

UFG
UNIVERSIDADE
FEDERAL DE GOIÁS

**Apache Flink para processamento
em tempo real de dados IoT na
nuvem**

Glossário

Glossário

IoT (Internet das Coisas):

- Expansão para objetos do dia a dia, não se limitando apenas a computadores.

Aplicações Vislumbradas para IoT:

- Relação com o conceito de "Big Data", envolvendo grande volume de dados gerados e coletados continuamente.

Apache Flink:

- Mecanismo de código aberto para processamento distribuído de dados ilimitados e limitados.
- Foco em streaming, baixa latência, execução de cálculos na memória e alta disponibilidade.

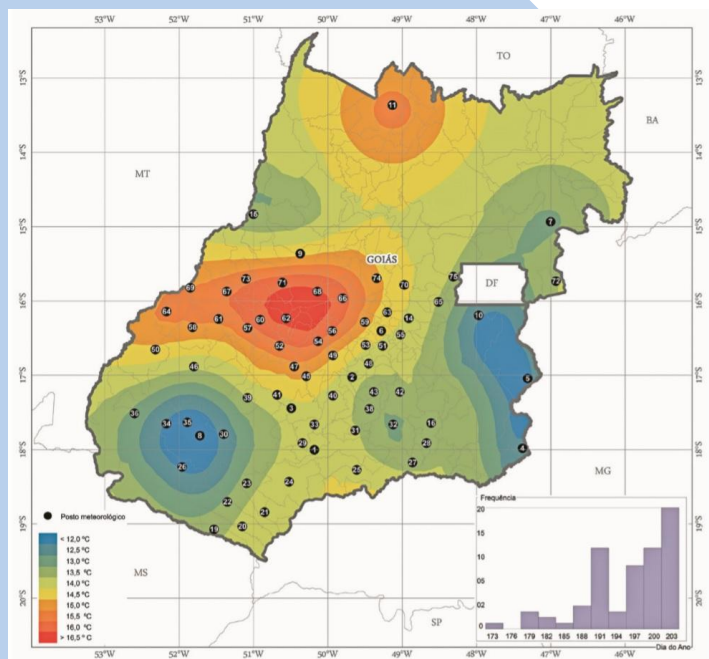


Proposta de Solução

Características

Objetivo

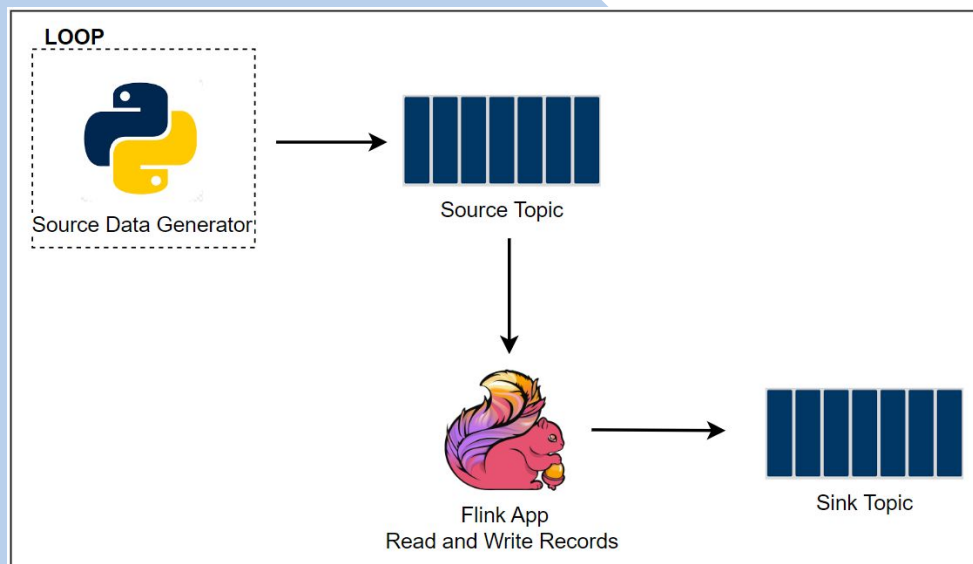
A construção de um sistema em tempo real para análise e geração de métricas baseadas em dados de temperatura residenciais, coletados periodicamente por sensores



