

Track 4 | **Cloud**

Building Serverless Platform on Private Cloud

최규황 (Samsung SDS)

Asaf Somekh (Founder & CEO, Iguazio)

AGENDA

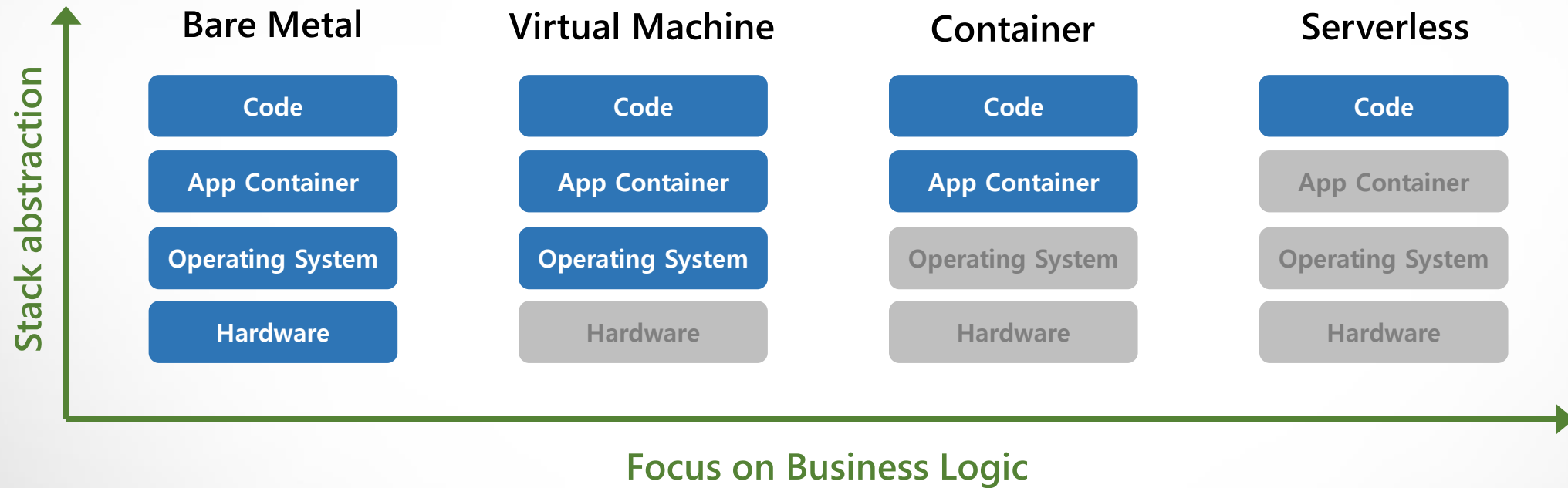
1. State of Serverless
2. Introducing Nuclio
3. Serverless Platform on Private Cloud

1

State of Serverless

The Evolution of Serverless

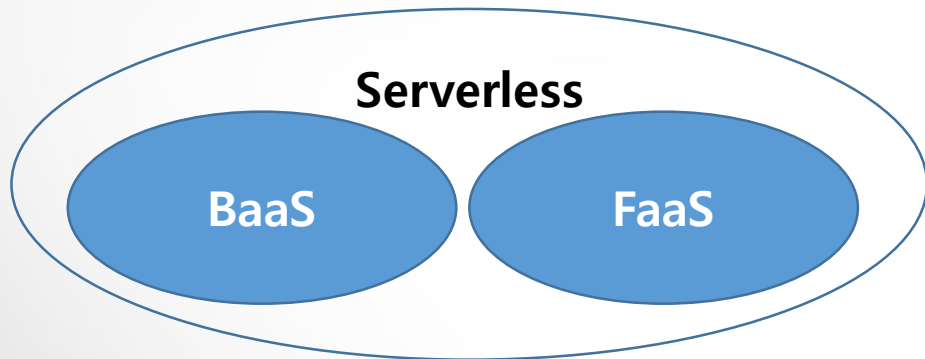
Allowing developers to focus on the business logic rather than the infrastructure



