

Data Ingestion With Event Hub

Data Science Dojo

Typical Event Processing



Applications



Cloud Gateways
(WebAPIs)



Scalable
Event Broker



External
Data Sources



Web/Thick
Client Dashboards



Devices



Field Gateways



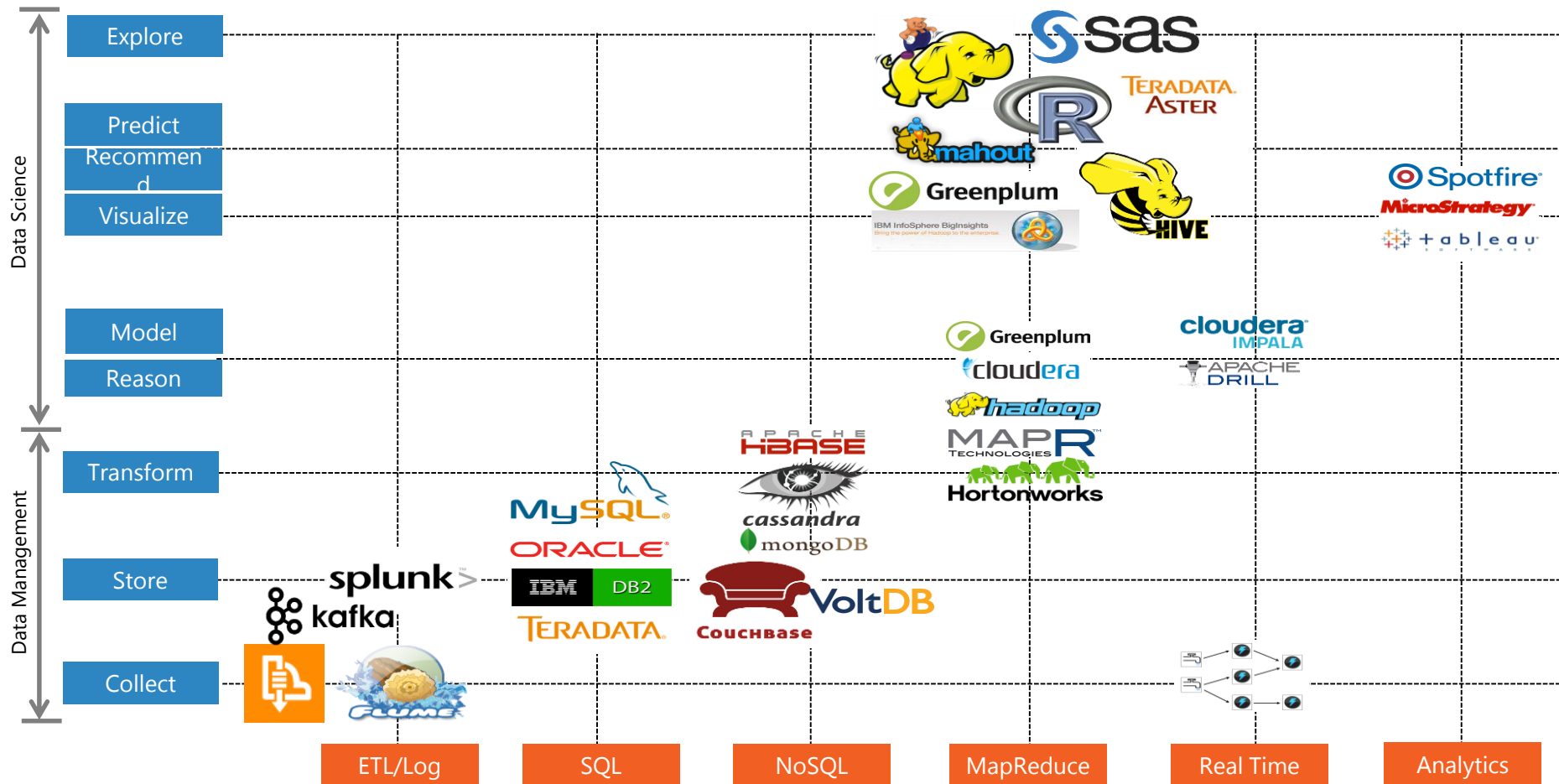
Search and Query



Data Analytics

data science dojo

unleash the data scientist in you



The Post Office & Shipping Centers

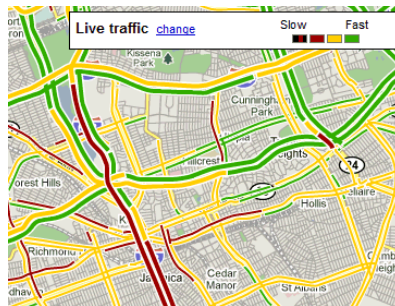


- Tracks address changes
- Tries again tomorrow if send failed
- Holds packages in short term
 - Too many failed deliveries
 - Vacations
- Reduces complexity through specialization
- Optimized to send, receive, and temporarily house packages

