



Azure Container Apps

Goodbye Kubernetes?



Geert Baeke
Cloud-native architect and Microsoft Azure MVP | Microsoft Incubator and Cloud-Native Architect @ De Cronos Groep St-Niklaas, Flemish Region, Belgium - [Contact info](#)

Geert Baeke
@GeertBaeke 2.34K subscribers 69 videos
I post tech videos on all things cloud and cloud-native. Focus on Azure and... >

[HOME](#) [VIDEOS](#) [PLAYLISTS](#) [COMMUNITY](#) [CHANNELS](#) [ABOUT](#) [SEARCH](#)

Videos ► Play all

Thumbnail	Title	Views	Published
	OpenAI Tweet Generator on Azure Container Apps	38 views	20 hours ago • 1 VPH
	AKS Dev Extension for VS Code	224 views	5 months ago
	Publish Your AKS Services with Azure Private Link and...	1.8K views	8 months ago
	Improve your development inner loop with Draft v2 for AKS	358 views	9 months ago

Popular videos ► Play all

Thumbnail	Title	Views	Published
	Creating a multi-stage YAML pipeline in Azure DevOps	21K views	3 years ago
	Getting started with Kubernetes service accounts	20K views	2 years ago
	Deploying Helm Charts with Azure DevOps	11K views	2 years ago
	Supercharge your Kubernetes deployments with Flux v2 a...	10K views	2 years ago

ome
xplore
otifications
essages
bookmarks
sts
profile

Geert Baeke
7,926 Tweets

←

Geert Baeke
@GeertBaeke

| Technology at the intersection of cloud-native infra
| Microsoft Azure MVP
| Check out my YouTube channel: youtube.com/geertbaeke
📍 Belgium ⚡ blog.baeke.info 📅 Joined March

Resources



<https://github.com/gbaeke/aca-openai>



<https://aca.super.site>

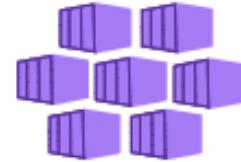




Container
Instances



App Services



Kubernetes
Services



Virtual
Machines

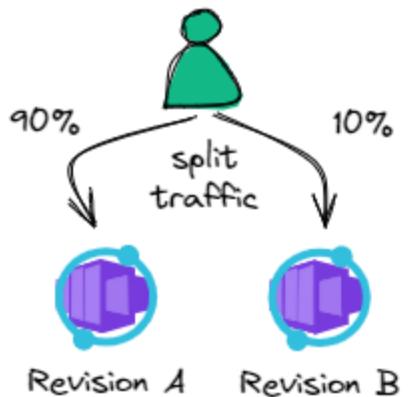


Container Apps

- Serverless CaaS for microservices
- Focus on apps, not infrastructure
- Scale dynamically based on events
- Built on open source

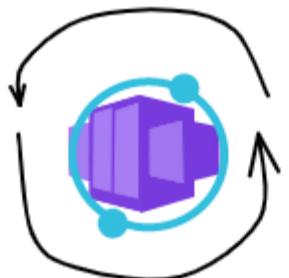


APIs and web apps



Scale based on concurrent HTTP requests
Default = 10

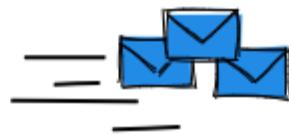
Background Processing



Continuously running background process

Scale based on CPU and memory (Keda scalers)

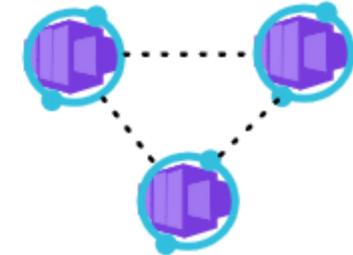
Event Processing



Reading messages from a topic or queue

Scale based on messages in the topic/queue (Keda scalers)