Serverless Architecture

서비스 개발자가 인프라를 관리할 필요 없이 애플리케이션과 서비스를 구축하고 실행하는 방식

Service type



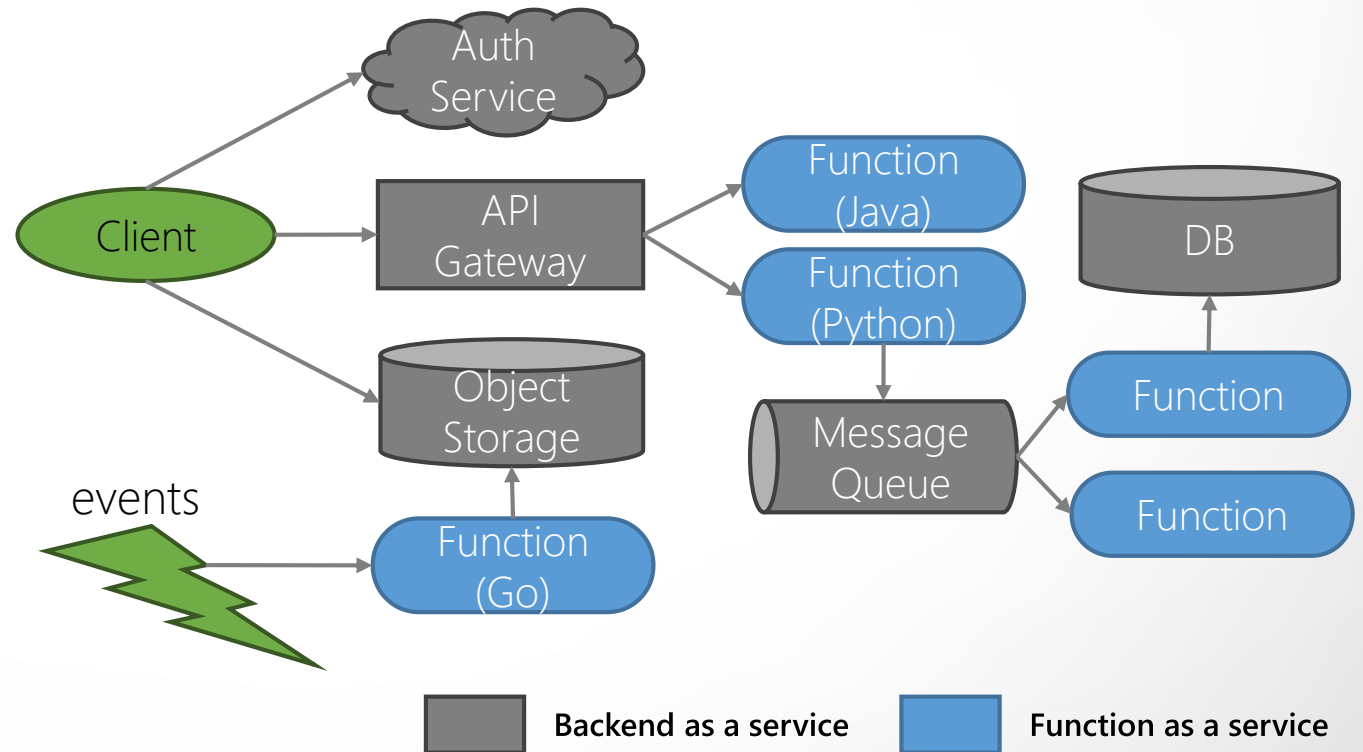
➤ BaaS (Backend as a Service)

- Queue, DB, 인증 등 Backend 기능을 Service 로 제공

➤ FaaS (Function as a Service)

- Function 형태의 애플리케이션을 개발하고 실행하는 환경을 서비스로 제공

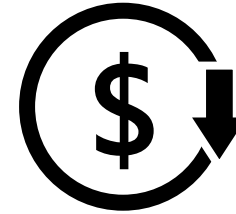
Example of Serverless Architecture



Why FaaS ?



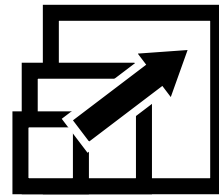
Fully managed service



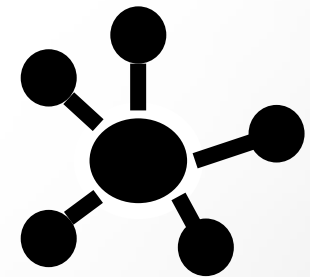
Only pay for
Resources you use



Enhance developer
productivity

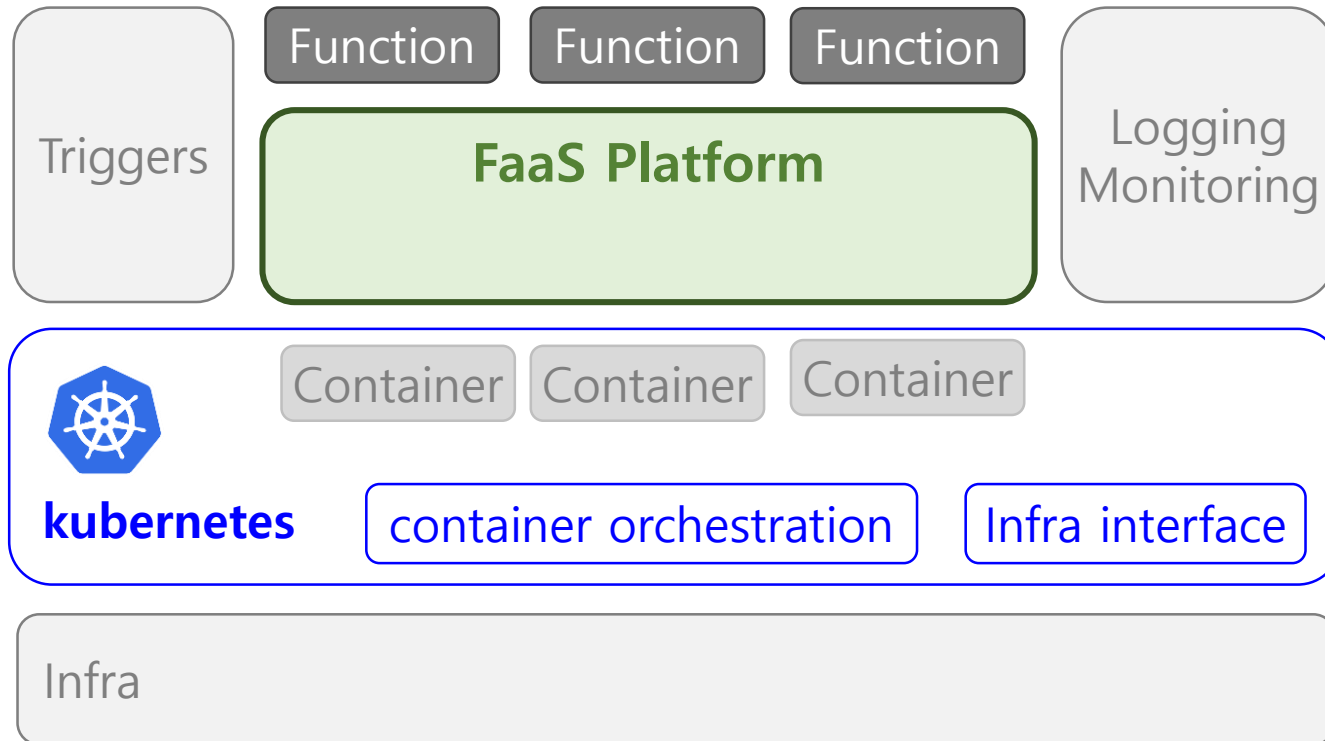


Scale flexibly



Seamless Integration

FaaS Platform on Kubernetes



- Kubernetes는 Platform을 위한 Platform
- Container orchestration
- Portability
- Kubernetes's scheduler, service discovery, networking, volume mount, ...
- Easy to Integrated with other Kubernetes based platform services