Hypothetical Scenario



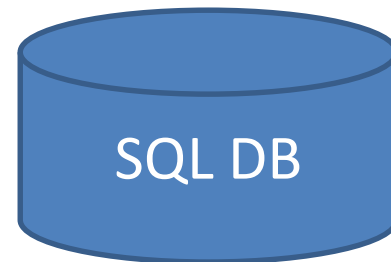
Bureau of
Transportation
Statistics



Sends Traffic Data



Server: ABC, Port 123

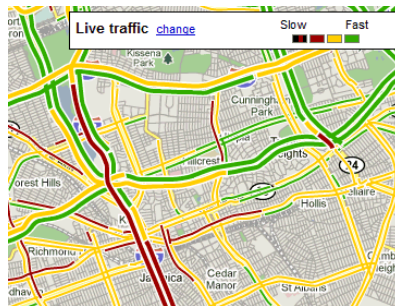


Your Company

Things Change



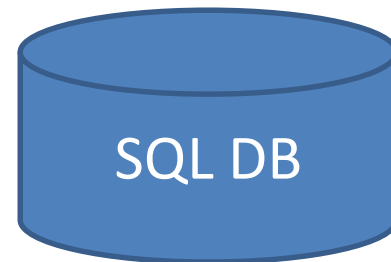
Bureau of
Transportation
Statistics



Sends Traffic Data



WAS Server: ABC, Port 123
Now Server: DEF, Port 456

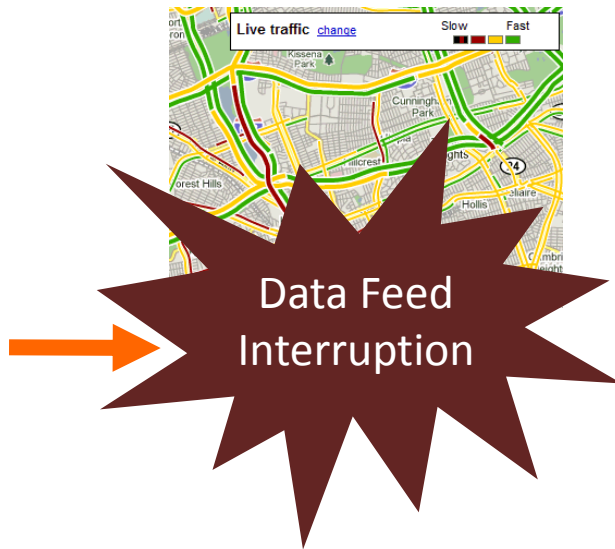


Your Company

Interruption Occurs

