Table of Contents

[Reference Architecture 2](#_Toc55330486)

[AKS Production checklist 2](#_Toc55330487)

[Secure Baseline Architecture 2](#_Toc55330488)

[Github Samples 2](#_Toc55330489)

[Links 2](#_Toc55330490)

[Windows Container and Private cluster Announcements 4](#_Toc55330491)

[Videos and YouTube Channels 4](#_Toc55330492)

[Hands-On 4](#_Toc55330493)

[Istio Support? 4](#_Toc55330494)

[How to deploy Nginx controller without HELM 4](#_Toc55330495)

[App Modernization on AKS – A practical approach 5](#_Toc55330496)

[How can I configure overprovisioning with Cluster Autoscaler? 5](#_Toc55330497)

[Containers extension in Windows Admin Center 5](#_Toc55330498)

[AKS Concepts 5](#_Toc55330499)

[What is a service?​ 6](#_Toc55330500)

[What is a Ingress?​ 6](#_Toc55330501)

[What is a Ingress Controller​ 6](#_Toc55330502)

[How does Ingress Controller Work? 6](#_Toc55330503)

[​What is Application Gateway Ingress Controller (AGIC)? 6](#_Toc55330504)

[AKS for Windows 7](#_Toc55330505)

[How do you select values for the horizontal pod autoscaler? 8](#_Toc55330506)

[How do you create and use Kubernetes custom resources? 9](#_Toc55330507)

[SQL Server container for Windows 9](#_Toc55330508)

[Persistent volume claims 10](#_Toc55330509)

# Reference Architecture

<https://docs.microsoft.com/en-us/azure/architecture/reference-architectures/containers/aks-microservices/aks-microservices>

# AKS Production checklist

<https://www.the-aks-checklist.com/>

# Secure Baseline Architecture

<https://docs.microsoft.com/en-us/azure/architecture/reference-architectures/containers/aks/secure-baseline-aks>

# Github Samples

<https://github.com/Azure/sg-aks-workshop>

<https://github.com/azure/aks-production-patterns>

<https://github.com/retaildevcrews/secure-aks-infra>

# Links

FAQ - <https://docs.microsoft.com/azure/aks/faq>

Kubernetes Best Practices – <https://aka.ms/aks/bestpractices>​

Reference Architecture - <https://docs.microsoft.com/en-us/azure/architecture/solution-ideas/articles/aks-api-first>

Taint and Toleration - <https://docs.microsoft.com/en-us/azure/aks/operator-best-practices-advanced-scheduler>

Headless service - <https://itnext.io/how-to-access-kubernetes-applications-using-services-c437a4c940ad>

SQL Server in container - <https://docs.microsoft.com/en-us/sql/linux/tutorial-sql-server-containers-kubernetes?view=sql-server-ver15>

HPA Scale - <https://docs.microsoft.com/en-us/azure/aks/tutorial-kubernetes-scale>

Cluster Scale - <https://docs.microsoft.com/en-us/azure/aks/cluster-autoscaler>

Network policies - <https://docs.microsoft.com/en-us/azure/aks/use-network-policies>

Network concepts - <https://docs.microsoft.com/en-us/azure/aks/concepts-network>

API Managements and AKS - <https://docs.microsoft.com/en-us/azure/api-management/api-management-kubernetes>

Application gateway Ingress controller - <https://github.com/Azure/application-gateway-kubernetes-ingress>

AGIC - <https://roykim.ca/2020/02/09/understanding-ingress-controllers-and-azure-app-gateway-for-azure-kubernetes-part-1-intro/>

APIM - <https://pumpingco.de/blog/control-ingress-to-aks-with-azure-api-management/>

Continuous Deployment - <https://azure.github.io/application-gateway-kubernetes-ingress/how-tos/continuous-deployment/>

Image and container level security

<https://docs.microsoft.com/en-us/azure/security-center/azure-container-registry-integration>

Node and cluster level security

Use AAD integrated clusters with Kubernetes RBAC enabled

<https://docs.microsoft.com/en-us/azure/aks/managed-aad>

Use private clusters with egress through a NVA

<https://docs.microsoft.com/en-us/azure/aks/private-clusters>

Deploy kured to perform automatic reboots if needed as a result of daily automatic security patching