CNCF Serverless Landscape

CNCF Serverless Landscape
2019-11-11T03:46:35Z d8965f9

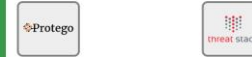
See the serverless interactive display at s.cncf.io

Greyed logos are not open source

Tools



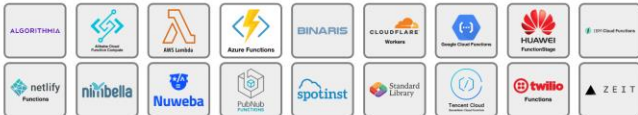
Security



Framework



Hosted Platform



Installable Platform



s.cncf.io

Serverless computing refers to a new model of cloud native computing, enabled by architectures that do not require server management to build and run applications. This landscape illustrates a finer-grained deployment model where applications, bundled as one or more functions, are uploaded to a platform and then executed, scaled, and billed in response to the exact demand needed at the moment.



Cloud Native Landscape



2

Introducing Nuclio

Iguazio's Serverless Open Source Project

Introducing Nuclio

High Performance Serverless Framework:

- **Fast**
 - High concurrency
 - Low latency
- **Open Source**
 - <https://nuclio.io/>
 - <https://github.com/nuclio/nuclio>
- **Built for Kubernetes**
- **Runs Everywhere**
 - Cloud , On-Prem, Edge, Hybrid, Laptop



- Focus on your app, automate maintenance

Nuclio lets you focus on the business logic and not the infrastructure

Nuclio and Kubernetes : A Great Match



- Abstraction, no Docker files or YAMLs
- Automated dev and ops flow
- Everything is a Kubernetes resource

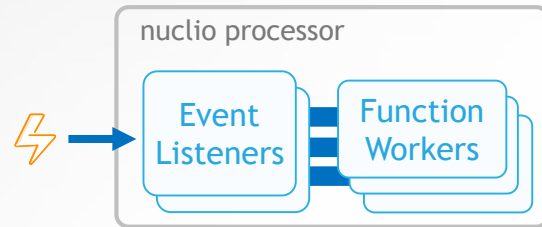


- Cloud independent APIs, on-prem, edge
- Auto-scaling + abstract infrastructure
- Fastest evolving features and community
- One platform for serverless and services

Introducing Nuclio – Iguazio’s Serverless Project

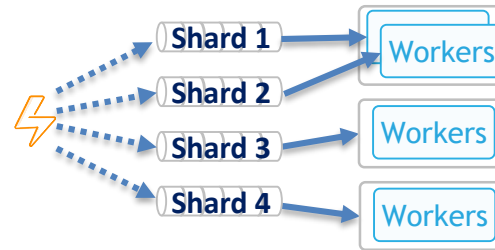
Taking Serverless to Data Intensive Applications

Extreme Performance



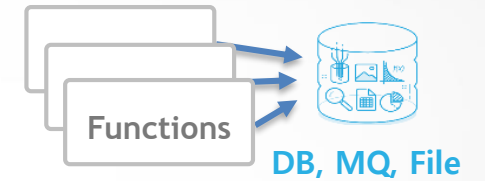
- Non-blocking, parallel
- Zero copy, buffer reuse
- Up to 400K events/sec/proc
- GPU optimizations

Advanced Data & AI Features



- Auto-rebalance, checkpoints
- Any source: Kafka, NATS, Kinesis, event-hub, iguazio, pub/sub, Rabbit MQ, Cron, ..
- Jupyter, NVIDIA Rapids integration

Statefulness



- Data bindings
- Shared volumes
- Context cache

Join the Nuclio community at

<https://nuclio.io/>

<https://github.com/nuclio/nuclio>

Accelerate ETL and Streaming with Nuclio

Simple code, automated DevOps, any source!
(e.g. read JSON Stream + aggregate + dump to Parquet)

```
def init_context(context):  
    os.makedirs(sink, exist_ok=True)  
  
def handler(context, event):  
    add_log_to_batch(context, event.body)  
  
    if len(batch) > batch_len:  
        df = _batch_to_df(context)  
        if not df.empty:  
            df = df.groupby(['log_ip']).agg({'feconn': 'mean',  
                                           'beconn': 'mean',  
                                           'time_backend_response': 'max',  
                                           'time_backend_response': 'mean',  
                                           'time_queue': 'mean',  
                                           'time_duration': 'mean',  
                                           'time_request': 'mean',  
                                           'time_backend_connect': 'mean'  
                                           })  
  
        df_to_parquet(df)  
        reset_batch()
```