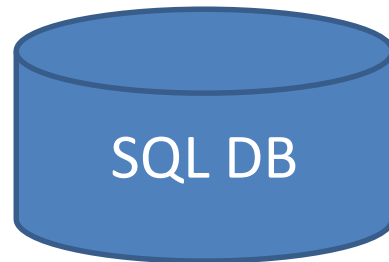


Bureau of
Transportation
Statistics



WAS Server: ABC, Port 123

Now Server: DEF, Port 456



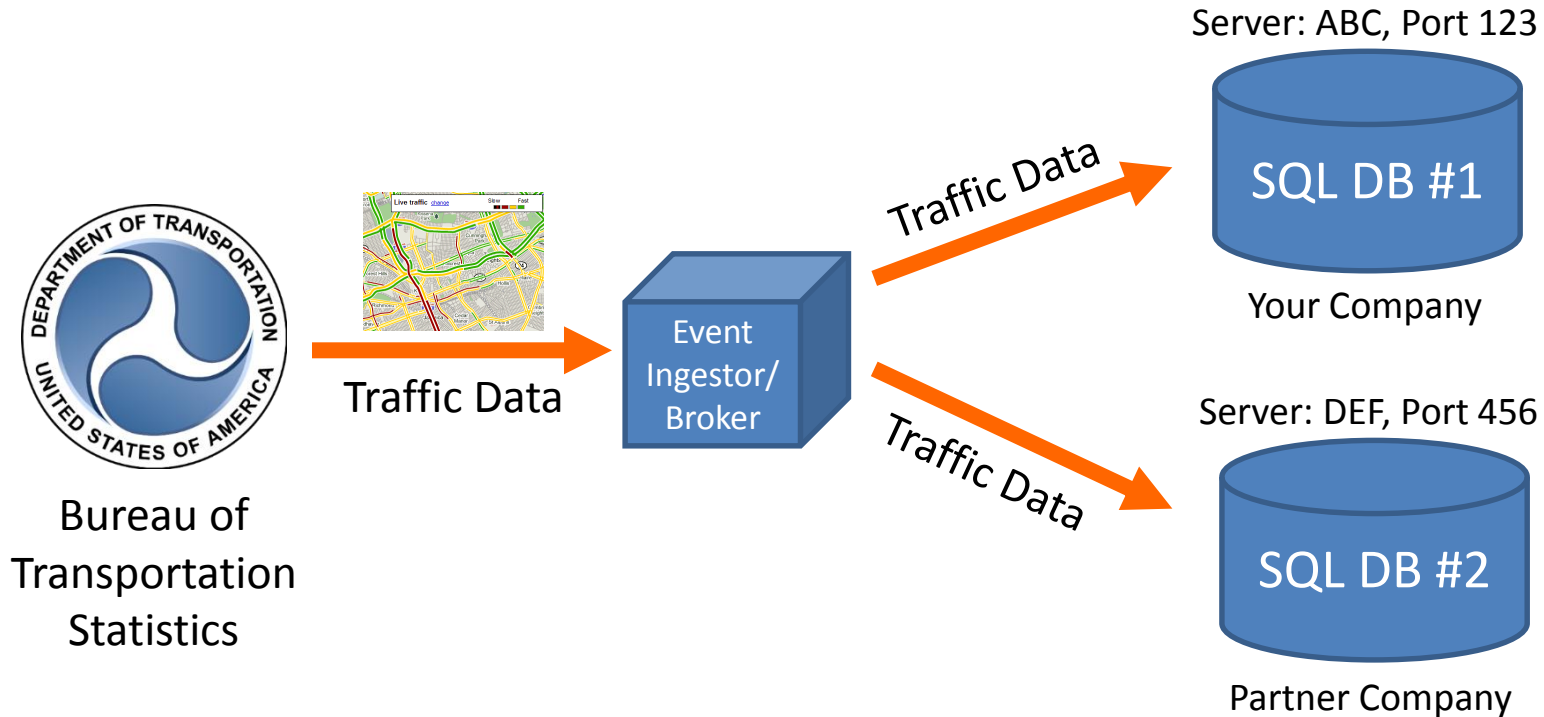
Your Company

Reactive Remedy

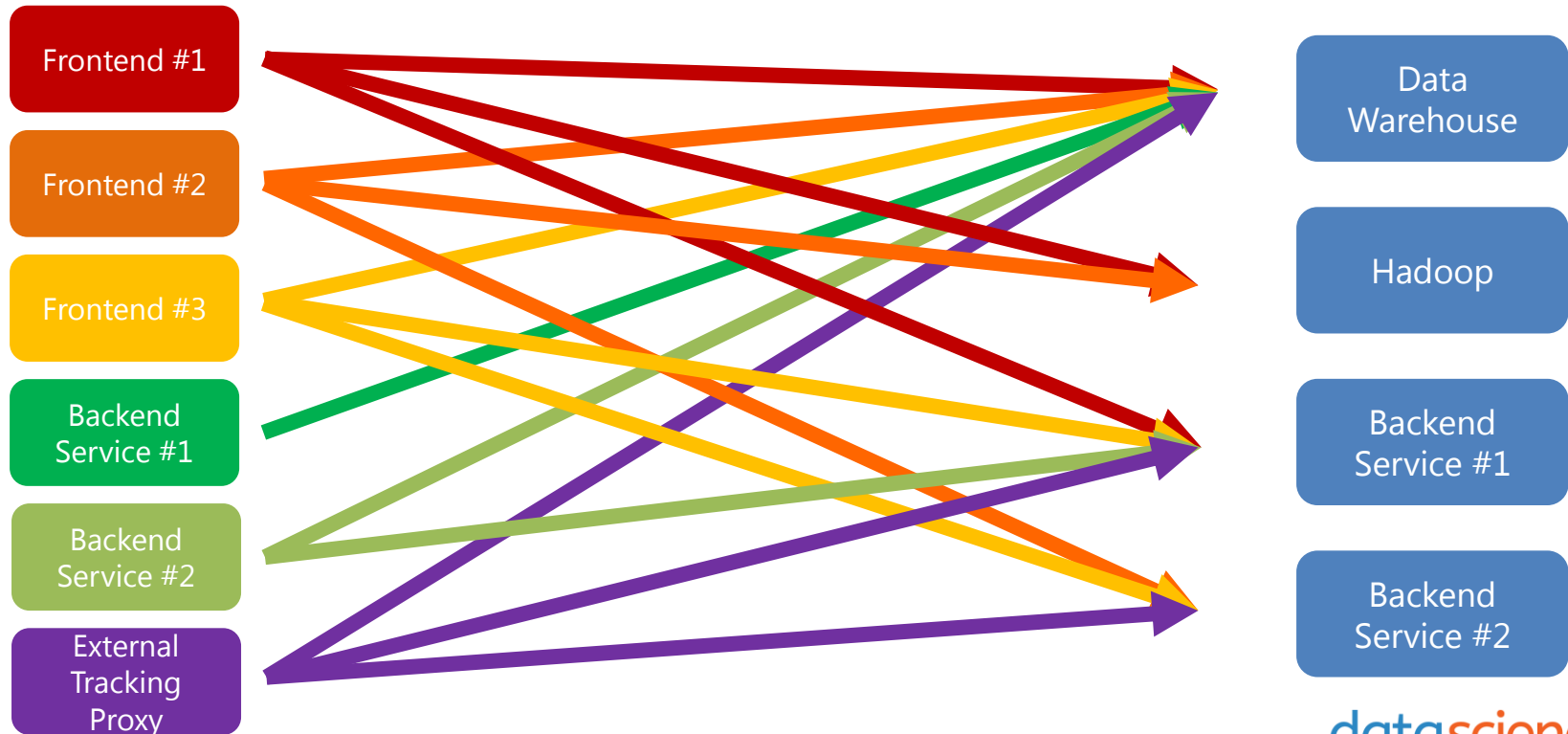
- Call Bureau of Transportation Statistics to change their outbound data to funnel into server DEF and port 456
- Bureau makes support ticket
 - Your request enters their task queue
- 4 days later... ports are changed



