Different Ways of Working with Kafka on Azure





L. Sud

Z

(arrest.

22

des and



XPOAR

co de México

82

AISEI

Pesos

Pesos

5

周刊

Veinte Pesos

(A)

Sos

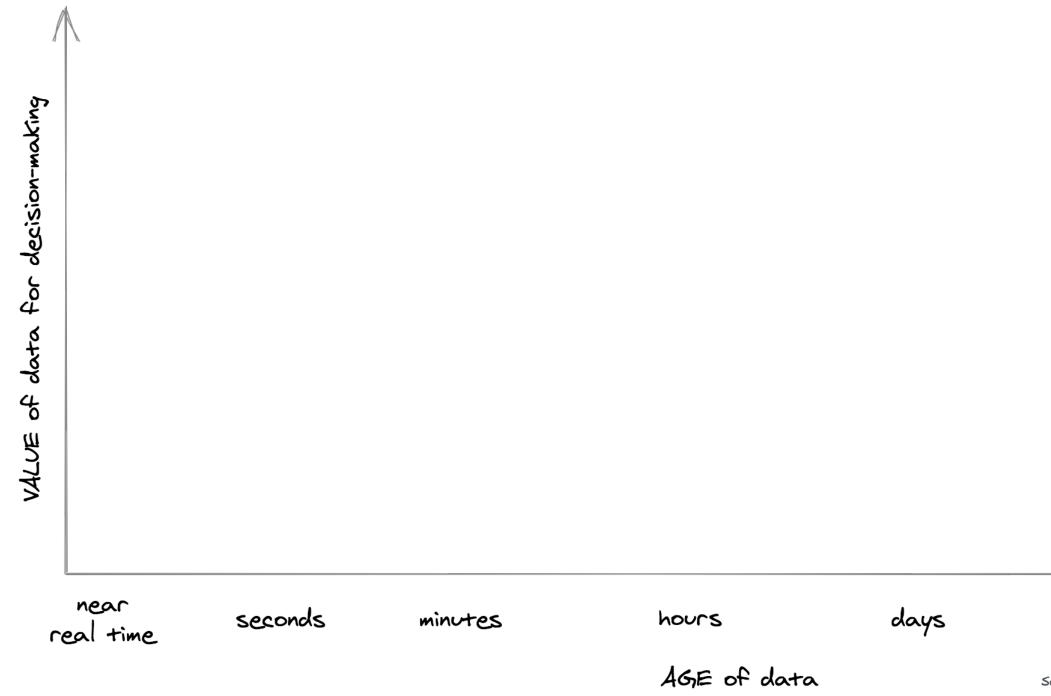