**28X
Faster!**

Python
with
Nuclio

500 MB/s



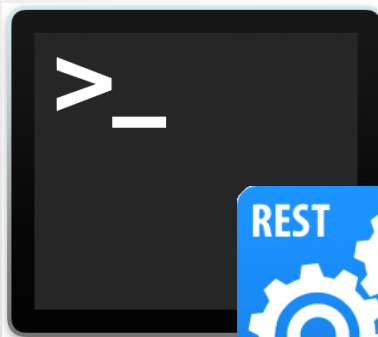
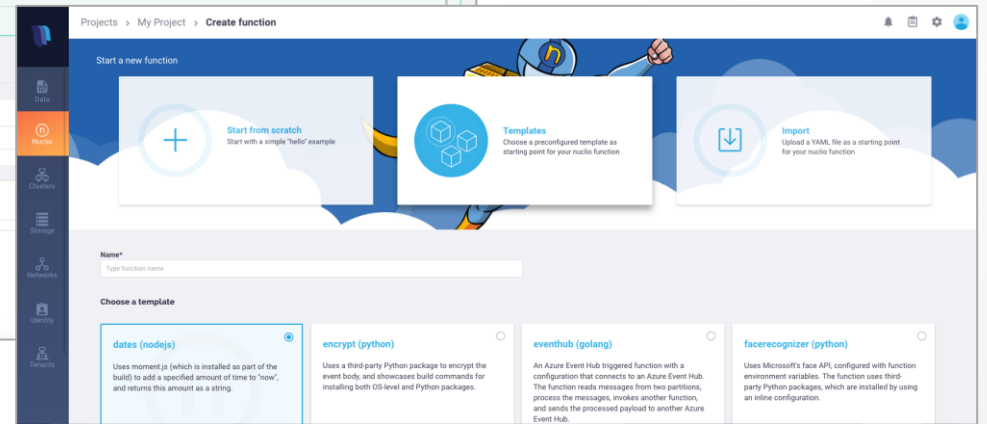
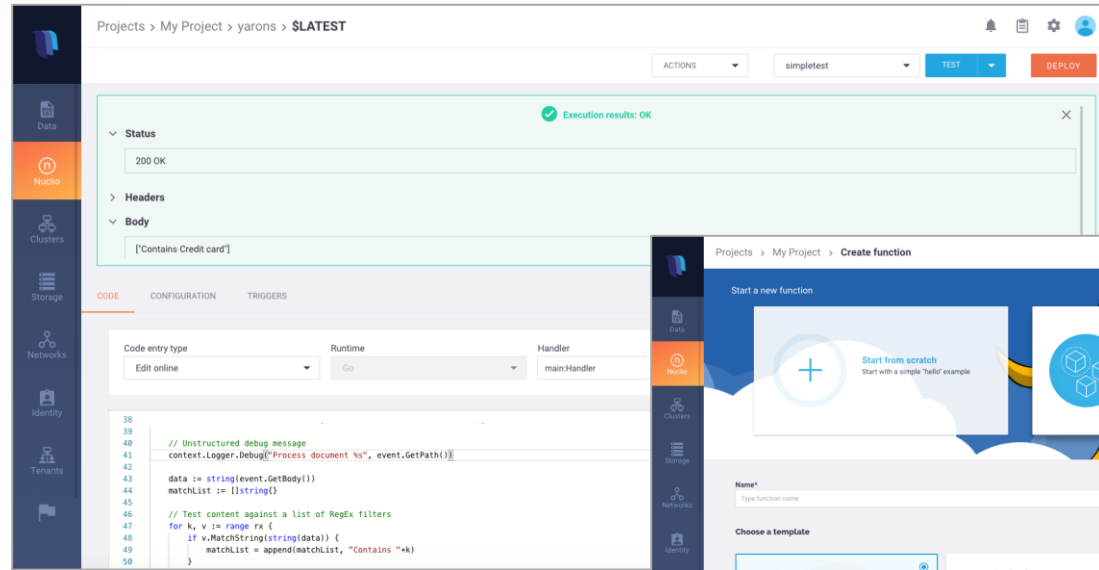
Simple
Python

18 MB/s

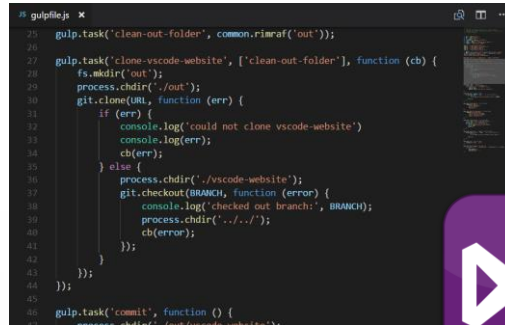
Simple Development and Operation Experience