<https://docs.microsoft.com/en-us/azure/aks/node-updates-kured>

Disable the dashboard

<https://docs.microsoft.com/en-us/azure/aks/kubernetes-dashboard>

Use Azure Security Center with AKS

[https://docs.microsoft.com/en-us/azure/security-center/recommendations-reference#recs-containers](https://docs.microsoft.com/en-us/azure/security-center/recommendations-reference)

Use Azure Policy (public preview) with AKS

<https://docs.microsoft.com/en-us/azure/governance/policy/concepts/policy-for-kubernetes>

Pod level security

Use AAD Pod Identity

<https://github.com/Azure/aad-pod-identity>

Workload level security

Use Kubernetes namespaces

Use network policy to control east-west traffic in the cluster

<https://docs.microsoft.com/en-us/azure/aks/use-network-policies>

Use the Secrets Store CSI driver to mount secrets into Kubernetes pods

<https://docs.microsoft.com/en-us/azure/key-vault/general/key-vault-integrate-kubernetes>

# Windows Container and Private cluster Announcements

<https://azure.microsoft.com/en-us/blog/announcing-the-general-availability-of-windows-server-containers-and-private-clusters-for-azure-kubernetes-service/>

# Videos and YouTube Channels

Lachlan Evenson (aka Crocodile Dundee Of Kubernetes) [https://www.youtube.com/channel/UCC5NsnXM2lE6kKfJKdQgsRQ](https://nam06.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.youtube.com%2Fchannel%2FUCC5NsnXM2lE6kKfJKdQgsRQ&data=01%7C01%7CAkshay.Nimbalkar%40microsoft.com%7C0dae892217c94734bb2108d6d7bb5af7%7C72f988bf86f141af91ab2d7cd011db47%7C1&sdata=fs6Dz3PLBk%2B9q%2BVXJIGLQ8267ouqatb5eZbEwpWt%2BJg%3D&reserved=0)

TGI Kubernetes by Heptio - [https://www.youtube.com/channel/UCjQU5ZI2mHswy7OOsii\_URg/videos](https://nam06.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.youtube.com%2Fchannel%2FUCjQU5ZI2mHswy7OOsii_URg%2Fvideos&data=01%7C01%7CAkshay.Nimbalkar%40microsoft.com%7C0dae892217c94734bb2108d6d7bb5af7%7C72f988bf86f141af91ab2d7cd011db47%7C1&sdata=2yff%2BJhrO1nk%2B5hic2Pnf9MExSs0bNSPF2vGYw%2BO8I4%3D&reserved=0)

Linux Foundation CKA (Must Register With MSFT For Free Access) [https://lms.quickstart.com](https://nam06.safelinks.protection.outlook.com/?url=https%3A%2F%2Flms.quickstart.com&data=01%7C01%7CAkshay.Nimbalkar%40microsoft.com%7C0dae892217c94734bb2108d6d7bb5af7%7C72f988bf86f141af91ab2d7cd011db47%7C1&sdata=iEoMgP0CDIcFskYZosOayRt2mSxS2cbzT5kX%2BMmxGZc%3D&reserved=0)

# Hands-On

Kubernetes BootCamp - [https://kubernetesbootcamp.github.io/kubernetes-bootcamp/index.html](https://nam06.safelinks.protection.outlook.com/?url=https%3A%2F%2Fkubernetesbootcamp.github.io%2Fkubernetes-bootcamp%2Findex.html&data=01%7C01%7CAkshay.Nimbalkar%40microsoft.com%7C0dae892217c94734bb2108d6d7bb5af7%7C72f988bf86f141af91ab2d7cd011db47%7C1&sdata=EaWm81bgJF3CYEpaoOLqVZLk40GxaXsEcB0PN4dBqmw%3D&reserved=0)

Katacoda Labs -[https://katacoda.com/courses/kubernetes](https://nam06.safelinks.protection.outlook.com/?url=https%3A%2F%2Fkatacoda.com%2Fcourses%2Fkubernetes&data=01%7C01%7CAkshay.Nimbalkar%40microsoft.com%7C0dae892217c94734bb2108d6d7bb5af7%7C72f988bf86f141af91ab2d7cd011db47%7C1&sdata=0BzcVwL1dsSU1fHK2osK%2Fpcn3iLydx4c6YzbRJ3OgqU%3D&reserved=0)