Metodologia

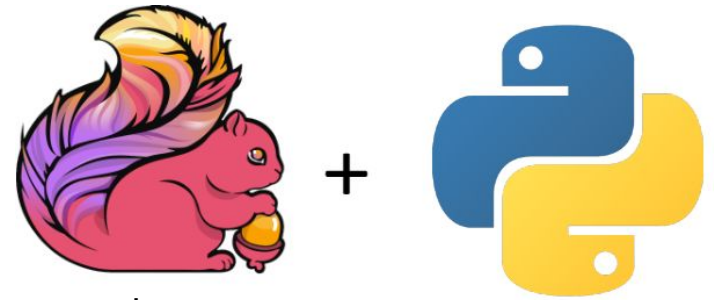
Aplicação

A construção de um pipeline na biblioteca de Python, PyFlink, que seja capaz de receber os dados coletados, analisá-los sistematicamente, gerar métricas em cima das coletas e ser capaz de apresentar dashboards de simples entendimento para informar o usuário da interface.



Apache Flink

Características



Tolerância a Falhas no Apache Flink:

- Mecanismos: pontos de verificação automáticos e sob demanda.
- Assincronia: Captura de snapshot sem interromper o processamento

Benefícios:

- Processamento de conjuntos ilimitados/limitados na mesma interface.
- Execução em grande escala, paralela em várias máquinas.
- Desempenho na memória, cálculos acessando dados locais.
- Consistência de estado exatamente uma vez após falhas.
- Conectores variados para sistemas populares.
- Abstração flexível na interface de programação.

Intersecção entre o Apache Flink

Aplicações comuns:

- Aplicações de análise de dados.
- Aplicações de pipelines de dados, transformando dados entre armazenamentos.
- Captura de picos de temperatura em tempo real por meio de sensores.

Conexão entre IoT e Apache Flink:

- Viabilidade da IoT relacionada ao avanço específico do protocolo da internet, atribuindo a cada equipamento um "endereço IP" próprio.
- Apache Flink desenvolvido para captura de dados em tempo real e processamento contínuo, destaca-se em ambientes conectados a sensores.

Apache Flink como Mecanismo de Processamento de Fluxo:

- Aplicações consistem em gráficos de fluxo de dados
- Ingestão de dados de uma ou mais fontes para um ou mais destinos.

IoT

Características

Programação:

- Apache Flink pode ser programado em várias linguagens, incluindo SQL de streaming, Python, Java, e outras linguagens JVM como Kotlin.

Funcionalidades e Tecnologias na IoT:

- Surgimento de novas expressões como smart buildings, smart cities, smart transport, smart grid, smart energy, smart health, entre outras.

IoT e Ambiente de Local

Atributos dos Objetos na IoT:

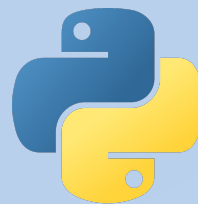
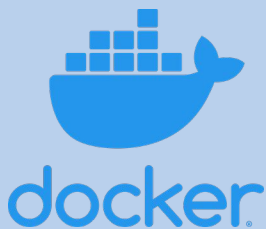
- Conectar objetos a uma rede potencializa novas aplicações, criando a Internet das Coisas.

Arquitetura Básica dos Objetos na IoT:

- Quatro unidades: processamento/memória, comunicação, energia, sensores/atuadores.

Instalação de Tecnologias Localmente:

- Instalação de Python, Docker e Apache Flink.
- Docker atua como camada de virtualização, facilitando a configuração do ambiente de desenvolvimento.



Apache Flink para processamento em tempo real de dados IoT na nuvem

Processamento de Dados com PyFlink:

- PyFlink é usado para processar dados de temperatura.
- O script Python `temperature_msg.py` mapeia registros, calcula a soma dos valores das transações para cada ID.

