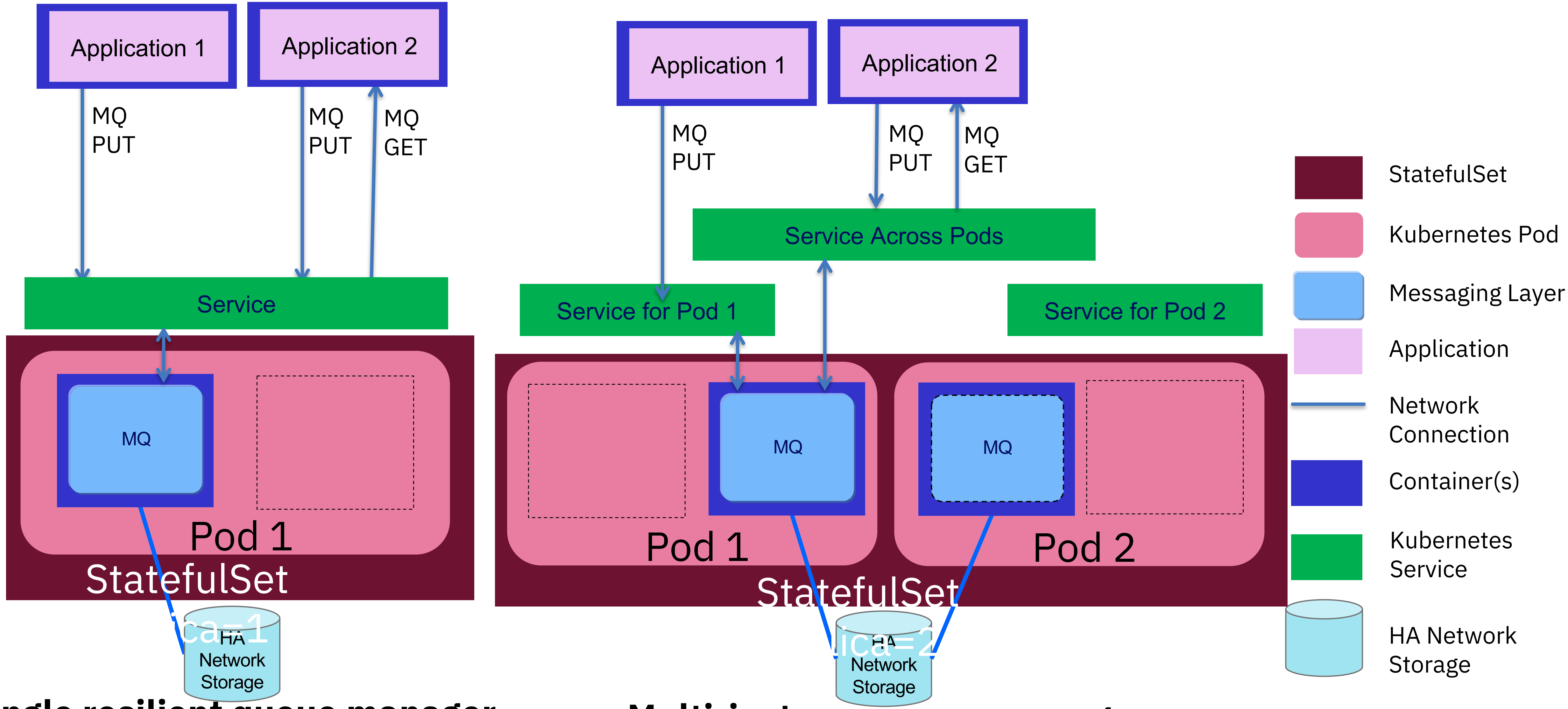
Native HA Replicated data queue manager



- Native HA

Foundational: Containerization

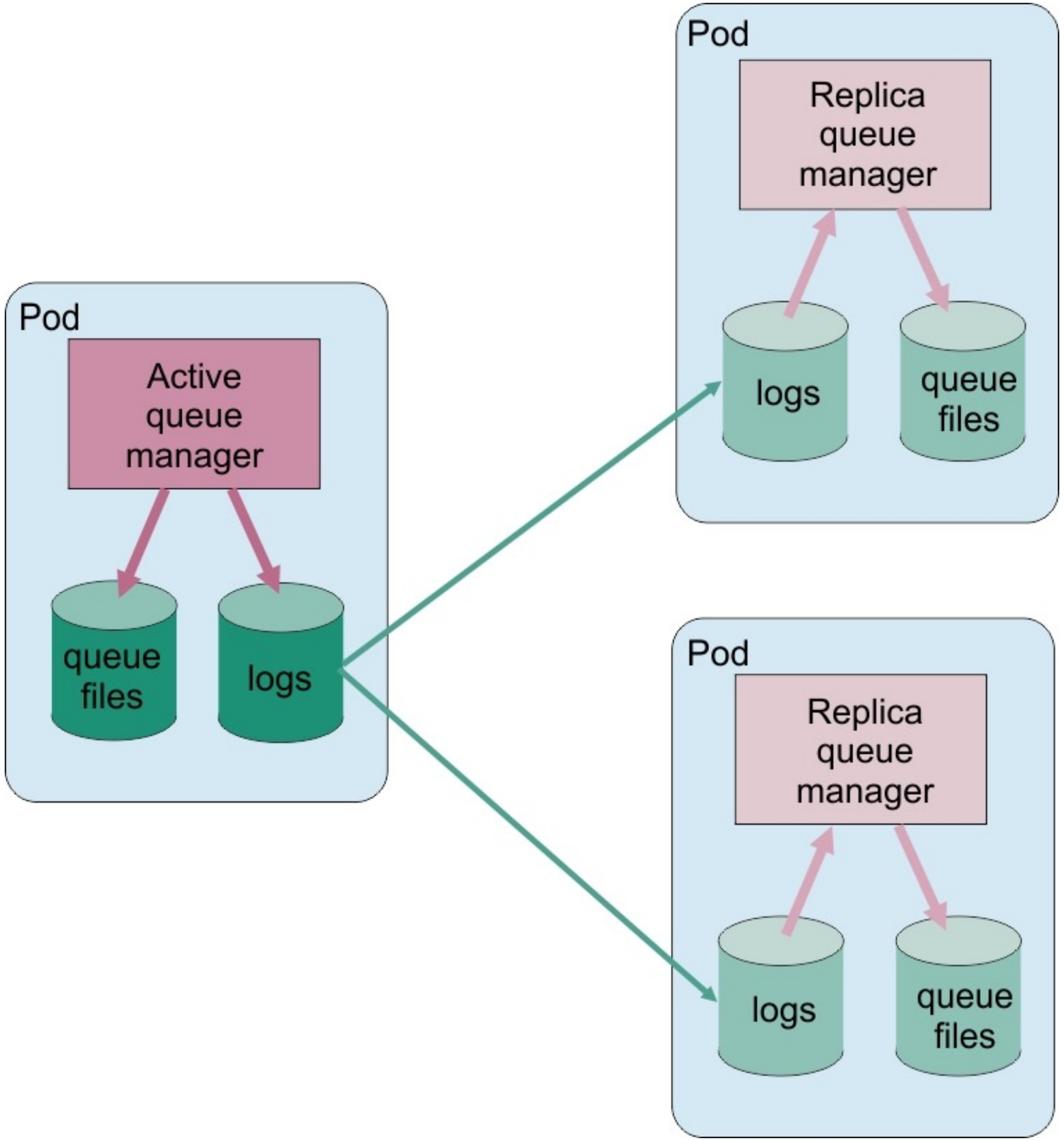
Diving into the MQ HA topologies



Single resilient queue manager

Multi-instance queue manager

Native HA configuration



A Kubernetes Service is used to route TCP/IP client connections to the current active instance, which is identified as being the only pod which is ready for network traffic. This happens without the need for the client application to be aware of the different instances.

A queue manager can only become the active instance on a pod that has quorum. The queue manager cannot become active on a pod that is not connected to at least one other pod, so there can never be two active instances at the same time

If an active pod fails, and subsequently recovers, it can rejoin the group in a replica role.