ollars







Diminishing Value of Data

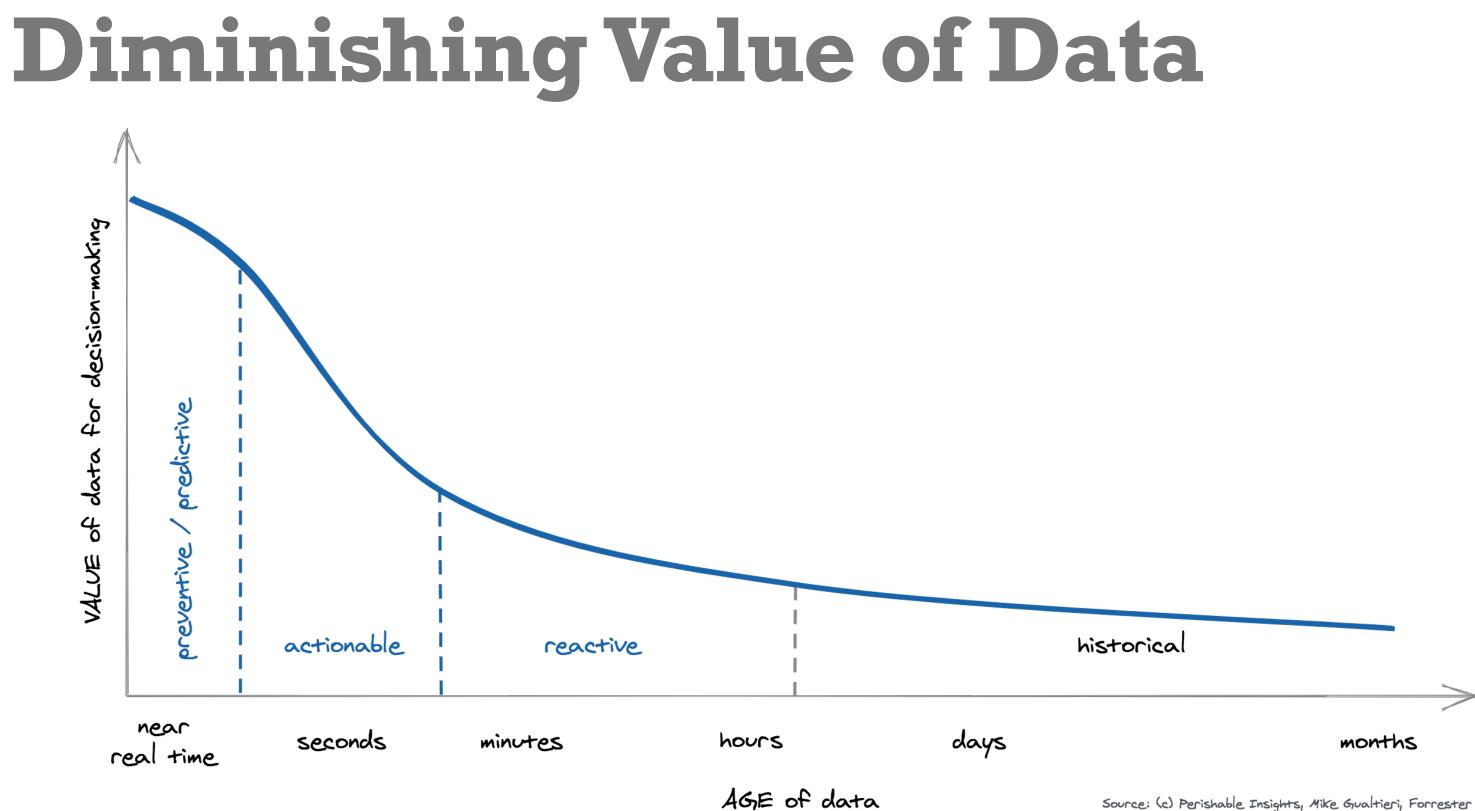


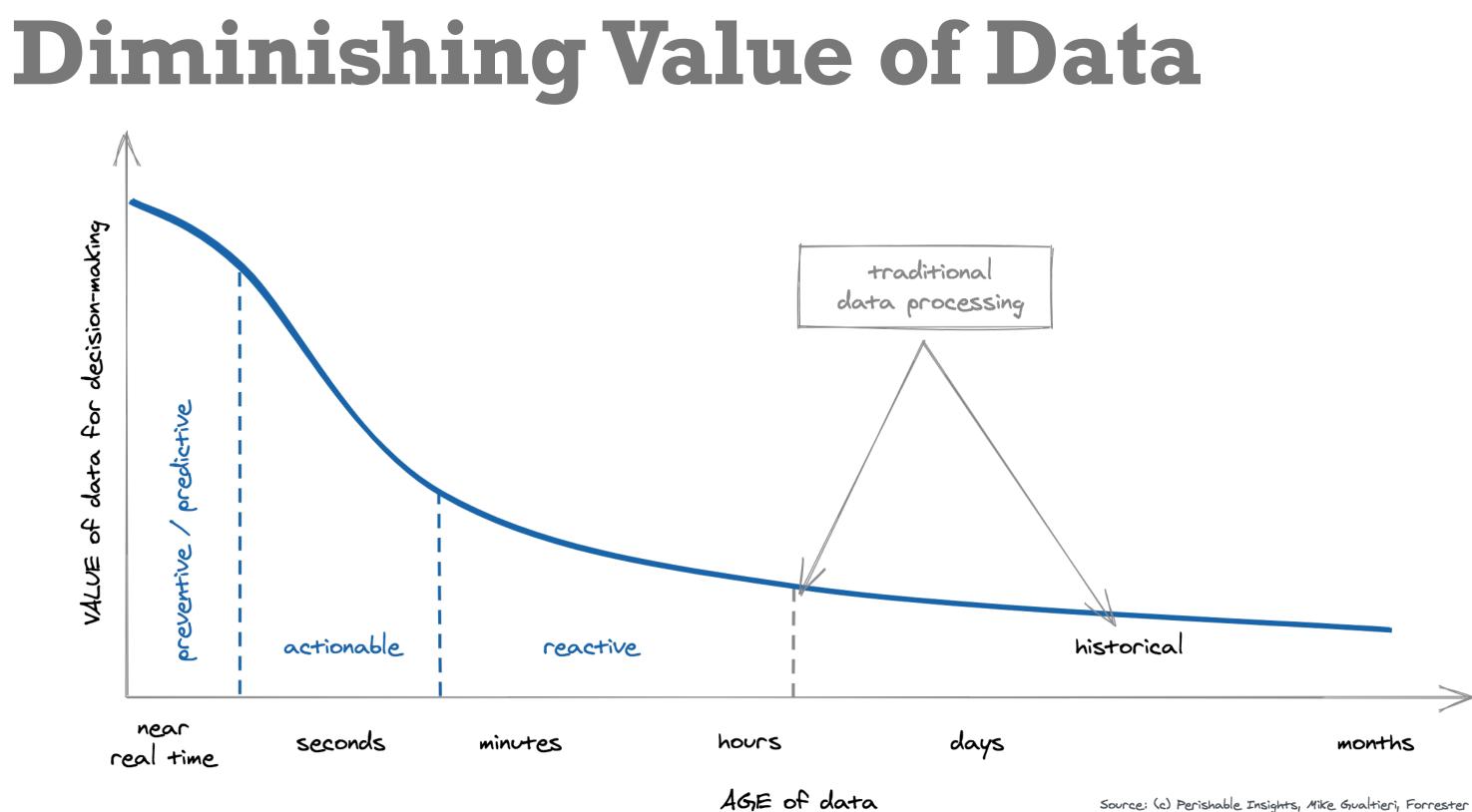
@hpgrahsl | @Azure #GlobalAzure, 16th April 2021, Linz - Austria