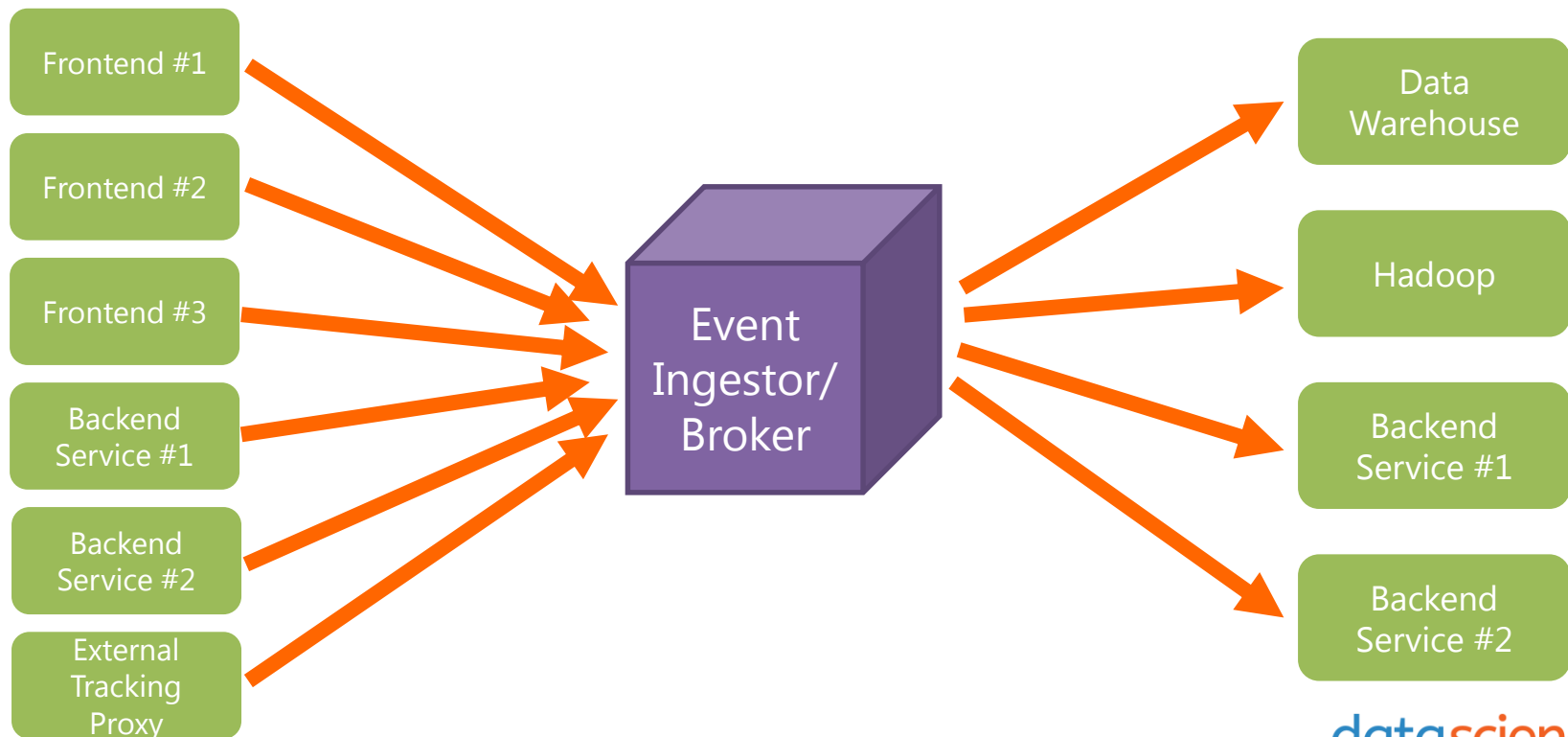
Preventative Solution: Middleware



Data Pipeline Complexity at LinkedIn



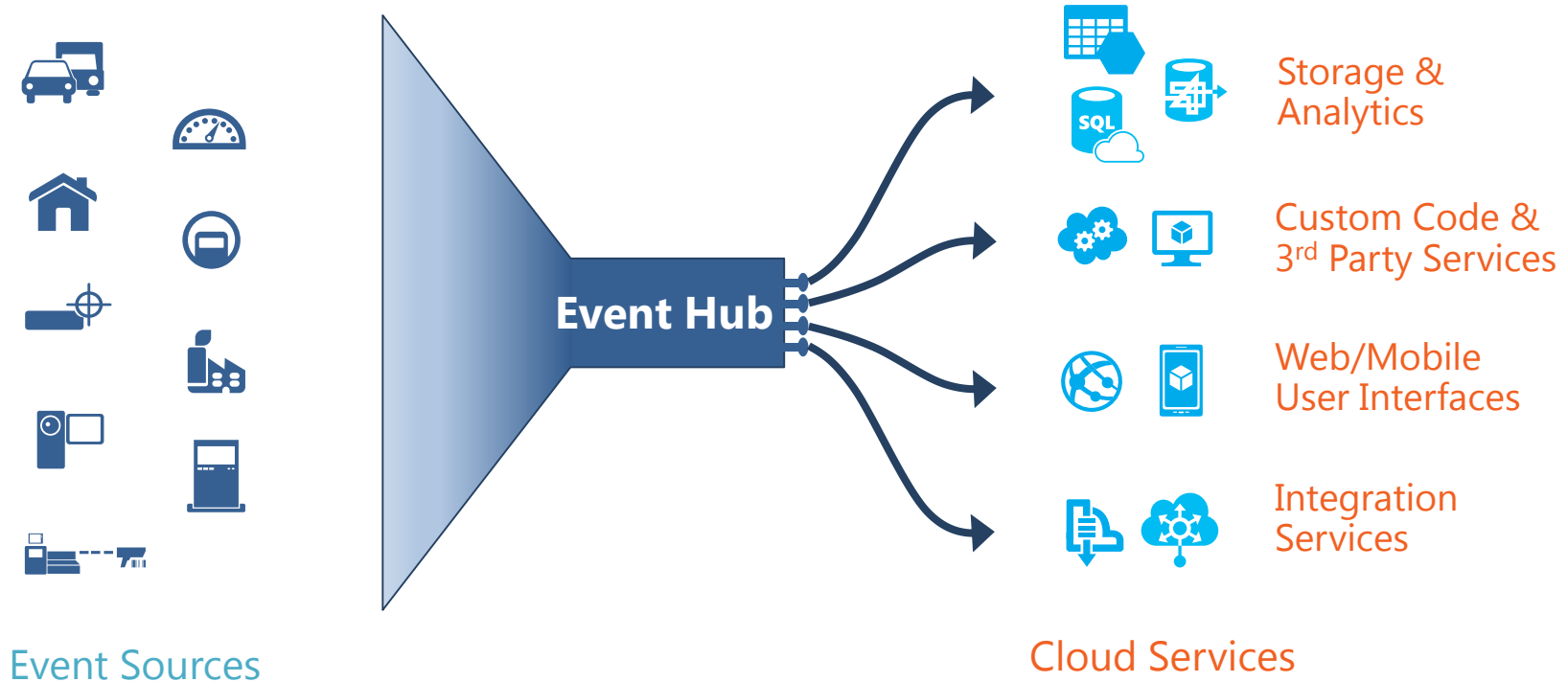
Data Pipeline Complexity at LinkedIn



