Containerizing Integration workloads on ICP

Sandeep Chellingi
Prolifics

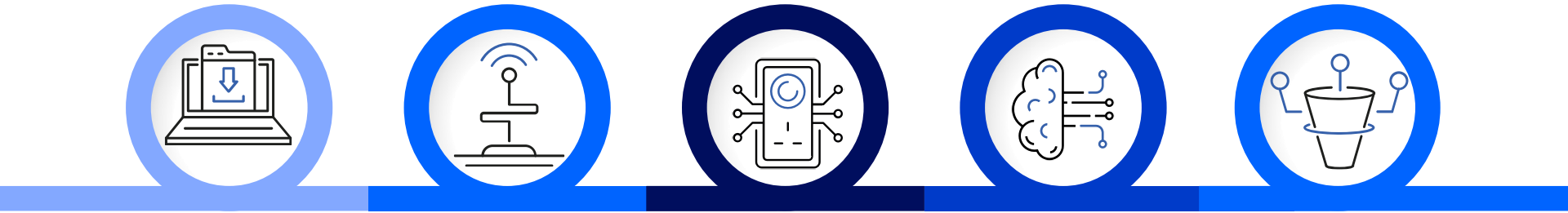
IBM **Cloud Integration**

To drive digital transformation,

organizations must tap into an ever-growing set of applications, processes and information sources across multiple clouds and on-premise…

… all which significantly expand the enterprise's ***need for modern integration capabilities***