Self-Service UI

- Code, debug
- Configure
- Manage
- App market



CLI and
REST API



VS Code
plug-in



Central Logging
and Monitoring



Kubeflow – ML Pipelines Portability on Kubernetes



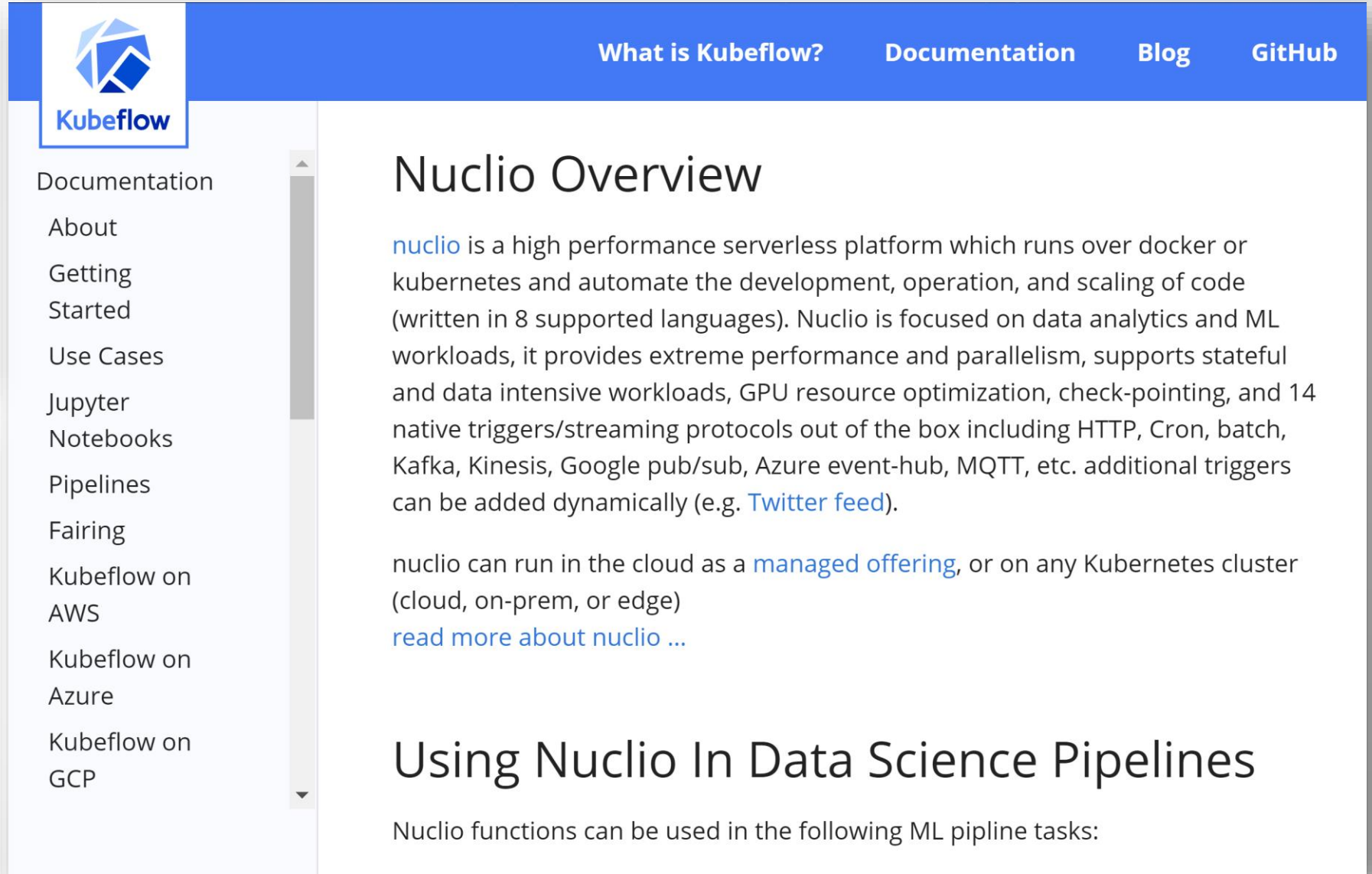
Making deployments of ML workflows on Kubernetes simple, portable and scalable

Strong industry momentum

Already integrated in Iguazio Data Science Platform



Kubeflow Adopts Nuclio for Serverless ML Pipelines

A screenshot of the Nuclio website. The top navigation bar is blue with white text for "What is Kubeflow?", "Documentation", "Blog", and "GitHub". The left sidebar is white with a blue "Kubeflow" logo and a list of links: "Documentation", "About", "Getting Started", "Use Cases", "Jupyter Notebooks", "Pipelines", "Fairing", "Kubeflow on AWS", "Kubeflow on Azure", and "Kubeflow on GCP". The main content area has a white background with a blue header "Nuclio Overview". The text describes Nuclio as a high-performance serverless platform for data analytics and ML, listing various triggers and protocols. It also mentions that Nuclio can run in the cloud as a managed offering or on any Kubernetes cluster. At the bottom, there is a section titled "Using Nuclio In Data Science Pipelines" with the text "Nuclio functions can be used in the following ML pipeline tasks:".

Nuclio Overview

[nuclio](#) is a high performance serverless platform which runs over docker or kubernetes and automate the development, operation, and scaling of code (written in 8 supported languages). Nuclio is focused on data analytics and ML workloads, it provides extreme performance and parallelism, supports stateful and data intensive workloads, GPU resource optimization, check-pointing, and 14 native triggers/streaming protocols out of the box including HTTP, Cron, batch, Kafka, Kinesis, Google pub/sub, Azure event-hub, MQTT, etc. additional triggers can be added dynamically (e.g. [Twitter feed](#)).

nuclio can run in the cloud as a [managed offering](#), or on any Kubernetes cluster (cloud, on-prem, or edge)
[read more about nuclio ...](#)

Using Nuclio In Data Science Pipelines

Nuclio functions can be used in the following ML pipeline tasks:

New Capabilities in Nuclio

Extending Nuclio Beyond Functions

Nuclio Serverless Applications

Pre-tested Application & Notebook Templates for accelerating development and integration

