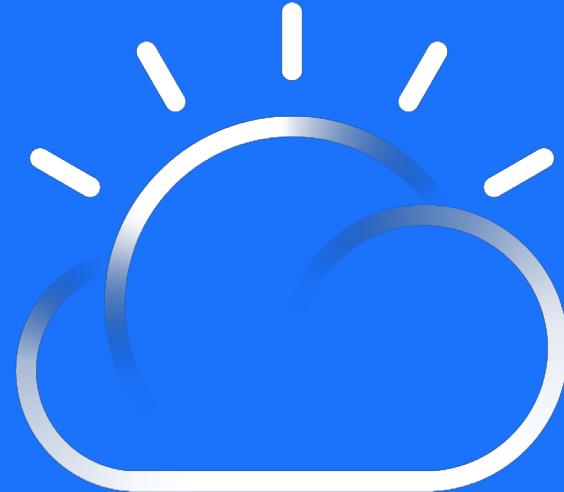


API Connect Deployment and Migration Best Practices



IBM Cloud

IBM

Please note

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice and 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. Information about potential future products may not be incorporated into any contract.

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

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including 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 results similar to those stated here.

Please note: Use API Connect Deployment white paper for detailed aspect of each option

Agenda

- **API Connect v2018 Overview**
- API Connect – Non-High Availability
- API Connect – High Availability
 - Quorum Requirements
 - Active / Passive
 - Active / Active
- Sample Deployment

v2018: Scale & Flexibility with Multi-Cloud Design

API Connect

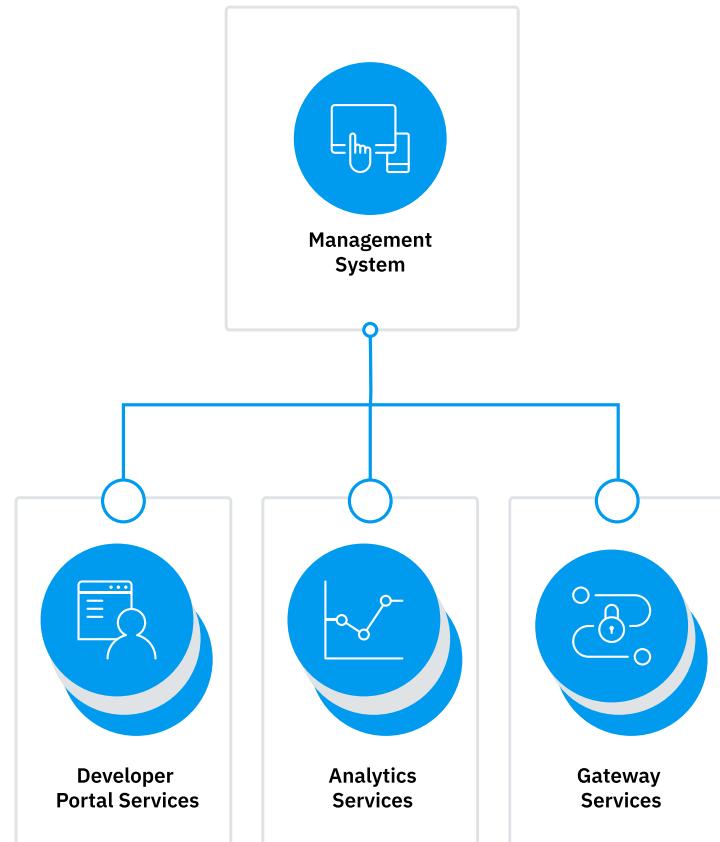
- **Unmatched scalability and performance** via new Microservices & container based design
- **Independently & automatically scale** API Connect 30+ microservices based on SLAs
- **DevOps ready** with 400+ REST APIs & CLIs for CI/CD
- **Run virtually anywhere** with agnostic Container/K8s & VMware support
- **Native OS developer toolkit** & enhanced UX to create APIs in minutes

API Gateway (DataPower)

- **Enterprise proven gateway** distributed, clustered & co-located with provider backends
- **Up to 5X performance** using new API Gateway service
- **Improved operational resiliency** during API manager maintenance & unavailability periods
- **Automatically scale gateways** without manual registration in API manager node
- **Autonomic gateway cluster management** to self-heal without manual intervention
- **Up to 2X Performance** with X2 appliance

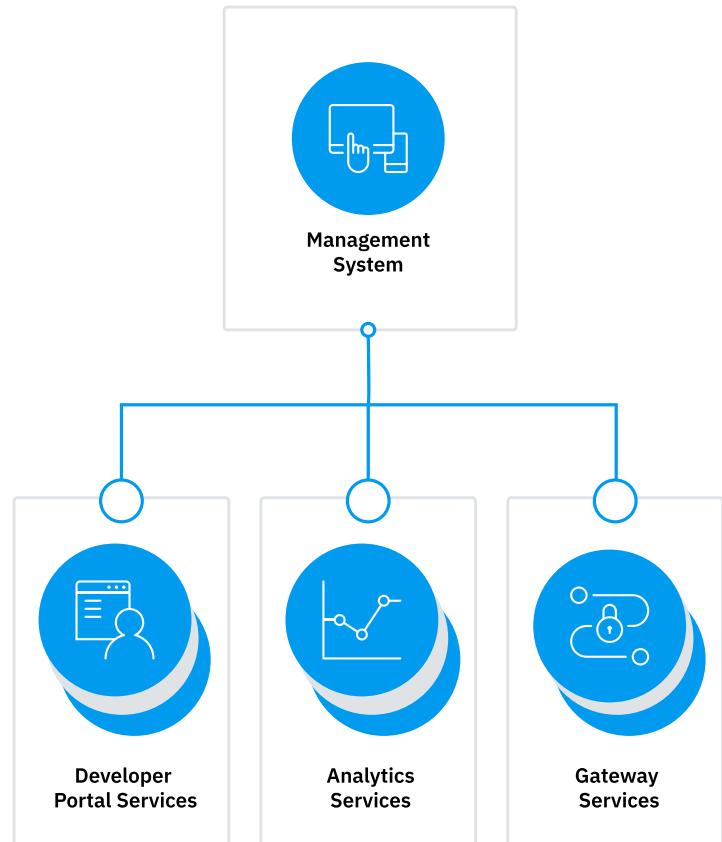
API Connect v2018 Components

- **Standalone API Analytics** component to scale independently based on API project growth
- **Zero to N portal clusters** can be configured to a API Connect deployment to align with API project growth
- **Native install of APIC toolkit** for enhanced user experience
- **V5 to V2018 Upgrade** through automated migration scripts with a parallel stack setup following modern software practices

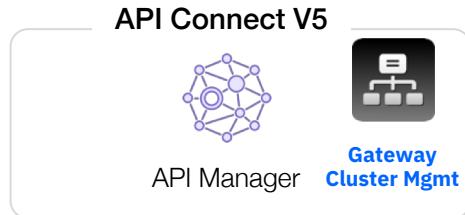


API Connect v2018 Components

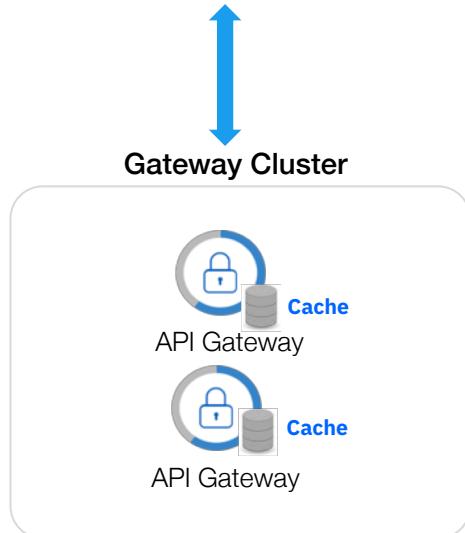
- Single Manager Cluster per API Connect Cloud, as it is the brain of the API Management system
- Manager can span multiple Availability Zones, giving flexibility in deployment scenarios
- Multiple Portal, Analytics and Gateway Cluster per Cloud, and are scoped to an Availability Zones
- API Connect Cloud defined as 1 APIM, with N Component Services, and most customers have 2+ Cloud environment (Development, Staging, Prod, etc.)



