

Streaming Analytics

with NiFi & Python

Mario Barrientos
Septiembre 2020

**Mario Francisco
Barrientos Alejos**

mbarrientos@martinexsa.com

Computer Science Engineer
3+ años de experiencia en Big Data
& Analytics

Cloudera Certified Administrator
Cloudera Spark Certified developer
HDP Certified Developer



Common use cases

- **Data Movement**

Optimize resources, on-premise to cloud.

- **Single view 360° / view of customer**

Ingest, transform, combine data into data lakes.

- **Stream Processing**

Process multiple streams of data into different endpoints.

- **Optimize log collection & analytics**

Optimize log analytics solutions.

- **Gain key insights with Streaming Analytics**

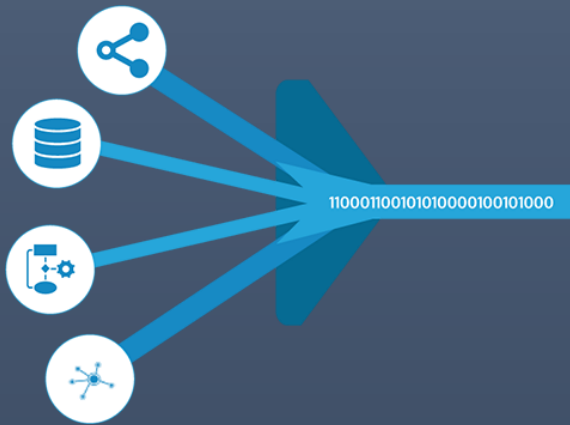
Analyzing streaming data with ML models.

- **Capture IoT data**

Ingest sensor data from IoT devices.