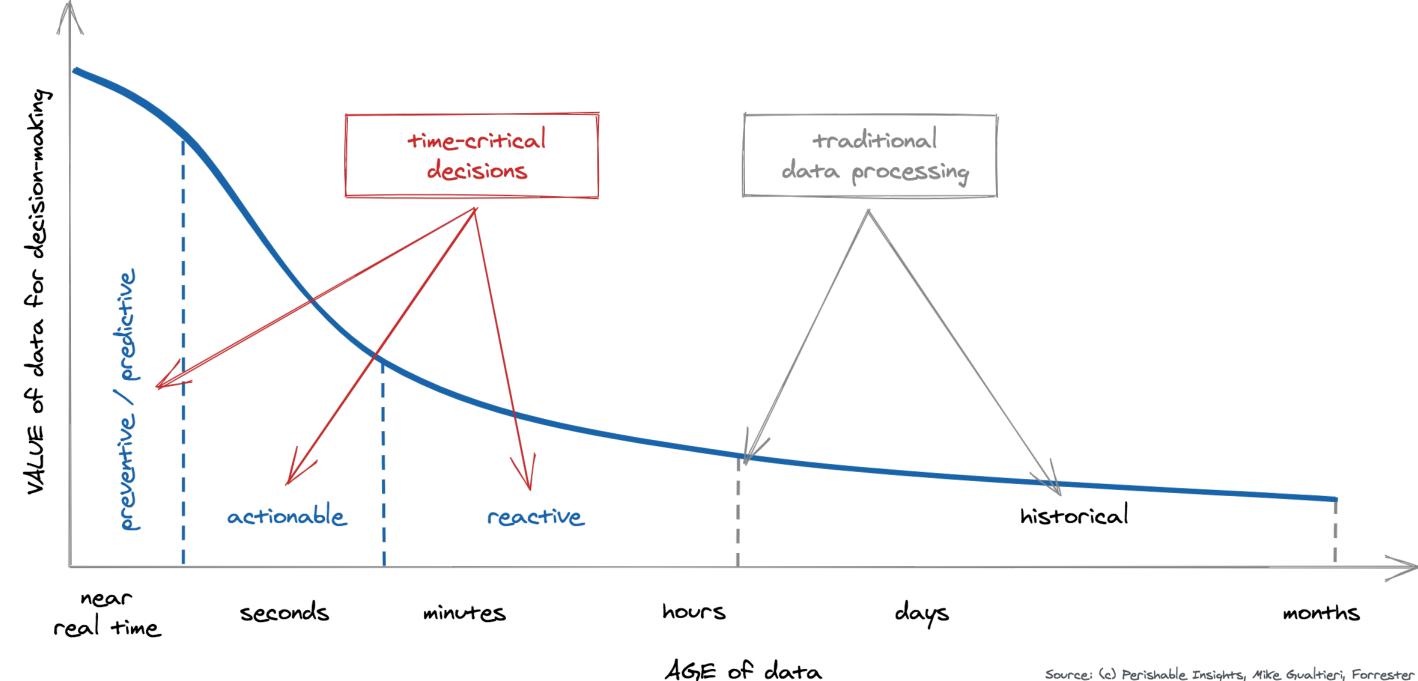


Source: (c) Perishable Insights, Mike Gualtieri, Forrester





Diminishing Value of Data



@hpgrahsl | @Azure #GlobalAzure, 16th April 2021, Linz - Austria



9

Hans-Peter Grahsl



- based in Graz, Austria

- **MongoDB Champion**

@hpgrahsl | @Azure #GlobalAzure, 16th April 2021, Linz - Austria

technical trainer at NETCONOMY

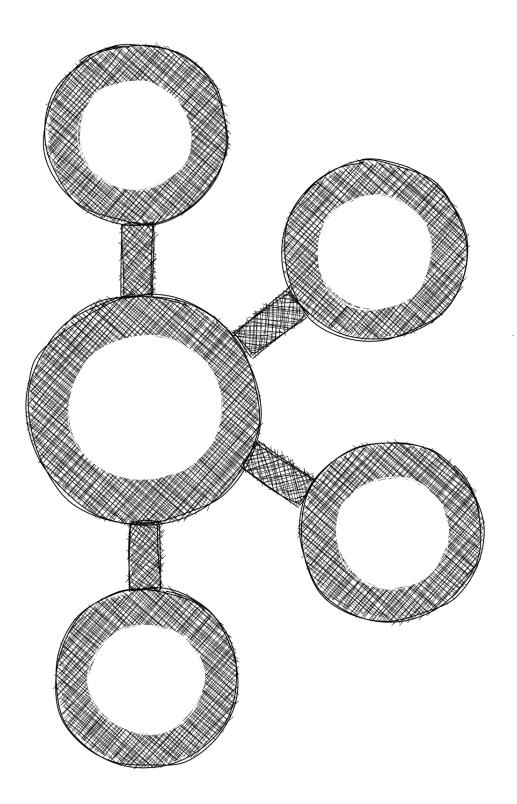
independent engineer & consultant

Confluent Community Catalyst

Stream Processing

Leata processing that is designed with infinite data sets in mind."