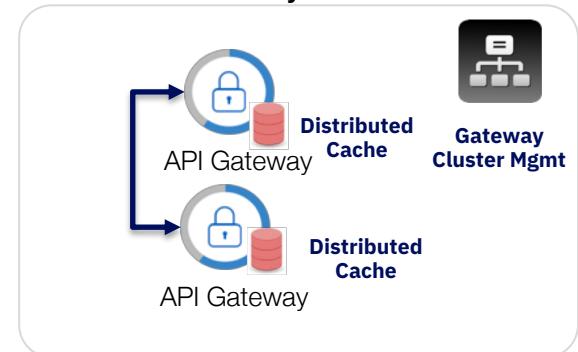
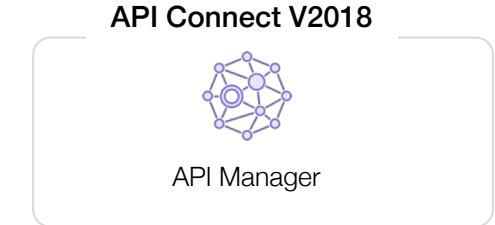
Multi-cloud scalable API Manager & Gateway Architecture



- **Gateway Cluster Management** moved from API Manager in V5 to API Gateway in V2018

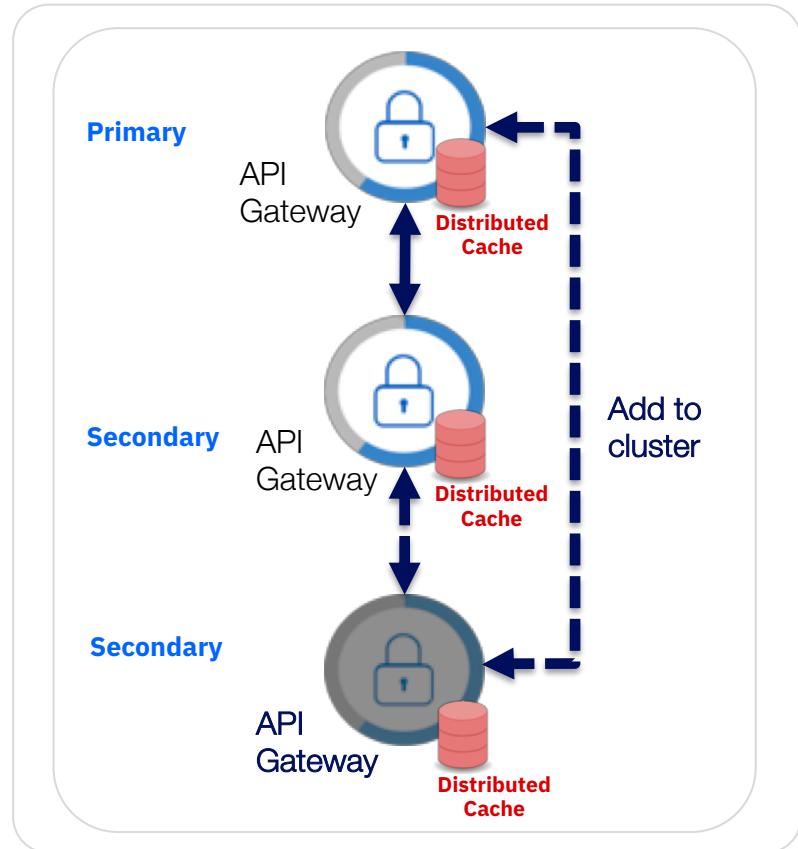


- **API Data Replication** performed between API Gateway instances in V2018, whereas in V5 it was between API Manager and each API Gateway instance
- **Reduces runtime dependency** between the API Gateway and API Manager in V2018



Dynamic & Autonomous Gateway Cluster Management

- Gateways can scale without manual registration in API manager node, unleashing exciting new topologies
- Autonomic Gateway cluster management will allow Gateways to self-heal without manual intervention
- Gateways contact peers within cluster to obtain runtime information
 - Rate Limit
 - API definitions
 - Subscriptions



APIC v2018: System Requirements & Minimum Configuration

VMware					Containers					
IBM Provides: OVA Images, InstallAssist					IBM Provides: Docker Images, InstallAssist					
Component	System Requirements	CPU Minimum	Memory Minimum (GB)	Disk Space Minimum (GB)		System Requirements	CPU Minimum (Worker#)	Memory Minimum (GB)	Disk Space Minimum (GB)	
API Manager	1. VMware ESXi 6.0 or ESXi 6.5*	4	16	250		1. Ubuntu 16.04	3	16	250	
DataPower API Gateway		4	8	32		2. Docker 17.03 or 18.03	4	8	32	
Analytics		4	16	200		3. Kubernetes 1.8.x, 1.9.x*	1.5	32	200	
Developer Portal		4	8	50		4. Helm**	1.5	8	50	
Total		16	48	532		5. Docker Repository***	10	64	532	
* If you are running on a Windows machine, you must have an ISO file creator to run InstallAssist (mkisofs, etc.)						• Kubernetes must <u>not</u> equal 1.8.9 or 1.9.4 and must be >1.8.4				
** Internally the OVA runs Ubuntu 16.04, Kubernetes 1.8.8, and Docker 18.03						• # On a single shared worker node for all subsystems				
* Kubernetes must include:						1. Ingress Controller with SSL Passthrough				
** Helm version must be compatible with Kubernetes version						2. Support for persistent volumes (shared block storage)				
*** Must have Docker on machine downloading APIC files										

[Up-to Date Information Kept in Software Compatibility Reports](#)

Agenda

- API Connect v2018 Overview
- **API Connect – Non-High Availability**
- API Connect – High Availability
 - Quorum Requirements
 - Active / Passive
 - Active / Active
- Sample Deployment

Standalone: Non-HA Single Instance Deployment

- One APIM cloud with Single Instance of each component
- Can be deployed on the same physical machine or can be a hybrid cloud setup
- Typical setup for smaller non-production environments and PoCs

Gateway Instance



Analytics Instance



Portal Instance



Manager Instance



Agenda

- API Connect v2018 Overview
- API Connect – Non-High Availability
- **API Connect – High Availability**
 - Quorum Requirements
 - Active / Passive
 - Active / Active
- Sample Deployment

High Availability in 2018

What is new with API Connect 2018?

Major advances in application development, deployment and management have led companies to begin pursuing multi-cloud application strategies

API Connect 2018 Ships Kubernetes in OVA Installations, giving customers some of the benefits of a cloud-native solution without having to install their own Kubernetes environment. This however comes with additional high availability requirements

API Connect leverages Kubernetes and other cloud scale technologies (data persistence such as Cassandra) to achieve the scalability and reliability needs for modern API management multi-cloud platforms. For E.g. if database within portal fails, traffic would be directed to remaining members and the failed node would be auto-restarted to support future traffic

Kubernetes & Underlying APIC Components Technology Requires Quorum, and without quorum the services will begin behaving abnormally. Quorum requirements calculated as such: Node Failure tolerance of $(N-1)/2$, Where N is number of instances or nodes in cluster



Gateway



Analytics



Manager

Portal
© 2018 IBM Corporation

High Availability in v5

- Cannot dynamically scale
- Slow upgrade process
- Gateways reliant on manager for gateway configurations
- Does not require Quorum

High Availability in v2018

- Dynamically scale
- Drastically reduced upgrade time
- Gateways self manage configurations at cluster level
- Requires Quorum

- Bottlenecks manager instance
- Does not promote remote gateway deployments
- Only 1 Analytics Cluster per API Connect Cloud

- Separated from manager instance
- Optimized for remote gateway deployments, by deploying analytics next to gateway to reduce latencies
- Deploy multiple analytics clusters per API Connect Cloud

- No true active/ active set up
- Cloud dissociation (split-brain scenarios)
- Impacted by analytics functionality

- True Active/ Active cluster configurations
- Quorum avoids cloud dissociation
- Better performance and stability with analytics removed

- Recommended to have 3+ Portal instances
- Only support 1 Portal Cluster per API Connect Cloud

- No changes from v5
- Deploy multiple portal clusters per API Connect Cloud

Agenda

- API Connect v2018 Overview
- API Connect – Non-High Availability
- **API Connect – High Availability**
 - Quorum Requirements
 - Active / Passive
 - Active / Active
- Sample Deployment



What is Quorum?



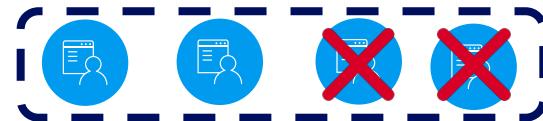
Cluster has Quorum



Cluster does not have Quorum



Cluster has Quorum



Cluster does not have Quorum



Cluster has Quorum



Cluster has Quorum



Cluster does not have Quorum

Odd numbers are better for Quorum!

