

# Streaming PI System Data and Advanced Analytics in Google Cloud

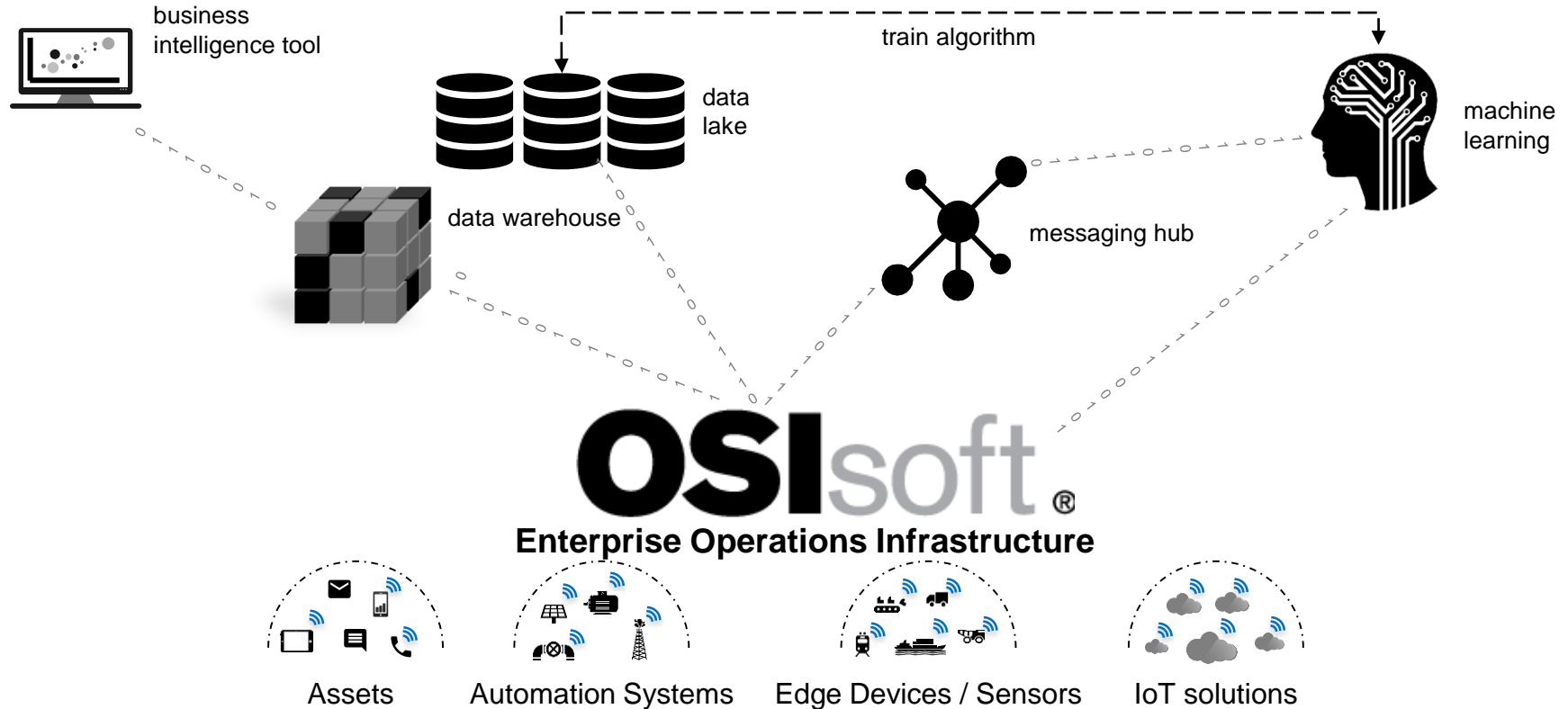
Max Podkolzin

Engineering Manager, OSIsoft

# Agenda:

- PI Integrator overview
- General Best Practices
- Scaling and Performance recommendations
- Data Modeling best practices
- Best Practices for streaming views
- PI Integrator for GCP Overview
- Google Pub/Sub Demo

# Accelerate and operationalize advanced analytics



**PI Integrators** *speed the process that brings trustworthy data to many unique analytics tools*










# General Best Practices

- Install the PI Integrator on its own dedicated server
- Install the latest versions of the PI Asset Framework and PI Data Archive
- Install the PI Integrator in close network proximity to the PI Data Archive
- Consider an aggregation server.
- Performance of the backend PI AF Server SQL database is a key performance factor for the PI Integrator because of the time spent with PI AF-related processing.