Azure Citadel - [https://azurecitadel.github.io/](https://nam06.safelinks.protection.outlook.com/?url=https%3A%2F%2Fazurecitadel.github.io%2F&data=01%7C01%7CAkshay.Nimbalkar%40microsoft.com%7C0dae892217c94734bb2108d6d7bb5af7%7C72f988bf86f141af91ab2d7cd011db47%7C1&sdata=CHfsw%2Fpc5WDqe11RstnG9VNeK0%2FL9KWS5zCheApeMX0%3D&reserved=0)

# Istio Support?

Helm templates for Istio deployment to the AKS clusters.

[https://istio.io/docs/setup/kubernetes/platform-setup/azure/](https://istio.io/docs/setup/kubernetes/platform-setup/azure/%20)

<https://docs.microsoft.com/en-us/azure/aks/servicemesh-istio-install?pivots=client-operating-system-windows>

# How to deploy Nginx controller without HELM

<https://dzone.com/articles/nginx-ingress-controller-configuration-in-aks>

<https://cloudblogs.microsoft.com/opensource/2018/11/27/tutorial-azure-devops-setup-cicd-pipeline-kubernetes-docker-helm/>

# App Modernization on AKS – A practical approach

<https://aka.ms/EVT504AL-OnDemand>

<https://azure.microsoft.com/en-us/resources/videos/azure-friday-intro-to-rudr-a-kubernetes-implementation-of-the-open-application-model/>

# How can I configure overprovisioning with Cluster Autoscaler?

<https://github.com/kubernetes/autoscaler/blob/master/cluster-autoscaler/FAQ.md#how-can-i-configure-overprovisioning-with-cluster-autoscaler>

# [Containers extension in Windows Admin Center](https://techcommunity.microsoft.com/t5/containers/announcement-new-updates-to-the-containers-extension-in-windows/ba-p/1449476)

A screenshot of a social media post

Description automatically generated

The above screenshot is from the Containers extension on Windows Admin Center. With this tool, customers can containerize their applications without having to write the docker file, and then are able to push the image to an external registry, such as ACR.

We have some details on the tool [here](https://nam06.safelinks.protection.outlook.com/?url=https%3A%2F%2Ftechcommunity.microsoft.com%2Ft5%2Fcontainers%2Fannouncement-new-updates-to-the-containers-extension-in-windows%2Fba-p%2F1449476&data=02%7C01%7CAkshay.Nimbalkar%40microsoft.com%7C61ce5c738b1e4d02e04108d8462eaef0%7C72f988bf86f141af91ab2d7cd011db47%7C1%7C0%7C637336513626533799&sdata=1oyKM3nuUuqA5xkZgR12slNQMuOW1VCqNzUqKqwXGi4%3D&reserved=0). We recently released an update to support [Web Deploy](https://nam06.safelinks.protection.outlook.com/?url=https%3A%2F%2Ftechcommunity.microsoft.com%2Ft5%2Fcontainers%2Fannouncement-adding-webdeploy-support-to-container-image%2Fba-p%2F1579144&data=02%7C01%7CAkshay.Nimbalkar%40microsoft.com%7C61ce5c738b1e4d02e04108d8462eaef0%7C72f988bf86f141af91ab2d7cd011db47%7C1%7C0%7C637336513626533799&sdata=S57o7JdmjjvqCjLoVUh8jRGZos8unuR5BP3DtY5I%2FvQ%3D&reserved=0). I also started a video series to show the end to end experience from WS2012R2 on-prem to AKS. Video 1 is available [here](https://nam06.safelinks.protection.outlook.com/?url=https%3A%2F%2Ftechcommunity.microsoft.com%2Ft5%2Fitops-talk-blog%2Fjourney-of-an-app-how-to-extract-a-web-app-from-iis-to-create-a%2Fba-p%2F1590189&data=02%7C01%7CAkshay.Nimbalkar%40microsoft.com%7C61ce5c738b1e4d02e04108d8462eaef0%7C72f988bf86f141af91ab2d7cd011db47%7C1%7C0%7C637336513626543793&sdata=kAd5%2Fe7MoRgrVk3%2FxCZZBOcfMV7ua2GOgWUhUirAraI%3D&reserved=0).