The need for Integration increases



SaaS Adoption

IoT

Process Automation

AI

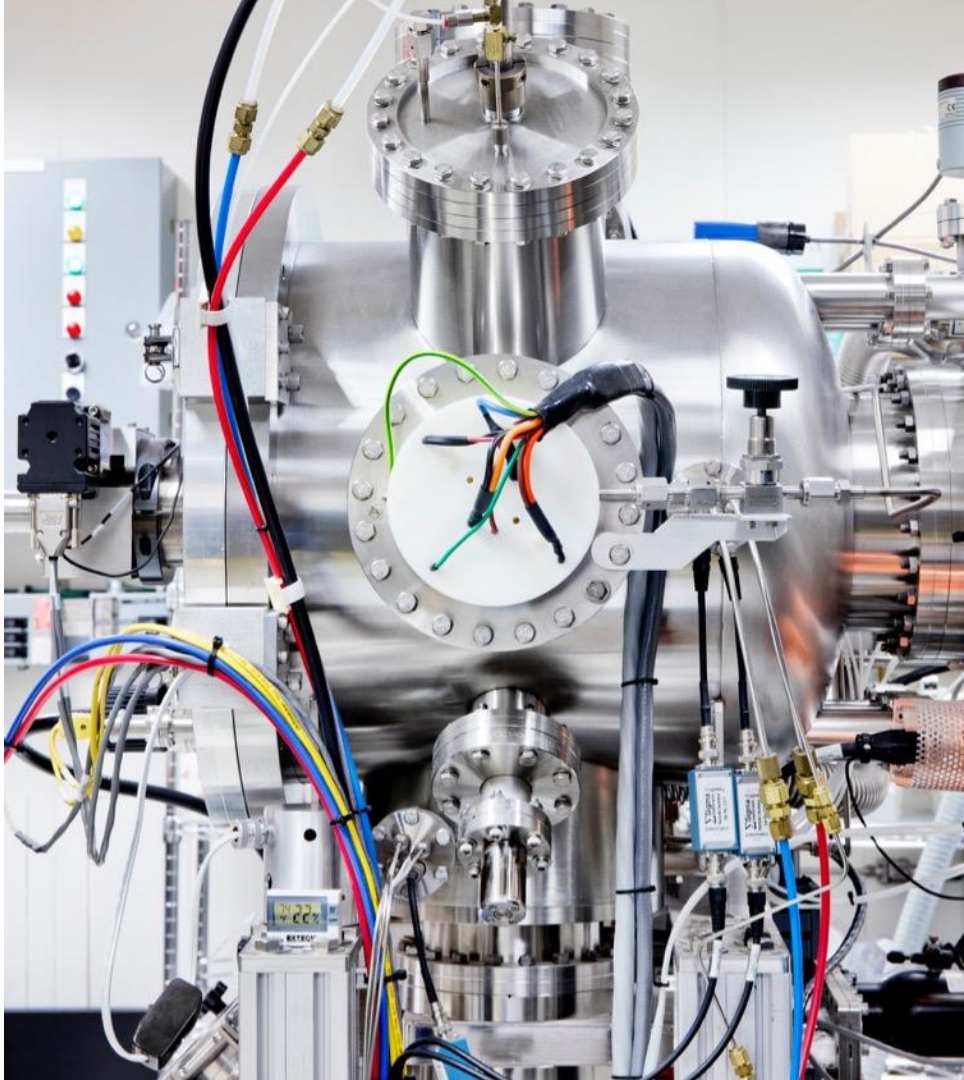
Micro services Design

Integration has become an obstacle to success because traditional, centralized and systematic integration approaches cannot cope with the volume and pace of business innovation.

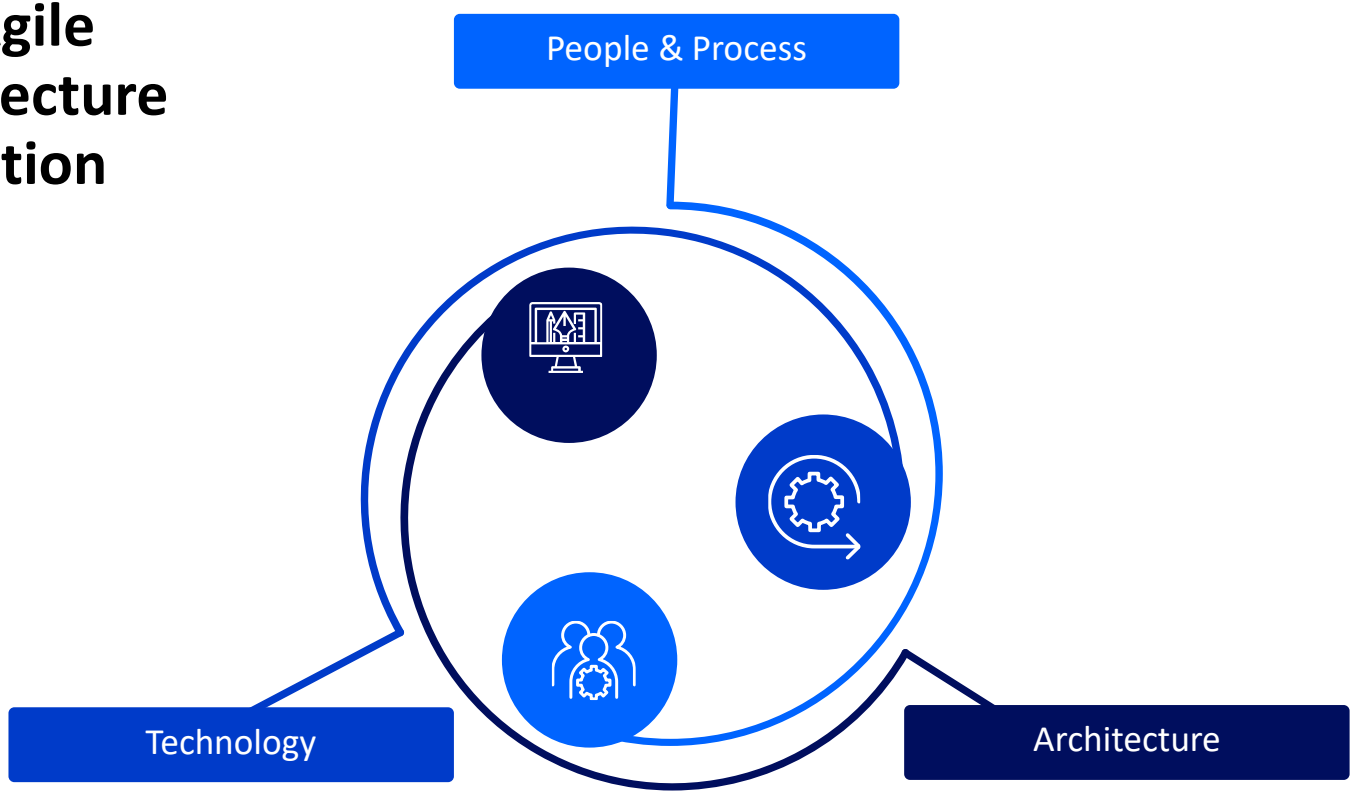
Source: Gartner-Modernizing Integration Strategies and Infrastructure Primer for 2018