Tyler Akidau

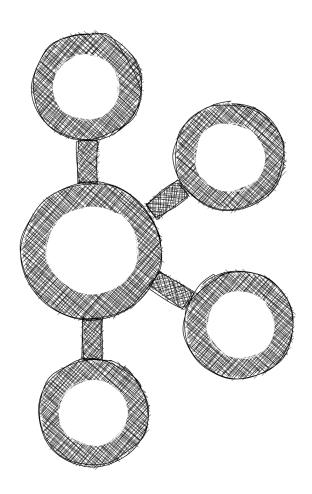


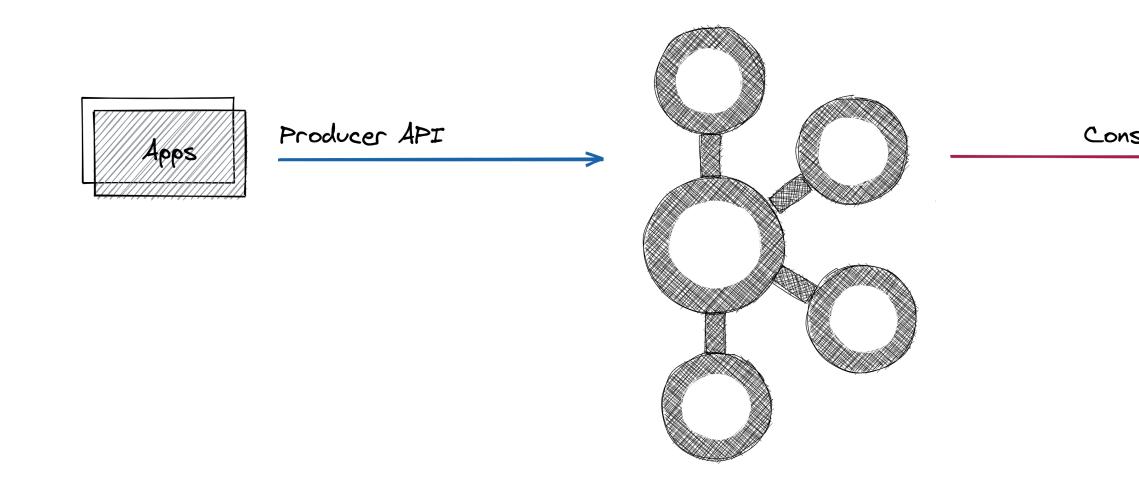


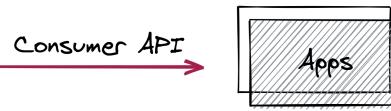
messaging **integration processing**

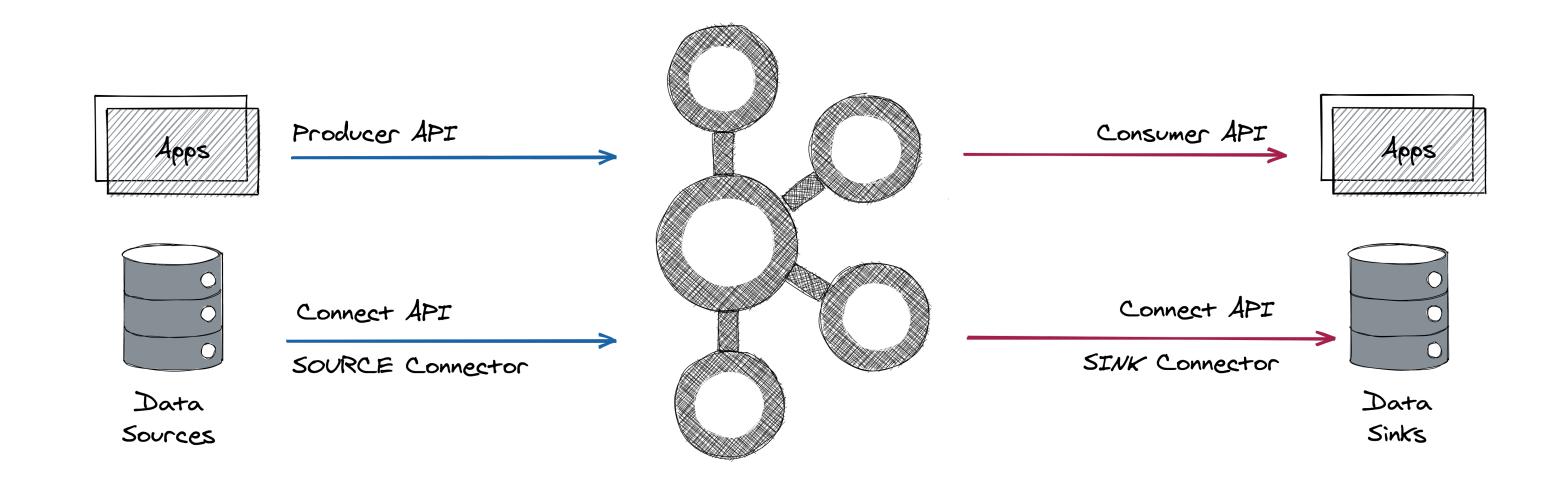
plus storage