Cluster can scale to even number of nodes under increased load, but better to always have odd number of members

Controlled behavior- If Quorum is lost



Gateway

- NO downtime for your APIs - Gateway will continue to serve traffic
- Additional API Connect configuration changes would be constrained



Manager

- API Manager will continue to serve traffic and will support all read transactions
- New creation of resources and deployment would be constrained



Analytics

- Continue to view analytics data that is already captured
- New ingestion data from the gateway would be constrained

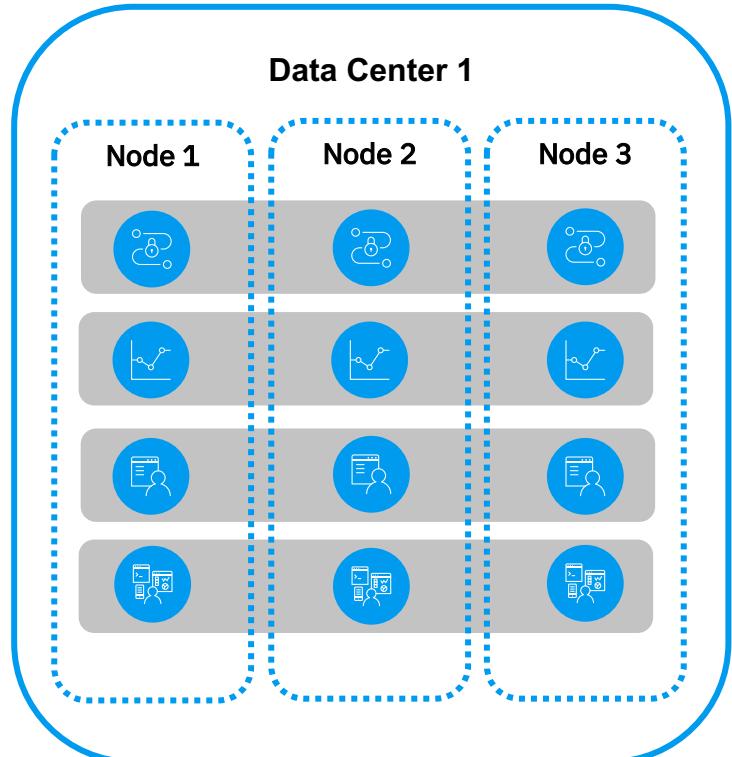


Portal

- Same as v5, site traffic would not be accepted

Single Datacenter – 3 instance HA Deployment

- One APIM Cloud with 3 instances minimum for HA based on quorum
 - $(N-1)/2$ dictates that this HA set up can handle a single node failure
 - If more than 1 node or 1 instance fails then the application begins behaving abnormally
- Nodes represent either Physical Machine or VMs

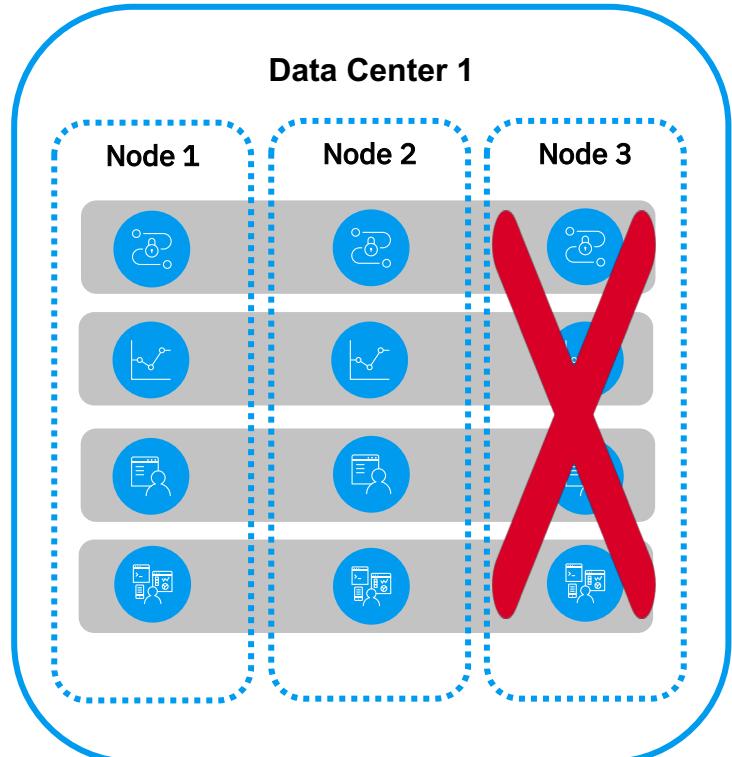


**Nodes represent physical machine or VMs

© IBM Corporation

Single Datacenter – 3 instance HA Deployment

- Scenario depicts either a failure of node 3, or a failure of the instances on node 3 or a loss connection from node 3
- **Quorum is maintained** between instances running on Node 1 & 2

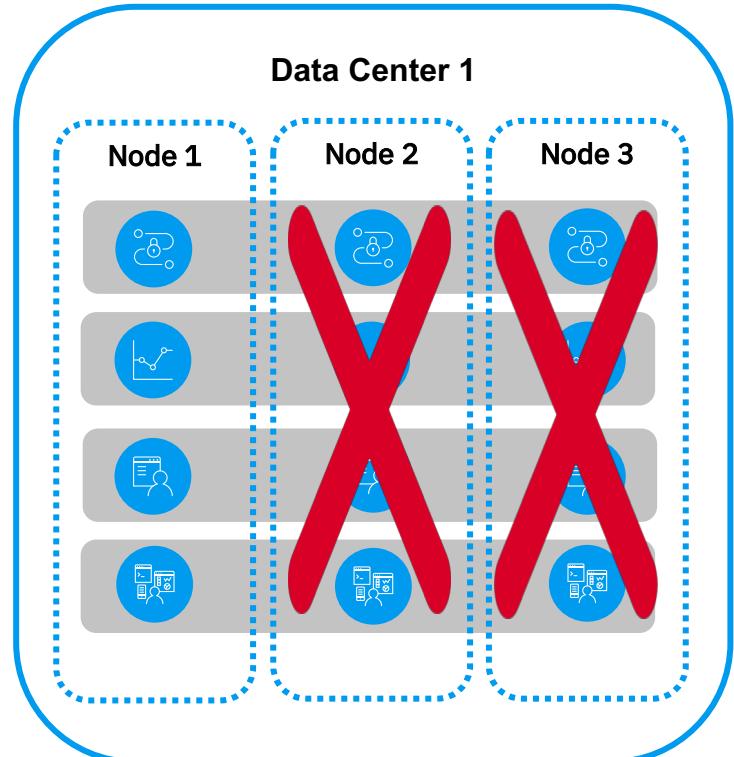


**Nodes represent physical machine or VMs

© IBM Corporation

Single Datacenter – 3 instance HA Deployment

- Scenario depicts either a failure of node 2 & 3, or a failure of the instances on node 2 & 3, or a loss connection from nodes 2 & 3
- **Quorum is lost** and controlled behavior kicks-in



**Nodes represent physical machine or VMs
© IBM Corporation

Agenda

- API Connect v2018 Overview
- API Connect – Non-High Availability
- API Connect – High Availability
 - Quorum Requirements
 - **Active / Passive**
 - Active / Active
- Sample Deployment

High Availability + Disaster Recovery (DR)

2 DC, Active/ Passive, HA Deployment

IBM Sub-Capacity Licensing

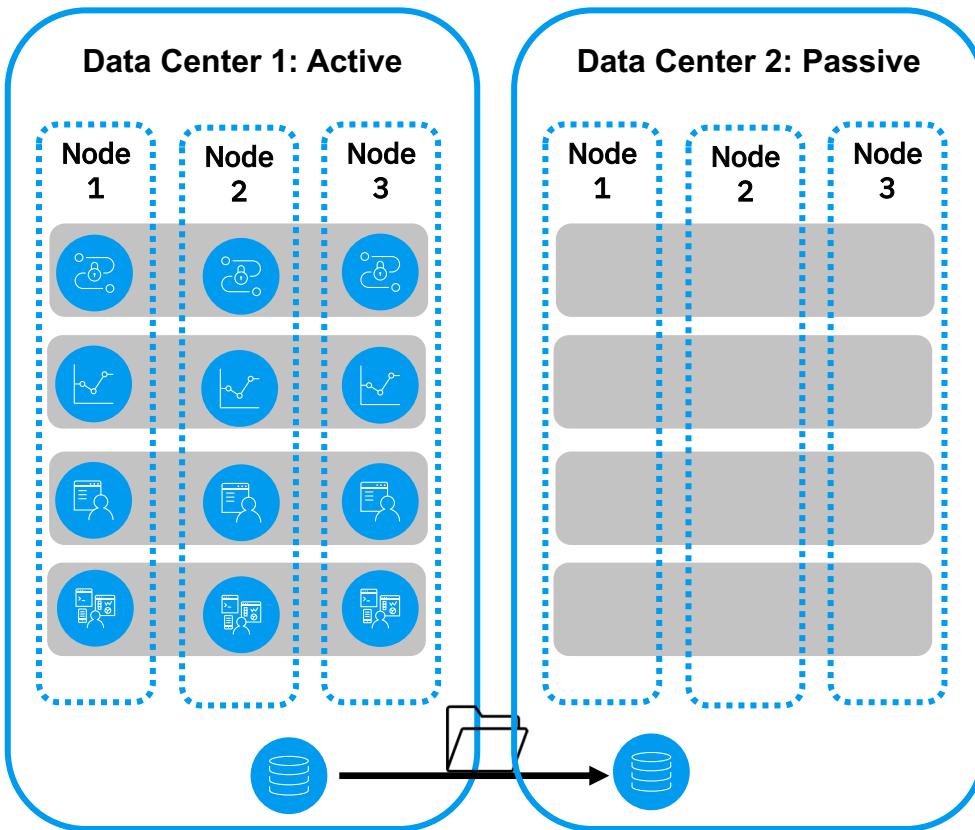
“In the case of a program or system configuration that is designed to support a high availability environment by using various techniques (for example, duplexing, mirroring of files or transactions, maintaining a heartbeat, or active linking with another machine, program, database, or other resource), **the program is considered to be doing work in both the warm and hot situation and license entitlements must be acquired.**”