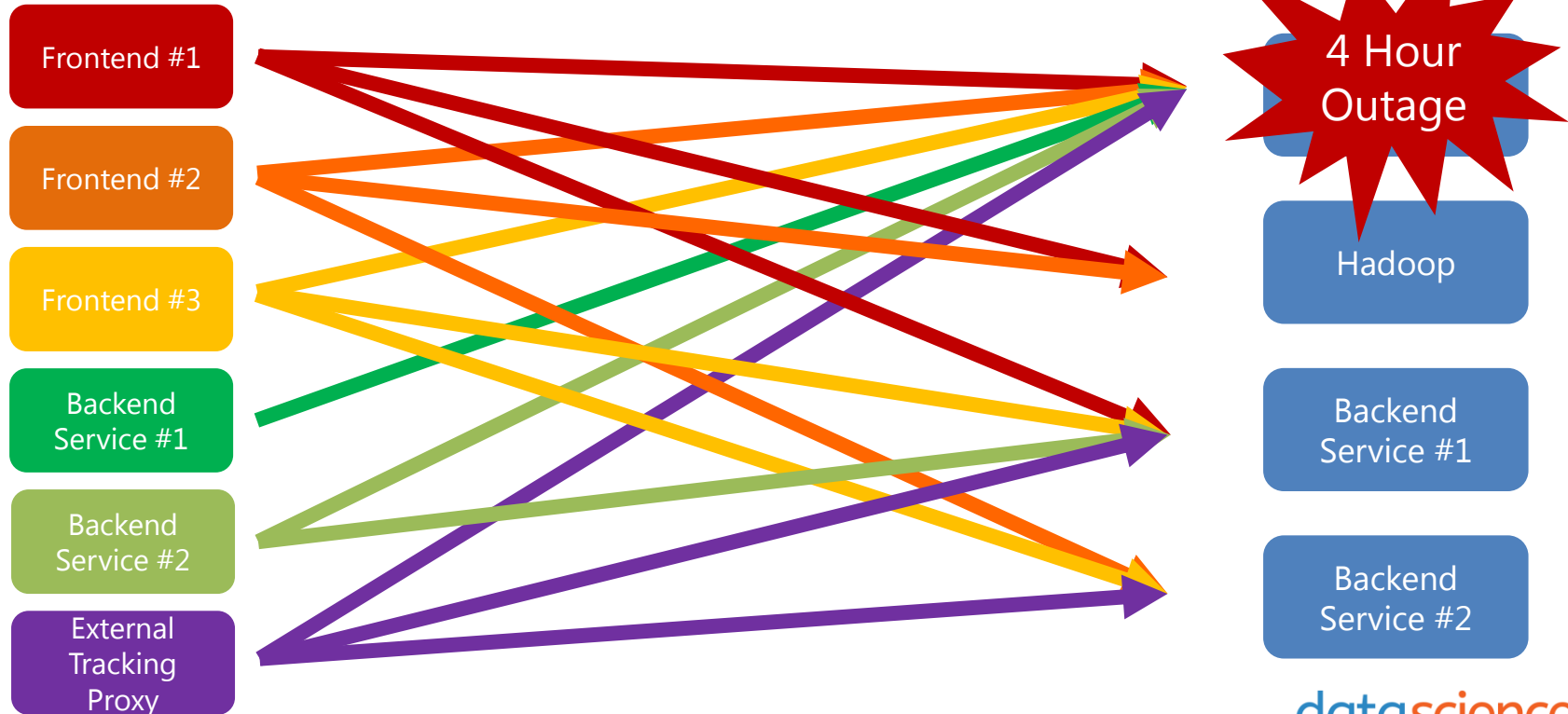
Popular Event Brokers



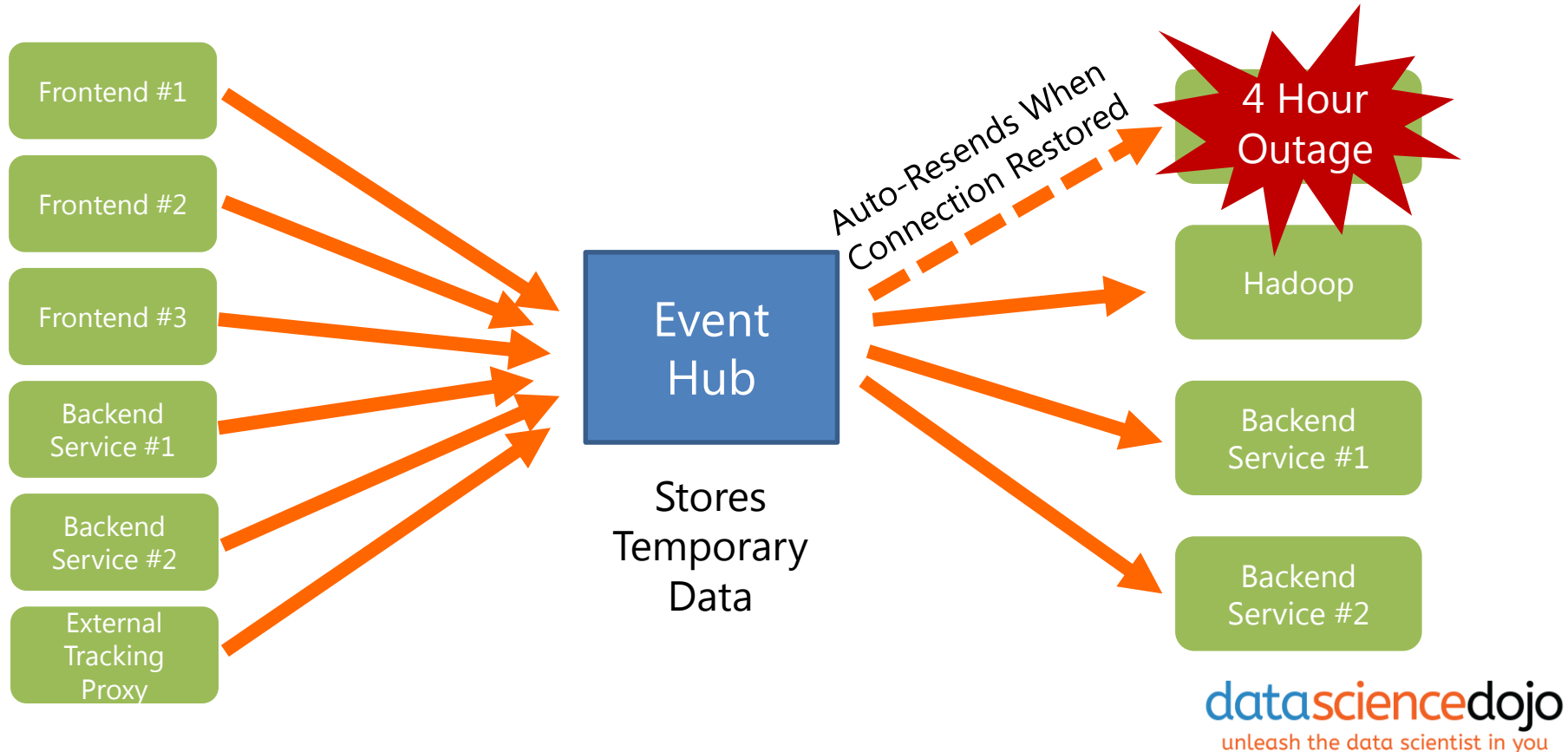
Event Hub for IoT: Big Data Ingestion



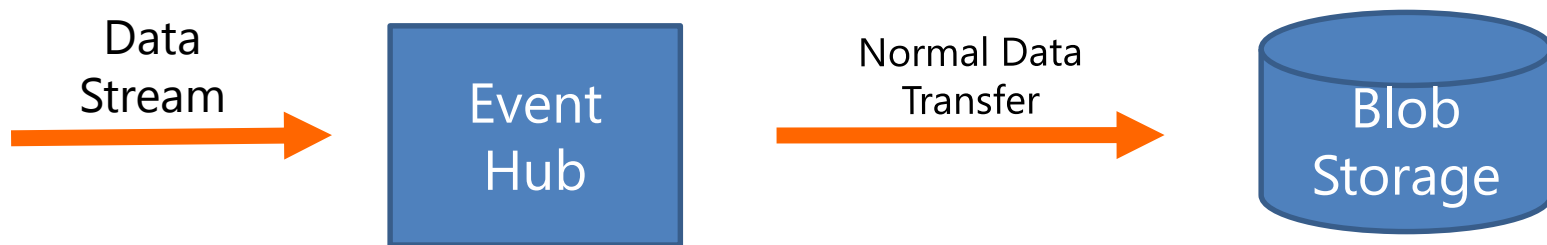
Server Down



