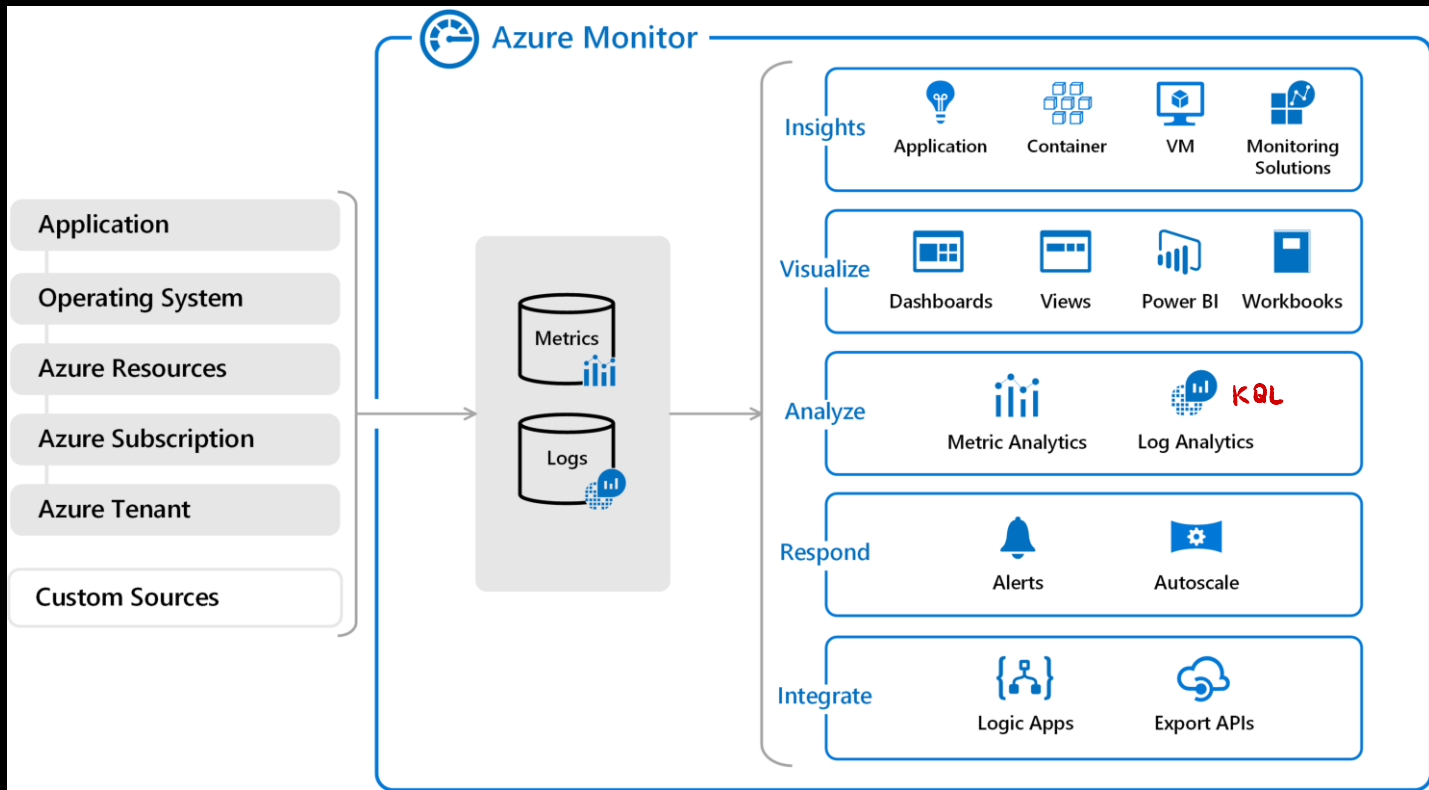




Azure Monitor Service





Metrics Explorer



QUERY w/ KQL (KUSTO QUERY LANGUAGE)



☒ 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



* REVIEW IN PORTAL

Azure Blob/ Data Lake Monitoring

IN STORAGE ACCOUNT



- Blob and Data Lake have same inbuilt 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

CATS

OVERVIEW

FAILURES

PERFORMANCE

AVAILABILITY

CAPACITY

LOG CONTAINERS ACCESSED
THROUGH STORAGE EXPLORER (\$1095)

CUSTOMIZABLE

WORKBOOKS (TEMPLATES OR CREATE)

Monitoring

METRICS

LOGS

ALERT RULES

DIAGNOSTICS

QUERY
PLANS

Query activity

Set an alert

CREATE
RULES

SCOPE
CONDITION
ACTION

WHAT IS
HAPPENING?