Backup	Entitlements Required
Hot	Yes
Warm	**Yes
Cold	No

**Based on definitions of “Doing Work” APIC does not qualify for Warm Passive systems. Entitlements are needed for Warm set-ups
See bottom of page 8 of [IBM Software Licensing Guide](#)

Two Datacenter – Active HA + Passive for DR

- Services are highly available in Data Center 1
- Data Center 2 is pre-configured and ready to have instances of the services deployed
- Periodic cronjobs are scheduled to backup data from each of the services
 - These backups are synced from DC1 to DC2



**Nodes represent physical machine or VMs

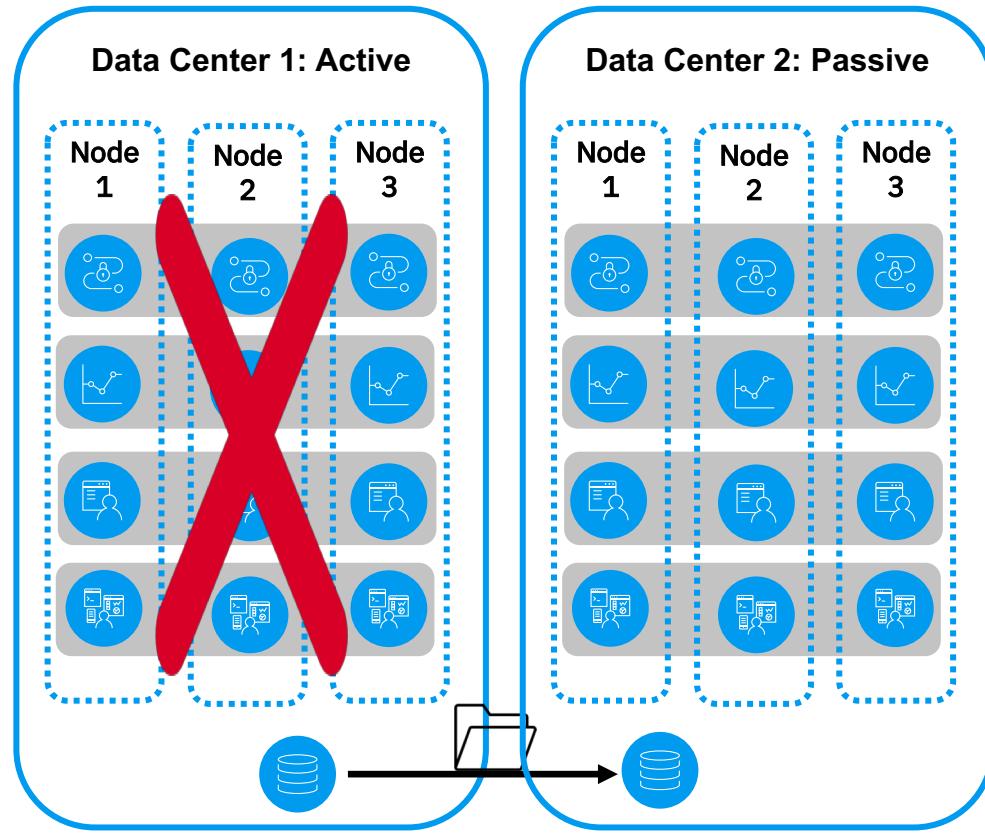
© IBM Corporation

24

High Availability + Disaster Recovery (DR)

2 DC, Active/ Passive, CA Deployment

- In the event of DC 1 failure, scripts can be executed in DC 2 that will begin deploying instances of the APIC services
- This gives users an answer to catastrophic events that cause datacenter failure
- However.... DC 2 can only recover last back-up
 - Any data not backed up is lost when failover to DC 2
- Additionally the customer would need to wait until the infrastructure and software ready before operations could resume

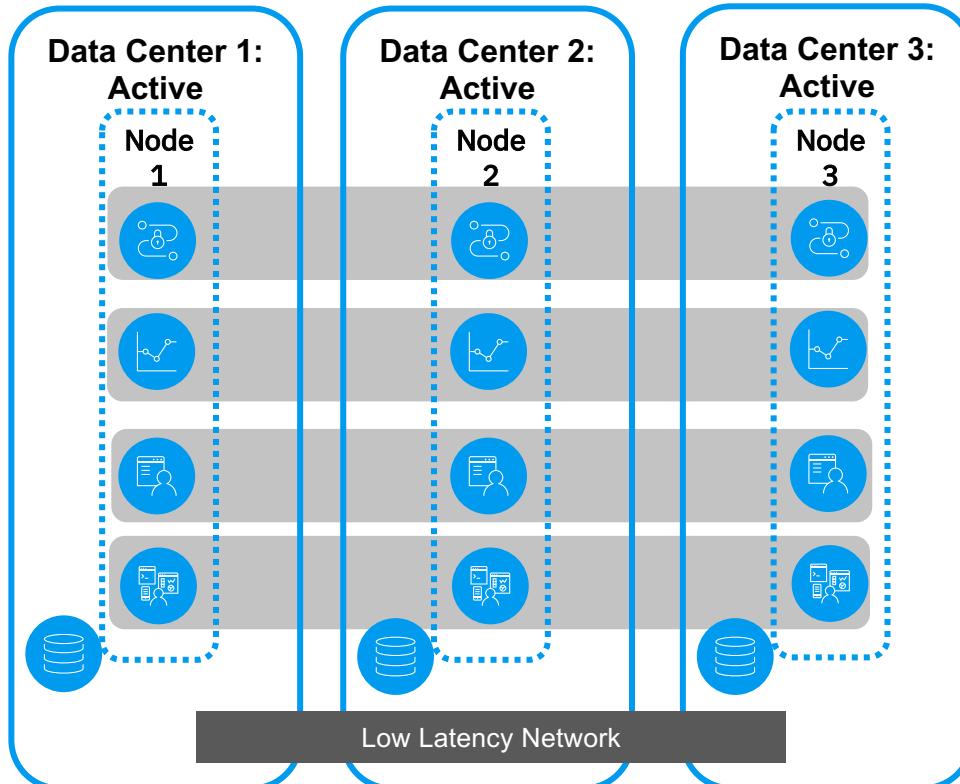


Agenda

- API Connect v2018 Overview
- API Connect – Non-High Availability
- API Connect – High Availability
 - Quorum Requirements
 - Active / Passive
 - **Active / Active**
- Sample Deployment

Three Datacenter – High Availability

- Can tolerate failure node failure, connection drop, or DC failure
- Optionally deploy additional instances in any DC as long as quorum maintained in event of a failure in DC
- Cluster can scale to even number of nodes under increased load, but better to always have odd number of members
- Provides complete datacenter fault tolerance



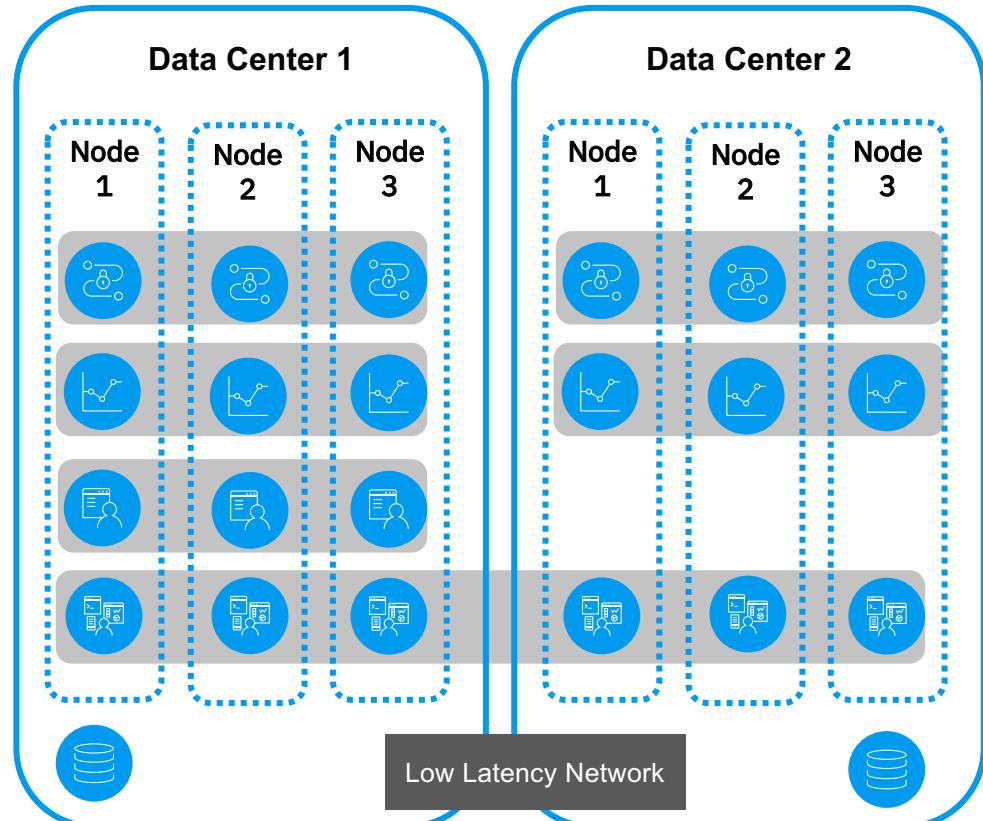
**Nodes represent physical machine or VMs