AKS Concepts

## What is a service?​

Logical set of Pods and a policy by which to access them. The set of Pods targeted by a Service is (usually) determined by a Label Selector.​ Services provide important features that are standardized across the cluster: load-balancing (L4), service discovery between applications, and features to support zero-downtime application deployments.​

## What is a Ingress?​

A Kubernetes API that manages external access to the services in the cluster​

* Supports HTTP and HTTPs​
* Path and Subdomain based routing​
* SSL Termination​
* Save on Public IPs​

## What is a Ingress Controller​

A daemon, deployed as a Kubernetes Pod, that watches the Ingress Endpoint for updates. Its job is to satisfy requests for ingresses. ​Some popular Ingress Controllers include nginx, Traefik and HAProxy.​

## How does Ingress Controller Work?

A screenshot of a cell phone

Description automatically generated

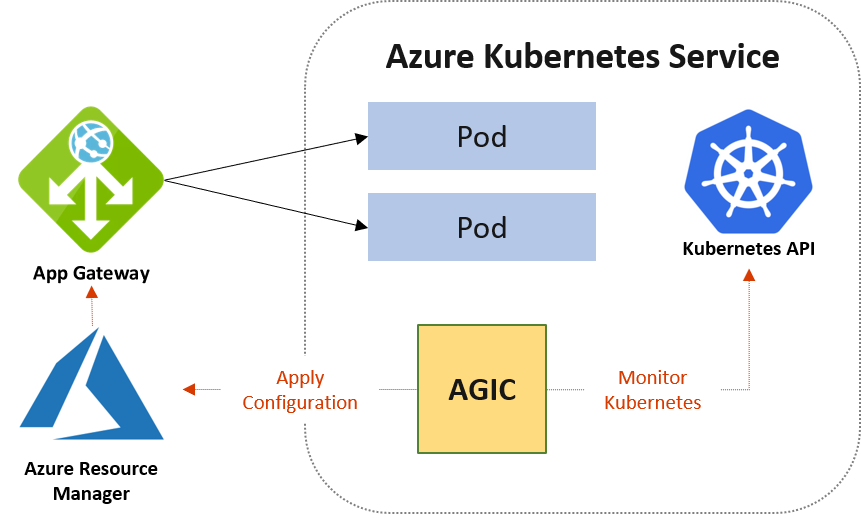
## ​What is Application Gateway Ingress Controller (AGIC)?

Application Gateway Ingress Controller (AGIC) is a Kubernetes application, which makes it possible for [Azure Kubernetes Service (AKS)](https://azure.microsoft.com/en-us/services/kubernetes-service/) customers to leverage Azure's native [Application Gateway](https://azure.microsoft.com/en-us/services/application-gateway/) L7 load-balancer to expose cloud software to the Internet. AGIC monitors the Kubernetes cluster it is hosted on and continuously updates an App Gateway, so that selected services are exposed to the Internet.

The Ingress Controller runs in its own pod on the customer’s AKS. AGIC monitors a subset of Kubernetes Resources for changes. The state of the AKS cluster is translated to App Gateway specific configuration and applied to the [Azure Resource Manager (ARM)](https://docs.microsoft.com/en-us/azure/azure-resource-manager/resource-group-overview).

AGIC is configured via the Kubernetes [Ingress resource](http://kubernetes.io/docs/user-guide/ingress/), along with Service and Deployments/Pods. It provides a number of features, leveraging Azure’s native App Gateway L7 load balancer. To name a few:

* URL routing
* Cookie-based affinity
* SSL termination
* End-to-end SSL
* Support for public, private, and hybrid web sites
* Integrated web application firewall



## AKS for Windows

Windows Server container support in the Azure Kubernetes Service is now GA. With this preview, you can:

* Lift and shift Windows applications to run on AKS
* Seamlessly manage Windows and Linux applications through a single unified API
* Mix Windows and Linux applications in the same Kubernetes cluster – with consistent monitoring experience and deployment pipelines

Now you can get the best of managed Kubernetes for all of your workloads whether they’re in Windows, Linux or both

References:

<https://azure.microsoft.com/en-us/updates/azure-kubernetes-service-aks-now-supports-windows-server-containers-3/>

<https://docs.microsoft.com/en-us/azure/aks/windows-container-cli>

<https://docs.microsoft.com/en-us/virtualization/windowscontainers/about/>

<https://kubernetes.io/blog/2019/04/01/kubernetes-v1.14-delivers-production-level-support-for-windows-nodes-and-windows-containers/>

## How do you select values for the horizontal pod autoscaler?

When your application has lot of demand you will need to scale-up or down based on different factors like number of requests, or many other specific counters. Of the most interesting feature of a microservice architecture is scaling, that it is more granular than on a classical deployment, providing us the flexibility to scale only the components that are on stress. HPA allows us to optimize how we consume our computational resources.

