IBM Message Hub

James Bennett

IBM’s statements regarding its plans, directions and intent are subject to change or withdrawal without notice at IBM’s sole discretion. Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision. The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

- **IBM Confidential**. Unless specifically advised otherwise, you should assume that all the information in this presentation (whether given in writing or orally) is IBM Confidential and restrict access to this information in accordance with the confidentiality terms in place between your organization and IBM.
- **Content Authority**. The workshops, sessions and materials have been prepared by IBM or the session speakers and reflect their own views. They are provided for informational purposes only, and are neither intended to, nor shall have the effect of being, legal or other guidance or advice to any participant. While efforts were made to verify the completeness and accuracy of the information contained in this presentation, it is provided AS-IS without warranty of any kind, express or implied. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this presentation or any other materials. Nothing contained in this presentation is intended to, nor shall have the effect of, creating any warranties or representations from IBM or its suppliers or licensors, or altering the terms and conditions of the applicable license agreement governing the use of IBM software.
- **Performance**. Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.
- **Customer Examples**. Any customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer. Nothing contained in these materials is intended to, nor shall have the effect of, stating or implying that any activities undertaken by you will result in any specific sales, revenue growth or other results.
- **Availability**. References in this presentation to IBM products, programs, or services do not imply that they will be available in all countries in which IBM operates.

Trademark acknowledgements

- IBM, the IBM logo are trademarks of International Business Machines Corporation, registered in many jurisdictions.
- Other company, product and service names may be trademarks, registered marks or service marks of their respective owners. A current list of IBM trademarks is available on the web at "Copyright and trademark information" ibm.com/legal/copytrade.shtml

IBM Watson and Cloud Platform

Application

DsX

IoT

Virtual agent

AI

Conversation	Discovery	Compare + Comply	Knowledge Query	Tone Analysis	Personality Insights
Visual Recognition	Speech	Document Conversion	Nat. Language Understanding	Nat. Language Classifier	+ more...

Data

Ingest → Enrich → Store → Analyze → Apply

Public, Private, Licensed

The Watson Data Platform

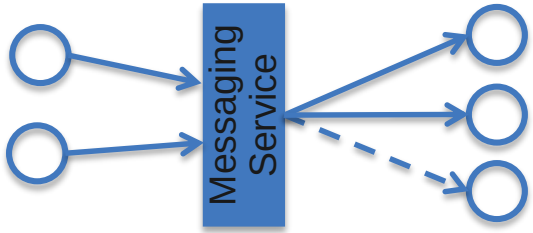
Cloud

Dev Services	Containers	Messaging	Blockchain	Logging	+ more...
Infrastructure	Storage	Compute	Physical Network	Infrastructure Mgmt	+ more...

Worker Offload

Intensive work offloaded and distributed amongst worker processes to be performed asynchronously

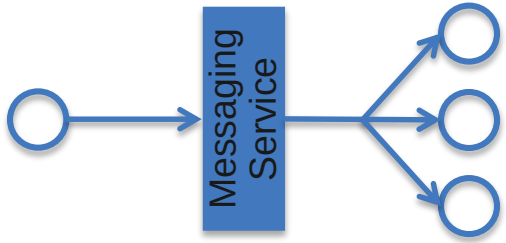
- Processing images or videos
- Performing text analytics



Event-Driven

Take one or more actions when something interesting happens

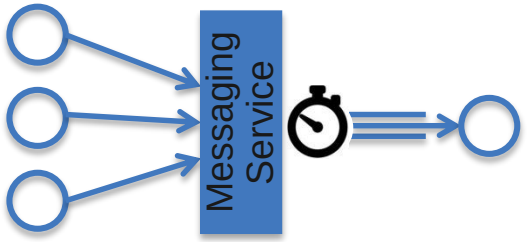
- Email logs and update dashboards when build finishes
- Upload videos once finished transcoding



Delayed Processing

Schedule a task to happen at a specific point in time

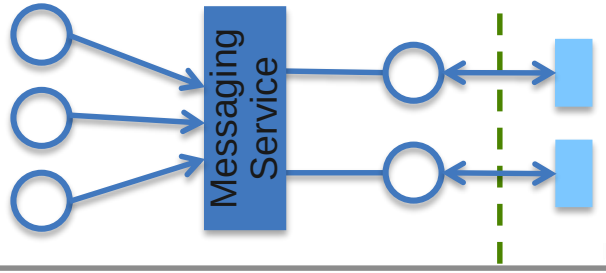
- Run detailed reports when app use is low
- Generate end of day summary