© IBM Corporation

27

Two Datacenter – Multiple Service Configuration

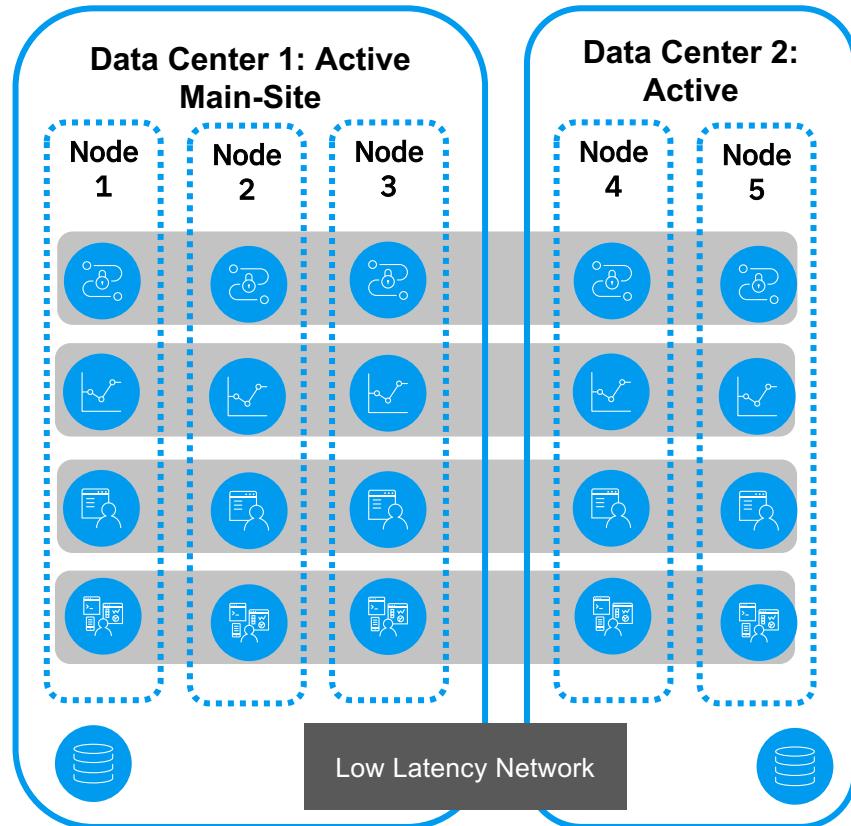
- If DC 2 fails then
 - All services maintain functionality
- If DC 1 fails then
 - Gateway, Analytics, Manager maintain functionality
 - Controlled behavior for portal kicks-in



**Nodes represent physical machine or VMs

Two Datacenter – Main-site Configuration

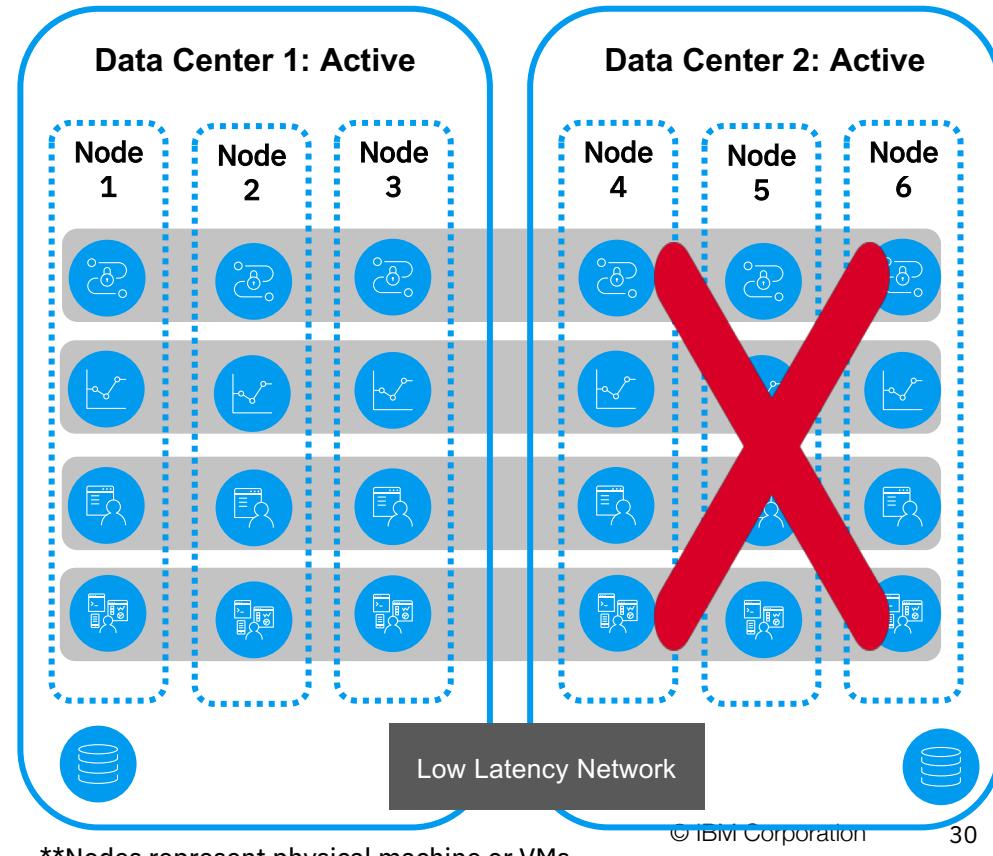
- If DC 2 fails then
 - All services maintain functionality
- If DC 1 fails then
 - Quorum is lost and controlled behavior kicks-in
- Provide flexibility to scale and spray load across data centers



**Nodes represent physical machine or VMs

Two Datacenter – Odd number of nodes in each

- If any Datacenter fails then
 - Quorum is lost and controlled behavior kicks-in
 - Does not provide any additional value in terms of HA, but does provide scale



Agenda

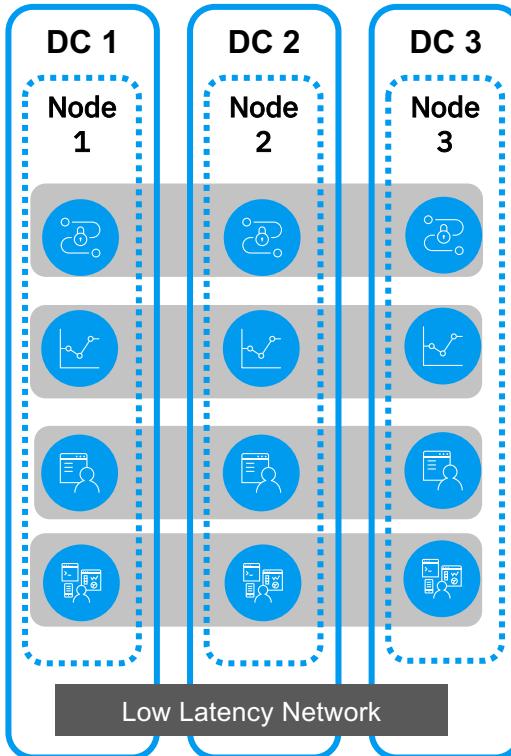
- API Connect v2018 Overview
- API Connect – Non-High Availability
- API Connect – High Availability
 - Quorum Requirements
 - Active / Passive
 - Active / Active
- **Sample Deployment**

Sample Deployment

- Have internal set of DCs already running internal apps
- Add APIM layer to these APPs
- Use 3rd party cloud to deploy new set of apps & services for external parties to consume
- No direct access to any components running in internal network from anyone outside company
- Separate Portal for internal and external API consumers
- Separate analytics services for security and networking latency reasons
- Single APIM layer for both internal and externally facing environment