**HPA Yaml File**

apiVersion: autoscaling/v1

kind: HorizontalPodAutoscaler

metadata:

name: azure-vote-back-hpa

spec:

maxReplicas: 10 # define max replica count

minReplicas: 3 # define min replica count

scaleTargetRef:

apiVersion: apps/v1

kind: Deployment

name: azure-vote-back

targetCPUUtilizationPercentage: 50 # target CPU utilization

apiVersion: autoscaling/v1

kind: HorizontalPodAutoscaler

metadata:

name: azure-vote-front-hpa

spec:

maxReplicas: 10 # define max replica count

minReplicas: 3 # define min replica count

scaleTargetRef:

apiVersion: apps/v1

kind: Deployment

name: azure-vote-front

targetCPUUtilizationPercentage: 50 # target CPU utilization

References:

<https://docs.microsoft.com/en-us/azure/architecture/reference-architectures/microservices/aks#scalability-considerations>

<https://docs.microsoft.com/en-us/azure/aks/tutorial-kubernetes-scale>

<https://docs.microsoft.com/en-us/azure/aks/cluster-autoscaler>

## How do you create and use Kubernetes custom resources?

References:

<https://github.com/Azure/kubernetes-keyvault-flexvol>

<https://github.com/operator-framework/getting-started>

## SQL Server container for Windows

References:

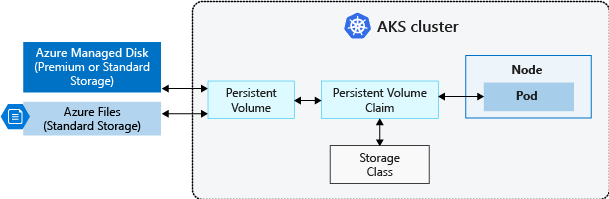
<https://github.com/microsoft/mssql-docker>

<https://hub.docker.com/_/microsoft-mssql-server?tab=description>

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-elastic-scale-introduction>

## Persistent volume claims

A PersistentVolumeClaim requests either Disk or File storage of a particular StorageClass, access mode, and size. The Kubernetes API server can dynamically provision the underlying storage resource in Azure if there is no existing resource to fulfill the claim based on the defined StorageClass. The pod definition includes the volume mount once the volume has been connected to the pod.



A PersistentVolume is bound to a PersistentVolumeClaim once an available storage resource has been assigned to the pod requesting it. There is a 1:1 mapping of persistent volumes to claims.

The following example YAML manifest shows a persistent volume claim that uses the managed-premium StorageClass and requests a Disk 5Gi in size:

YAMLCopy

apiVersion: v1

kind: PersistentVolumeClaim

metadata:

name: azure-managed-disk

spec:

accessModes:

- ReadWriteOnce

storageClassName: managed-premium

resources:

requests:

storage: 5Gi

When you create a pod definition, the persistent volume claim is specified to request the desired storage. You also then specify the volumeMount for your applications to read and write data. The following example YAML manifest shows how the previous persistent volume claim can be used to mount a volume at /mnt/azure:

YAMLCopy

kind: Pod

apiVersion: v1

metadata:

name: nginx

spec:

containers:

- name: myfrontend

image: nginx

volumeMounts:

- mountPath: "/mnt/azure"

name: volume

volumes:

- name: volume

persistentVolumeClaim:

claimName: azure-managed-disk

References:

<https://docs.microsoft.com/en-us/azure/aks/concepts-storage#persistent-volume-claims>

<https://kubernetes.io/docs/concepts/storage/persistent-volumes/>

<https://docs.microsoft.com/en-us/sql/linux/tutorial-sql-server-containers-kubernetes?view=sql-server-ver15>