Configuração na Amazon EC2:

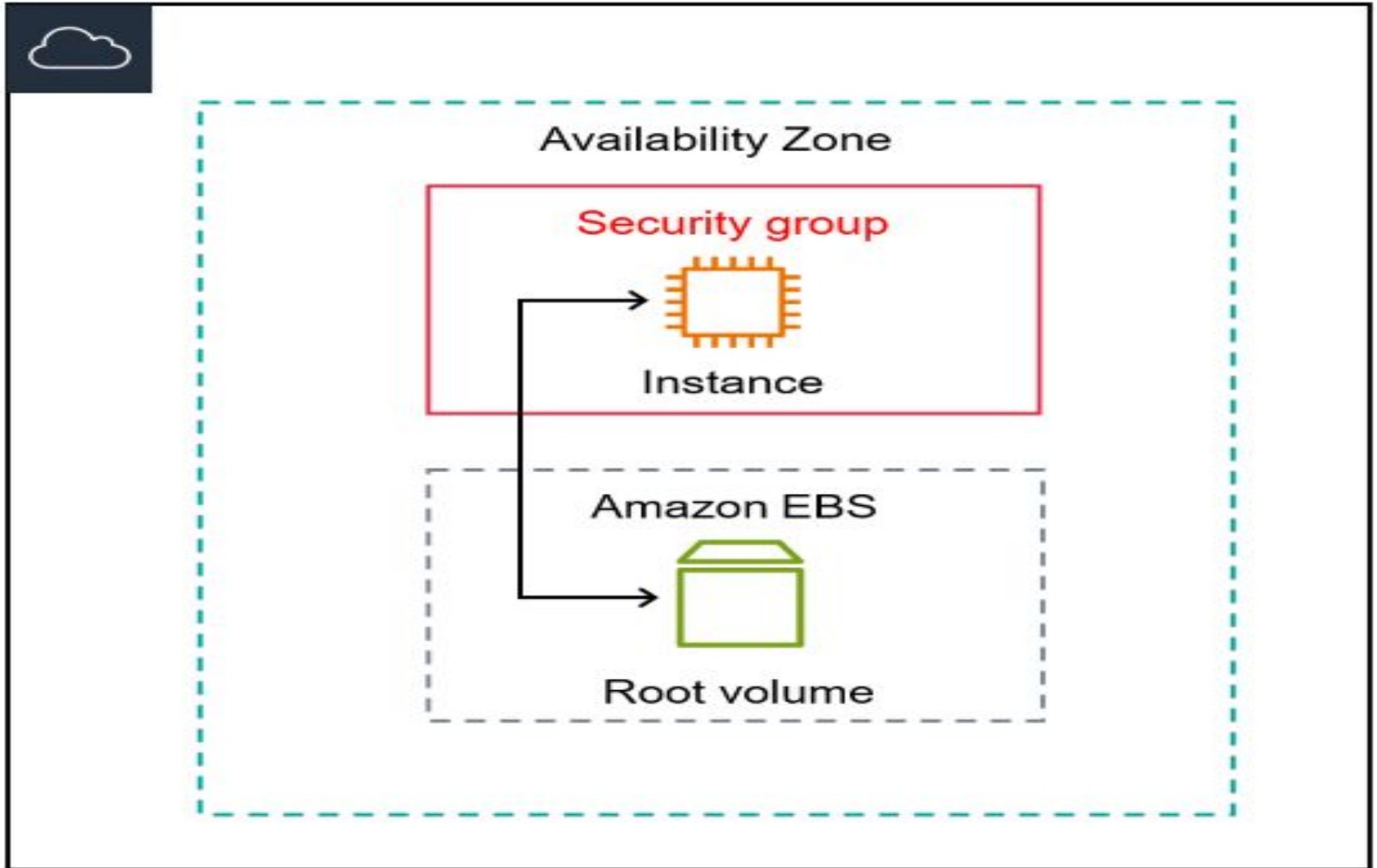
- Construção de uma máquina virtual na Amazon EC2 com uma imagem Docker.
- Download de imagens das tecnologias mencionadas.