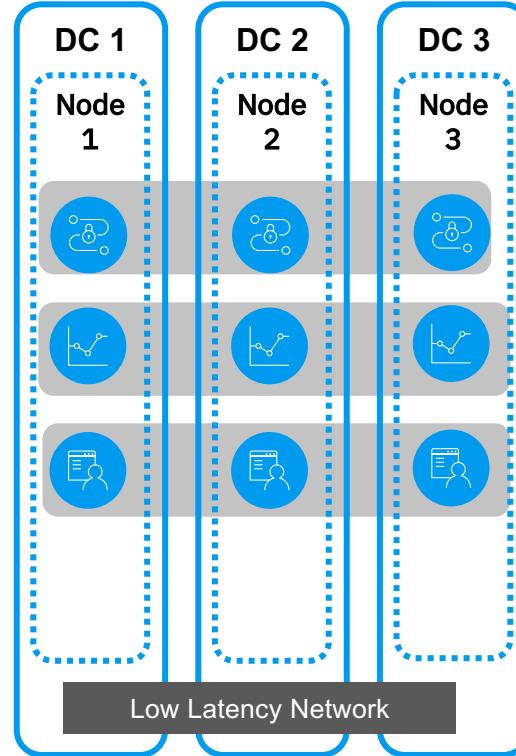
Customer Managed DCs

Internal Traffic



3rd Party Cloud (AWS)

3rd Party Traffic/ BP integrations

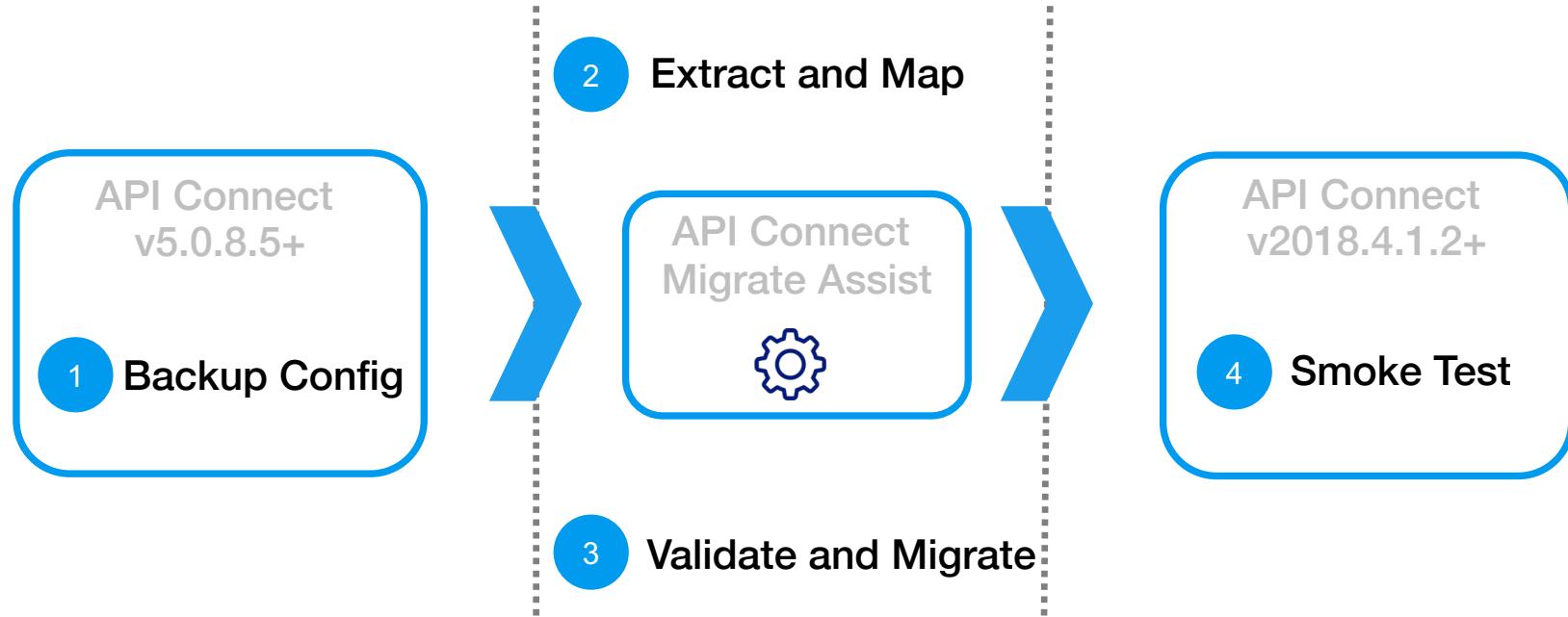


Migrate to this Next Generation API Platform

- No production downtime required for migration
- Run current APIs as-is on the new architecture
- Flexibility to adopt new deployment topology with easy to use Migrate Assist
- Native install of Migrate Assist* for enhanced user experience



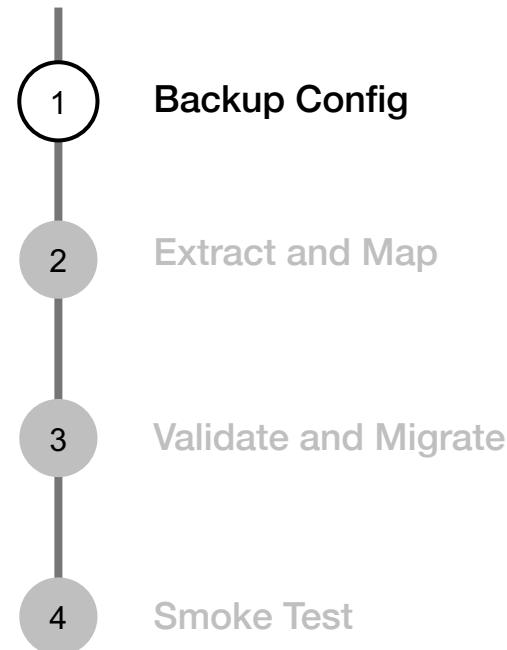
Migration Details



Simple 4 step migration process guided through easy to use
Migration tooling with flexibility to map to customer future needs

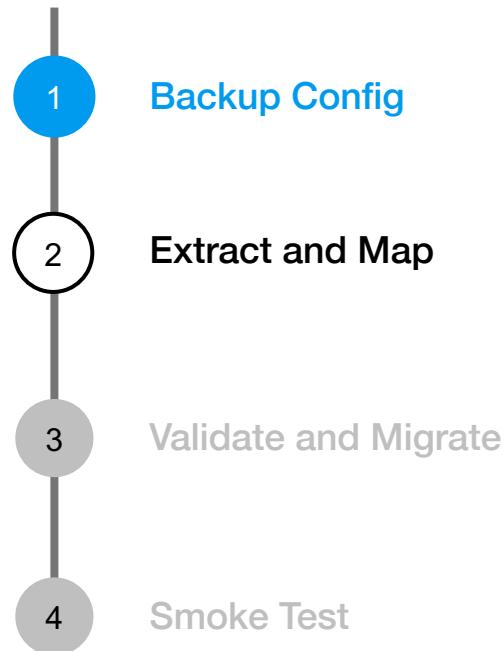
ABCs of Migration: 1 of 4

- **Export configuration** on v5.0.8.5+ as a backup including APIs/Products etc. using the enhanced CLI
- **(Optional)** If using Portal Delegated User Registry (PDUR), use Developer Portal CLI to export data



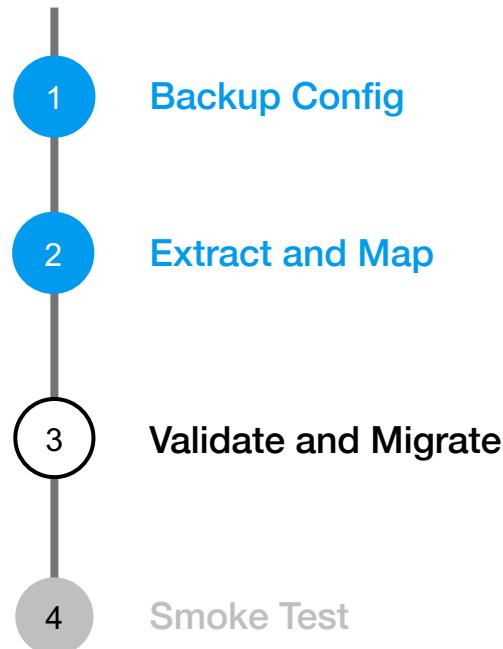
ABCs of Migration: 2 of 4

- **Extract the backup** to a human-readable yaml files using Migrate Assist
- **Flexibility to map to new resources** to reflect any changes on v2018 setup topology through mapping files, if desired



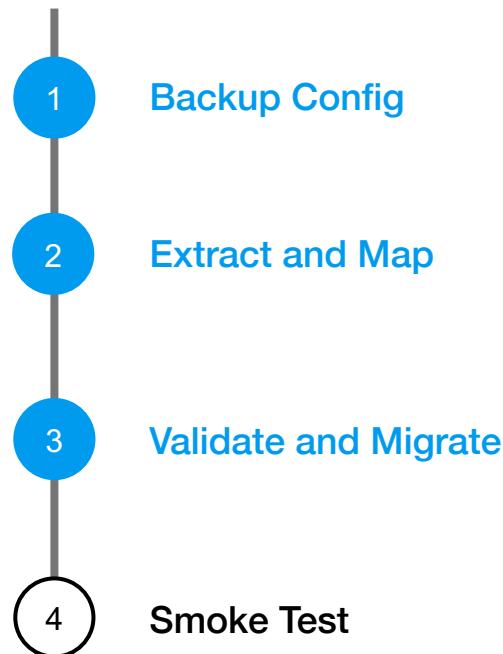
ABCs of Migration: 3 of 4

- Validate v2018 prerequisites are in place
- Validate the mapping and prerequisites using Migrate Assist
- Migrate / Load the APIs and the configuration to v2018 setup



ABCs of Migration: 4 of 4

- **Smoke test** the API endpoints to ensure successful migration
- **Finalize testing** by validating user access, on-boarding etc.