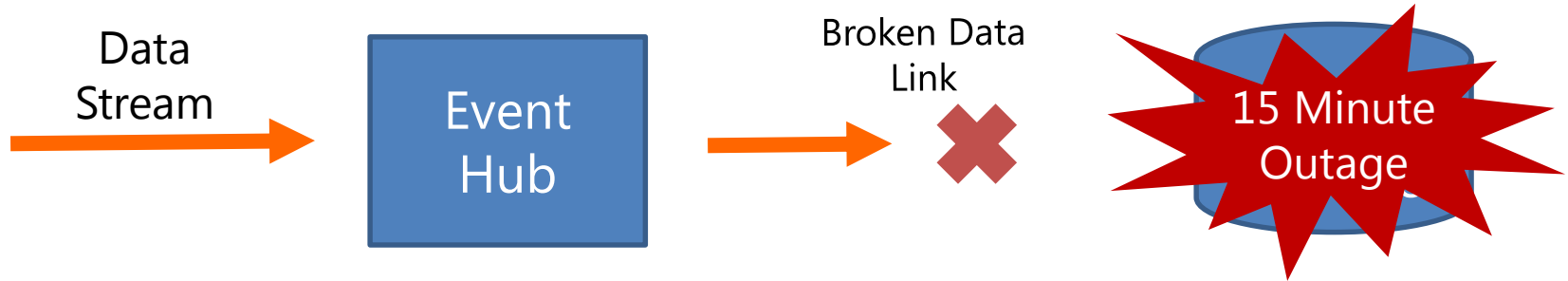
Temporary Storage



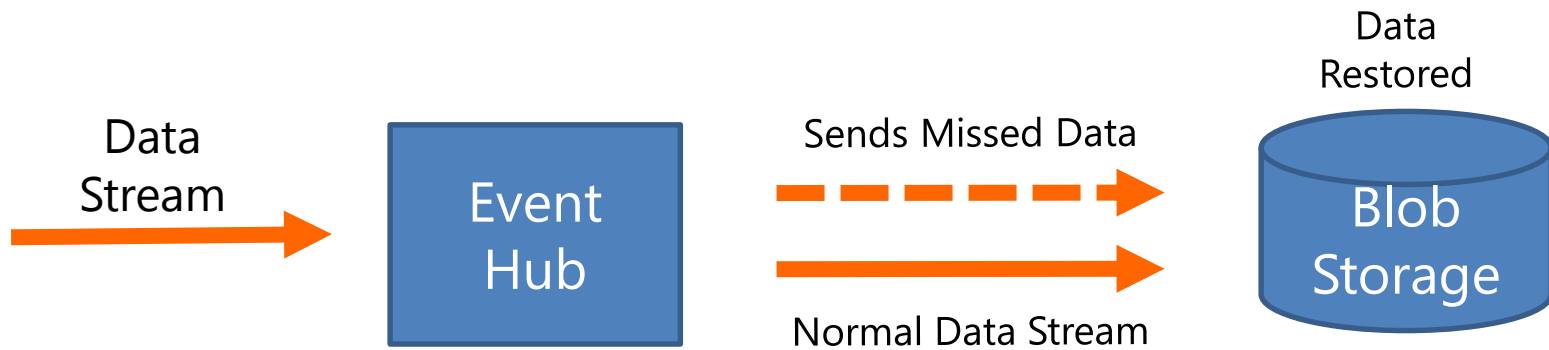
Demo: Normal Scenario



Demo: Output Downage



Demo: Output Restored

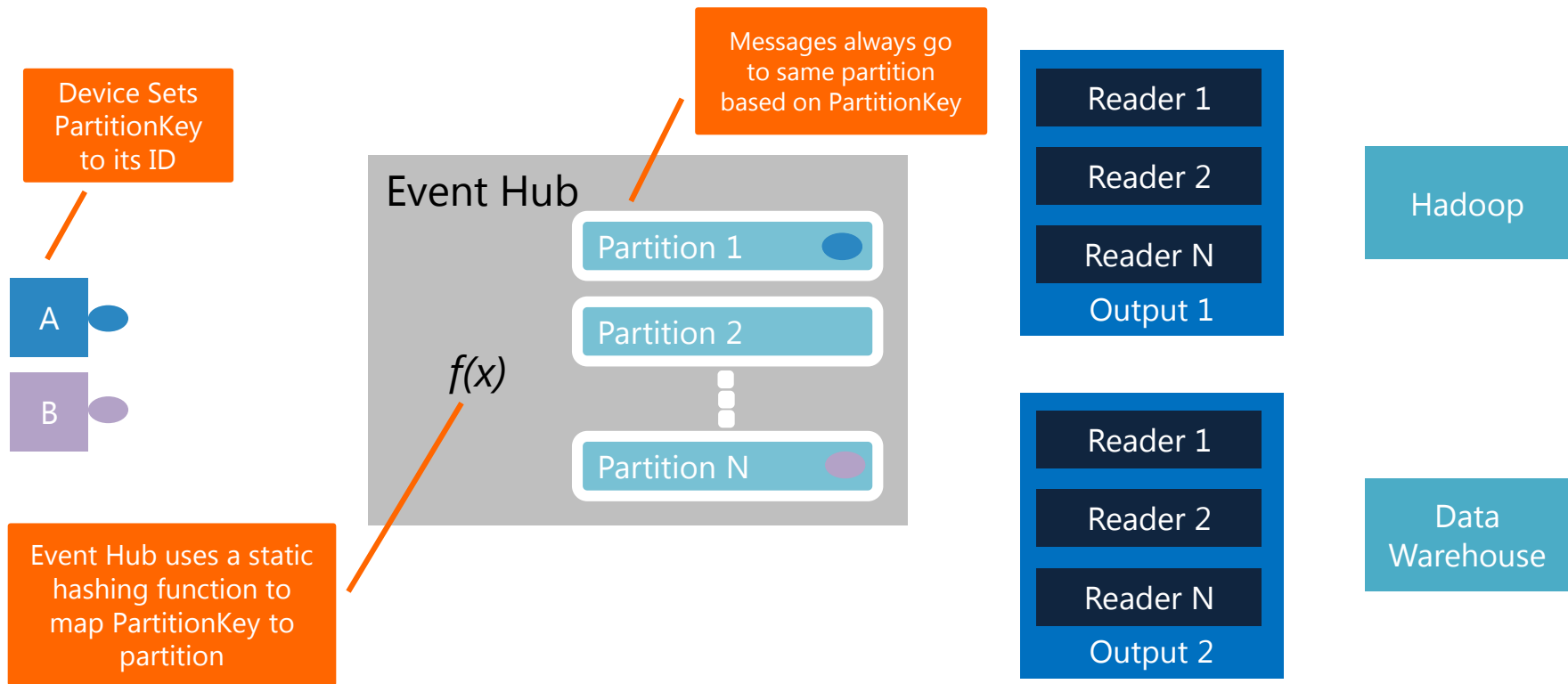


The Post Office



- Tracks address changes
- Tries again tomorrow if send failed
- Holds packages in short term
 - Too many failed deliveries
 - Vacations
- Reduces complexity through specialization

