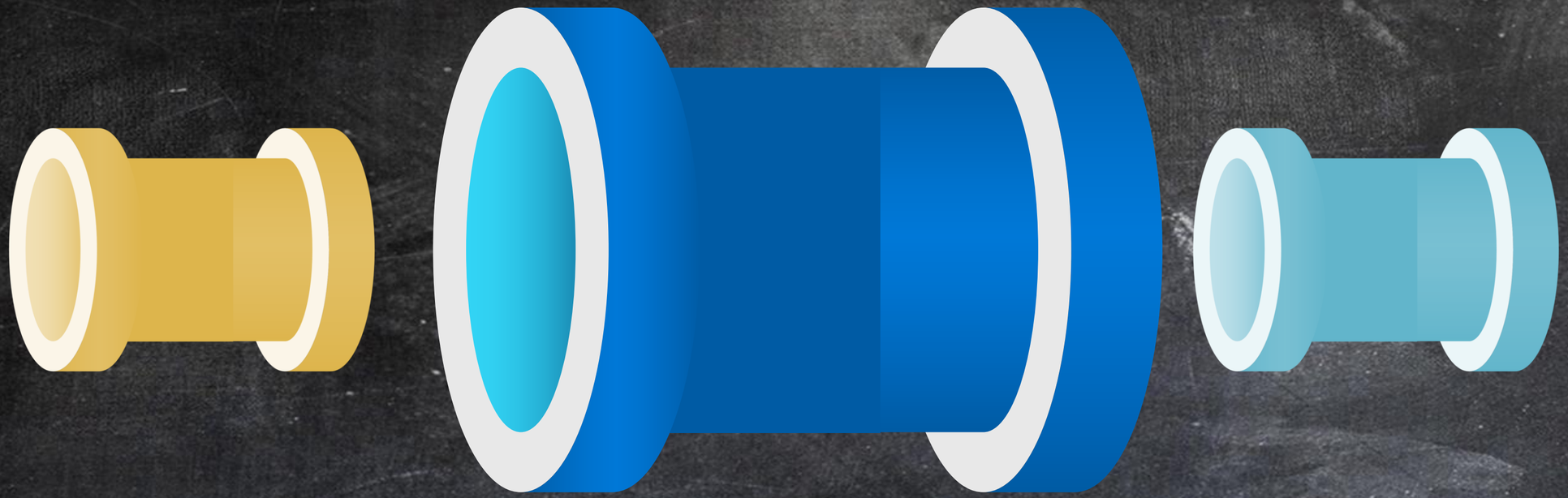
Cloud Formations