What is and What is not Migrated

Full Auto Migration

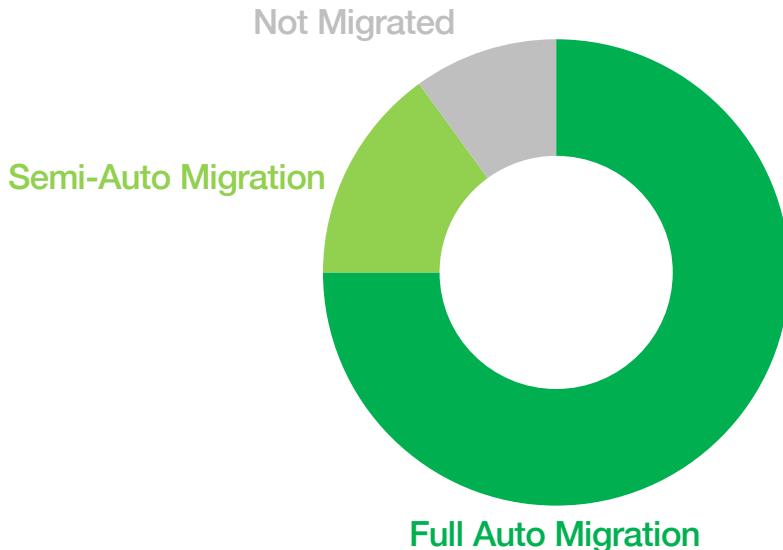
- Provider Organization artifacts including APIs, Products, Consumers, subscriptions etc.

Semi-Auto Migration

- Customization like gateway extensions, PDUR, custom portal themes etc.

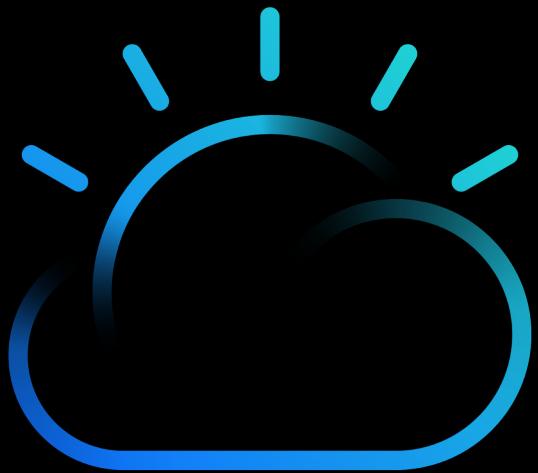
Not Migrated

- Infrastructure configuration such as server definitions, transient data such as invitations etc.

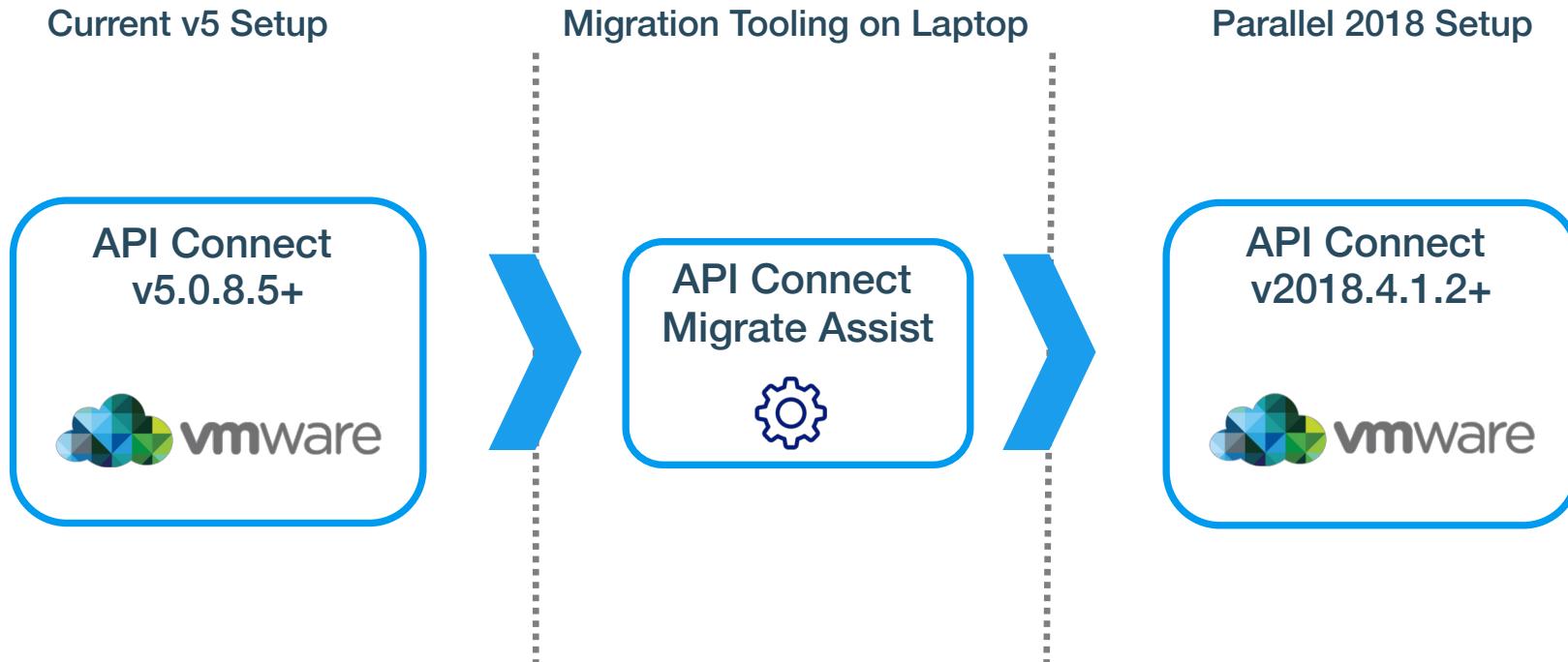


[Learn More](#) on Migration FAQs

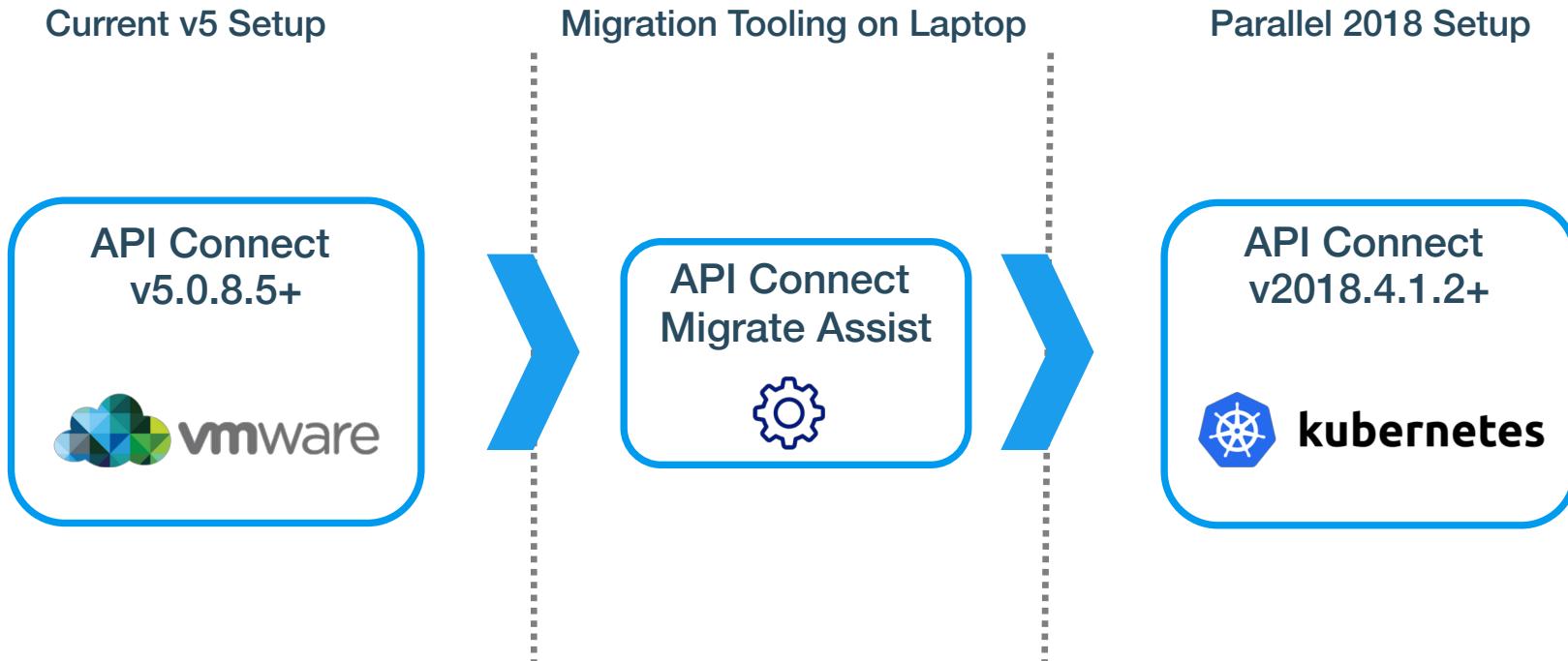
Thank You



Overview of Migration: OVA to OVA



Overview of Migration: OVA to K8s



→ Clients can migrate from an v5 OVA to a v2018 k8s setup using the migration tooling