# Scaling and Performance Recommendations

[Overview](#)[Log](#)[Security](#)[View Configuration](#)[Statistics](#)

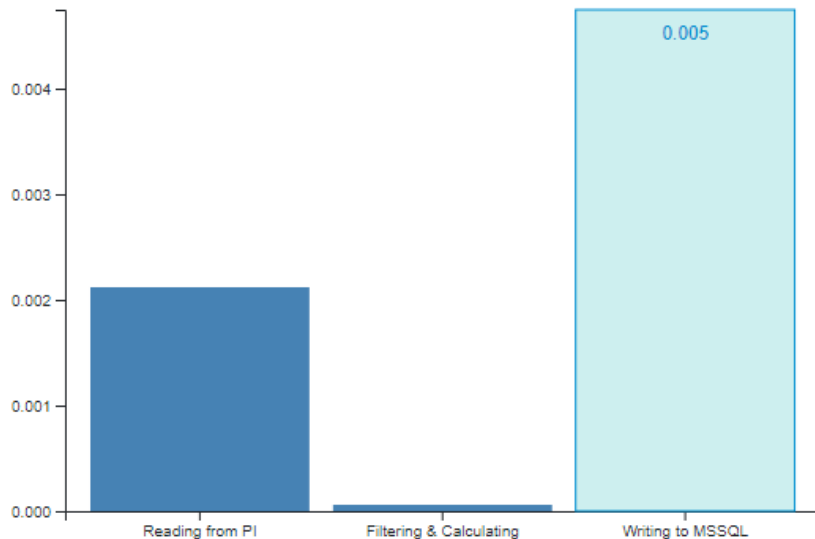
## Run History

Run Instances	<input checked="" type="radio"/> Duration	<input type="radio"/> ...	<input type="radio"/> ...	<input type="radio"/> ...
	seconds ▼			
 4/1/20 3:10 PM	0.097	481	0	0
 4/1/20 3:15 PM	0.063	5	0	0
 4/1/20 3:20 PM	0.047	5	0	0
 4/1/20 3:25 PM	0.047	5	0	0
 4/1/20 3:30 PM	0.078	5	0	0
 4/1/20 3:35 PM	0.063	5	0	0
 4/1/20 3:40 PM	0.063	5	0	0
<b>Total</b>	<b>1197</b>	<b>75.17</b>	<b>6,461</b>	<b>0</b>

[View Logs for Run](#)[Download Selected Report](#)[Enable Full Reporting](#)

## Duration (seconds)

4/1/20 3:20 PM

[Show All Runs](#)

# Scaling and Performance Recommendations

- Minimize using the PI Integrator to generate and publish data on the fly
- Consider using filters for bad/null values.
- Separate views for static/contextual or slow-moving data and faster-updating or industrial instrumentation data (core PI System data).
- Strategize and prioritize your data requests to limit the size and scope of the data sets that are processed and generated.

# Best Practices for Data Modeling within PI AF

- **Equipment-oriented**

- Easier starting point for generating data sets used with a BI tool.
- Contains unique attributes for the columns of that table.
- Less favorable for systems where measurements and attributes are continuously changing.

