

Wide World of Logs

Monitoring in the cloud



Agenda

1

Introductions

2

Problem Background

3

Sample of Solution

4

Azure Log Analytics

5

What is Kusto Query Language (KQL)?

6

Demo

7

Cost Considerations

8

Summary and Questions

About Me



1 Data Strategist

2 Multi-Cloud Architect

3 APAC AI Organiser

4 Loves Science Fiction Books

5 Loves Sports (Cricket)



Contact Me



[/in/sarathboppudi](https://www.linkedin.com/in/sarathboppudi)



[@BoppudiSarath](https://twitter.com/BoppudiSarath)



<https://sarathboppudi.com>



<https://github.com/sandman153>



[/in/sarathboppudi](https://www.linkedin.com/in/sarathboppudi)



[@BoppudiSarath](https://twitter.com/BoppudiSarath)



<https://sarathboppudi.com>



<https://github.com/sandman153>

Problem Statement

Greenfields Project

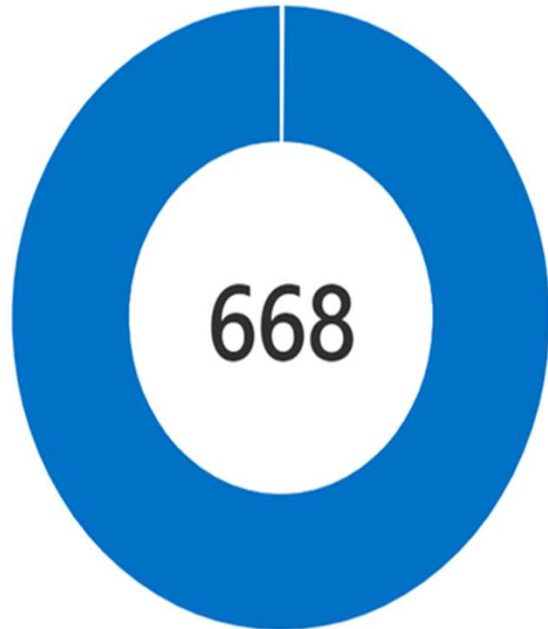
Number of resources to develop solution is small

Minimal timeframe available to build a solution





















Repeatable process

Minimize Costs

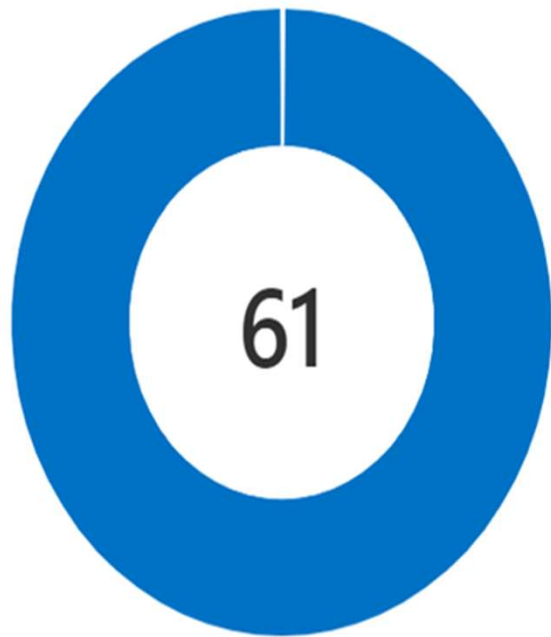
Pipelines Run by Data Factory



Pipelines Run Trend




























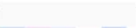
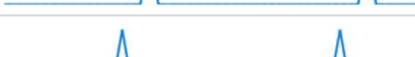





ResourceId	↑↓	Count	↑↓	Trend	Status	↑↓	PipelineName
DATAFACTORY-[REDACTED]		6			 Queued		SendProcessingEmail
DATAFACTORY-[REDACTED]		6			 Succeeded		SendProcessingEmail
DATAFACTORY-[REDACTED]		91			 Queued		Refresh Power BI Dataset
DATAFACTORY-[REDACTED]		88			 Succeeded		Refresh Power BI Dataset
DATAFACTORY-[REDACTED]		2			 Failed		Refresh Power BI Dataset
DATAFACTORY-[REDACTED]		77			 Queued		Refresh Power BI Dataflow
DATAFACTORY-[REDACTED]		76			 Succeeded		Refresh Power BI Dataflow
DATAFACTORY-[REDACTED]		1			 Failed		Refresh Power BI Dataflow
DATAFACTORY-[REDACTED]		2			 Succeeded		Migrate Roster Data For [REDACTED]
DATAFACTORY-[REDACTED]		2			 Queued		Migrate Roster Data For [REDACTED]