Agenda

- Migration Overview
- **ABC's Of Migration**
- Must Knows of Migration
- Demo

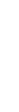
Agenda

- Migration Overview
- ABC's Of Migration
- **Must Knows of Migration**
- Demo

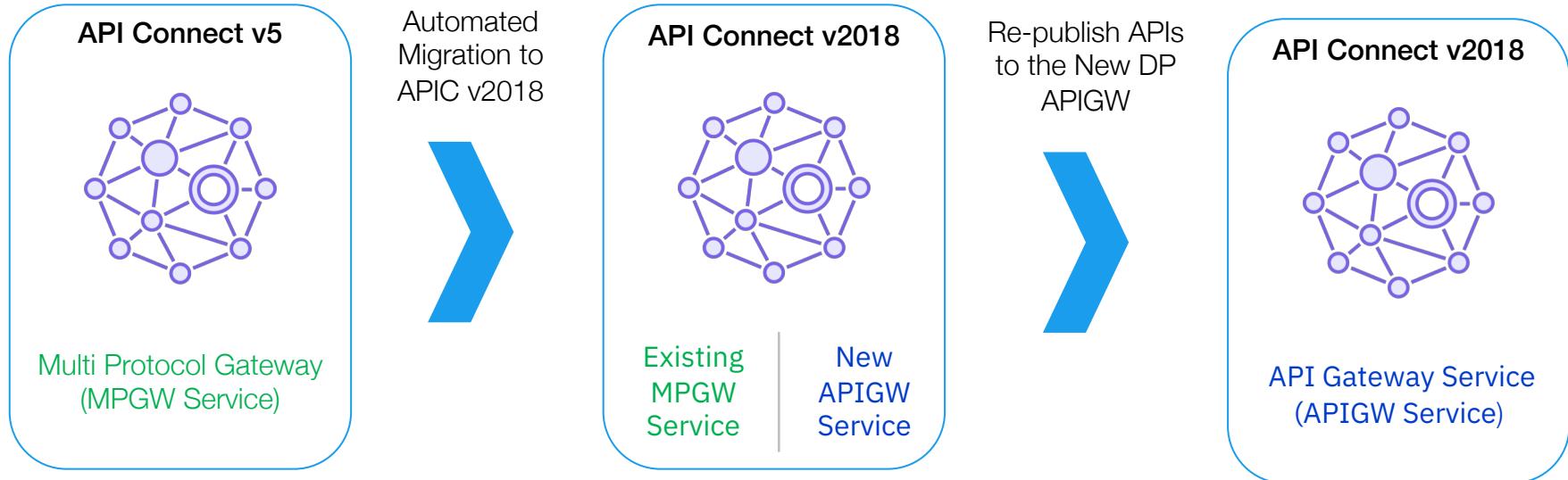
Migration Must Knows - Gateway

- **v5 supports Multi-Protocol Gateway** (MPGW) service as the primary gateway service
- **v2018 supports two Gateway Services**
 - New API Gateway (APIGW) service, built for API workloads
 - Multi-Protocol Gateway (MPGW) service / DataPower Gateway (DPGW) service
- **OOTB API Portability to MPGW/DPGW** service on v2018 for APIs deployed on v5

Gateway



Migration Must Knows - Gateway



- **Migration tooling will automatically move APIs as-is to v2018 MPGW**
- **Migration Tooling will NOT migrate APIs to the New APIGW Service.** Clients will update policy versions / Gateway scripts in assembly and re-publish to API Gateway Service

Migration Must Knows - Portal

- v2018 Developer Portal based on Drupal v8
 - Drupal codebase are different and requires reconfiguration for portal customizations
 - v2018 built-in theme generator accelerates this process
- **User identities are centrally managed in API Manager** including consumers on-boarded using Social Identity providers
 - No PDUR on Portal
 - PDUR data is exported and provided as input during migration and loaded in to API Manager

Gateway



Portal



Migration Must Knows - Analytics

- Analytics data is not migrated by Migration Tooling
 - Default retention is for 90 days
 - Analytics data from v5 would not be migrated to v2018, but would continue to be available for use on v5 setup and through offloaded system
- **3rd Party Offload for Analytics** data retention via supported third-party systems on v5
 - Options include HTTP, Elasticsearch, Apache Kafka and Syslog

Gateway



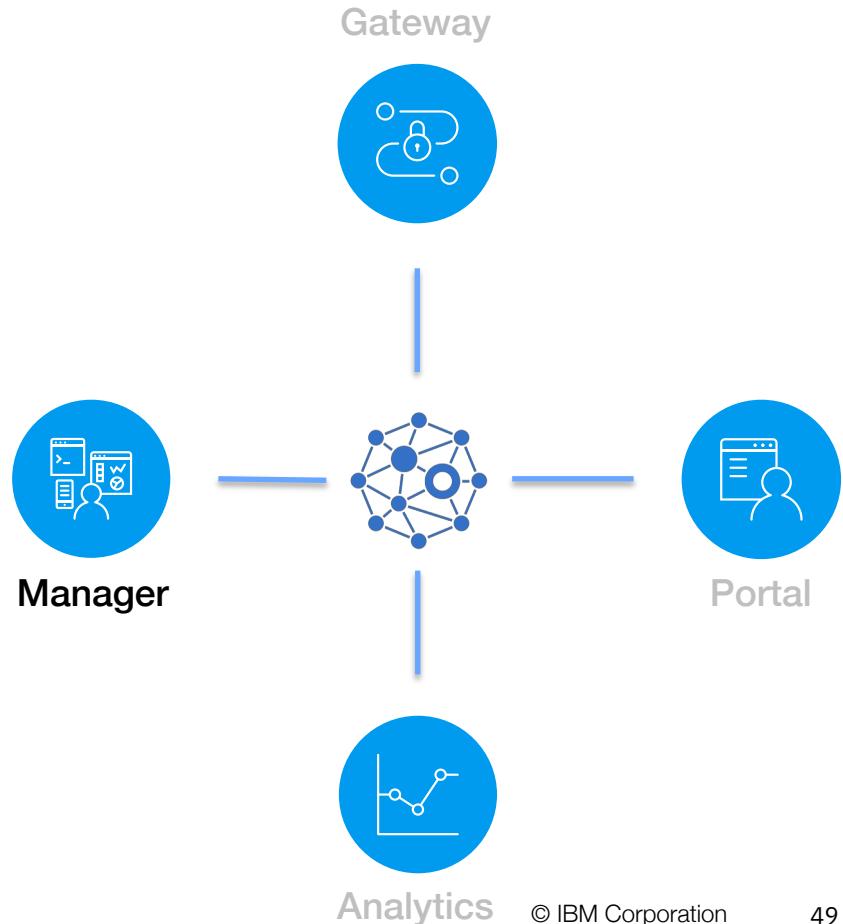
Portal



Analytics

Migration Must Knows - Manager

- **Resolve transient data** prior to backup (i.e. complete pending tasks/invitations)
 - Lifecycle approval, subscription approval tasks
 - Pending invitations for user on-boarding (Consumer / Provider)
- **Migration adjusts resource names to adhere to supported v2018 characters**
 - Applies to API property names and catalog names
 - Supported characters include A-Z, a-z, 0-9, _, and –



IBM Services Offering to Guide Customers through Migration

Offering	What's Included
<p>API Connect Version 2018x Migration Assessment Service Offering</p>	<ul style="list-style-type: none">• Guidance on migrating to v2018• Discovery of existing v5 deployment topologies• Identification of customizations and dependencies the deployed configuration• Assessment and action plan for what needs to be done to migrate to v2018
<p>API Connect Version 2018x Migration Implementation Service Offering</p>	<ul style="list-style-type: none">• Coverage of manual steps to migrate data artifacts native to APIC v5• Templates for input files, preparation of those files and running the migration script itself• Migration logs including a report of what artifacts have been successfully migrated