Key customer challenges



Sources variants



Real-time analytics



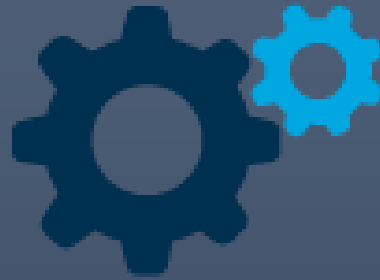
Visibility



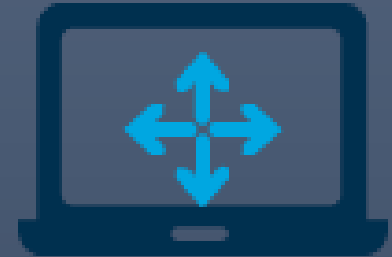
Data in motion



Ingest



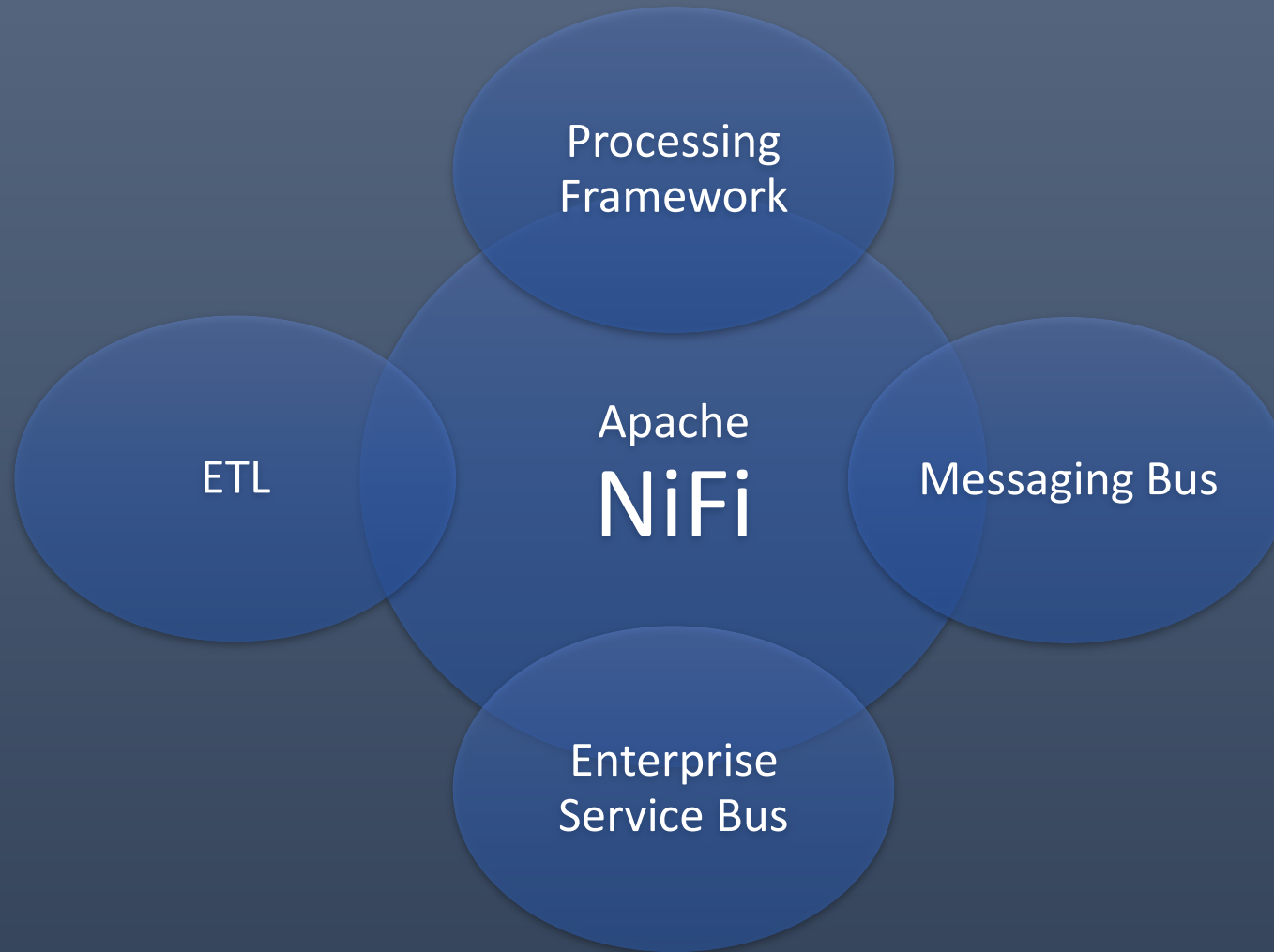
Process



Serve



NiFi Positioning

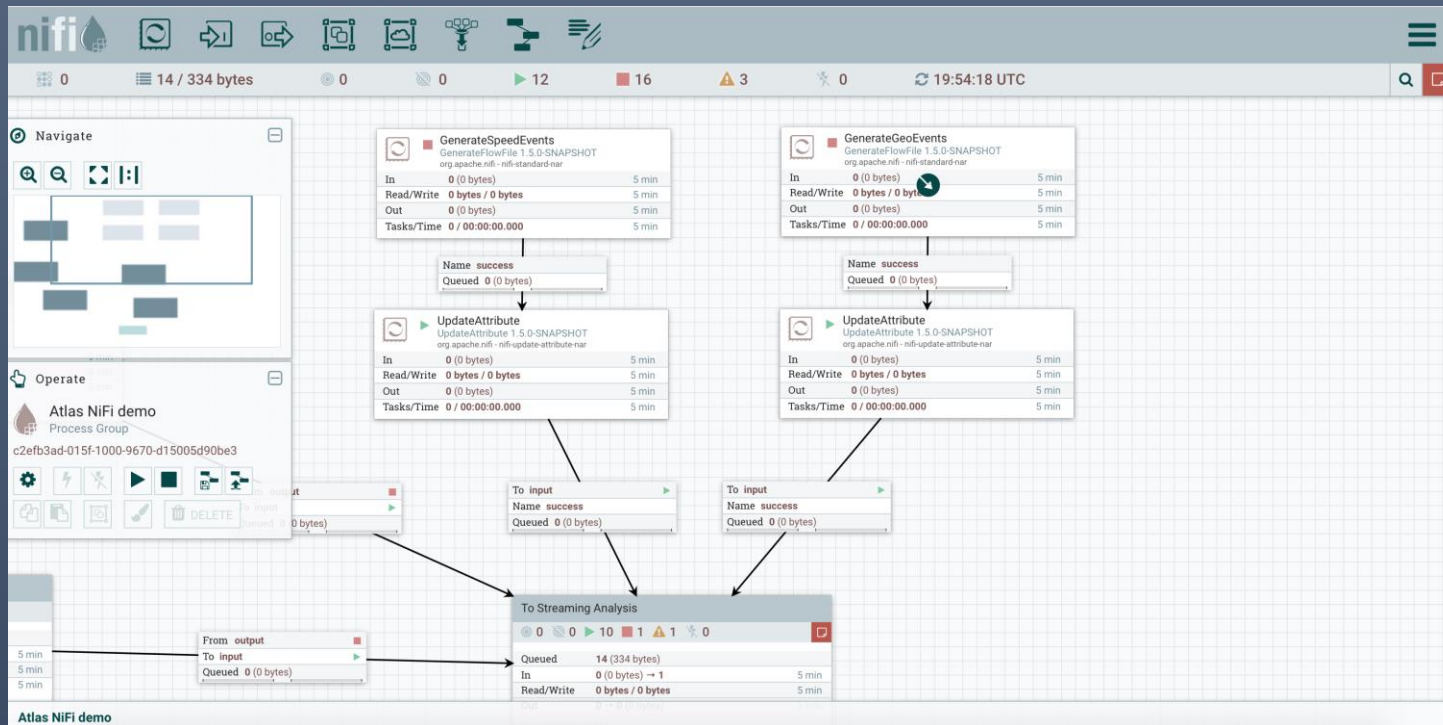


Flow Management

- Web-based user interface
- Highly configurable
- Out-of-the-box data provenance
- Secure