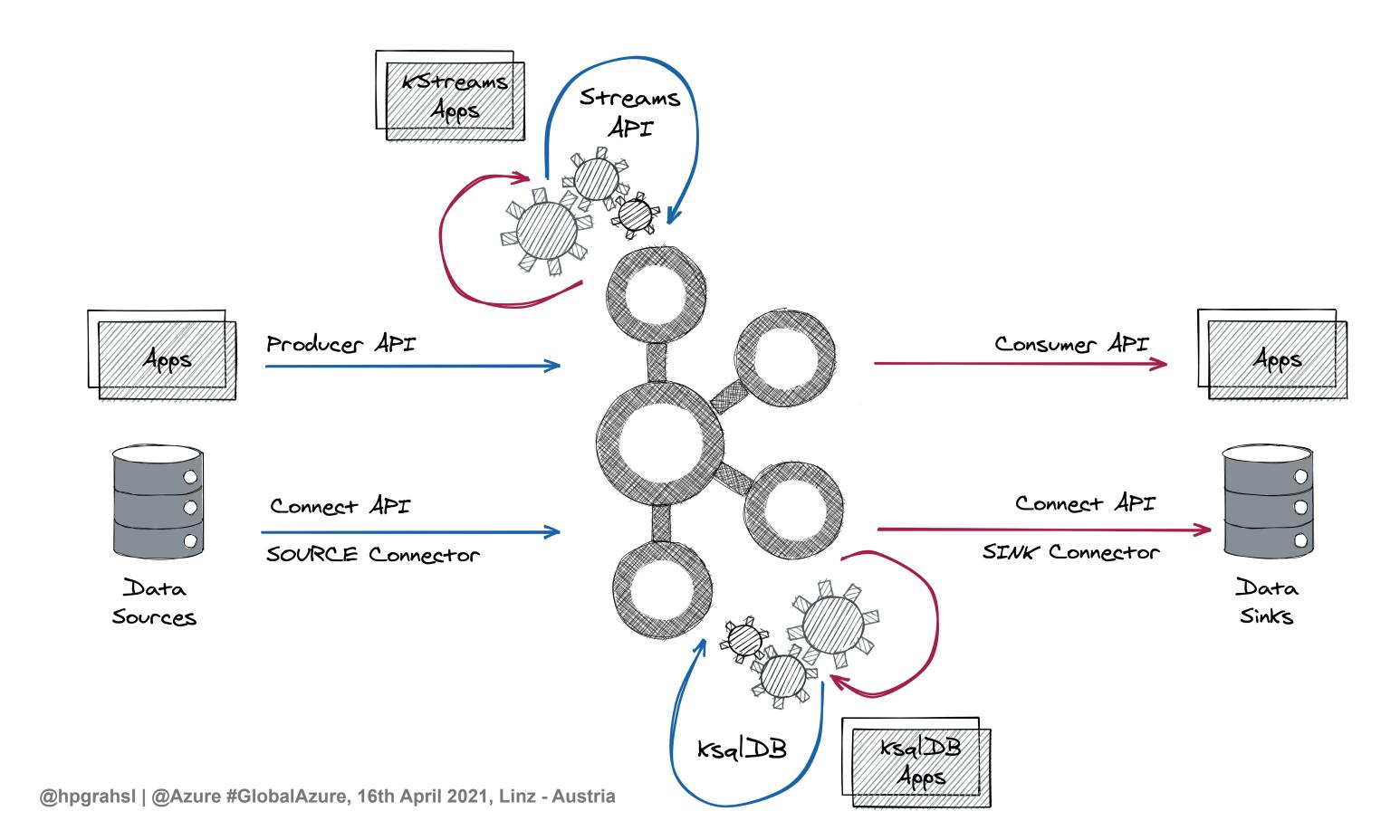
13



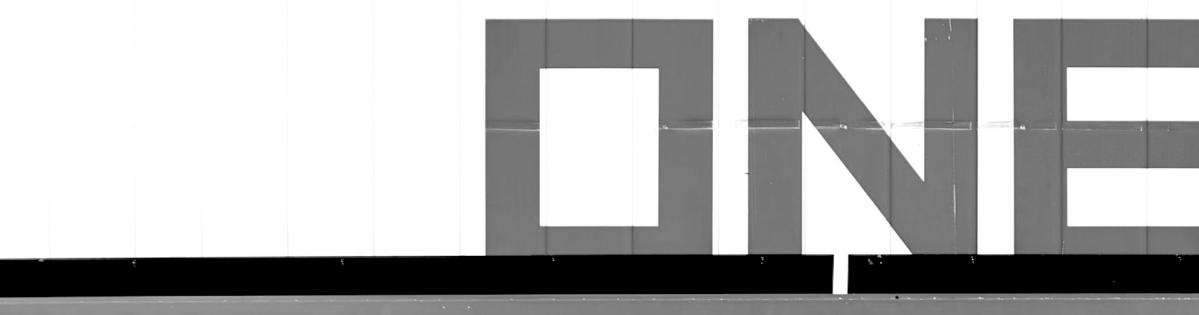












2327037

Kafka with

Avre HDHsehu

@hpgrahsl | @Azure #GlobalAzure, 16th April 2021, Linz - Austria





HDInsight Services "Family"

