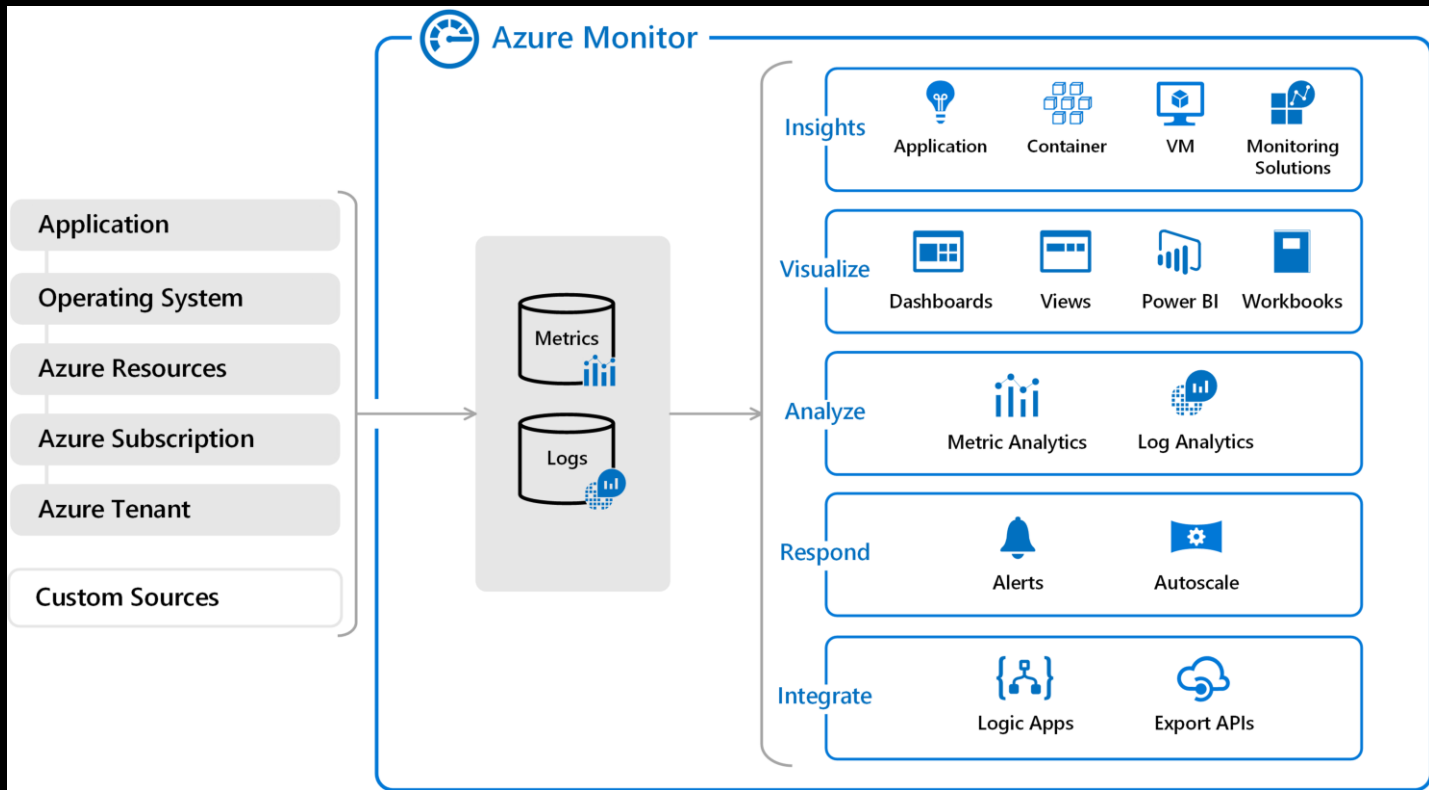




Azure Monitor Service





Metrics Explorer





☒ Show legacy language converter

Event | where EventLevelName == "Error" | project TimeGenerated, Computer, EventLevelName, Source, EventID

Advanced Analytics 00:00:00.339

4K Results ☒ List ☐ Table

Drag a column header and drop it here to group by that column:

TimeGenerated	Computer	EventLevelName	Source	EventID
7/17/2017 11:39:02 AM	srv01.contoso.com	Error	Microsoft Windows-L...	5873
7/17/2017 11:39:12 AM	srv01.contoso.com	Error	HealthService	4502
7/17/2017 11:39:12 AM	srv02.contoso.com	Error	HealthService	4502
7/17/2017 11:39:12 AM	srv01.contoso.com	Error	HealthService	4502
7/17/2017 11:39:12 AM	srv03.contoso.com	Error	HealthService	4502
7/17/2017 11:39:26 AM	srv03.contoso.com	Error	NRM Agent	100
7/17/2017 11:39:26 AM	srv02.contoso.com	Error	NRM Agent	100

Log Analytics



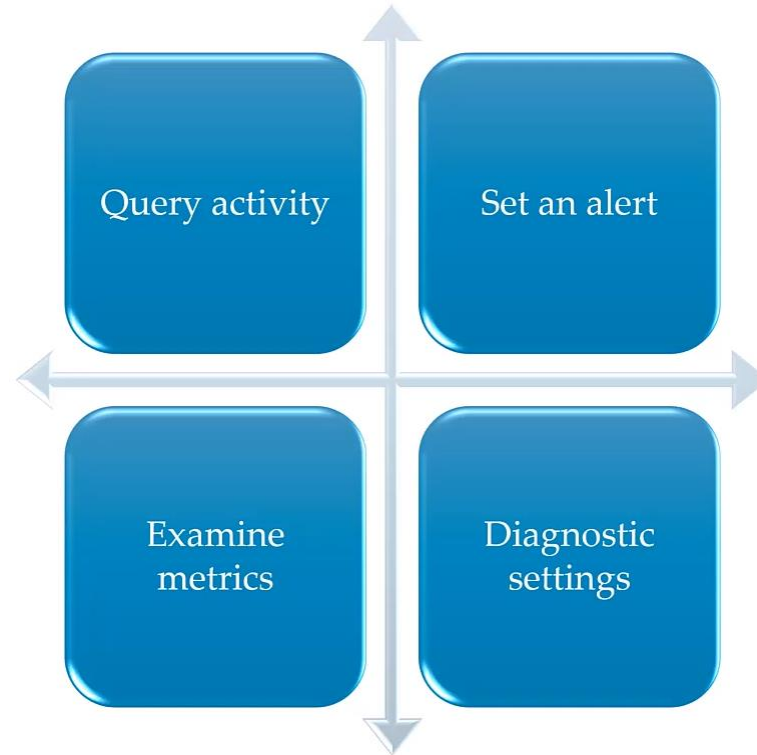
Azure Blob/ Data Lake Monitoring



Monitoring

- Blob and Data Lake have same inbuild monitoring tools
 - Because Data Lake is built on top of Blob
- Monitoring options
 - Insights/Workbooks
 - Metrics/Alerts
 - Classic Diagnostic settings
 - Logging aspects and retention period

Monitoring





Databricks - Monitoring



- **Ganglia**

- Built in to Databricks by default
- Default collection in every 15 min
- Can view snapshot of live data

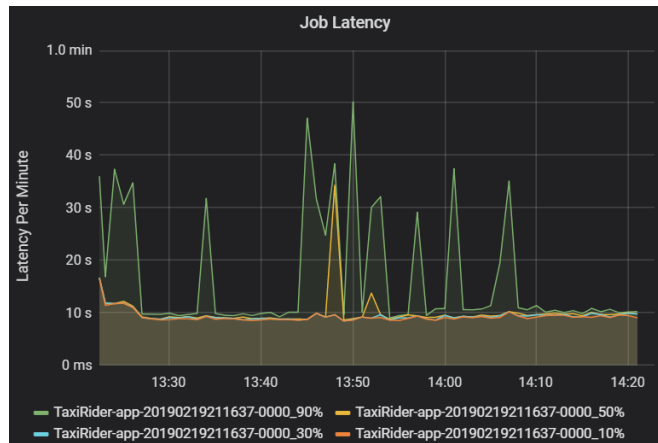
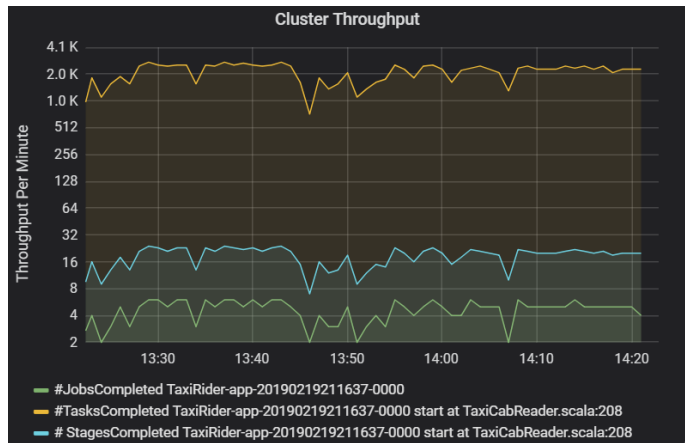
- **Azure Monitor**