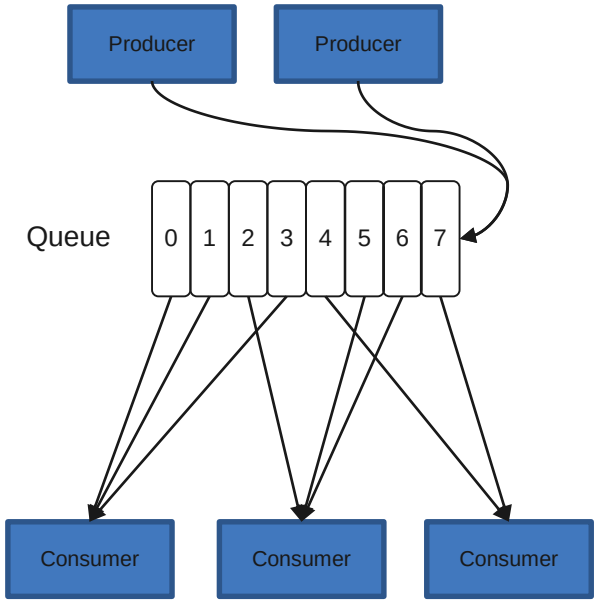
3rd Party Integration

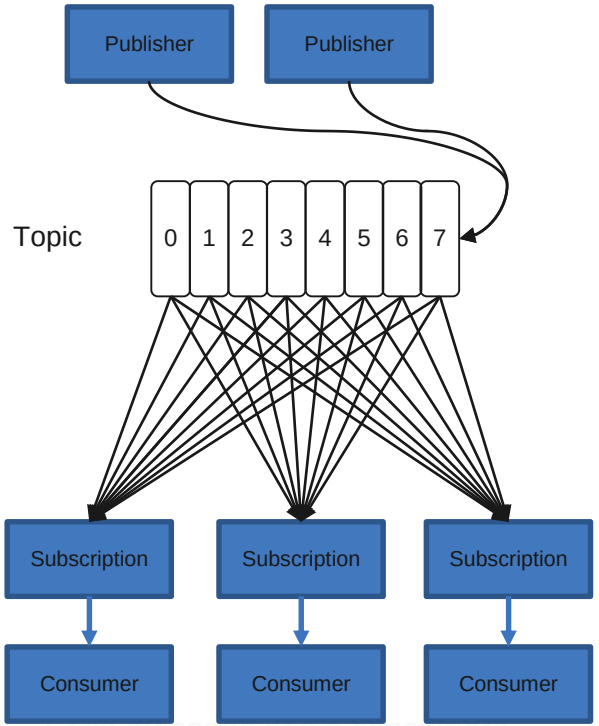
Ensure applications remain responsive even when 3rd party systems are not available or responding fast enough