How Integration has Changed

- Need for integration outside the data center
- Need to drive speed and efficiency in integration development while lowering cost
- Many different integration requirements that include traditional and modern integration



IBM can help you modernize with Agile Integration Architecture and Cloud Integration Platform



Source: Gartner-Modernizing Integration Strategies and Infrastructure Primer for 2018

Agile integration allows you to invest strategically

People & Process

Do less of this
Centrally controlled / governed limits innovation

Core IT

- Core IT applied to all challenges
- Limited resource pool
- Expensive resource
- Centralized IT and removed from day to day operations
- High-skilled but single threaded
- Lack business expertise

Do more of this
Innovate faster and closer to your line of business

Distributed

- Core IT now applied to only hardest problems
- Much larger resource pool
- Cost effective
- Decentralized and closer to the business decision makers
- Simpler tools provide success at lower skill level
- Are the business experts

Source: “A Developer's Guide To Forrester's Strategies For Integration And Digital Business Platforms” Forrester report.

Agile Integration Architecture

	Fine grained deployment	Decentralized Ownership	Cloud native infrastructure
	Improve build independence and production velocity	*Accelerate agility and innovation*	*Dynamic scalability and inherent resilience*
Application	Dependency free rapid integration delivery	Business autonomy for integration delivery	Scale and administer integrations with applications that live anywhere
API	Consumer centric exposure of business APIs	Self-administration of API exposure and subscription	Multi-platform cloud agnostic API management componentry
Messaging	Independent application centric messaging	Self-provisioning of messaging and event capabilities	Cloud scale inherently resilient multi-platform messaging

IBM's Modern Integration Platform

Technology

Connect to Data and Services anywhere

Expose as APIs, Microservices and Events

Build new integrations

Leverage existing integrations

Applications | Data Stores | Messaging | Event Streams | Data Streams

Management | Security | AI

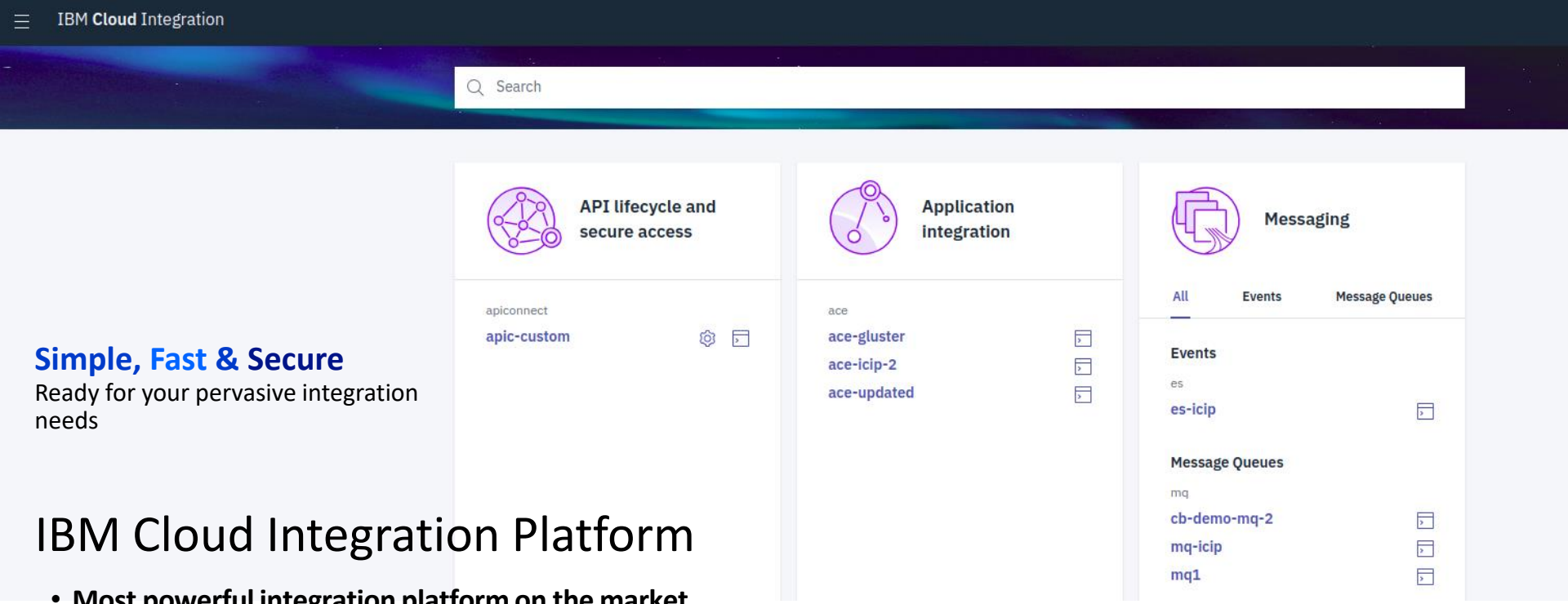
IBM Multicloud services (provided by IBM Cloud Private)