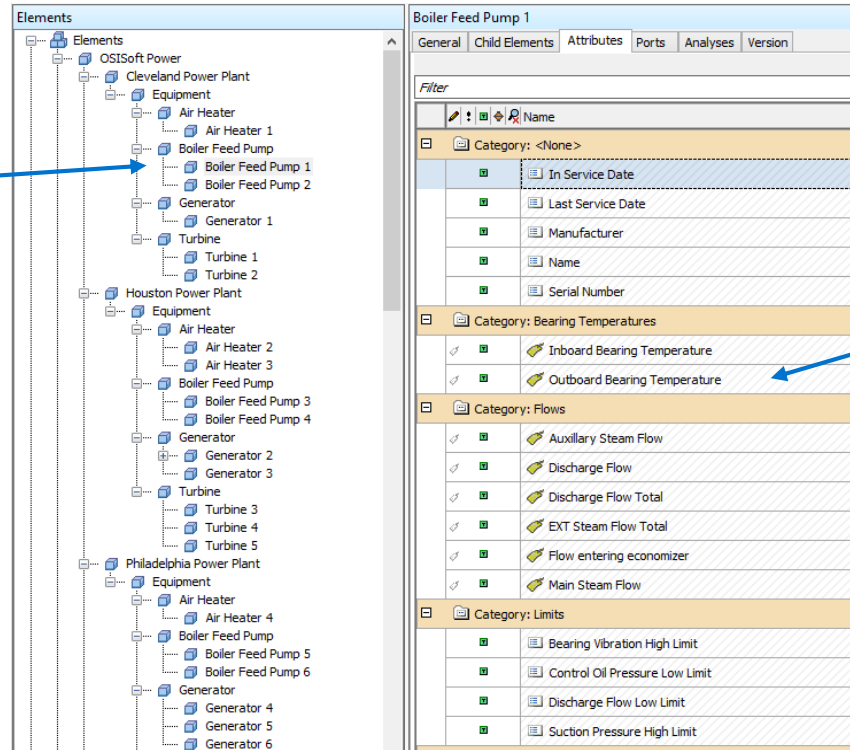
- **Measurement-oriented**

- Measurement-oriented model might be beneficial so that users do not have to republish data each time a measurement is added or removed.