Triggers Run by Data Factory

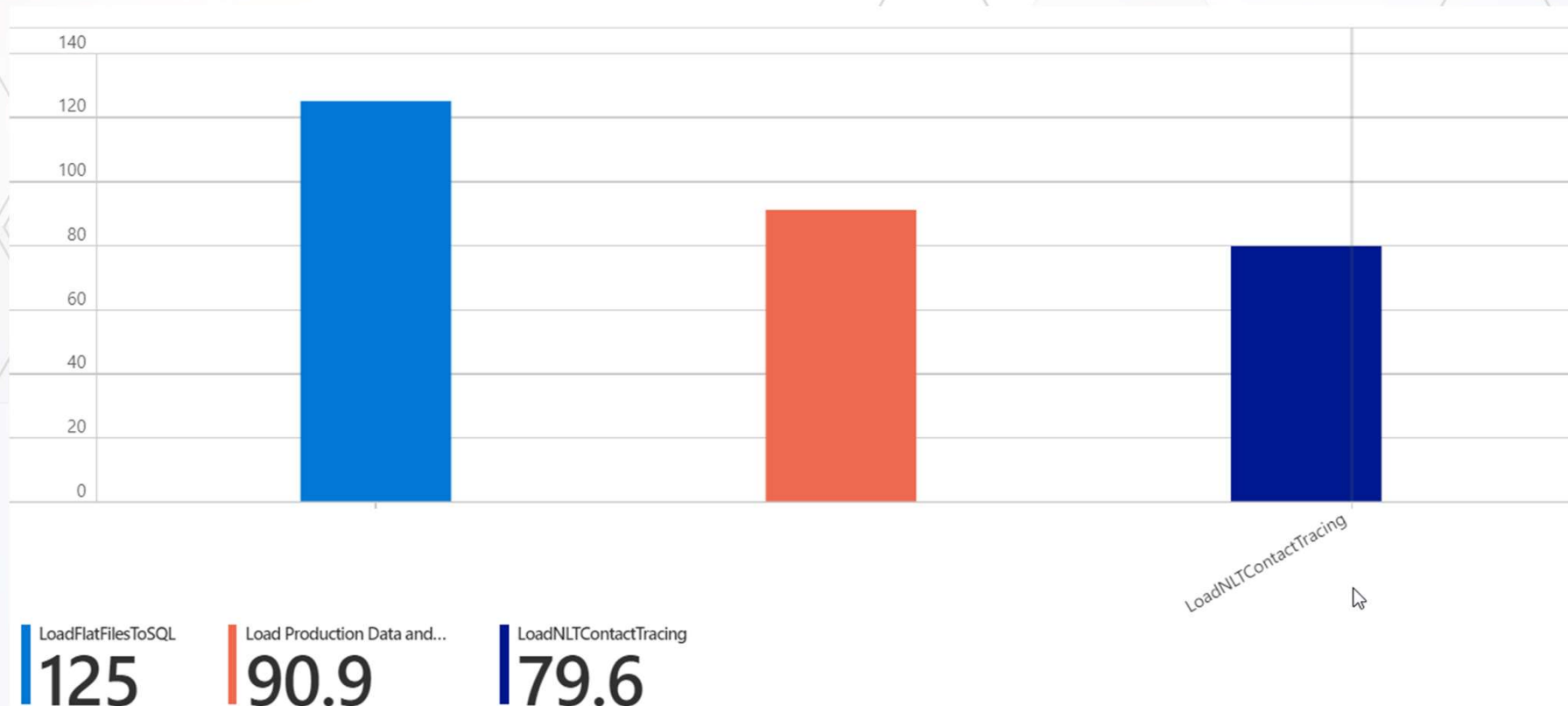


/SUBSCRIPTIONS/ 61 /RESOURCEGROUPS/ /PROVIDERS/MICROSOFT.DATAFACTORY/

Triggers Run Trend

ResourceId	↑↓	Count	↑↓	Trend	Status	↑↓	TriggerName
 DATAFACTORY- 		16			✔ Succeeded		Transaction Event Trigger
 DATAFACTORY- 		12			✔ Succeeded		DayRun
 DATAFACTORY- 		6			✔ Succeeded		Maintenance 
 DATAFACTORY- 		5			✔ Succeeded		Maintenance 
 DATAFACTORY- 		2			✔ Succeeded		4am Integration Run
 DATAFACTORY- 		2			✔ Succeeded		5am Daily Fact Refresh
 DATAFACTORY- 		2			✔ Succeeded		5am RBA Rate load
 DATAFACTORY- 		2			✔ Succeeded		7am Cluster Acquisition
 DATAFACTORY- 		2			✔ Succeeded		 Event Trigger 2
 DATAFACTORY- 		2			✔ Succeeded		 Schedule Event Trigger