Deploy to targeted audiences enabling easy consumption

Public, Private, and SaaS

What is the IBM Cloud Integration Platform?

- IBM Cloud Private 3.1.1, with pre-installed Helm charts for:
  - IBM API Connect v2018
  - IBM App Connect Enterprise v11
  - IBM MQ Advanced v9
  - IBM Event Streams v2018
  - IBM Aspera HSTS v1
- Platform navigator, bringing all instances together in a single unified experience



Simple, Fast & Secure
Ready for your pervasive integration needs

IBM Cloud Integration Platform

- **Most powerful integration platform on the market**
  NEW offering incorporating traditional and modern integration including APIs, App Integration, Message queuing, Event streams and Fast file transfer
- **Deploy wherever needed**
  Supports deployment on-premise or in any cloud
- **Enterprise grade**
  Secure, scalable modern architecture

Industry leading capabilities to accelerate business value

Simple

Fast

Secure

API Lifecycle
Unlock business data and assets as APIs

Data Integration
Understand, cleanse, transform and deliver quality data

Application Integration
Connect your cloud and on-prem applications

High Speed Transfer
Super fast & secure data transport across any cloud

Messaging & Events
Deliver msgs reliably with enterprise-grade messaging

Secure Access
Control access to vital resources wherever they are

Cloud Integration Platform Architecture

Cloud Integration Platform (CIP) deploys and manages instances of **Integration Services** running on a Kubernetes Infrastructure

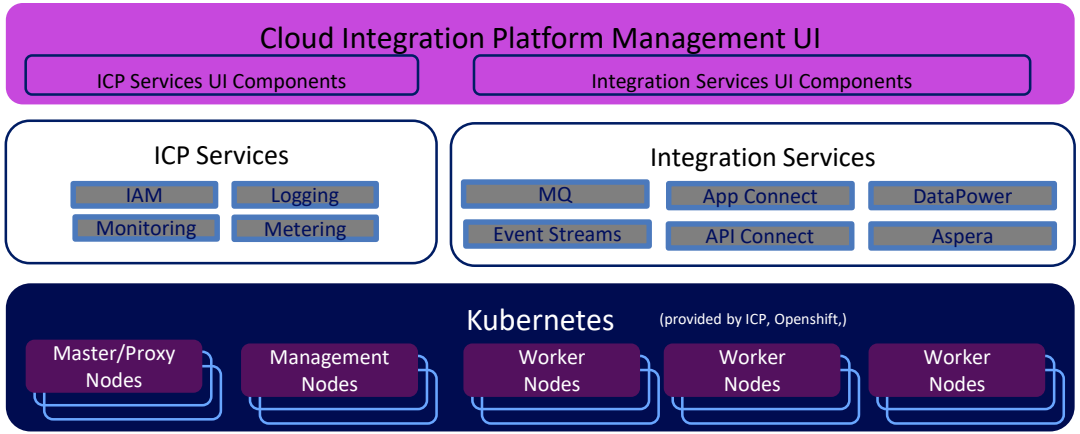
Instances of Integration services are deployed individually as needed to satisfy each use-case.