Real-time functions



ML Jobs



Model Serving

Adding Nuclio ML Jobs

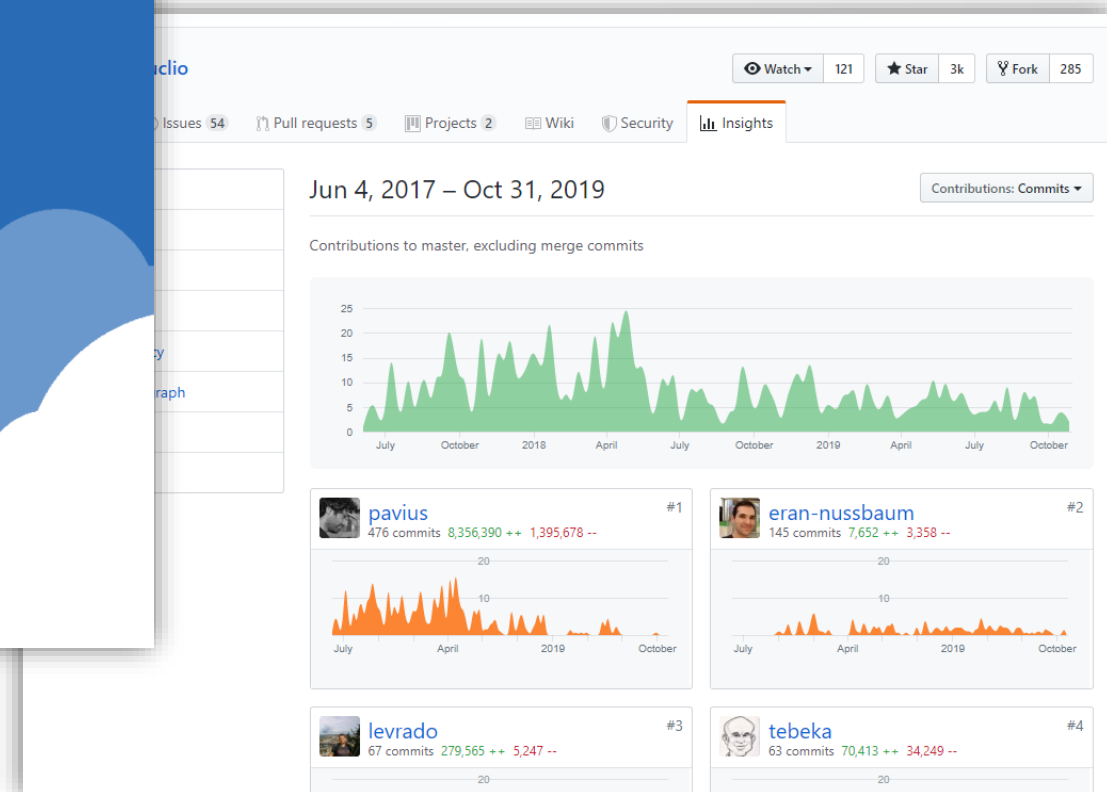
- Extending Nuclio architecture to support elastic ML Jobs (Spark, Dask, Horovod, ..)
- Zero DevOps from code to deployment
- Maximum performance and scalability
- Integrated logging, monitoring, templates, ..
- Glue-less integration with KubeFlow & Iguazio DB