- Updating existing CRM system
- Booking appointment



Queues vs topics





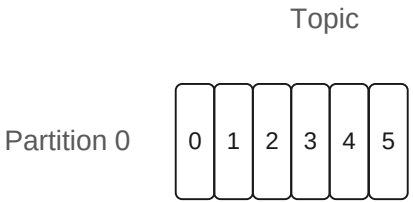
Each message has 1 consumer

Gives easy sharing among consumers

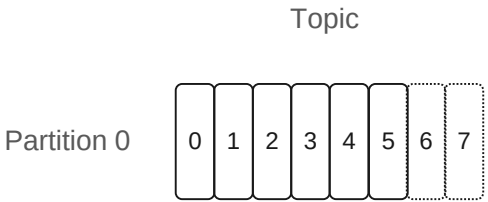
Each message has n consumers

Hybrid models such as shared subscriptions or consumer groups exist

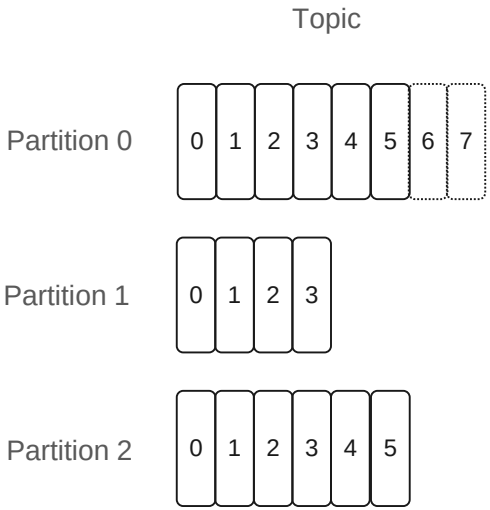
Kafka is: Built for Scale



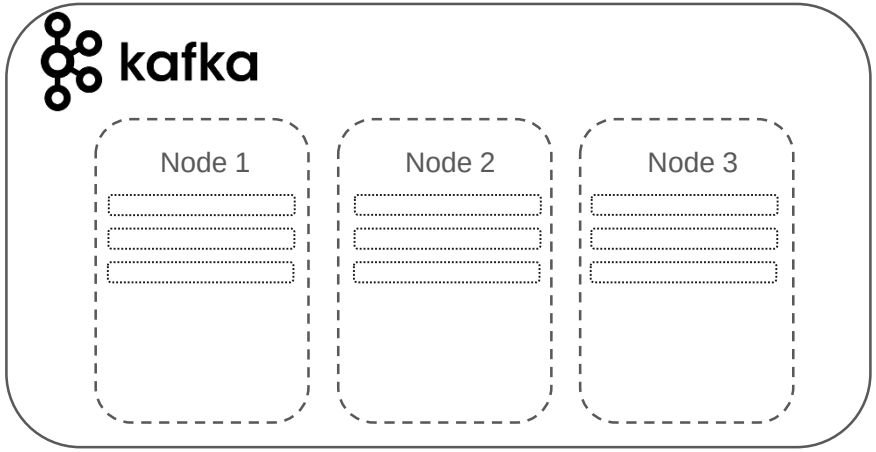
Kafka is: Built for Scale



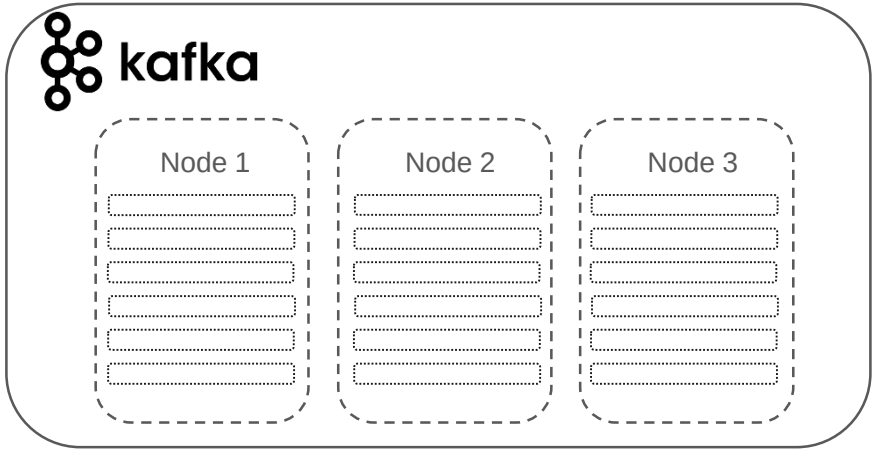
Kafka is: Built for Scale



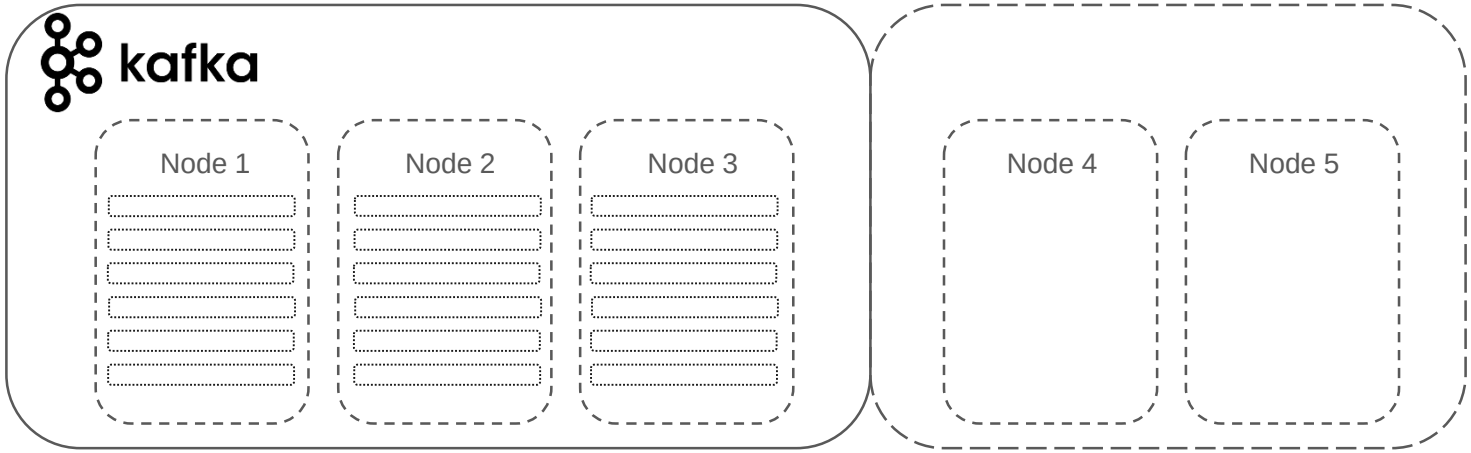
Kafka is: Built for Scale



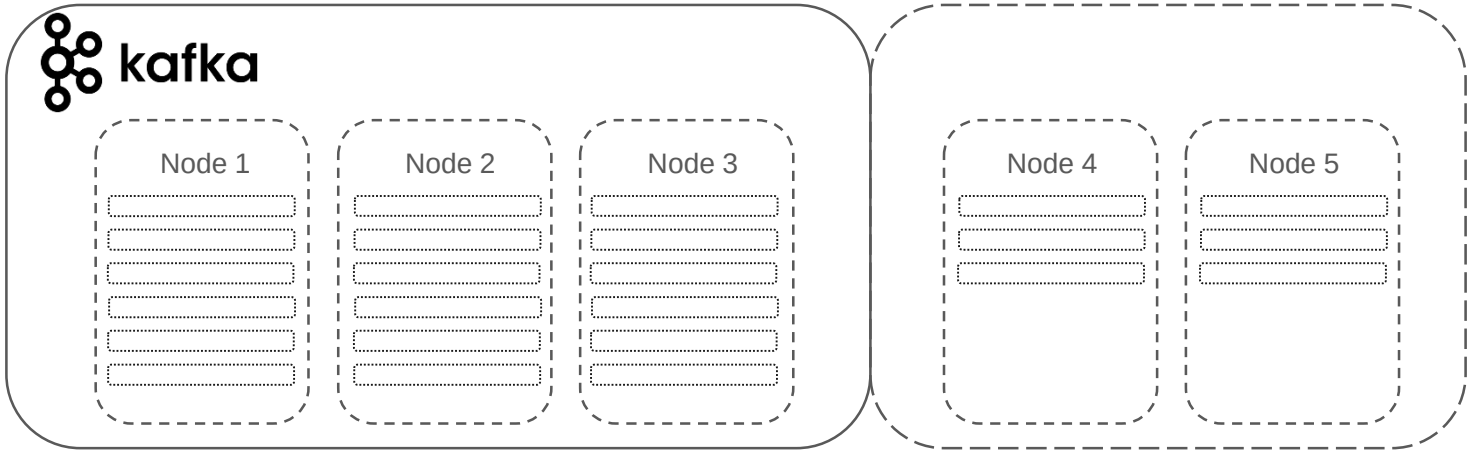