# Equipment-oriented model example

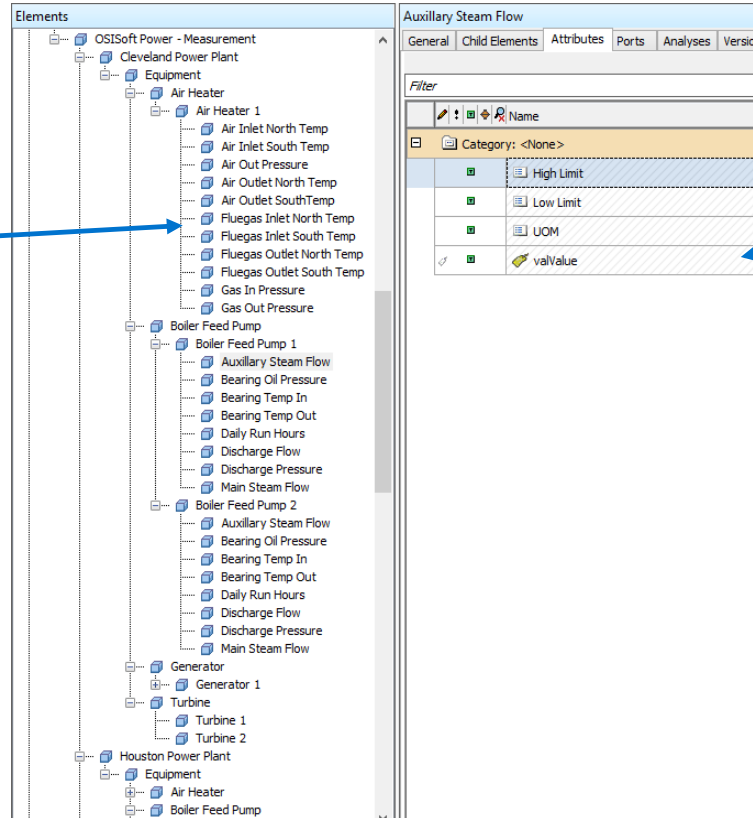
Leaf elements are  
equipment units



Attributes are  
measurements

# Measurement oriented model example

Leaf elements are  
Measurements



Leaf elements are  
Measurement Value  
and properties

# Best Practices for streaming views

- Consider increasing the number of worker nodes on the machine.
- Use a smaller number of views with more attributes in the search shape
- Consider increasing the time interval between scans for large scheduled streaming view
- Azure Event Hubs and Azure IoT Hub writers support specifying the message size and message batch timeout.

# Streaming Views: message triggering

Message Designer

Schema Options  
Syncing mode (flattened)

Message Trigger  
Trigger a new message when 2 key values change

Backfill Data  
Do not backfill data

Message Filters  
0 filters


☐ Trigger a message in regular time intervals: 1 minutes

☒ Trigger a new message when the key value(s) selected below have changed

☒ Trigger a message when **any** of the selected key values have changed

☐ Trigger a message when **all** of the selected key values have changed

Trigger on ☐ Archive Data ☒ Snapshot Data



Message Content

```
{
  "0081F2": "0081F2 (Name)",
  "AltT": "AltT (Value)",
  "Brng": "Brng (Value)",
  "CMsgs": "CMsgs (Value)",
  "Timestamp": "Timestamp",
  "Lat": "Lat (Value)",
  "Long": "Long (Value)",
  "CallSus": "CallSus (Value)"
}
```