Addressing The Entire Lifecycle of ML Pipelines with Serverless

Join the Nuclio Open Source Serverless Community



<https://nuclio.io/>



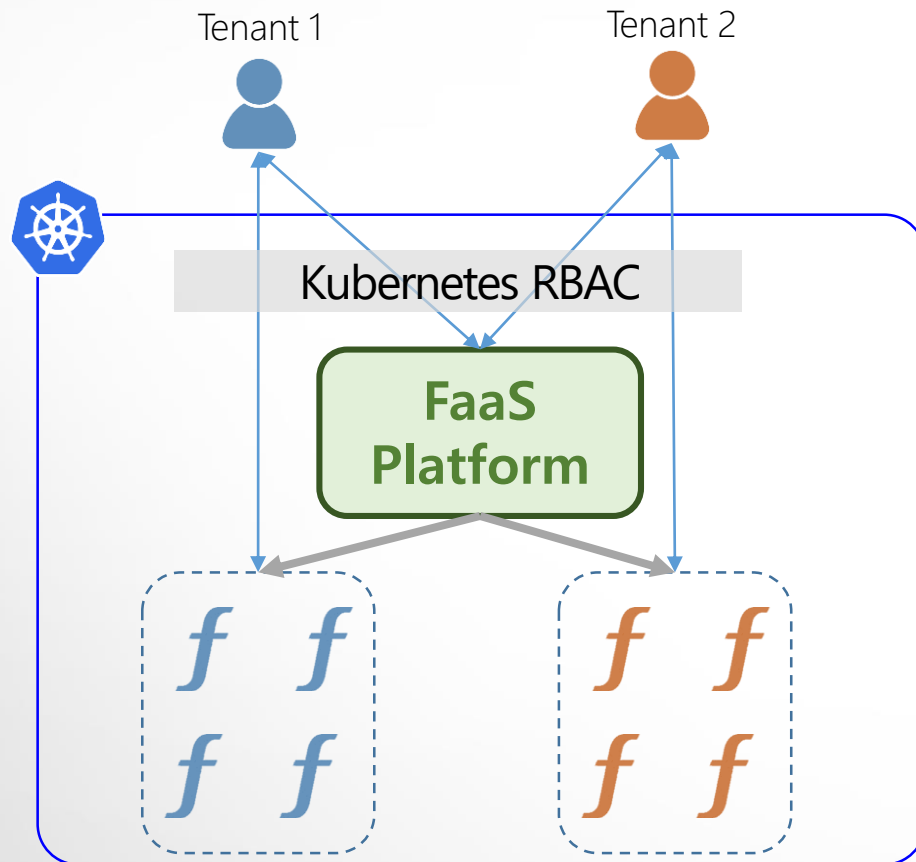
<https://github.com/nuclio/nuclio>

3

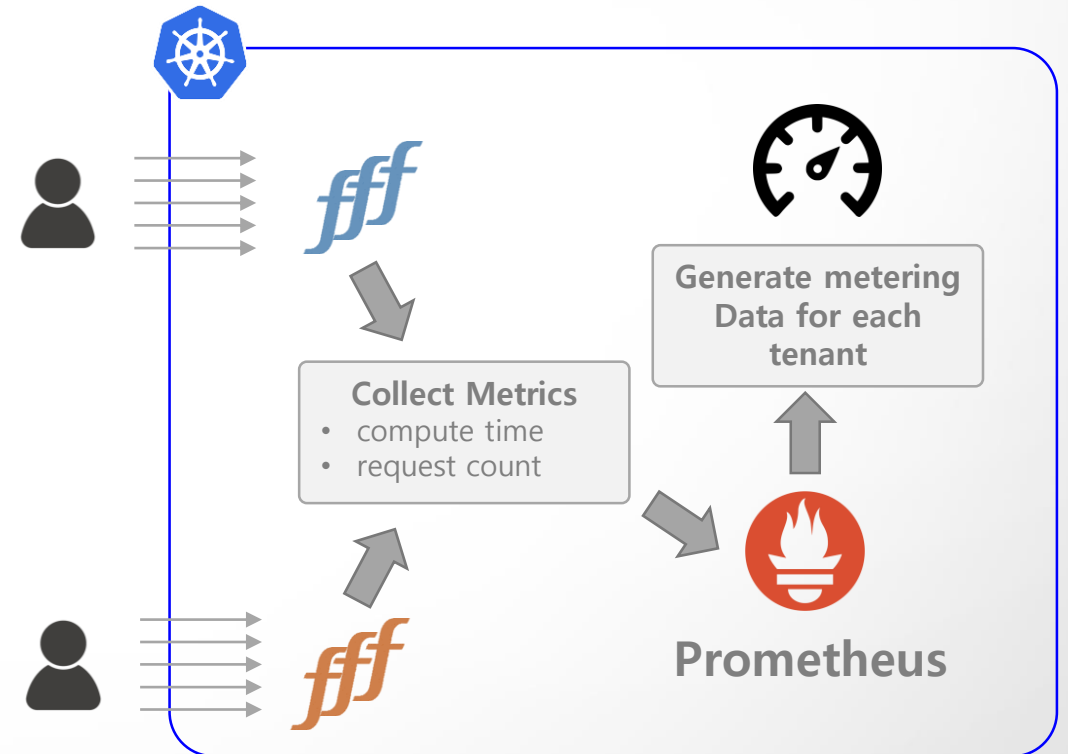
Serverless Platform on Private cloud

Challenges – On Private Cloud

1. Multi-tenancy

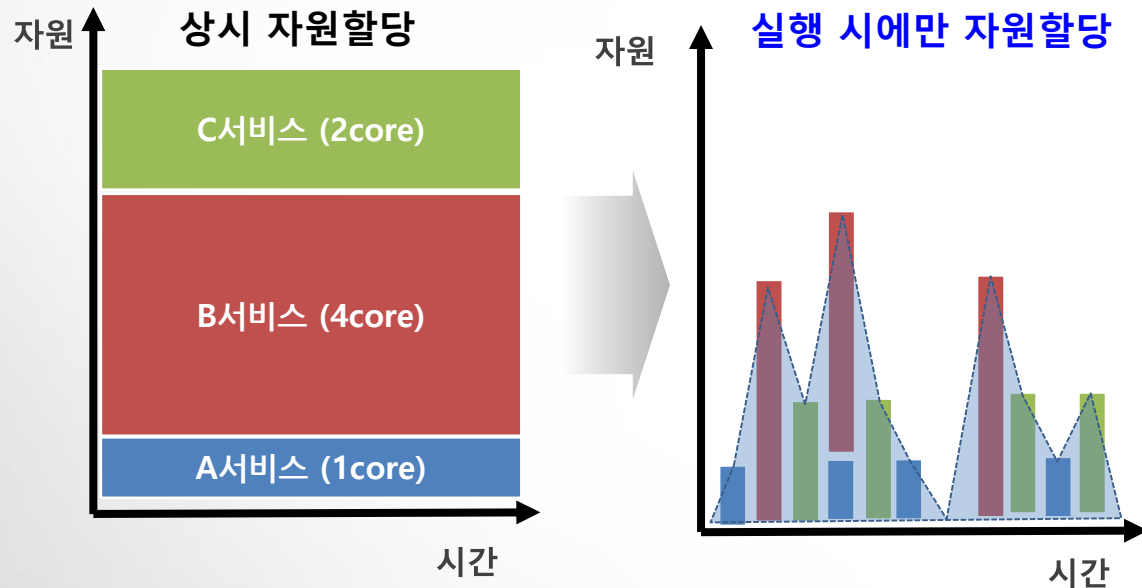


2. Metering



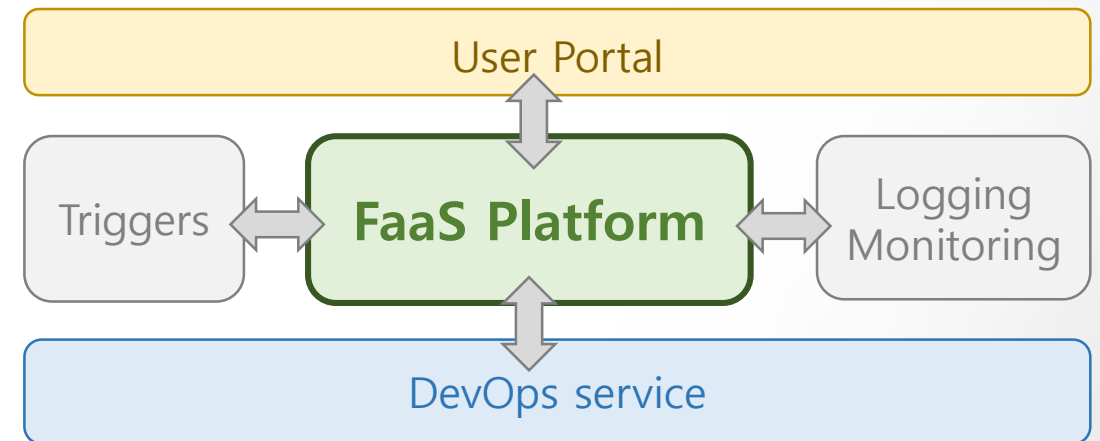
Challenges – On Private Cloud

3. Advanced Scheduling



- Usage based scale (cap setting required)
- Scale to zero when idle (cold start ?)