Native HA is a capability preview in IBM Cloud Pak® for Integration 2021.1.1, and is suitable for evaluation purposes only. For more information, see [Evaluating the Native HA feature in IBM Cloud Pak for Integration 2021.1.1]()

Event Streams High Availability

Event streams deploys Apache Kafka in a HA topology by default.

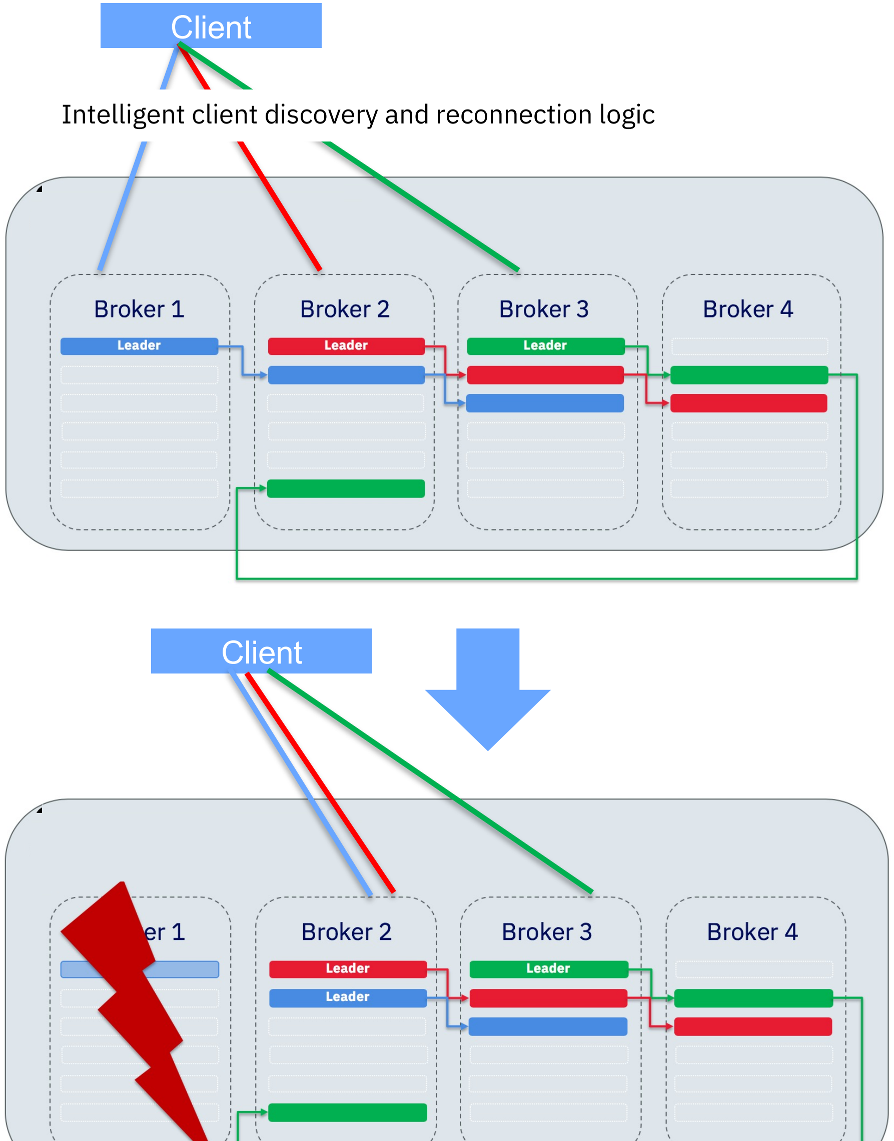
Chart deploys brokers as a stateful set with 3 members by default.

Kubernetes will schedule the pods to different nodes.

Kafka’s architecture is inherently HA. Client connection protocol handles discovery and failure.