Resultados e Conclusões

Dados obtidos e Oportunidades

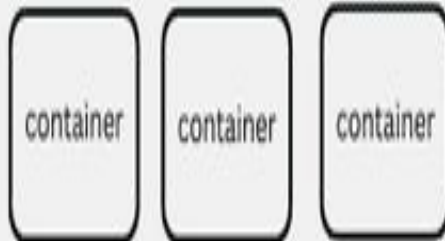
- Profunda quantidade amostral de dados obtida para geração de métricas importantes.
- Capacidade de análise aprofundada do Apache Flink para processamento em tempo real de dados IoT na nuvem.
- Resultados satisfatórios e métricas em conformidade com o espaço amostral, com eficácia de 100%.
- Oportunidade tecnológica para cenários que demandam processamento sem a necessidade de recursos locais.

Projeto Prático





Docker



Container Software

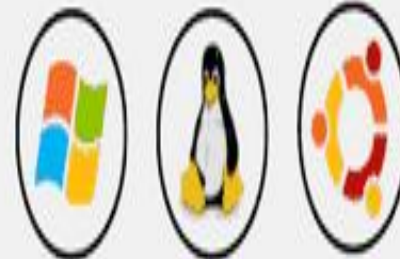
Operating System

Infrastructure
Host Server(s)



Virtual Machine

VM1 VM2 VM3



Hypervisor

Infrastructure
Host Server(s)



Overview

Jobs ^

Running Jobs

Completed Jobs

Task Managers

Job Manager

Submit New Job

Available Task Slots

1

Total Task Slots 2 | Task Managers 1

Running Jobs

1

Finished 0 | Canceled 0 | Failed 0

Running Job List

Job Name	Start Time	Duration	End Time	Tasks	Status
insert-into_default_catalog.default_database.es_sink	2020-09-28 10:29:40	7s	-	2 / 2	RUNNING

Completed Job List

Job Name	Start Time	Duration	End Time	Tasks	Status
----------	------------	----------	----------	-------	--------

 Home

Recently viewed >

 Kibana 

Discover

Dashboard

Canvas

Maps

Machine Learning

Visualize

 Observability 

Logs

Metrics

APM

Uptime

 close

page and improve our products and services, see our [Privacy Statement](#). To stop collection, [disable usage data here](#).

Dashboards

 Create dashboard

 Search...

 Title	Description	Actions
 payment_dashboard		

Rows per page: 20 



< 1 >

Help us improve the Elastic Stack

To learn about how usage data helps us manage and improve our products and services, see our [Privacy Statement](#). To stop collection, [disable usage data here](#).



Dismiss

Full screen Share Clone Edit


 