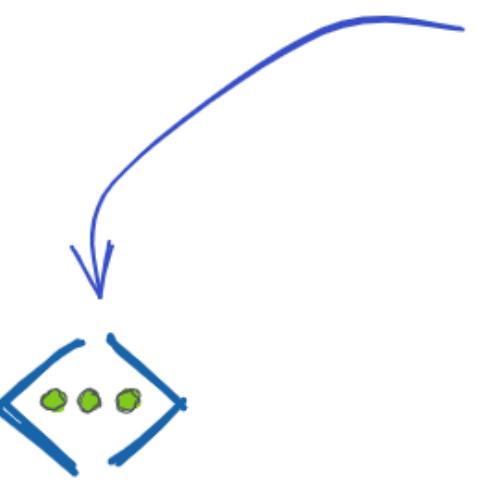
Microservices



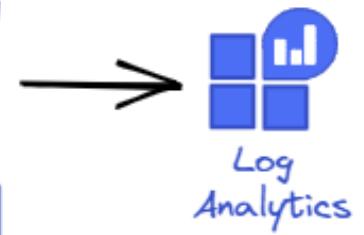
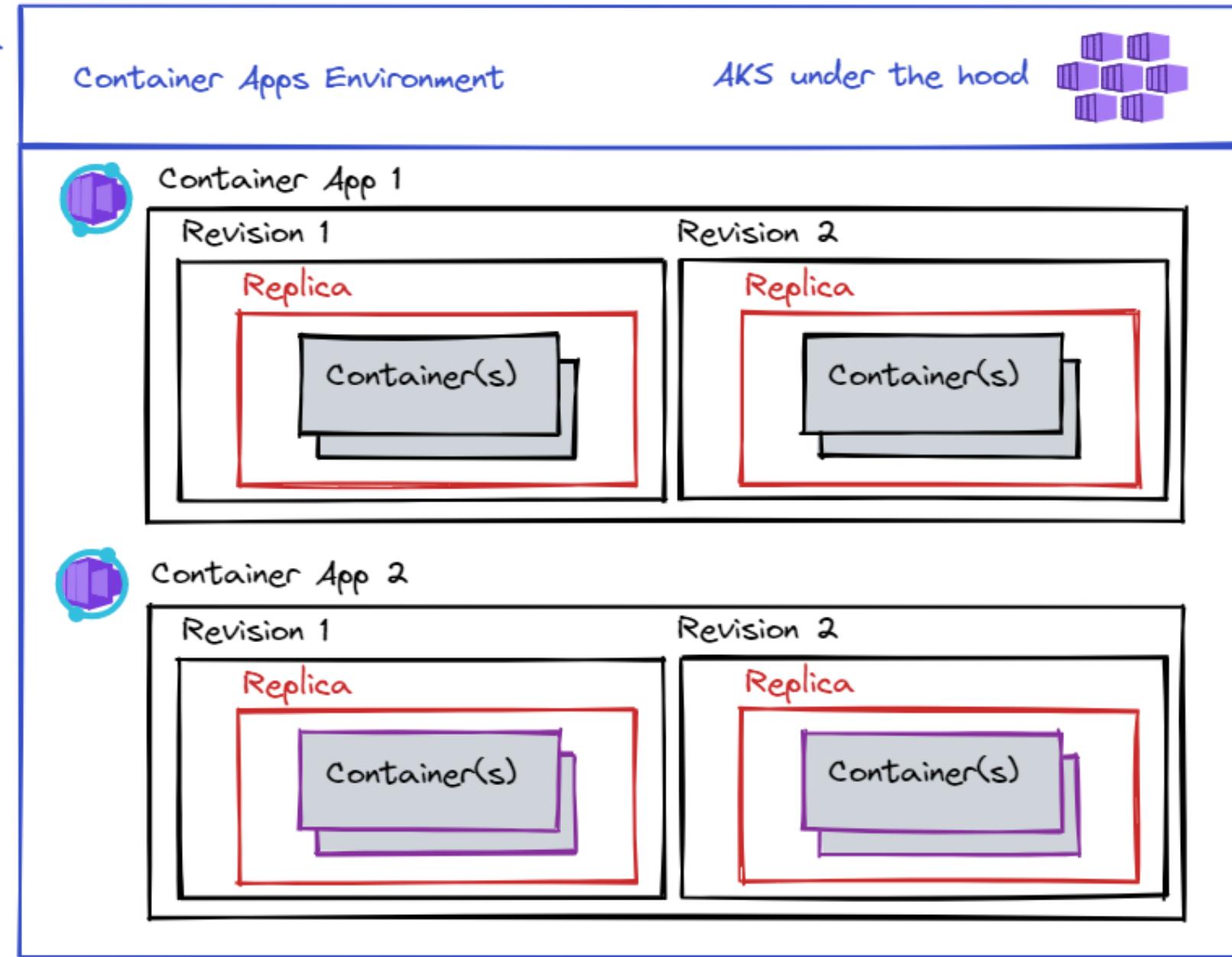
Microservices
Optional Dapr Integration



Scale based on any Keda scalers



Optional link
to VNET



Developer Productivity

```
az containerapp up -n appname -g rg \  
  --environment acaenv \  
  -l westeurope --source . \  
  --ingress external --target-port 8080
```





Automated Deployments

- ARM Templates
- Bicep
- Terraform
- Azure CLI



```
// deploy container apps environment
resource acaEnv 'Microsoft.App/managedEnvironments@2022-03-01' = {
    name: '${parPrefix}-aca-env'
    location: parLocation
    properties: {
        appLogsConfiguration: {
            destination: 'log-analytics'
            logAnalyticsConfiguration: {
                customerId: logAnalyticsWorkspace.properties.customerId
                sharedKey: listKeys(logAnalyticsWorkspace.id,
                    logAnalyticsWorkspace.apiVersion).primarySharedKey
            }
        }
        daprAIConnectionString: appInsights.properties.ConnectionString
    }
}
```

```
// deploy ui as a container app
resource webui 'Microsoft.App/containerApps@2022-06-01-preview' = {
    name: '${parPrefix}-webui'
    location: parLocation
    identity: {
        type: 'UserAssigned'
        userAssignedIdentities: {
            '${acaId.id}': {}
        }
    }
    properties: {
        managedEnvironmentId: acaEnv.id
        configuration: {
            activeRevisionsMode: 'Single'
            dapr: {
                ...
            }
            ingress: {
                ...
            }
            registries: [
                {
                    ...
                }
            ]
        }
        template: {
            containers: [
                {
                    image: '${acr.properties.loginServer}/openaiui:latest'
                    name: 'webui'
                }
            ]
        }
    }
}
```

```
    properties: {
      managedEnvironmentId: acaEnv.id
      configuration: {
        activeRevisionsMode: 'Single'
        dapr: {
          ...
        }
        ingress: {
          ...
        }
        registries: [
          {
            ...
          }
        ]
      }
      template: {
        containers: [
          {
            image: '${acr.properties.loginServer}/openaiui:latest'
            name: 'webui'
            env: [
              {
                name: 'INVOKE_URL'
                value: 'http://localhost:3500/v1.0/invoke/openai/method/generate'
              }
            ]
          }
        ]
      }
    }
  }
}
```