Event Hub, Stream Management



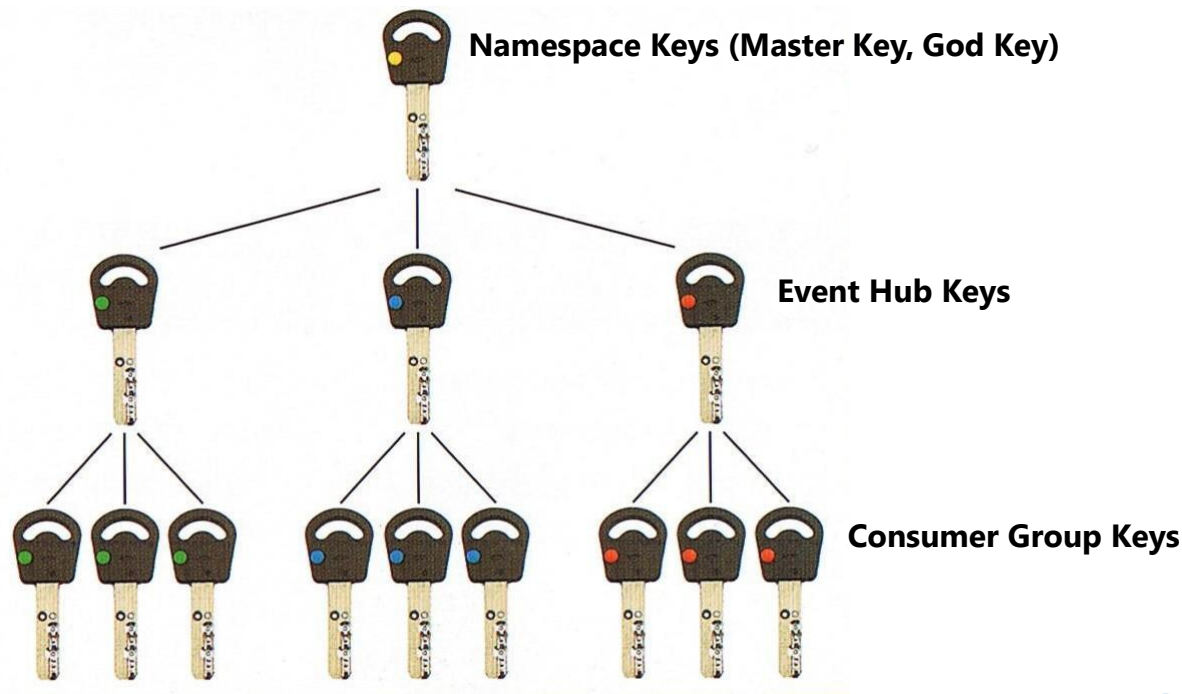
Service Bus Namespace

Service Bus Namespace

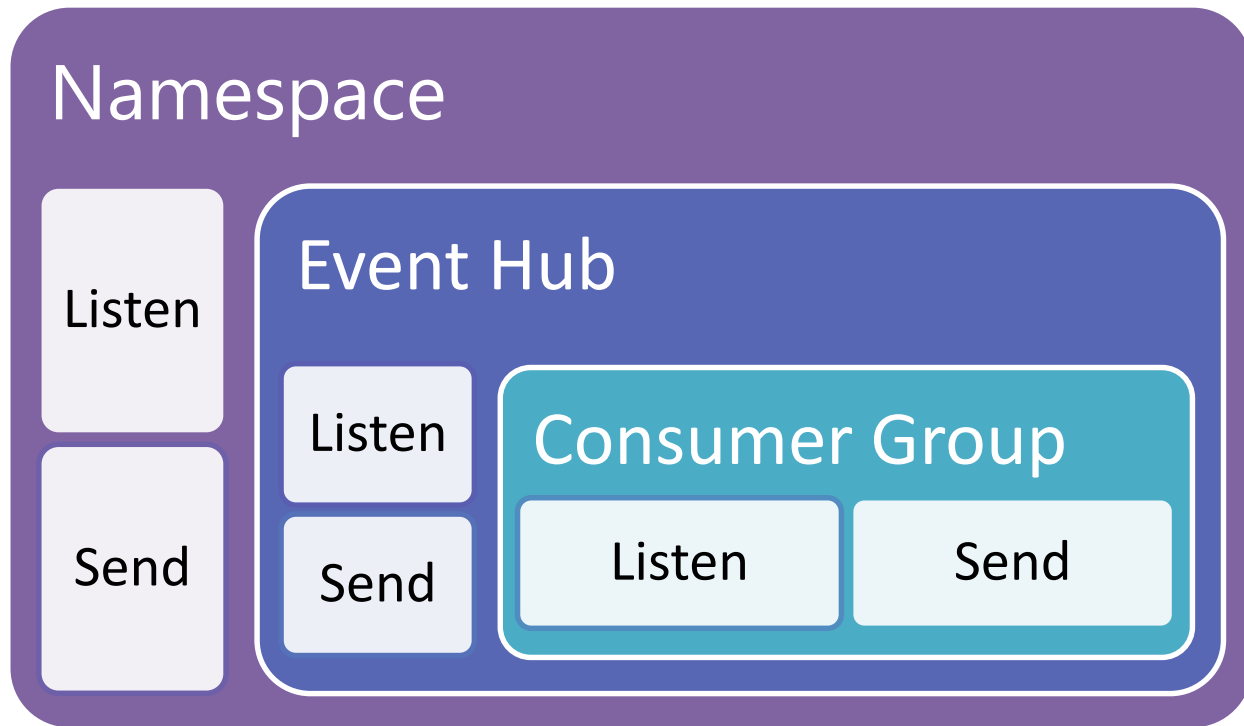
Event Hub 1

Event Hub 2

Access Rights, Policy, Keys



Access Rights



Access Rights



Hands-On Lab

Credit Card Transactions (swipes)



- Credit card transactions are usually done in batch as an end-of-the-day send.
- Stream process for insights now.
- US mainland transactions



Streaming to Event Hub

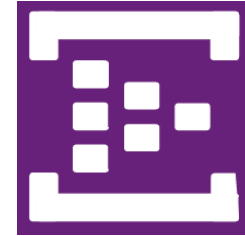


**Credit Card
Reader
(Synthetic)**

Swipes



**Message Broker
(DataScienceDojo's
Webpage)**



**Data Ingestor
(Azure Event Hub)**


The Streamer


- <http://dojodemos.azurewebsites.net/credit-card-streamer/>


Credit Card Streamer

This app will simulate the kind of data streams that banks would encounter, credit card swipe data. The app will generate synthetic data from a credit card transaction (swipe) and pushes it into a given Azure Event Hub as a JSON. The application logic for this app is written entirely in JavaScript so the speed and interval of the transactions is dependent on the processing power of the user device.

Event Hub Credentials

Event Hub Name (Need help? PDF Guide) 

Service Bus Namespace (Need help? PDF Guide) 

Shared Access Policy Name (Need help? PDF Guide) 

Output Preview

Display Format (Data is still sent as a JSON):

```
Successfully loaded database. Ready to simulate data.
```

The Data

```
{  
  "swipe_date":"2015-05-22T20:16:27.122Z",  
  "transaction_id":3127484,  
  "card_type":"VISA",  
  "card_number":"4913419738164560",  
  "expiration_month":"02",  
  "expiration_year":"18",  
  "cvv_code":"520",  
  "user_id":"972288",  
  "user_gender":"male",  
  "user_first_name":"Alexander",  
  "user_last_name":"Hamilton",  
  "merchant":"McDonald's",  
  "transaction_amount":13.64,  
  "balance":336.48,  
  "merchant_fee":.5,  
  "swipe_city":"New York",  
  "swipe_state":"New York",  
  "swip_city_state":"New York, NY",  
  "InstanceNo":1  
}
```

Inside the Event Hub



**Credit Card
Reader
(Synthetic)**

Swipes
→



**Message Broker
(DataScienceDojo's
Webpage)**

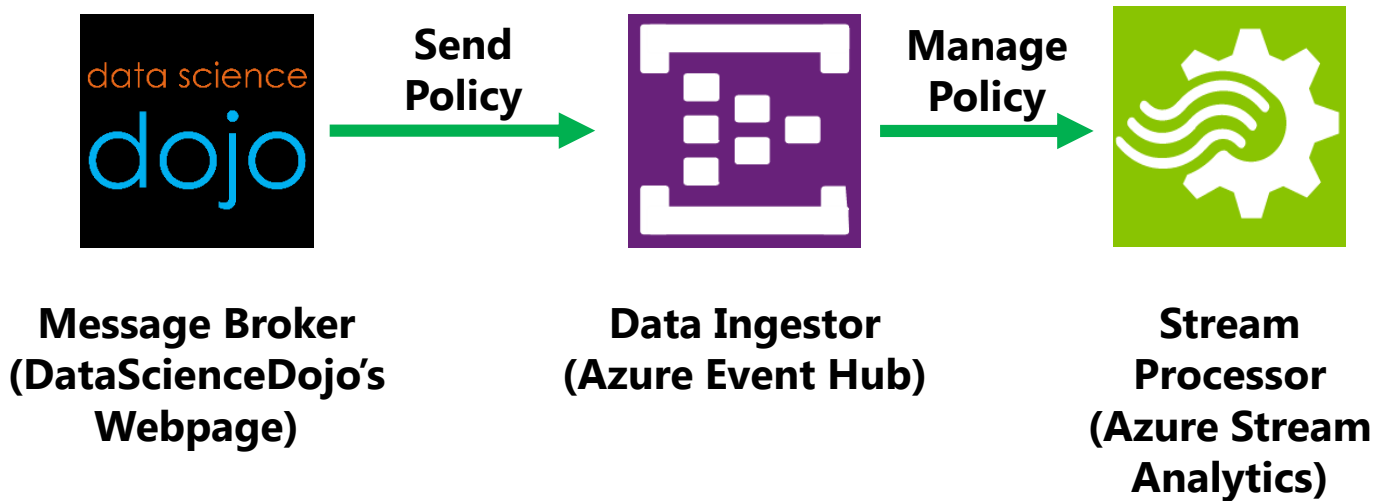
→

Service Bus Namespace:
YourNameSpace

Event Hub 1:

- Credit Card Swipes

Setting Policies



QUESTIONS