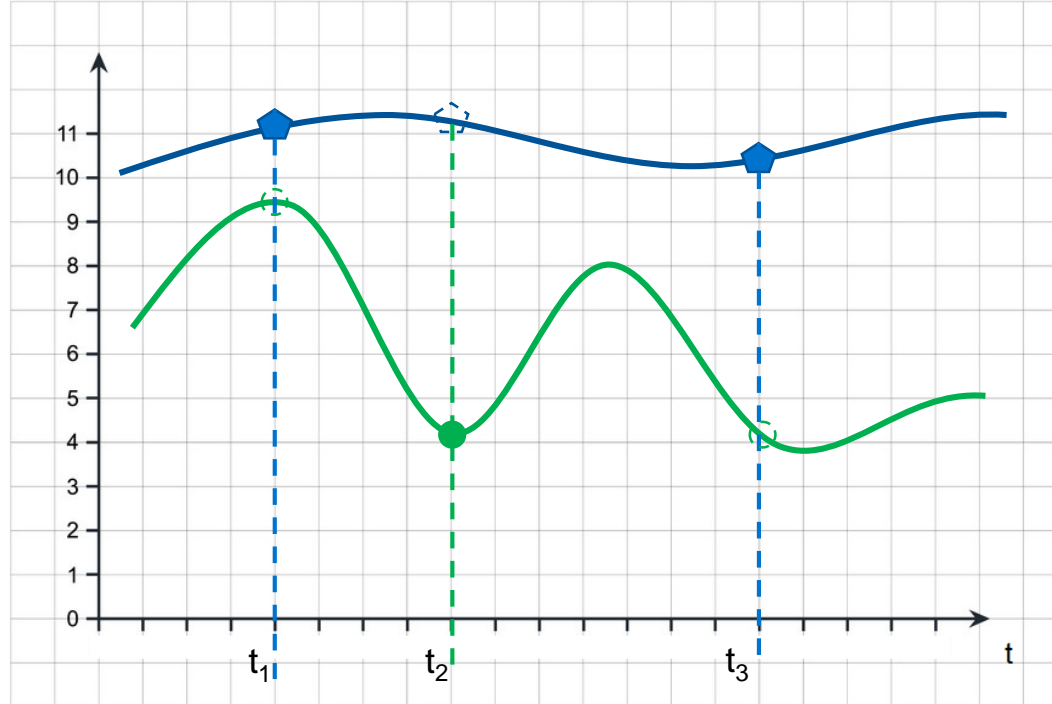
Event-triggering  
mode

Key Attributes

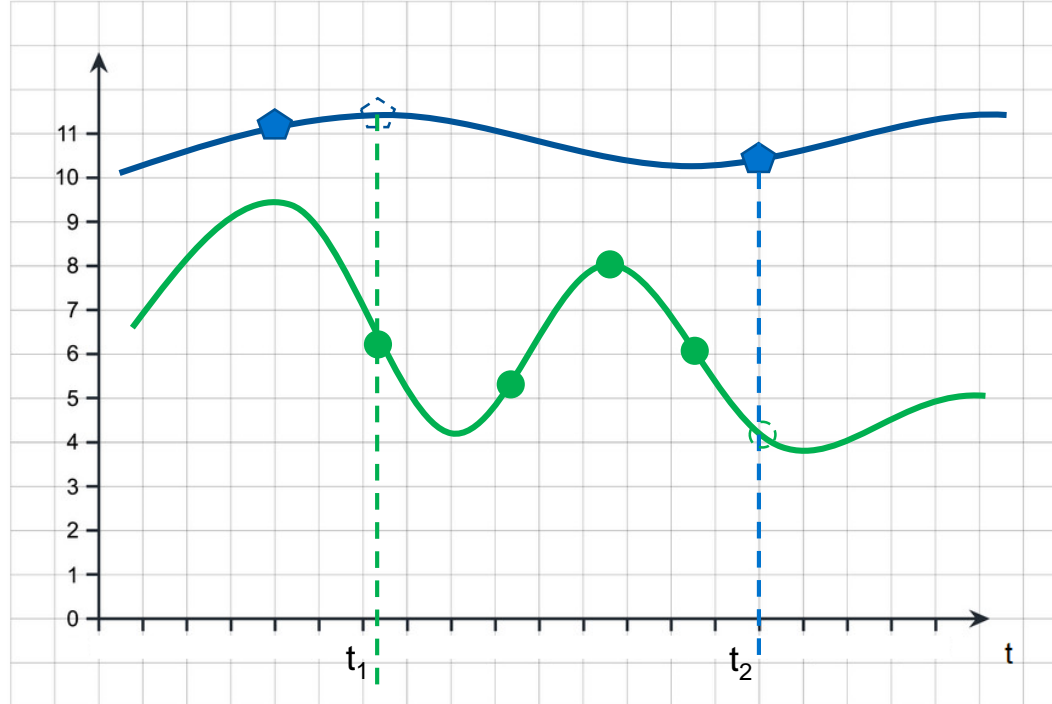
Key Attributes

Snapshot/Archive  
Triggering

# Trigger “When any”



# Trigger “When any”



# Streaming views: routing

Message Designer

Schema Options  
Free-form mode

Message Trigger  
Trigger a new message every 1 minutes

Backfill Data  
Do not backfill data

Message Filters  
0 filters

You are not syncing (free-form) your schema to the asset shape.

Import Schema  
Select Import Source  
Free-form

Select Schema Structure  
Free-form

Save Schema to Registry

```
{
  "N816NN": "N816NN (Name)",
  "Timestamp": "Timestamp",
  "Air Density at Altitude": "Air Density at Altitude (Value)",
  "Air Speed - Acceleration (ASI)": "Air Speed - Acceleration (ASI) (Value)",
  "Altitude - Indicated (AMSL)": "Altitude - Indicated (AMSL) (Value)",
  "Engine Fuel Flow 1": "Engine Fuel Flow 1 (Value)",
  "Engine Fuel Flow 2": "Engine Fuel Flow 2 (Value)",
  "Engine Fuel Pressure 1": "Engine Fuel Pressure 1 (Value)",
  "Engine Fuel Pressure 2": "Engine Fuel Pressure 2 (Value)",
  "View Id": ""
}
```