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**.

CIP Leverages the IBM Cloud Private **(ICP) services** which run on dedicated master/proxy and Management nodes in HA or non HA configurations.

The Management UI unifies the management UIs of the Integration Services and the ICP services.



Node Types

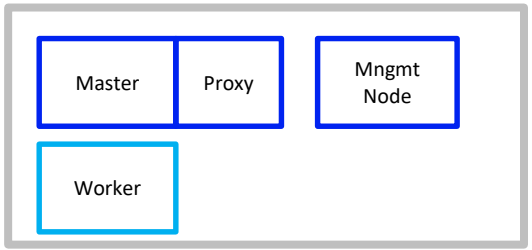
A kubernetes **Node** is a VM or bare metal machine which is part of a Kubernetes cluster.

Master Nodes run the services that control the cluster including the *etcd* database that stores the current state of the cluster.

Proxy Nodes transmit external requests to the services created inside your cluster.

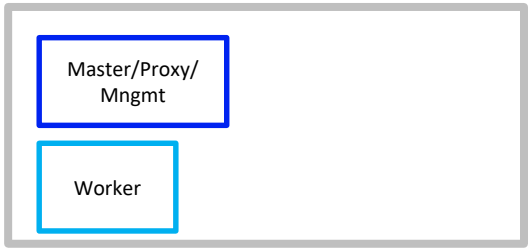
Management Nodes host management services such as monitoring, metering, and logging.

Worker Nodes run Integration Services. If Cloud Integration Platform is running in a cluster with other workloads then these also run on the worker nodes.



Cloud Integration Platform
Cluster Node types

Master, Proxy and Management node types can be combined together but it is recommended in production clusters to keep them separate so that excessive workload does not impact stability.



Cloud Integration Platform
Theoretical minimum cluster

The minimum *theoretical* configuration is therefore one Master/Proxy/Management 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.

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

Container Platform Availability Considerations

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.

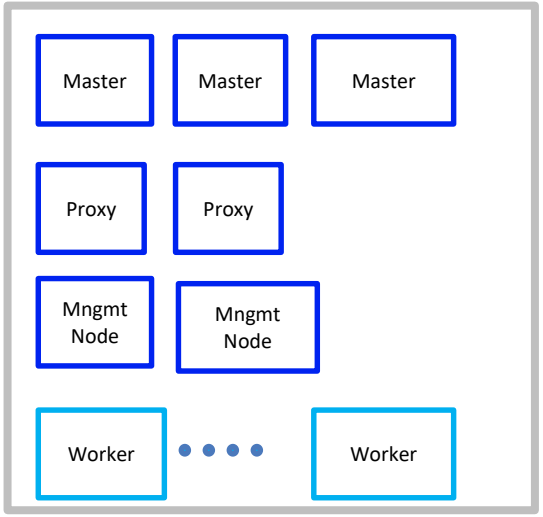
Proxy Nodes do not require a quorum so 2 or more Proxy nodes are needed for HA

Management Nodes do not require a quorum so 2 or more Proxy nodes are needed for HA

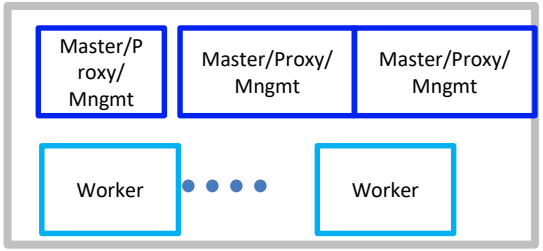
Worker Nodes run Integration Services. Depending on the integration services required 2 or more worker nodes may be needed for HA. (More detail in subsequent slides)

As previously noted of course Master/Proxy/Management nodes can be combined so a minimal configuration could be 3 Master/Proxy/Management nodes.