Kafka is: Built for Scale



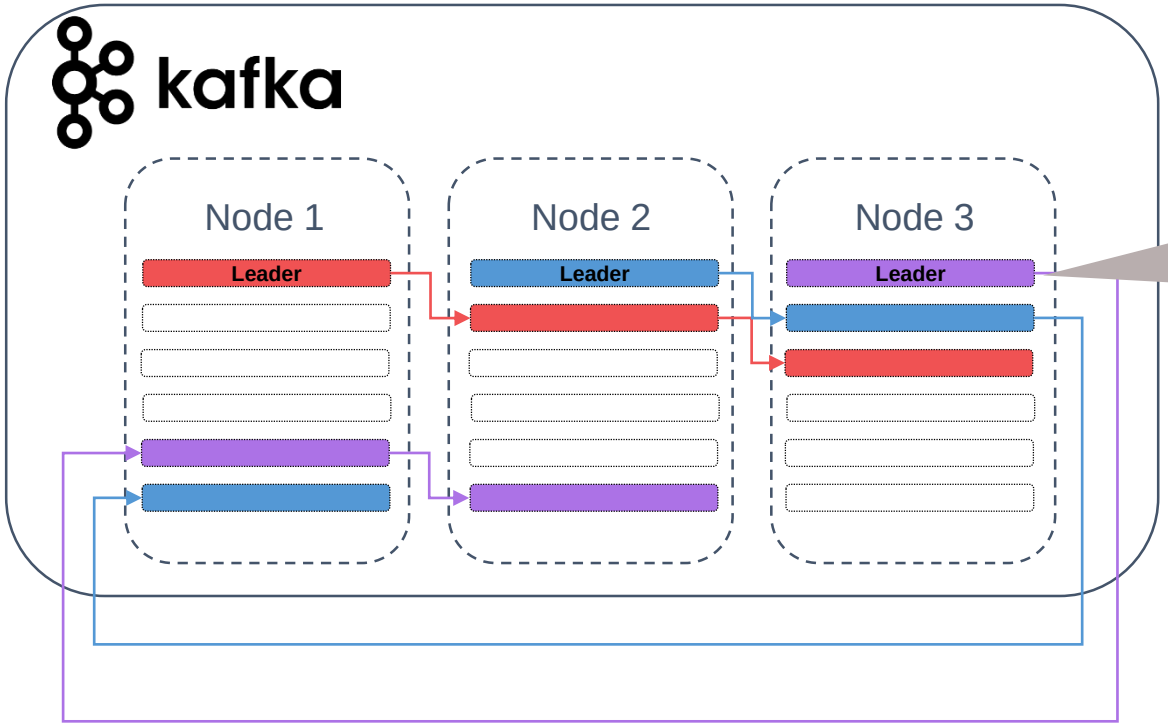
Kafka is: Built for Scale



Kafka is: Built for Scale

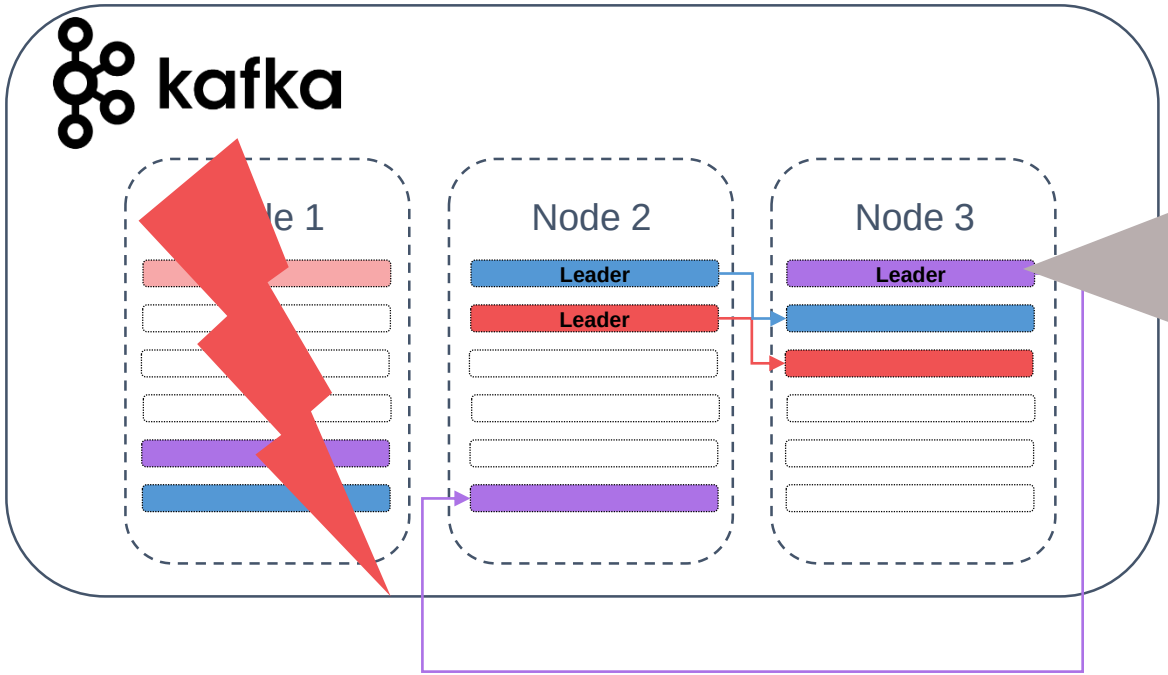


Kafka is: Built for Scale



Each partition has a leader responsible for replicating the messages, in order, to the other replicas.