Apache NiFi



Automate and manage the flow of data between systems

End to end, can collect, curate, analyze and act on data in real-time, on premises, or in the cloud

Drag-and-drop visual interface.

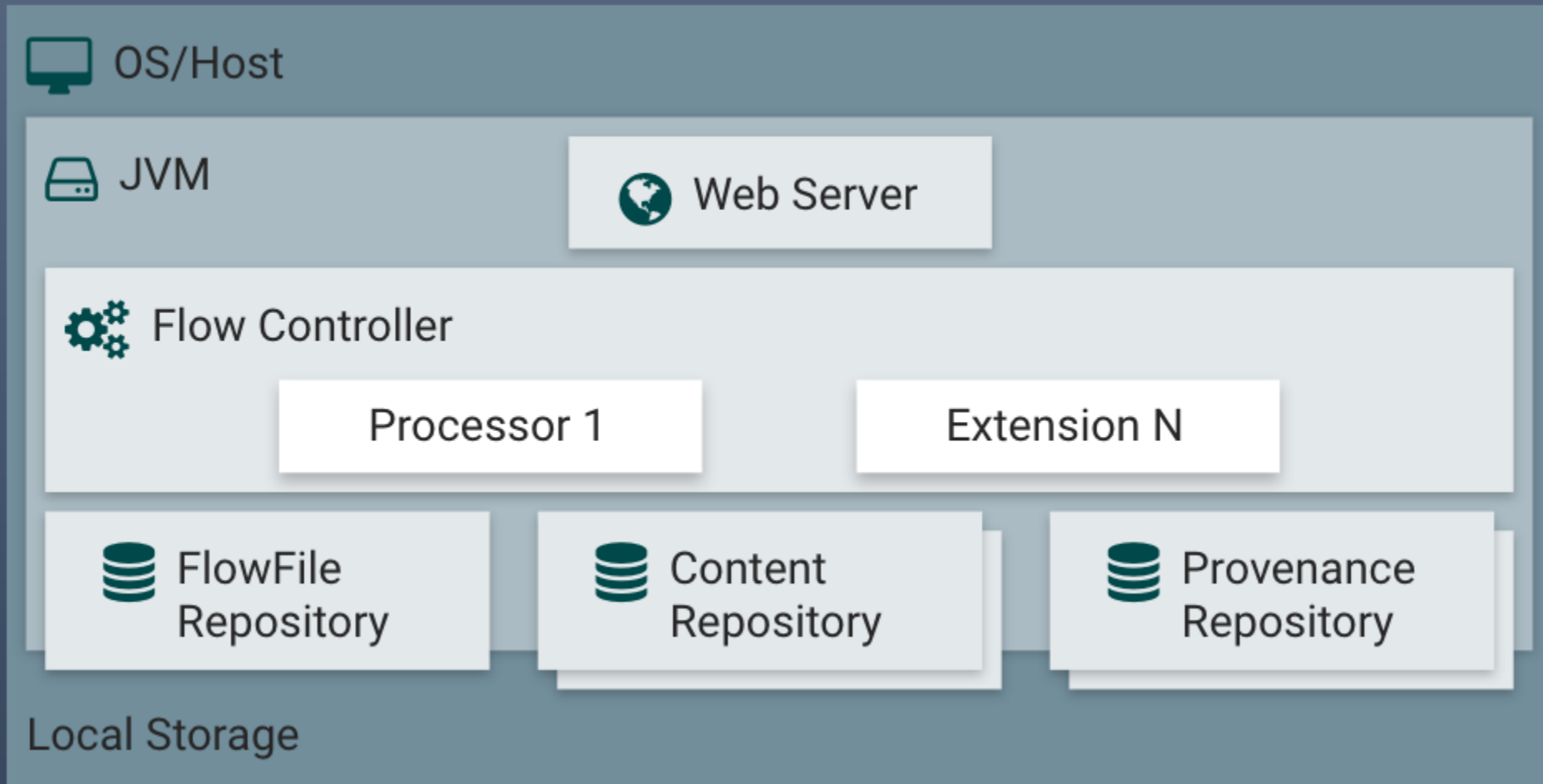
Flowfiles

Flowfile Processors

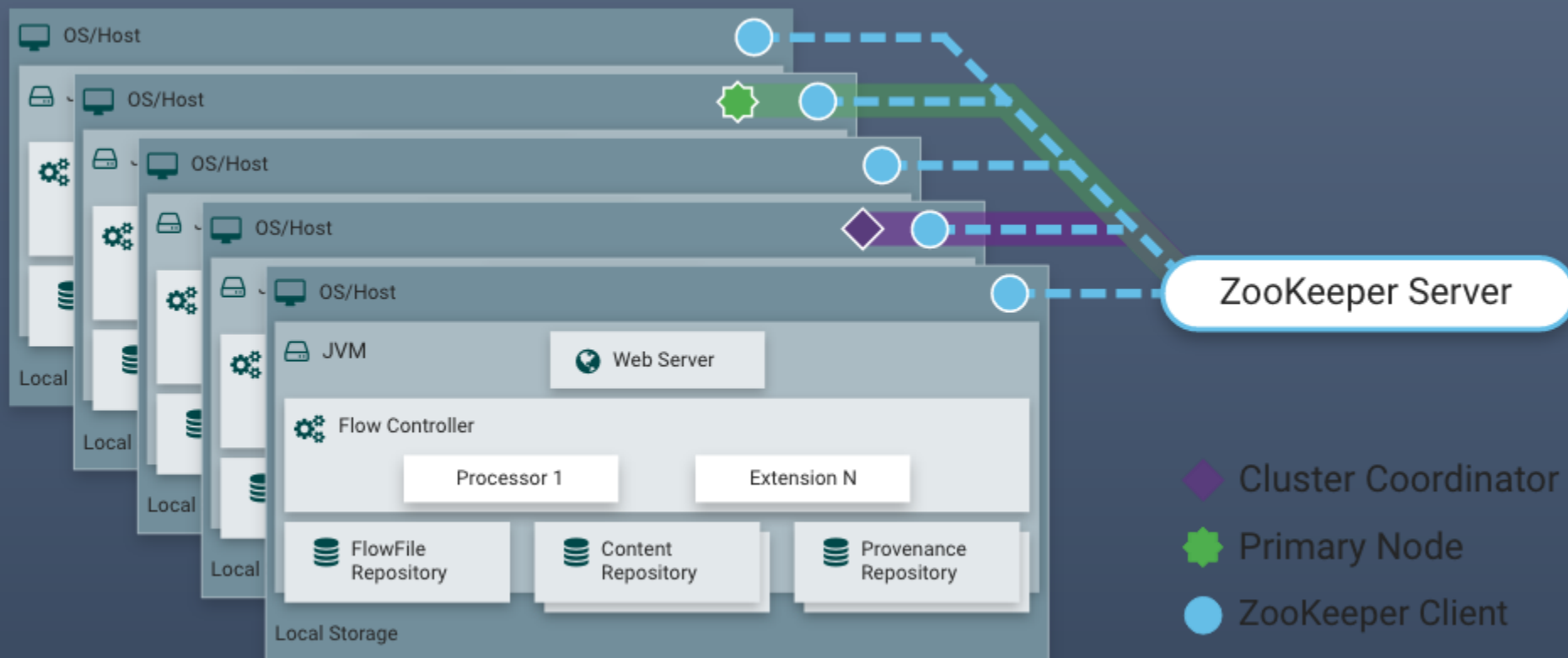
Connections



Apache NiFi – Architecture



Apache NiFi – Architecture (Cluster)



Apache NiFi

280+ PROCESSORS FOR DEEPER ECOSYSTEM INTEGRATION



NiFi flow design is like software development



Programming Language

- IDE
- Algorithm, develop
- Functions/Module
- Libraries, packages

Apache NiFi

- GUI
- Flow Design, drag & drop
- Process Groups (i/o)
- Templates

<https://nipyapi.readthedocs.io>

Apache NiFi + Python

Use Python inside NiFi

- Processors:
 - ExecuteScript
 - ExecuteStreamCommand

Processor Details

SETTINGS SCHEDULING PROPERTIES COMMENTS

Required field

Property	Value
Script Engine	python
Script File	No value set
Script Body	import cx_Oracle #Added for nifi START import java.io from ...
Module Directory	C:\Users\ahmadshe\AppData\Local\Programs\Python\Pyt...

Access NiFi through Python

- Low-level client API (NiPyAPI)
- Still Beta
- Access NiFi components through Python
- Import/Export versioned flow using NiFi Registry





DEMO

Demo