A topology that is often used is: 3 Master/Proxies + 2 Management + 3 Workers.



Cloud Integration Platform HA Cluster



Cloud Integration Platform minimal HA Cluster

Deployment Considerations

- **How many independent clusters do you need?**
  - Enterprises deploy multiple clusters for any/all of the following reasons:
    - Geographic availability – Ability to survive a regional outage
    - To separate organizations
    - To separate development, test and production.
    - To guard against errors by individual operators.
- **Are the Kubernetes nodes deployed into separate failure domains?**
  - Separate Availability zones in a public cloud.
  - Separate physical servers/racks in a data centre
    - What are the common points of failure?

High Availability for Integration services.

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

The Cloud Integration **SolutionPak** is composed from the **CloudPaks** making from the component products.

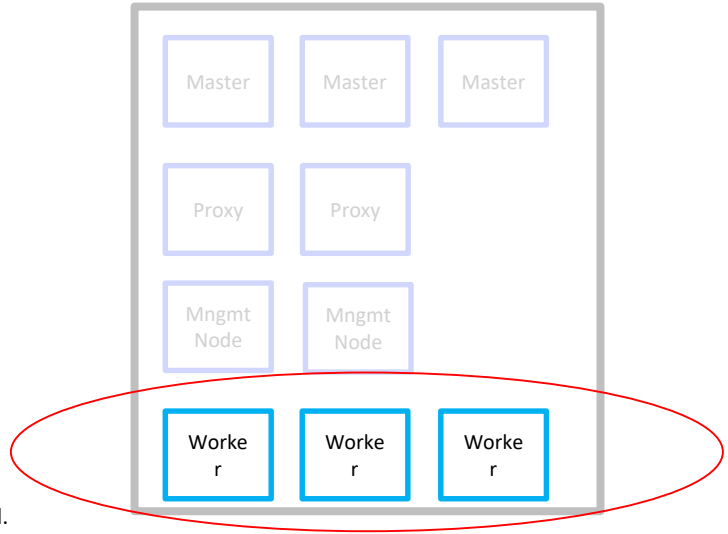
IBM cloudpaks are developed by the product development teams to embody best practice for deploying the product onto kubernetes as a secure, scalable **Highly Available** deployment.

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 solutionpak, 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.



Cloud Integration Platform HA Cluster

* Some of the component products require 2 worker nodes for active/standby syle 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.

Summary - High Availability for Integration.

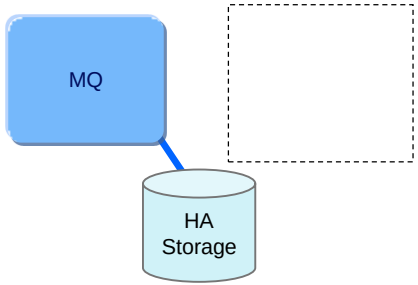
At a high level. All you need to do is deploy the cluster in a HA configuration with at least 3 workers.

Deploy the MQ services from the helm charts and they will be highly available by default.

Product	Approach(es) to HA	Minimum number of worker nodes in the cluster
MQ	Data Availability – Failover Service availability – Active/Active	2
Event Streams	Quorum	3
APIC	Quorum	3
App Connect	Stateless – Active/active Stateful – Failover	2
Aspera	Quorum	3
Datapower	Quorum	3