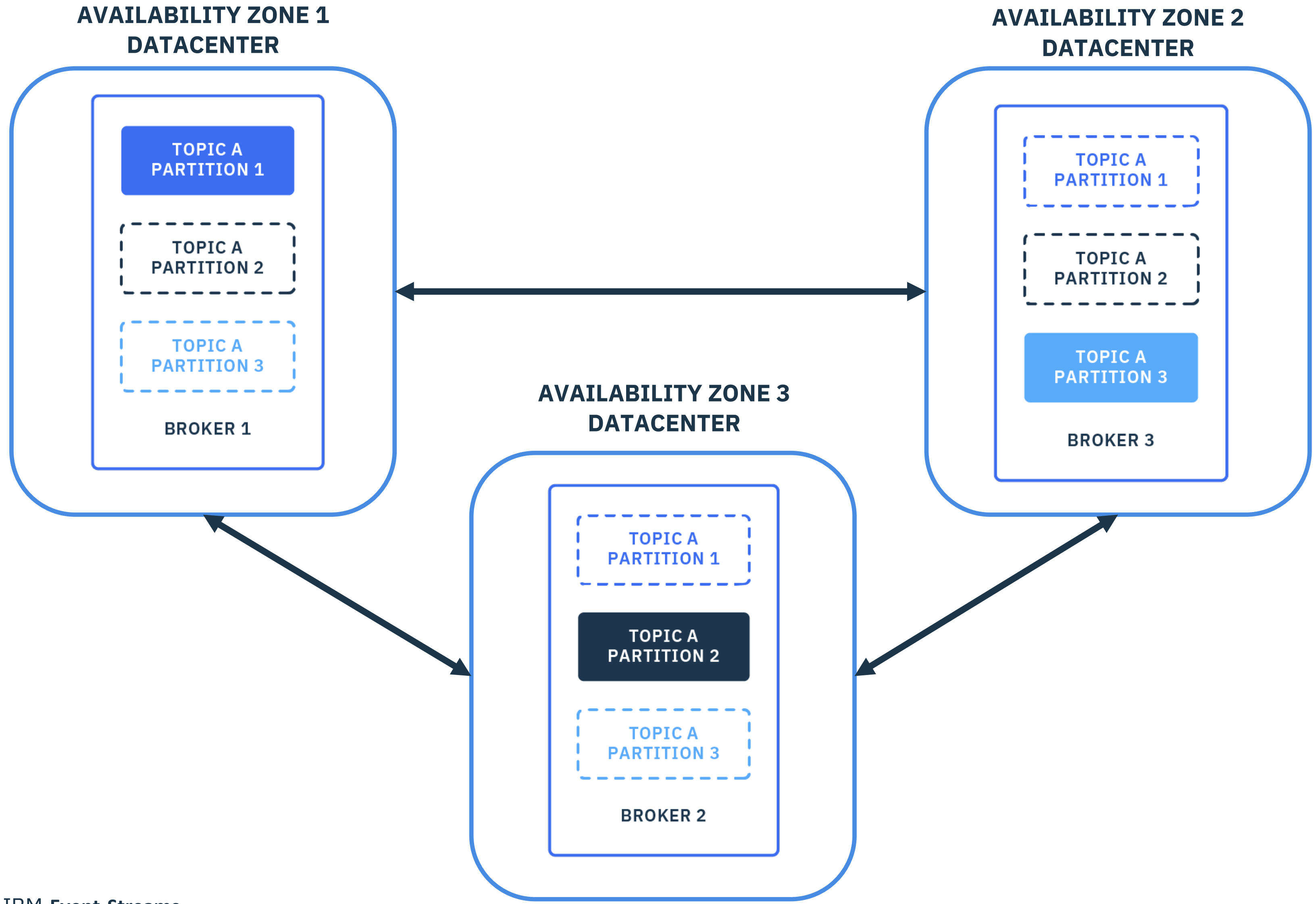
Default for Topics creation is 3 replicas with min in-sync copies=2

Event Streams can be deployed into Multiple AZs and will distribute replicas across AZs even if there are multiple brokers per AZ. (New in 2019.4.1)


Intelligent client discovery and reconnection logic

Event Streams	12[note3]	12[note3]	20.8[note3]	20.8[note3]	At least 3 shared Worker Nodes