- large-scale parallel batch processing
- general purpose data warehousing
- stream processing for IoT
- data science & machine learning



HDInsight Services "Family"







\$ kafka





@hpgrahsl | @Azure #GlobalAzure, 16th April 2021, Linz - Austria

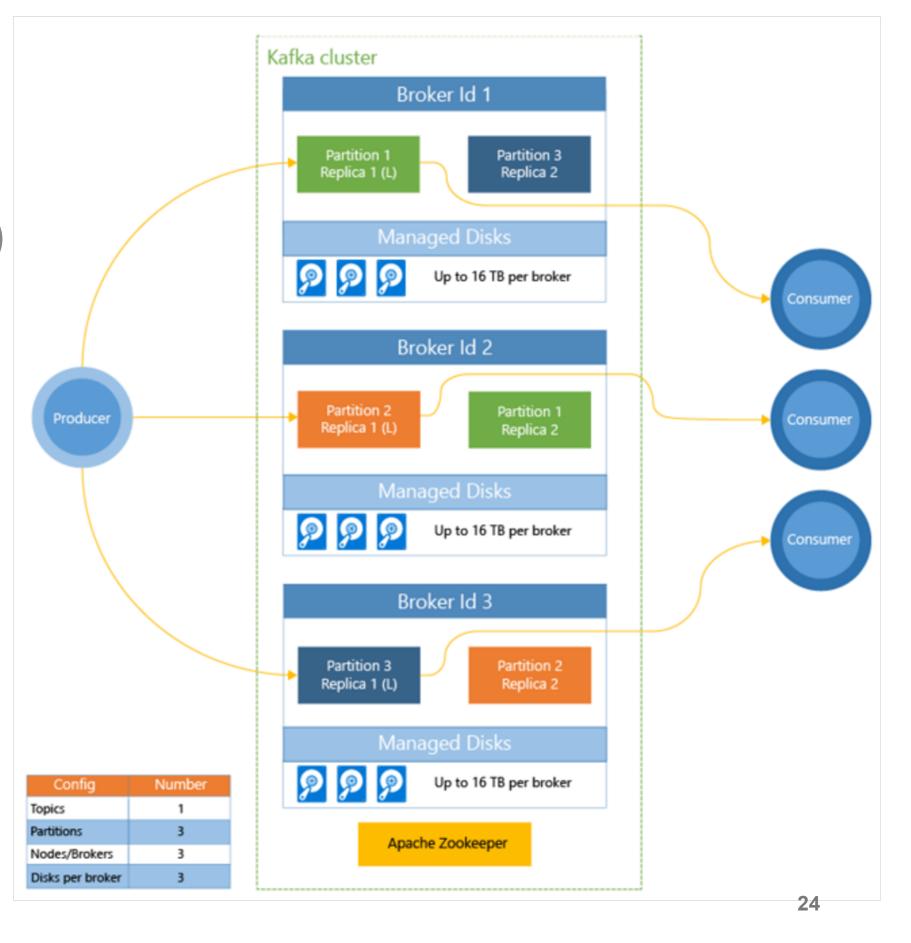


HBASE



HDInsight Apache Kafka®

- broker + zookeeper nodes
- managed disks / storage
- flexible provisioning
- 99.9% SLA uptime



? Client Access ?

Section -



? Client Access ?



when run in same VNet

with VNet peering + IP advertising

from on-premises with VPN gateway

by using Kafka REST proxy

Apache Kafka® HDInsight

• main benefits:

easy provisioning with flexible pricing

open-source Kafka components only

Supported by Microsoft SLAs

Apache Kafka® HDInsight

• main drawbacks:

Outdated version (Kafka 2.1.1)