- No Native support for Databricks
- Download and Install “Dropwizard metrics library”
 - Build a JAR file – Spark-listeners-loganalytics-1.0-SNAPSHOT.jar
 - This is available on GitHub
 - Then you can create gauges or counters in your application code
 - You can also use Log4j appender in the same library.
 - Then you can create a log4j.properties configuration file

Databricks - Monitoring

- **Grafana**

- Open source visualization platform
- 3 steps required:
 - Install Dropwizard metrics library in Databricks
 - Configure Azure Log Analytics workspace
 - Deploy and configure Grafana



Stream Analytics Service

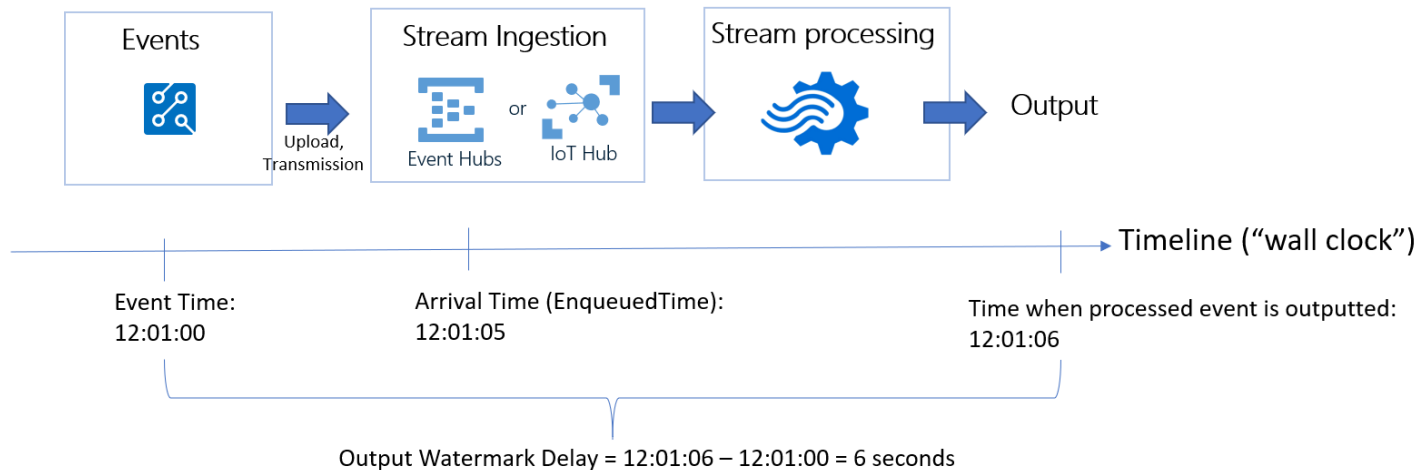


- **Jobs can be monitored**
 - Azure Portal
 - PowerShell
 - .NET SDK
 - Visual Studio
- **Important metrics**
 - SU% Utilization
 - Runtime Error
 - Watermark delay
 - Input deserialization error
 - Backlogged Input events
 - Data Conversion Errors

Watermark delay matrices

Simple case: no time window, late arrival and out-of-order policy set to 10 seconds

```
SELECT *  
FROM input TIMESTAMP BY eventTime
```



Watermark delay matrices