Enhanced resilience with clusters across multiple zones



Multi-Availability Zone

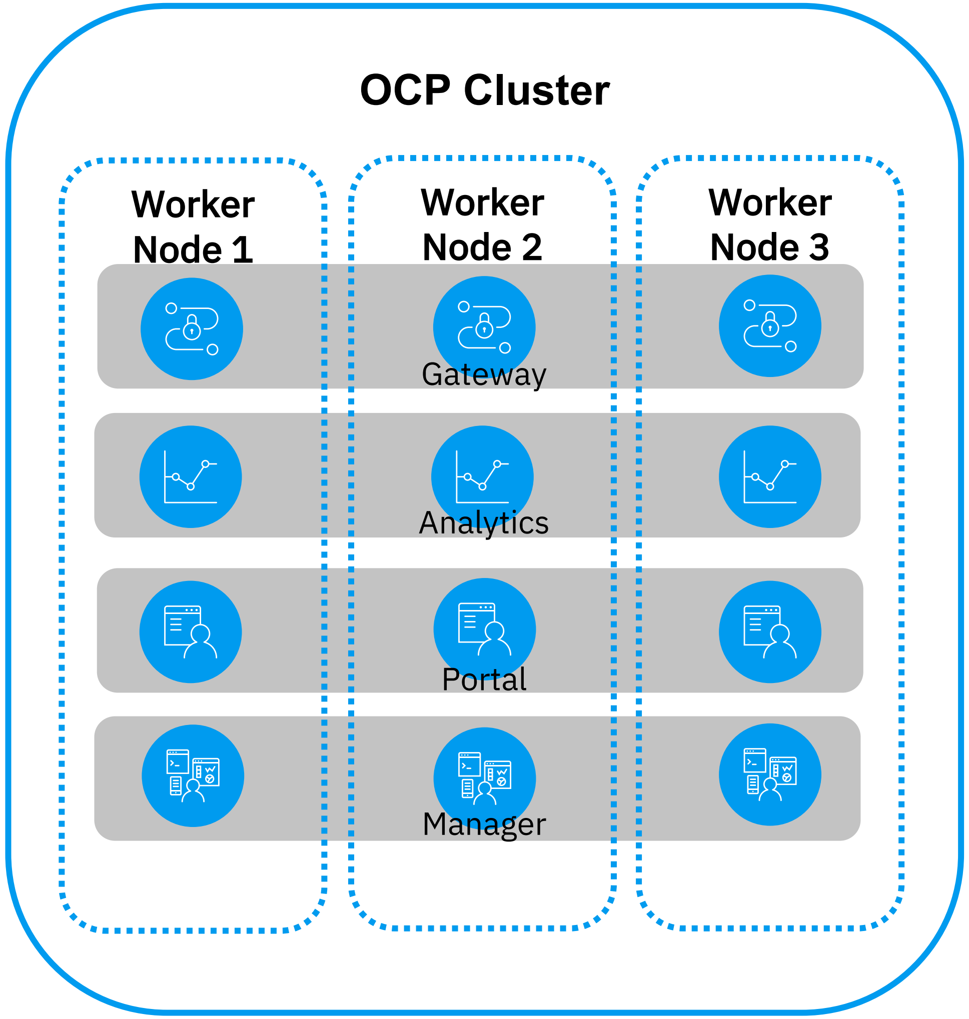
- Must have at least 3 zones
- Kafka brokers and ZooKeeper servers span across zones
- Can tolerate failure of a zone with no service degradation
- High-speed network with low latency between zones required (< 20ms)

API Connect High Availability

- API Connect requires 3 Worker nodes for HA.
- Chart deploys as HA by default:
  - global setting: mode = standard (dev mode is non HA)
  - Cassandra cluster size = 3
  - Gateway replica count = 3
- Kubernetes distributes the resulting resources across 3 (or more) nodes on the cluster for high availability.

API Connect whitepaper discusses multi-cluster, multi data centre and DR topics.