only "core" Kafka components

per default no external broker access



Azure Event Hubs for Kafka

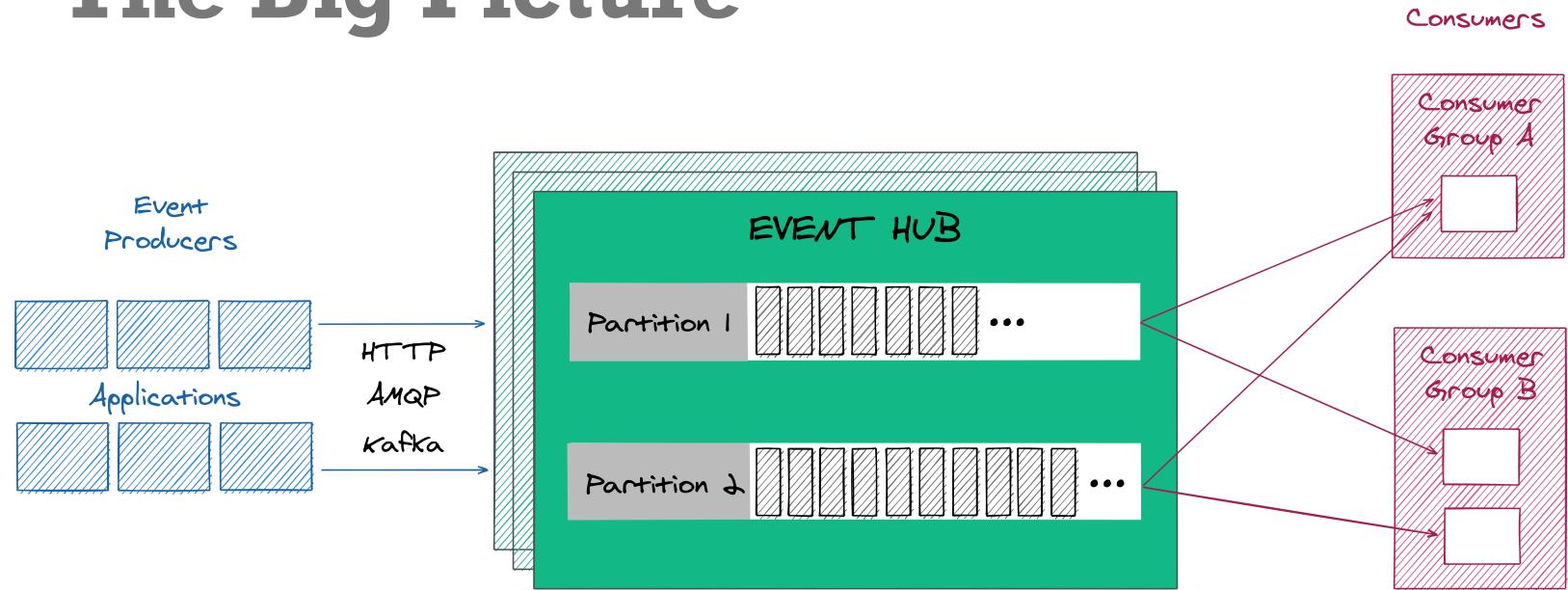
Azure Event Hubs

- fully-managed PaaS
- distributed event ingestion service
- supports auto-scaling capabilities



well-integrated with complementary services

The Big Picture



@hpgrahsl | @Azure #GlobalAzure, 16th April 2021, Linz - Austria

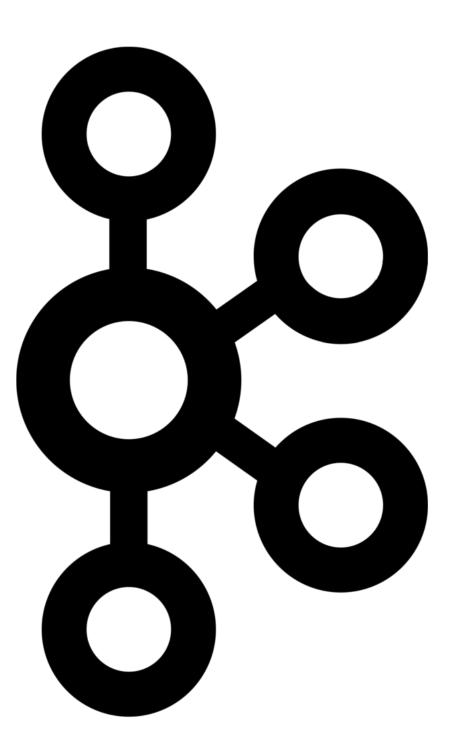
Event

Look-alikes

"Conceptually, Kafka and Event Hubs are very similar: they're both partitioned logs built for streaming data, whereby the client controls which part of the retained log it wants to read."

Event Hubs for Kafka

- overlay on top of Event Hubs
- protocol compatible with Kafka 1.0+
- transparent re-use (code + tools)
- migration benefits in both ways



same same

but different



The Virtual Promise...

"Update the connection string in configurations to point to the Kafka endpoint exposed by your event hub instead of pointing to your Kafka cluster. Then, you can start streaming events from your applications that use the Kafka protocol into Event Hubs."

The devil is in the details

Unsupported Kafka Features

Sidempotent producers & transactions

O compression of messages

Size-based retention or log compaction

O HTTP access via **Kafka REST proxy**

S Kafka Streams & ksqIDB connections

Customer Feedback

10 hubs (=topics) per namespace https://bit.ly/3dvQCA1

1 MB message size limit *https://bit.ly/3sQDIIN*

no Kafka Streams / ksqlDB connections https://bit.ly/3mi4hyu https://bit.ly/39Huu4s

Event Hubs for Kafka

• main benefits:

hybrid messaging scenarios OOTB

auto-inflate for elastic scaling

Variable Ware-native & Kafka-like'' experience

Event Hubs for Kafka

main drawbacks:

fundamental Kafka (protocol) features missing

Selected quotas & limits -> show-stoppers ?