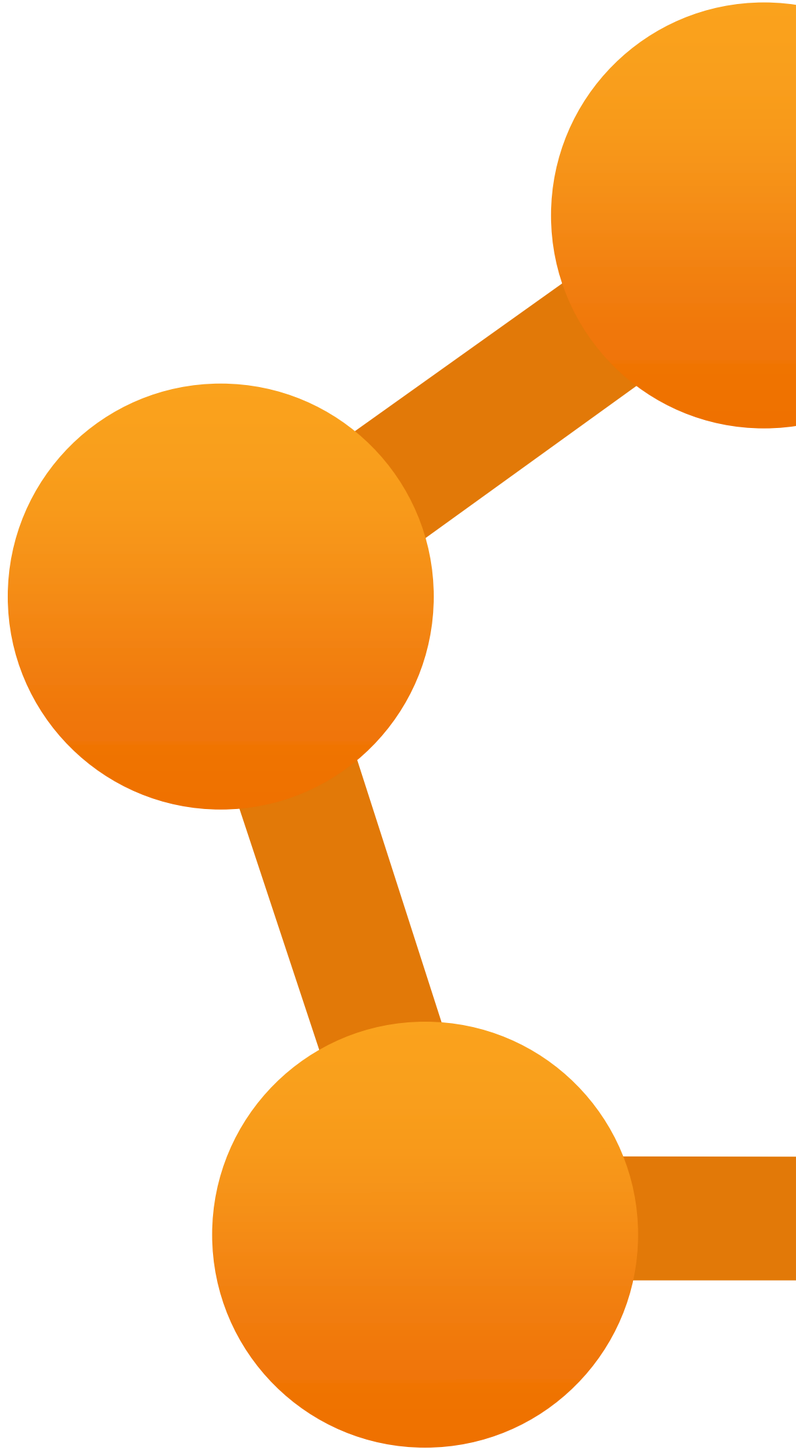
Proudly presents...