Top Pipelines by Duration



Activities Duration Trend

Search						
ResourceId	↑↓	Count	↑↓	Trend	Status	↑↓ ActivityName
DATAFACTORY-...		894			✓ Succeeded	Wait1
DATAFACTORY-...		577			✓ Queued	Get Power BI Dataflow status
DATAFACTORY-...		576			✓ Succeeded	Get Power BI Dataflow status
DATAFACTORY-...		315			✓ Queued	Get Power BI Dataset status
DATAFACTORY-...		315			✓ Succeeded	Get Power BI Dataset status
DATAFACTORY-...		168			✓ Queued	Get Secret from AKV
DATAFACTORY-...		168		 [2,5,0,0,5,0,8,0,0,5,0,1,15,1,2,7,11,0,7,6,0,5,2,2,5,0,0,5,0,0,9,0,0,5,0,1,15,4,2,7,11,0,7,6,0,5,2]	✓ Queued	Get ClientId from AKV
DATAFACTORY-...		168			✓ Queued	Get AAD Token
DATAFACTORY-...		168			✓ Succeeded	Get AAD Token

Azure Log Analytics

Tool in the cloud to run queries collected by Azure Monitor

There is no cost of Log Analytics agent, but charges are incurred for data ingested

Can use Azure Monitor to collect log information from various services

Azure Monitor Logs is based on Azure Data Explorer, and log queries are written in Kusto query language (KQL)

What is Kusto Query Language (KQL)?

Read-Only request to process data and return results

KQL uses schema entities similar to SQL (databases, tables and columns)

Example

```
ADFPipelineRun  
| where Start > ago(2d)
```


Max Pipeline Runs by Duration

ADFPipelineRun

| where Start > ago(2d)

| where Status != "InProgress"

| extend DurationInMinutes = round((End - Start)/1m, 2)

| summarize max(DurationInMinutes) by PipelineName

| top 3 by max_DurationInMinutes

Turn on Diagnostic Settings

log

- ☐ ActivityRuns
- ☐ PipelineRuns
- ☐ TriggerRuns
- ☐ SandboxPipelineRuns
- ☐ SandboxActivityRuns
- ☐ SSISPackageEventMessages
- ☐ SSISPackageExecutableStatistics

☐ Send to Log Analytics workspace

☐ Archive to a storage account

☐ Stream to an event hub

☐ Send to partner solution

Select All

Azure Log Analytics Data Capture

Data will be written to the logs as the pipelines are run

Wait overnight to see some numbers start to come through

A pre-built solution is available from the Azure Marketplace to analyse Azure Data Factory Logs

Demo

204.41	940.5
1,121.58	788.6
0.00	—
26,966.00	—
7,609.27	—
12,629.64	—
28,830.67	—

Demo Agenda

1 Power BI Audit Reports

2 How to write queries in KQL

3 Azure Data Factory KQL Queries - Pipelines

4 Azure Data Factory KQL Queries - Triggers

5 Azure Data Factory KQL Queries - Activities

6 Azure Data Factory KQL Queries - Complex Queries

7 Export

8 Create Power BI Report

Cost Consideration

Default pricing model is Pay-As-You-Go

Data volume is based on the size of data stored (GB)

Each workspace is charged as a separate service

To mitigate some of the cost you can set the data retention period

More information available from <https://bit.ly/3h48pla>

Summary

Comprehensive monitoring of ADF resources is possible for minimal effort

Solution built is versatile and can be used across different Microsoft Azure services

Solution can be prototyped rapidly

Improvements to the process can be made using this data and enhance it with Machine Learning



Any Questions?

A hand in a dark suit sleeve holds a glowing blue sphere. Inside the sphere, the words "THANK YOU" are written in a bold, white, sans-serif font. The sphere is surrounded by a complex, glowing blue energy field resembling a plasma or lightning pattern. The background is dark and filled with faint, glowing blue lines that look like a network or a map.

**THANK
YOU**

The End

Contact Me



[/in/sarathboppudi](https://www.linkedin.com/in/sarathboppudi)



[@BoppudiSarath](https://twitter.com/BoppudiSarath)



<https://sarathboppudi.com>



<https://github.com/sandman153>