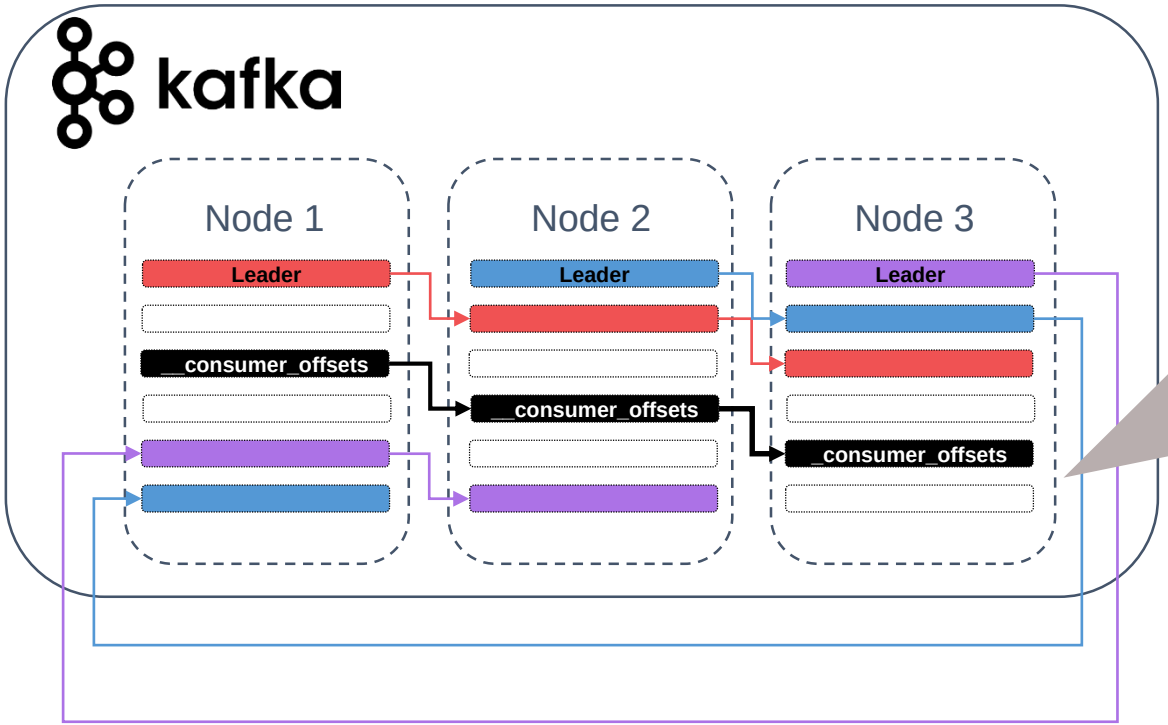
Kafka is: Built for Scale



Each partition has a leader responsible for replicating the messages, in order, to the other replicas.

Recovery from node failure is very fast. There is no data loss and no need to wait for propagation of changes to get consistency restored.

Kafka is: Built for Scale

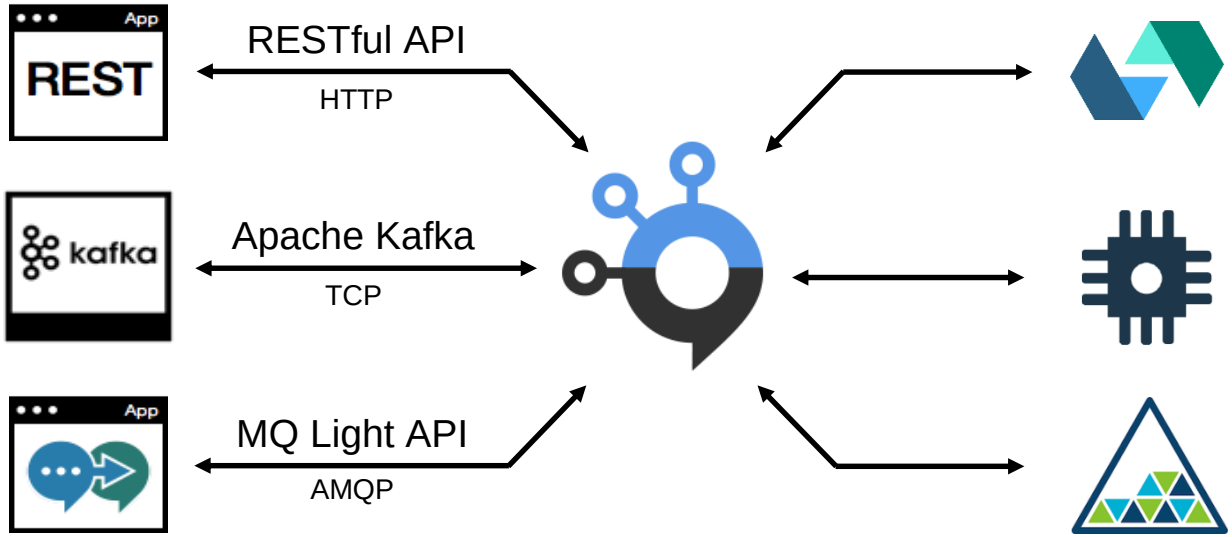


The __consumer_offsets topic contains the offset information for the consumers.

This topic is replicated in the usual way.

Actually, there are 50 partitions for this topic, so it's very wide.