IBM MQ High Availability in Cloud Pak / CIP

Single resilient queue manager

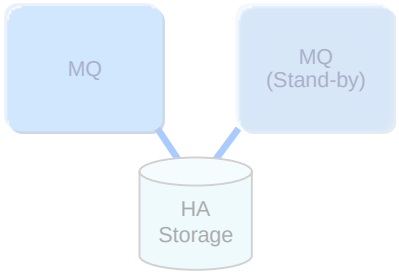


Platform Load Balancing

Platform HA

Current Approach

Multi-instance queue manager

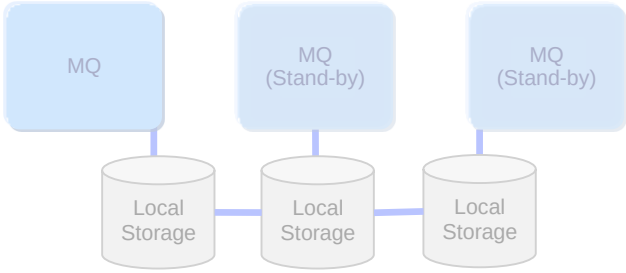


Client load balancing

IBM MQ Product HA

Investigating use of multi-instance within Kubernetes

Replicated data queue manager

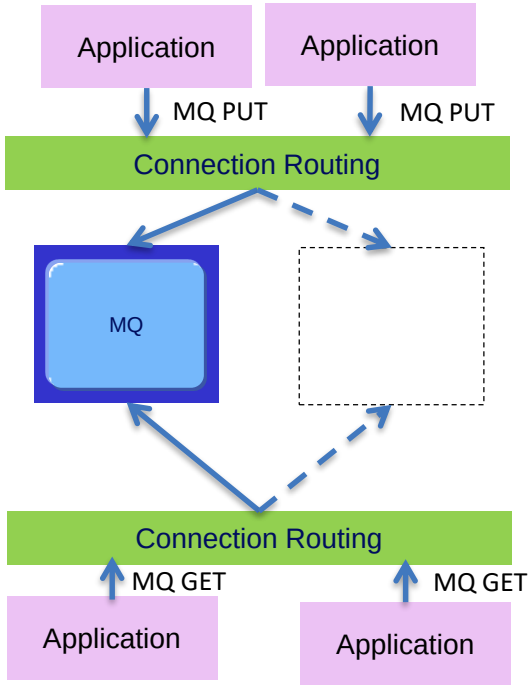


Client or Server load balancing

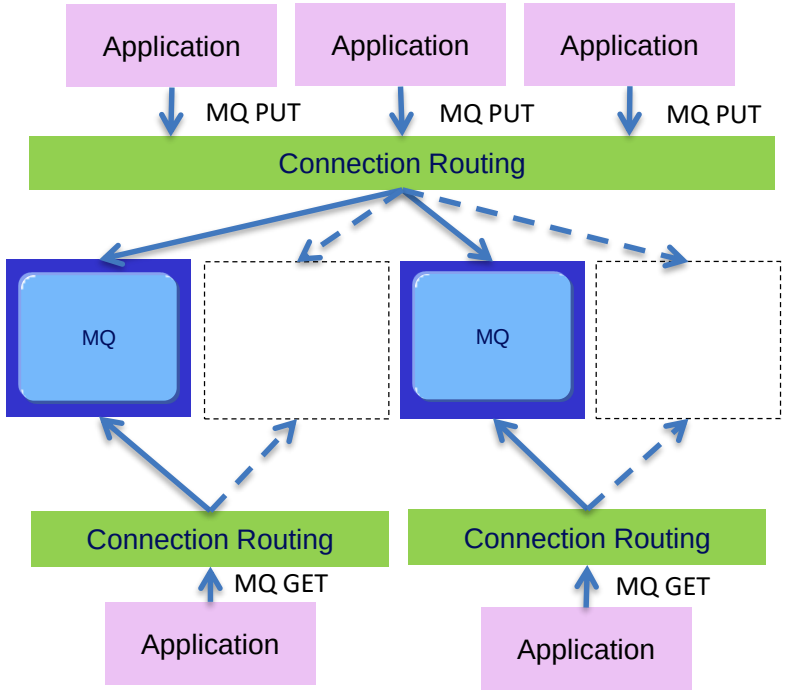
IBM MQ Product HA

Technology not supported in Kubernetes

Improving the MQ Service availability



Single resilient container



Multiple resilient containers

- ✓ Provide access to store messages, even in failover situations. Connection Routing provided to distribute the traffic across the available instances.
- ✓ Applications retrieving messages attach to the individual Queue Manager. Connection Routing provided to route traffic to container location.