Examine
metrics

Diagnostic
settings

WRITE METRICS/LOGS
TO STORAGE

- LOG ANALYTICS
- STORAGE ACCT
- EVENT HUB



REVIEW DATA FACTORY MONITOR

- ADF PORTAL
- AZURE PORTAL



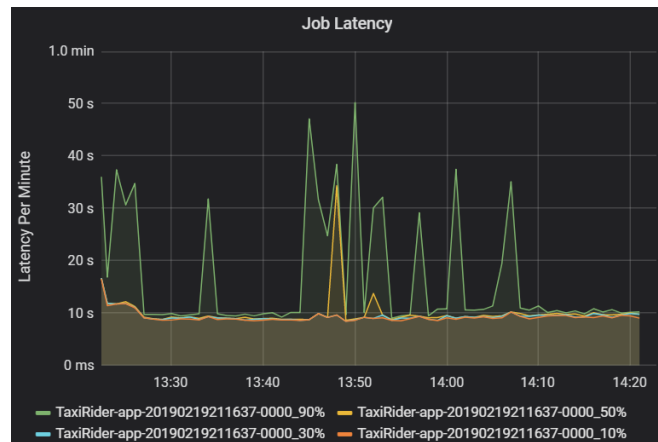
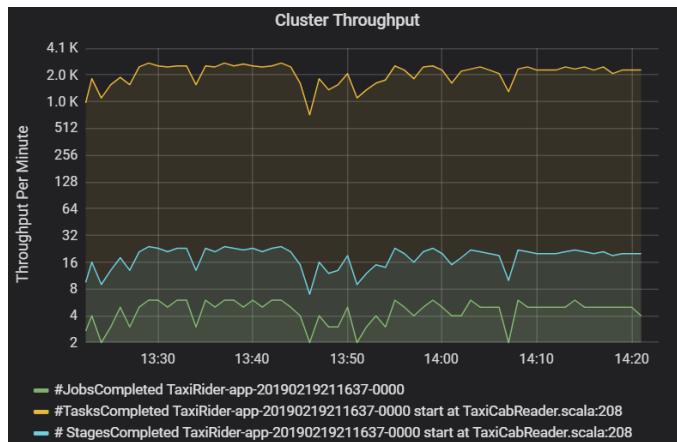


- **Ganglia** DB WORKSPACE > CLUSTER > METRICS LOGS
 - Built in to Databricks by default
 - Default collection in every 15 min
 - Can view snapshot of live data
- **Azure Monitor** (BEST SOLUTION)
 - No Native support for Databricks
 - Download and Install "Dropwizard metrics library"
 - Build a JAR 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 **THIRD - PARTY**
 - 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



IF SU% LOW, CAN
SCALE DOWN TO
SAVE ON COSTS

- Jobs can be monitored

- Azure Portal
- PowerShell
- .NET SDK
- Visual Studio

- Important metrics

- SU% Utilization **STREAMING UNITS**
- Runtime Error
- Watermark delay **JOB HEALTH SIGNALS**
- Input deserialization error **MALFORMED ERRORS**
- Backlogged Input events **BOTTLENECK**
- Data Conversion Errors

MONITORING SECTION

LOGS

METRICS

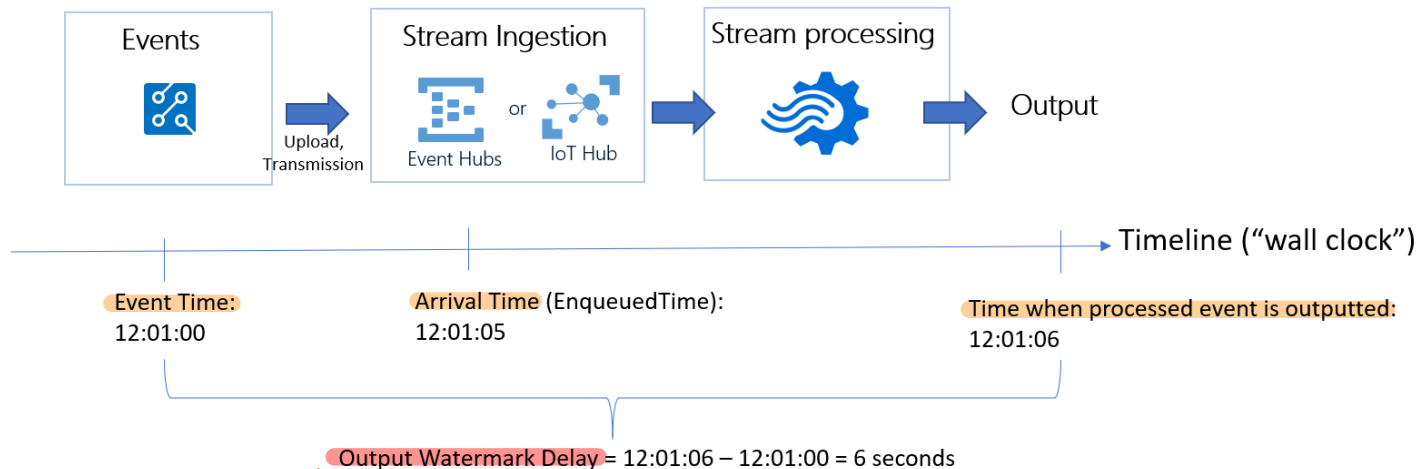
ALERT RULES

DIAGNOSTICS

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
```



SHOULD BE SMALL. INDICATES TIMELY STREAMS.

Watermark delay matrices

CAN SET ALERTS BASED ON WATERMARK THRESHOLDS