IBM Message Hub



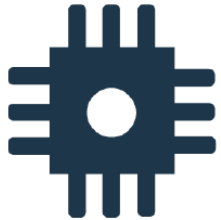
Message Hub messaging APIs

Apache Kafka producer/consumer	•Part of Apache Kafka open-source project •Provides best blend of function and performance •Java only
Other Kafka clients	•A wide range of open-source clients now available •Some use native code, others are pure implementations •C, C++, JavaScript, Python, .NET, golang
Confluent REST API	•Originally offered when the only Kafka client was Java •Implemented by the Confluent REST Proxy •A bit hard to consume and not very efficient
MQ Light API	•API supported by IBM MQ too •Nice feature set but really hard to implement on top of Kafka •Very limited scalability •Evaluating options here, deprecation being considered

Building: Message Hub + Public Cloud

- **Message Hub + Cloud Functions**
  - Monitor events received in Message Hub to trigger actions in Cloud Functions
  - Blog post

- **Message Hub + Watson IoT**
  - Push events from Watson IoT to Message Hub for use in real-time applications
  - Documentation

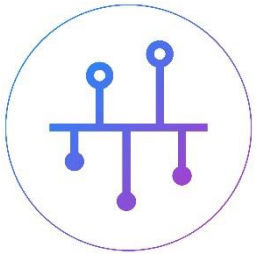


- **Message Hub + Object Storage**
  - Push events from Message Hub automatically into Object Storage for cost effective long term storage
  - Documentation

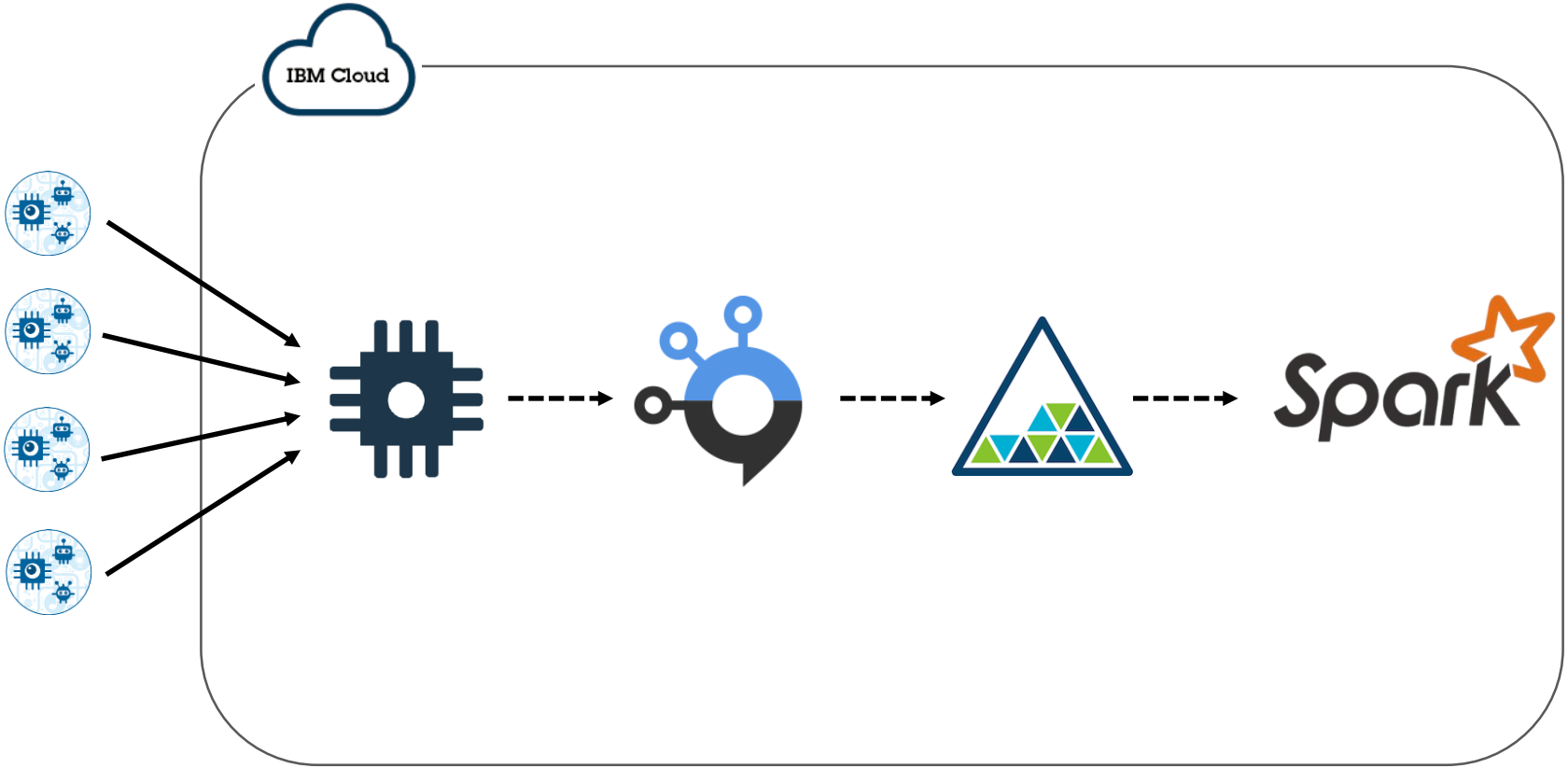
Building: Message Hub + Hybrid Cloud

- **Message Hub + MQ**
  - Consume events from MQ using our new bridge to allow other applications and services to consume directly from Message Hub
  - Documentation