Add Property to Schema

Edit Property

Data Property Time Property Static Value

Select Static Value Option

View ID

View Name  
Custom

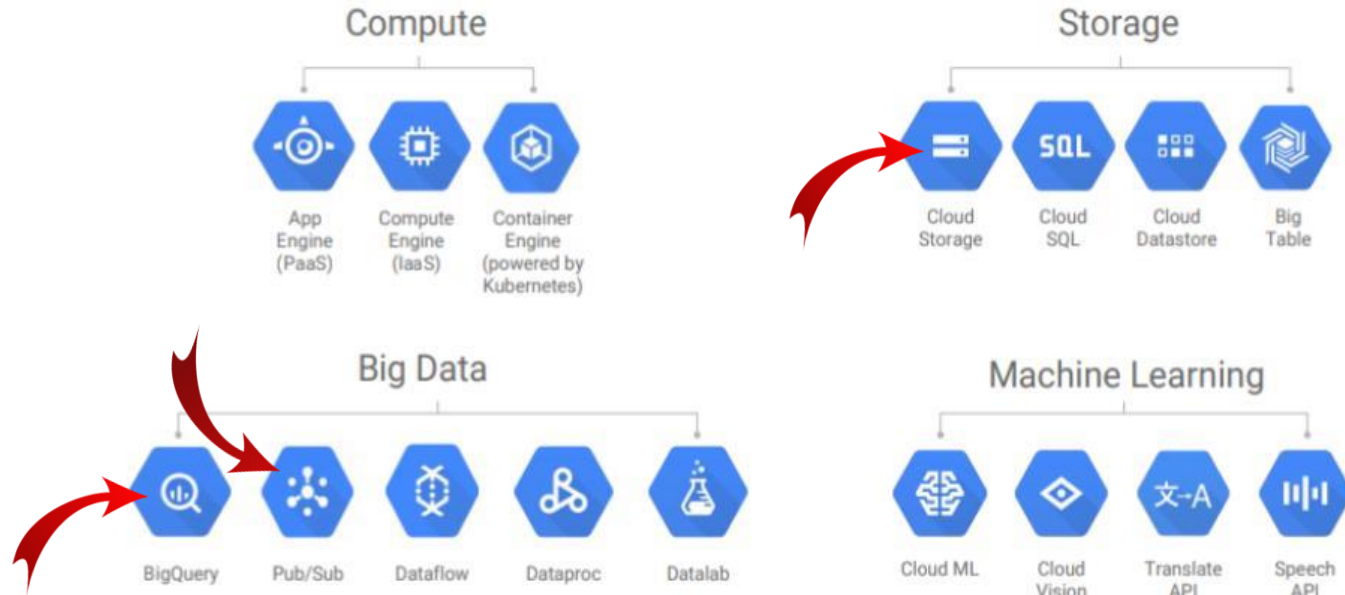
Property Name  
View Id

Property Data Content  
ef1ac2bc-2b33-4495-9b61-ca41fb7ddbdb

Cancel Update Property

```
{
  "N816NN": "N816NN",
  "Timestamp": "2020-03-04T07:52:32.4696113-08:00",
  "Air Density at Altitude": 0.4493381083011627,
  "Air Speed - Acceleration (ASI)": 0.5210391879081726,
  "Altitude - Indicated (AMSL)": 25013.078125,
  "Engine Fuel Flow 1": 0.7463105916976929,
  "Engine Fuel Flow 2": 0.7463105916976929,
  "Engine Fuel Pressure 1": 29.944839477539062,
  "Engine Fuel Pressure 2": 29.944839477539062,
  "View Id": "ef1ac2bc-2b33-4495-9b61-ca41fb7ddbdb"
}
```