Kafka Streams / ksqIDB clients unsupported



Confluent Cloud

- most complete and versatile service
- cloud-native with elastic scalability
- ready for hybrid & multi-cloud





Confluent Cloud

hosts fully-managed:

- Kafka Connect
- 100+ Connectors
- ksqlDB
- Schema Registry



Tiered Storage

- currently unique to Confluent Cloud
- *infinite* data growth
- retention time *unlimited*

BUT NO Azure Blob Storage yet 😓



provisioning options

directly at Confident Cloud

@hpgrahsl | @Azure #GlobalAzure, 16th April 2021, Linz - Austria

via Azure

Marketplace

46

Products > Apache Kafka® on Confluent Cloud™

•••••	Apache Kafka® on Confluent Cloud™ Confluent ★★★★★(0) Write a review	
Confluent Cloud	Overview Plans + Pricing Reviews	
GET IT NOW	Fully managed event streaming platform powered by Apache Kafka	CONFLUENT Nove : person : New cluster Chater name: p
Pricing information Starting at US\$0.00/month	Start your 3-month trial of Confluent Cloud with up to \$200 off on each of your first 3 monthly bills.	shater,0 Provider & region ()
Categories Analytics		- Brights*
Support Support Help Legal License Agreement Privacy Policy	Confluent Cloud alleviates the burden of managing Apache Kafka, Schema Registry, Connect, and ksqlDB so teams can effectively focus on modern app development and deliver immediate value with real-time analytics use cases.	Cluster type o Basic Standard Dedicated Availability o Multi same Stage same
	Start streaming with Confluent Cloud on Azure in seconds with on-demand provisioning of elastically scalable clusters, without the need for complex cluster configuration and sizing. To complete the serverless experience for Apache Kafka, all components have usage based pricing, and upgrades happen seamlessly without downtime.	
	With Confluent, your data is protected by industry-standard security features and our service reliability is backed by our enterprise-grade uptime SLA. We have been trusted in mission-critical use cases by customers of many types, from fast-	

growing startups to Fortune 100 industry leaders. To learn more about our cluster types & pricing go to https://www.confluent.io/confluent-cloud#features

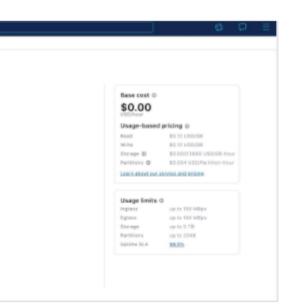
@hpgrahsl | @Azure #GlobalAzure, 16th April 2021, Linz - Austria

More \checkmark

ρ









Confluent Cloud on Azure

main benefits:

V fully-managed Kafka by its original creators

ready for hybrid- / multi-cloud







Confluent Cloud on Azure

main drawbacks:

compare pricing -> not cheap

underlying infra not customizable

higher degree of vendor dependence





Kubernetes

- open-source container orchestration
- deploying / managing / scaling
- CNCF graduate project



Butwhy



Azure Kubernetes Service





@hpgrahsl | @Azure #GlobalAzure, 16th April 2021, Linz - Austria

remaining challenges: Network Storage Security



OperatorHub.io

Welcome to OperatorHub.io

OperatorHub.io is a new home for the Kubernetes community to share Operators. Find an existing Operator or list your own today.

CATEGORIES

AI/Machine Learning

Application Runtime

Big Data

Cloud Provider

Database

Developer Tools

Drivers and plugins

Integration & Delivery

Logging & Tracing

Monitoring

Networking

OpenShift Optional

Security

Storage

Streaming & Messaging

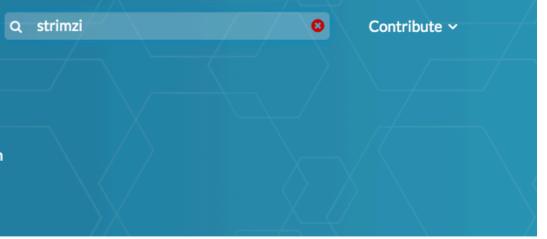
1 ITEMS



Strimzi provided by Strimzi

Strimzi provides a way to run an Apache Kafka cluster on Kubernetes or OpenShift in various deployment configurations.

@hpgrahsl | @Azure #GlobalAzure, 16th April 2021, Linz - Austria



view III List ∽ sort A-Z ∽





- Operators (cluster / topic / user)
- Kafka Connect + managed Connectors
- replication with MirrorMaker
- HTTP Bridge for Kafka
- Cruise Control cluster balancing



Kafka on AKS with Strimzi

main benefits:

k8s-native experience with built-in security

tweakable / customizable in various ways

ease of use for "non-ops-savvy folks" -> ME









Kafka on AKS with Strimzi

main drawbacks:

Kafka is OUR OWN responsibility

k8s knowledge despite "operator magic"

on Microsoft support offering





don't just roll the dice...

dig deeper & navigate further!

20 40

Hanas

.

http://bit.ly/kafka-ga21