- **Message Hub + IIB**
  - IIB supports native Kafka consumer and producer nodes, allowing you to produce to and consume from a Message Hub instance
  - Blog post



Example scenario: Ingesting IoT data



Message Hub: Why invest?

Run your own Kafka:

- Deploy Zookeeper cluster
- Deploy your own Kafka cluster
- Put systems in place to operate the service 24x7
- Keep up-to-date with the open source community
- Handle version-to-version migration and updates

Invest in Message Hub:

- One click deploy of a fully managed Kafka environment
- Additional custom integration points with the IBM Cloud Platform

Message Hub: Public vs Dedicated

Environment	Public	Dedicated
Cluster tenancy model	Multi-tenant	Single-tenant
Plan	Standard (Lite to be added soon)	Dedicated
Time to deploy	5 seconds	3 weeks
Partition limit	100 per tenant	100 per environment
Retained log data per partition	1GB	10GB
Throughput quotas	Yes	No
Ordering	Self-service	Sales team
Billing	Based on usage • Partitions/month • Messages sent/received	Subscription

Message Hub -- Futures

Our Goal: To deliver a world class managed Kafka platform

What do clients expect from Message Hub?

- High throughput
- High-capacity disk storage for log compaction
- Support for Kafka Streams library for streaming applications
- Support for Kafka Connect and the 70+ connectors to and from other systems
- Multi-Datacenter replication for HA/DR
- Flexible configuration – log retention

Today: What are we targeting?

“As a Bluemix User, I am looking to consume Message Hub. Consuming a multi tenant service with some performance limitations is not an issue for me”

Message Hub Public

”As a Bluemix User, I am looking to consume Message Hub. I have performance & other functional requirements which go beyond the public service. I would like isolated network and isolated compute by having the service deployed within my Bluemix Dedicated environment.”

Message Hub Dedicated

Future: What are we targeting?

“As a Bluemix User, I am looking to consume Message Hub. Consuming a multi-tenant service with some performance limitations is not an issue for me.”

Message Hub Public

”As a Bluemix User, I am looking to consume Message Hub. I have performance & other functional requirements which go beyond the public service. I am happy to pay a premium for a single-tenant version of the service on shared network and dedicated compute.”

Message Hub Premium

”As a Bluemix User, I am looking to consume Message Hub. I require more isolation than the Message Hub Premium plan. I am happy to pay a premium for a single-tenant version of the service on isolated/dedicated network and isolated/dedicated compute.”

Message Hub Premium++ (most closely aligned to Dedicated today)

What does this mean?

- Exposing the Kafka capability required by our clients but doing so in a manner which maintains strong service stability

- Continued contribution ensure the success of the underlying core Kafka OSS project

- A managed service that elastically scales in line with our clients growth

Core Message Hub

- Continued improvements to the end-end user experience
  - We have and will continue to heavily invest in samples to onboard new users with less pain
  - New Bridges UI is the first step on our journey for an improved GUI

- Exposing more metrics in our client facing dashboards

- Some example requirements from our early adopters:
  - The ability to increase the number of partitions in a topic after initial provision
  - The ability for clients to manage the keys utilised for encryption
  - IAM integration for per topic produce/consume policy enforcement

Supporting Connectivity

- Continue to recommend and contribute to best of breed OSS Kafka clients

- Respond to client demand for our next set of Message Hub Bridges
  - Dash..?
  - Cloudant..?
  - …?
- Upgrade current Bridge functionality to utilize Kafka Connect as the underlying technology

- Better support for bring your own connectors with the Kafka Connect API

Improving Message Hub Deployments

- Managing a highly available, cross DC, Kafka environment
  - Kafka MirrorMaker is a self configured option available today
  - Question we have asked ourselves: how can we make this as pain free as possible for our clients?

- Improving the configurability of our Premium/Dedicated deployments
  - Coming next: Capacity driven instead of defined partitions
    - You choose how you want to divide up your cluster
  - Future: Define the storage capacity

- How can we provide better service guarantees of a Public Message Hub instance?
  - Coming next: Implement quotas
  - Future: Availability Zones