4. Integrate with existing system



- Air-gapped environment
- Portal integration for self-service

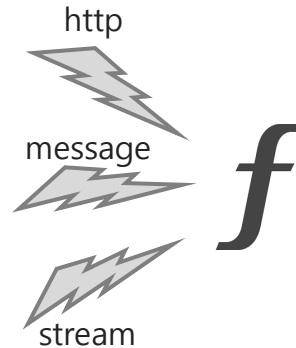
SDS Functions

SDS Function 서비스는 클라우드에서 함수 형태의 애플리케이션을 빠르게 만들어 실행하기 위한 솔루션으로, Python, Java, Javascript, Shell 개발 언어를 지원하며, 코드를 실행하기 위해 필요한 자원이 동적으로 할당하고 규모에 따라 유연하게 확장됨



코드 업로드

- Function 코딩/빌드/배포를 위한 콘솔
- 재활용 가능한 Function Template 제공
- Python, Java, Shell, Javascript



이벤트 설정
(Function 실행조건)

- Http, RabbitMQ, MQTT, NATS
- Kafka, Cron

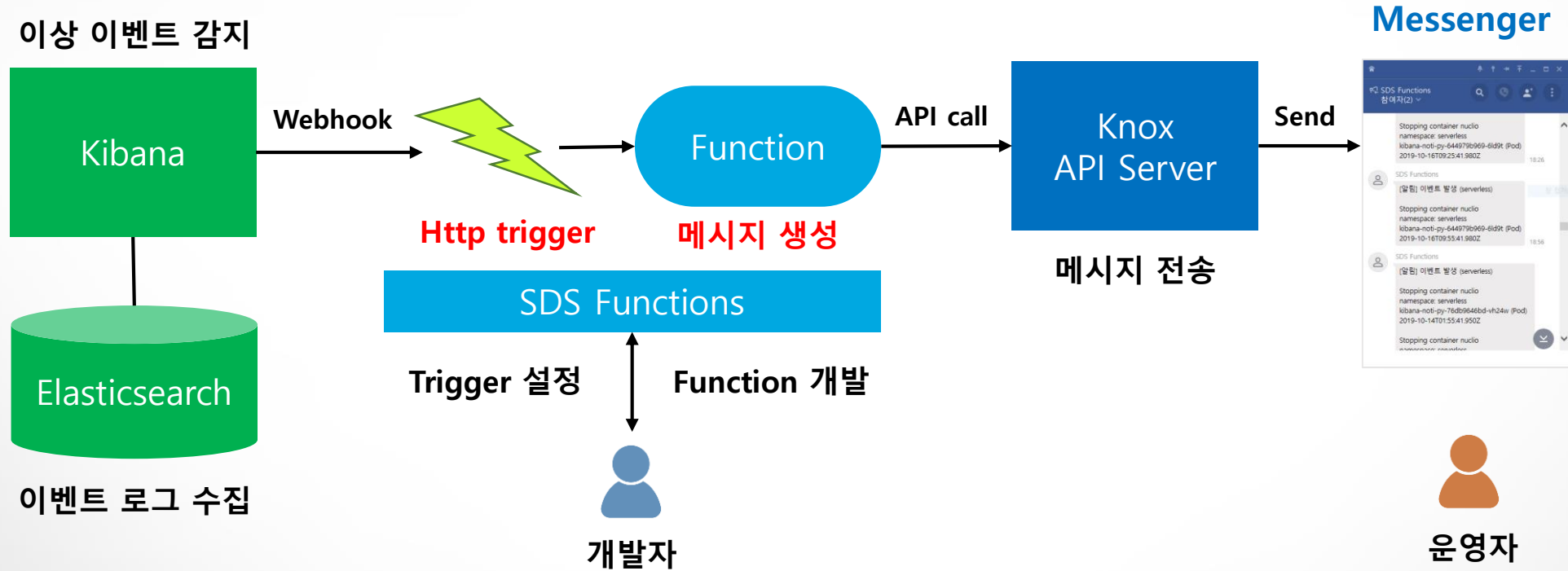


Function 실행
(규모에 따라 확장)

- Auto-scaling
- Support Cold start

이벤트 알림 서비스

사내 메신저 연계 **template** 을 활용 컨테이너 운영 시 이상 이벤트를 메신저로 알림



Road map

Knative



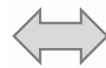
cloudevents spec version 1.0
(Oct 25,2019)



Knative build integration

Enhance Monitoring, Logging

Monitoring



Logging



elasticsearch

Serverless ML Pipeline



Integrate with other platform services

API G/W

Queue

DB

events





Thank You



The graphic features the text 'Q & A' in a clean, sans-serif font. The 'Q' and 'A' are white, while the ampersand is a vibrant lime green. To the right of the text, two orange triangles point towards a large, dark blue circle on the right side of the frame. The background is a dark blue gradient with a pattern of lighter blue circles on the left side.

Q & A

Partner Disrupt Foresee