Tweet Generator

Text:

Kubernetes

Sentiment:

Funny

Submit

John Doe

Just now

"Kubernetes? More like Kuber-knees! You'll be on the floor in no time trying to figure this one out!" #Kubernetes #DevOps

Container App Environment - No VNET - powered by



Tweet Generator

Text:
Kubernetes

Sentiment:
Funny

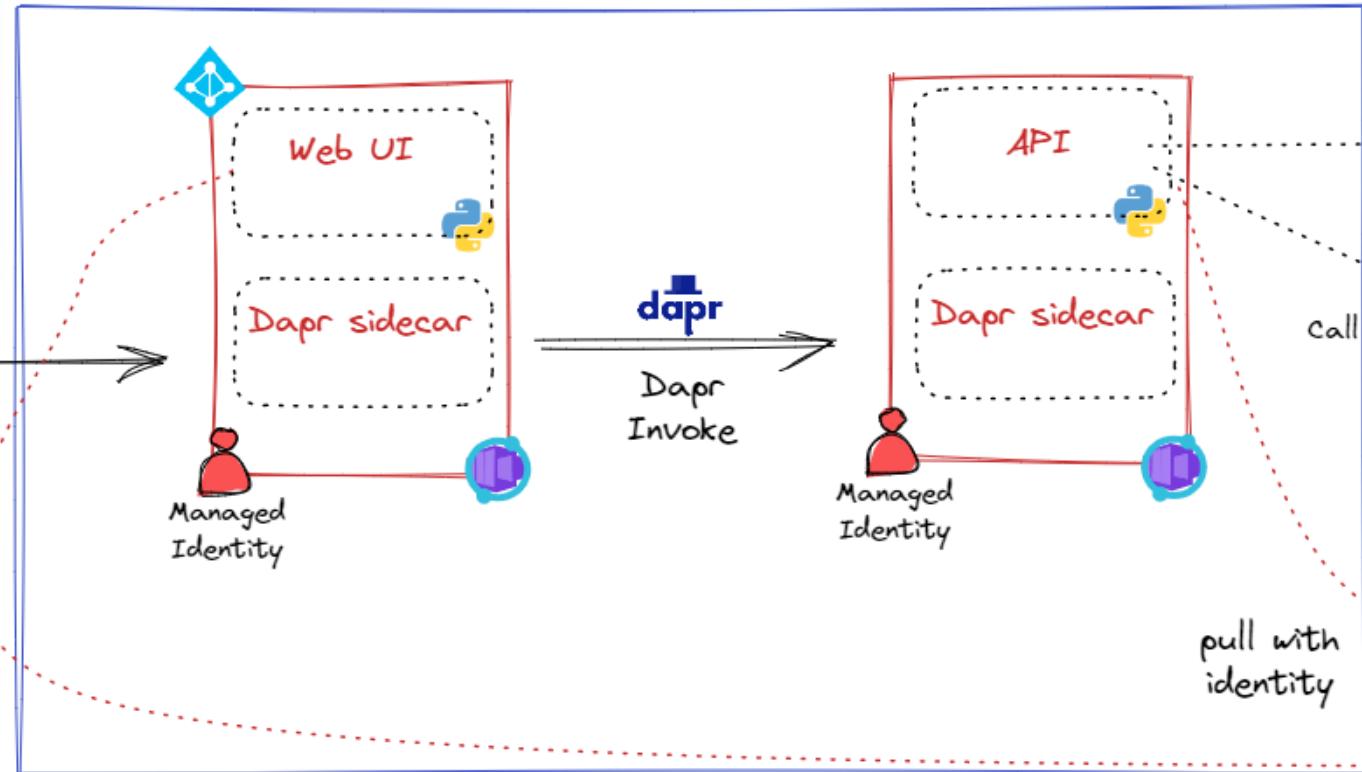
Submit

John Doe
Just now
Kubernetes: The 'Kube of Life' - it's like a puzzle, but you can fit so much more in the box!
#KubernetesHumor



HTTP(s)

Public IP



Linked to Log Analytics Workspace



Azure Key Vault



Retrieve Azure OpenAI secret

Azure OpenAI



call text-davinci-003 model

Azure Container Registry





Questions?