Event Streams High Availability

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

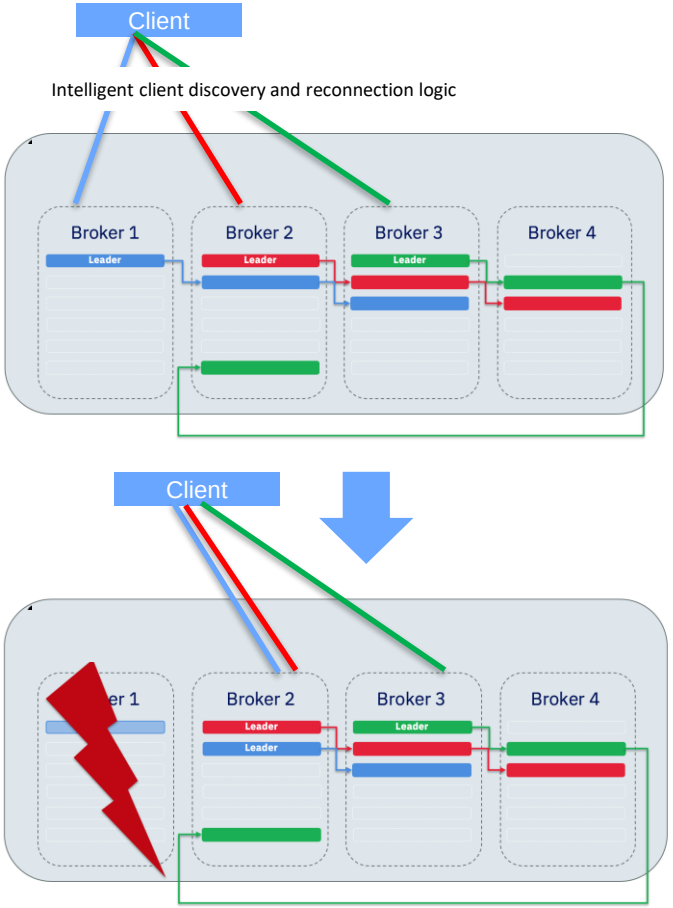
Chart deploys brokers as a stateful set with 3 members by default.

Kubernetes will schedule the pods to different nodes.

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

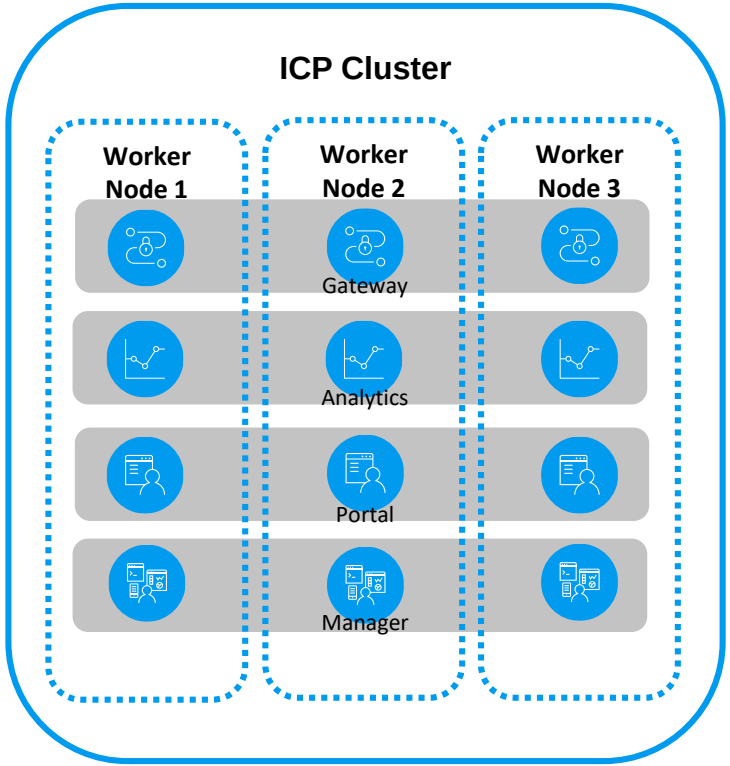
Default for Topics creation is 3 replicas with min in-sync copies=2

For Multi-AZ deployments number of brokers must match number of AZs today.



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

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).