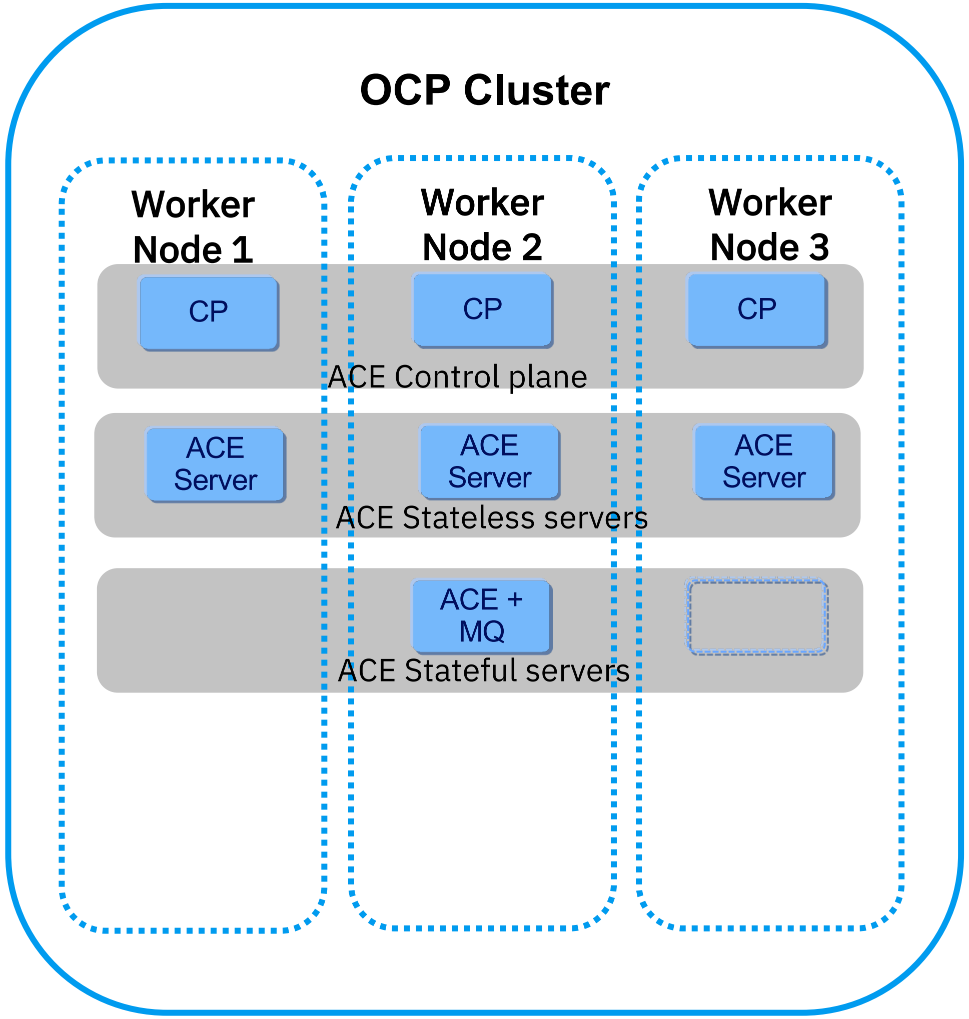
http://ibm.biz/APIC2018paper



IBM CONFIDENTIAL – INTERNAL AND BUSINESS PARTNER USE ONLY

App Connect High Availability

- ACE Control plane is HA (replicaset of 3) by default.
- Integration servers deployed without local MQ Queue Managers (QM) are stateless.
  - Deployed HA (replicaset of 3) by default.
  - BAR file is retrieved from control plane at startup.
- Integration servers deployed with a local QM are deployed like MQ.
  - Single Resilient Queue Manager (stateful set of 1).



IBM CONFIDENTIAL – INTERNAL AND BUSINESS PARTNER USE ONLY

Summary

Overview of IBM Cloud Pak for Integration Architecture

OpenShift High Availability considerations

Deployment considerations

Platform and Services Availability

MQ High Availability

Event Streams High Availability

API Connect High Availability

App Connect High Availability

IBM