# Google Cloud Platform Overview





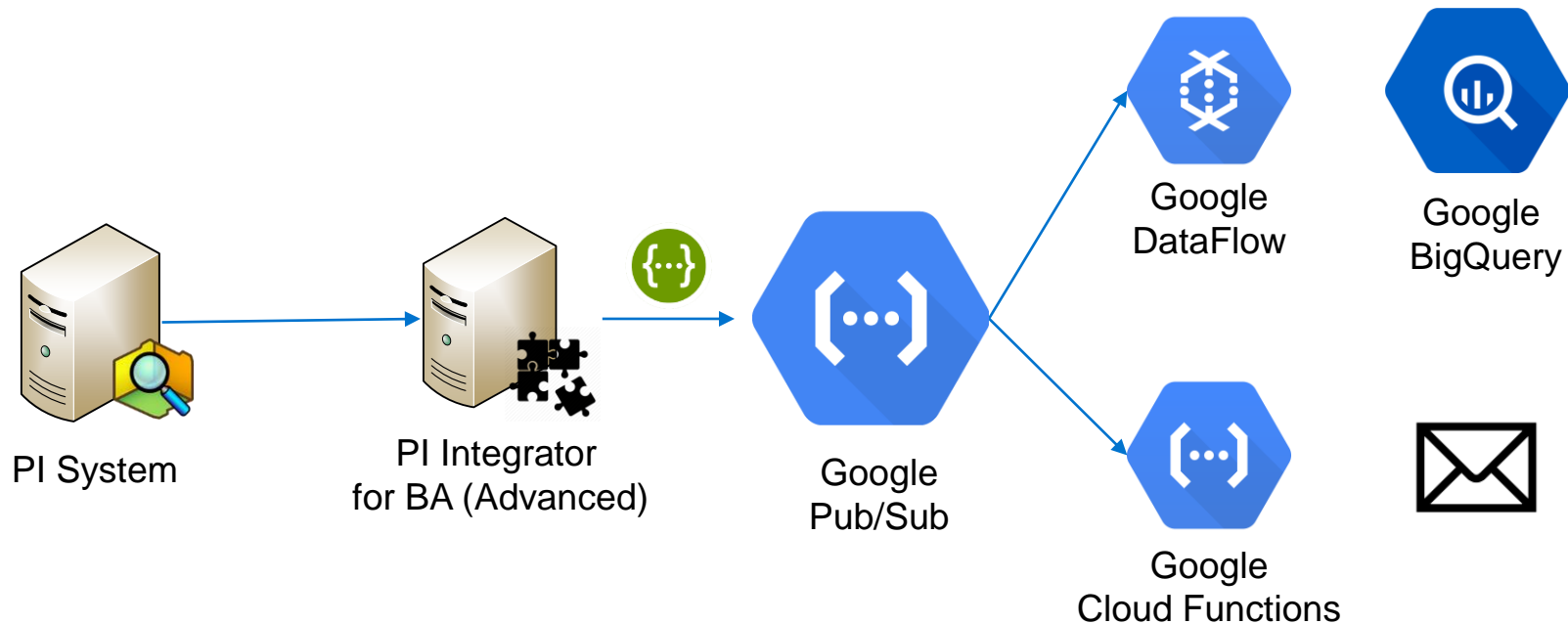
# Google Pub Sub Core concepts:

- **Topic:** A named resource to which messages are sent by publishers.
- **Subscription:** A named resource representing the stream of messages from a single, specific topic, to be delivered to the subscribing application..
- **Message:** The combination of data and (optional) attributes that a publisher sends to a topic and is eventually delivered to subscribers.
- **Message attribute:** A key-value pair that a publisher can define for a message.

# Google Pub Sub Key Use Cases:

- **Data streaming from various processes or devices.**
- Balancing workloads in network clusters.
- Implementing asynchronous workflows.
- Distributing event notifications.
- Refreshing distributed caches.
- Logging to multiple systems.
- ...

# Demo overview:



# DEMO

# Speakers



- Max Podkolzin
- Engineering Manager
- OSIsoft
- [mpodkolzin@osisoft.com](mailto:mpodkolzin@osisoft.com)

# Questions?

Please wait for  
the **microphone**

State your  
**name & company**



## Save the Date...



AMSTERDAM  
October 26-29, 2020