Search

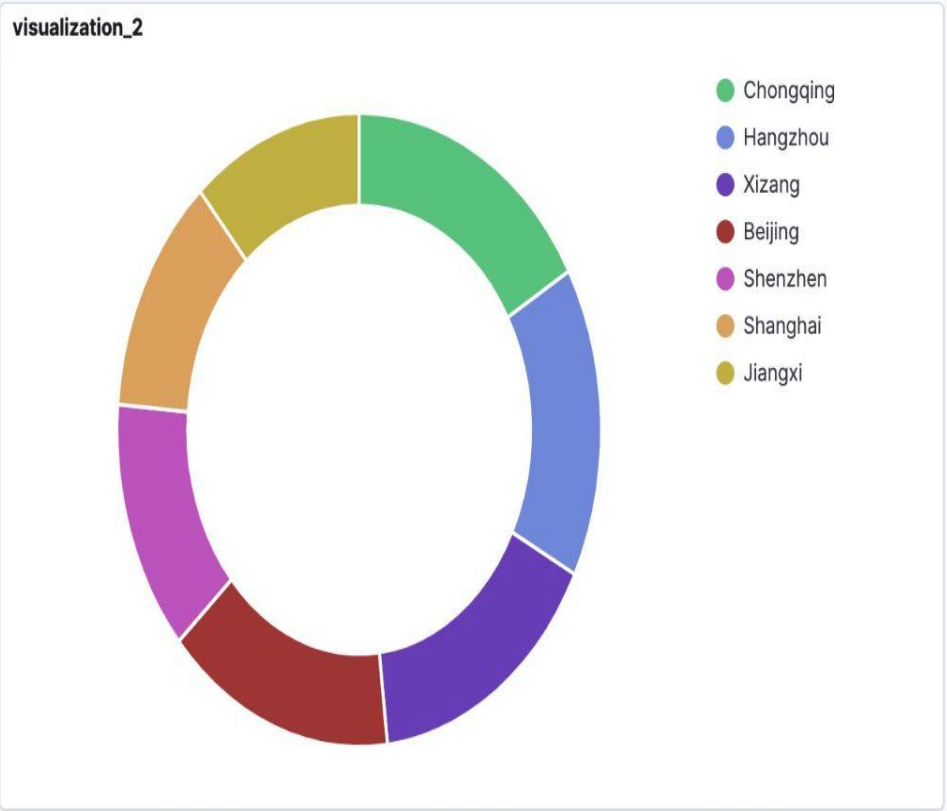
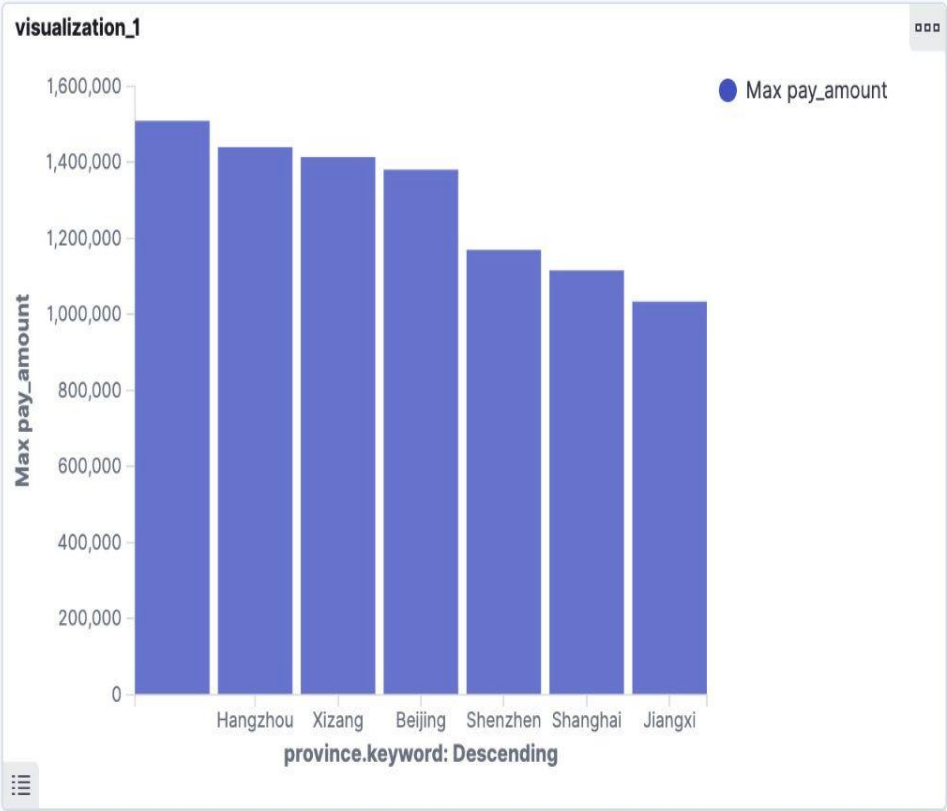
KQL

  Last 15 minutes

Show dates

 Refresh

+ Add filter



REFERÊNCIAS

- [1] Faccioni Filho, Mauro. "Internet das coisas." Unisul Virtual(2016)
- [2]<https://aws.amazon.com/pt/what-is/apache-flink/>
- [3] Santos, Bruno P., et al. "Internet das coisas: da teoria a prática." 'Minicursos SBRC-Simposio Brasileiro de Redes de Computadores e 'Sistemas Distribuidos 31 (2016): 16.
- [4] Santaella, Lucia, et al. "Desvelando a Internet das coisas." Revista GEMInIS 4.2 (2013): 19-32.
- [5]<https://github.com/apache/flink-playgrounds/tree/master/pyflinkwalkthrough>
- [6]<https://thingspeak.com/channels/public>
- [7]<https://medium.com/@daeynasvistas/a-iot-internet-das-coisas-surgiucomo-a-nova-gera>
- [8]<https://www.confluent.io/blog/apache-flink-stream-processing-use-caseswith-examples>

Obrigado(a)!