Integration Pipelines



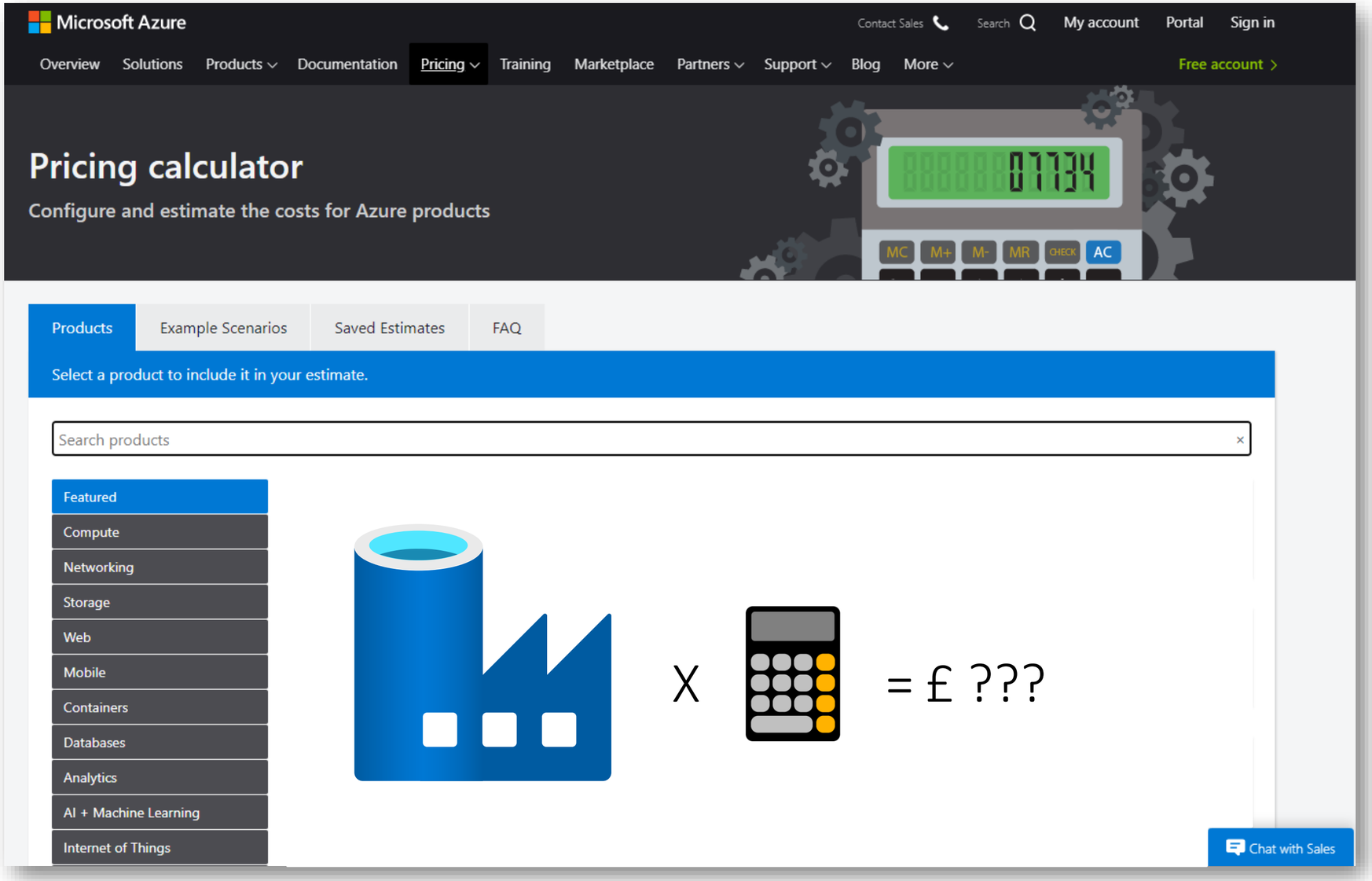
Module 12 – Final Thoughts

Running Costs



Cloud Formations

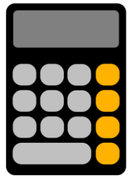
Azure Price Calculator – Data Factory



https://azure.microsoft.com/en-us/pricing/calculator/

Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd



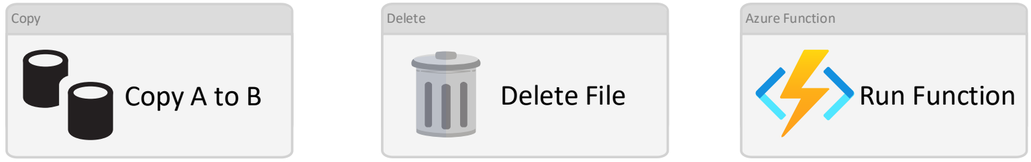
Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

Data Factory Pricing – IR Type

Based on 1 Hour Running

	Activity Runs	Data Handling Compute per Hour	Pipeline Activity Execution Duration per Hour	
Azure IR	x1000 £0.75	CPU £0.19	Int. £0.004	Ext. £0.00019
Hosted IR	x1000 £1.12	CPU £0.07	Int. £0.001	Ext. £0.0001



£2.13 per Month

* Prices quoted for UK South Azure Region charged in Pounds Sterling (£) captured on 26th Feb 2024.

Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

Data Factory Pricing – Data Flow & SSIS Compute

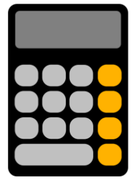
	Number of Compute Cores	Number of VM Instances	VM Uptime Per Hour	SQL Server Licence Type	
Data Flows	CPU CPU			N/A	8x Cores General Purpose 8 Hours a day £426.30 per Month
SSIS IR	CPU CPU	SQL SQL		• Standard • Enterprise	8x Cores Standard 8 Hours a day £443.97 per Month

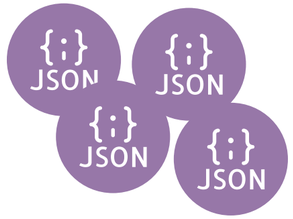
* Prices quoted for UK South Azure Region charged in Pounds Sterling (£) captured on 26th Feb 2024.

Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

Data Factory Pricing – Operations



	Entity Read/Write	Monitoring
Operations		
	Per 50,000 = £0.37	Per 50,000 = £0.19

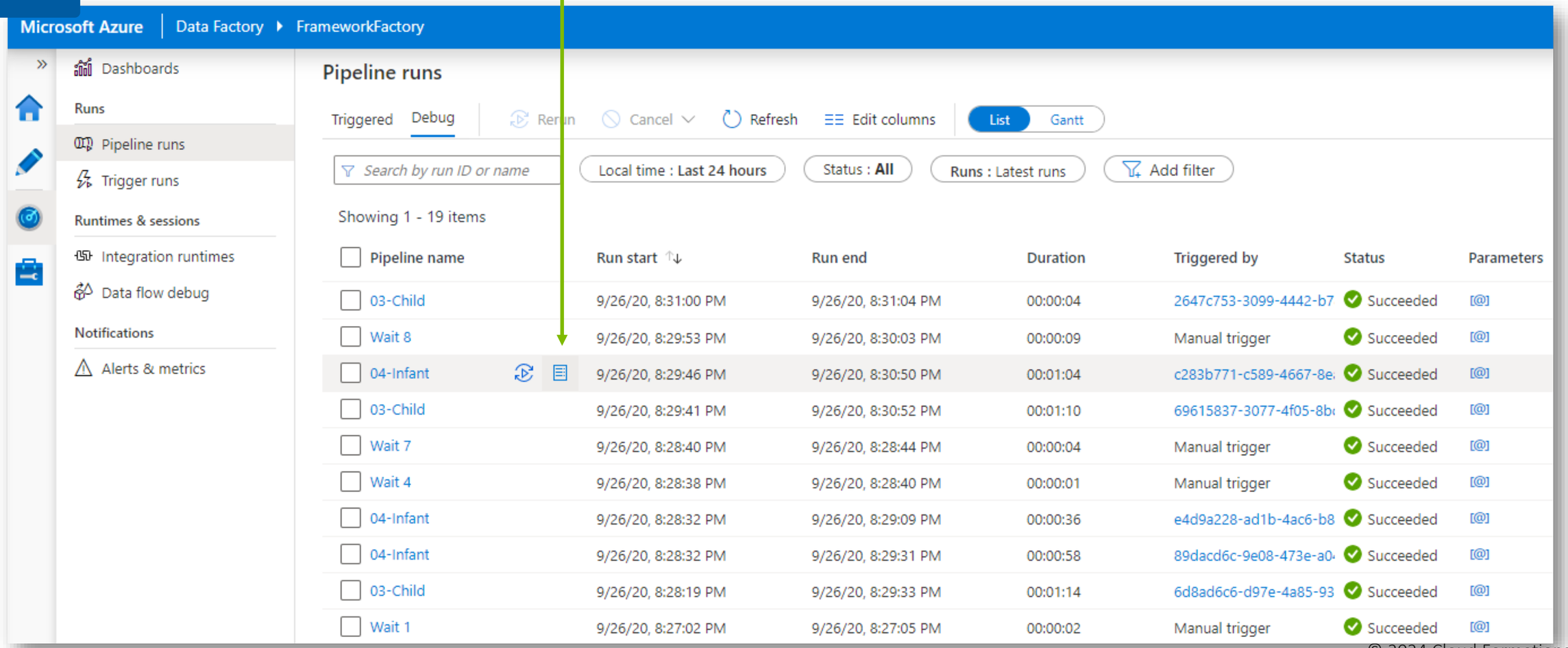
* Prices quoted for UK South Azure Region charged in Pounds Sterling (£) captured on 26th Feb 2024.

Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

How Much Does My Pipeline Cost?

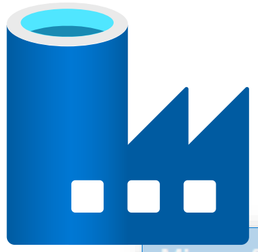
Consumption

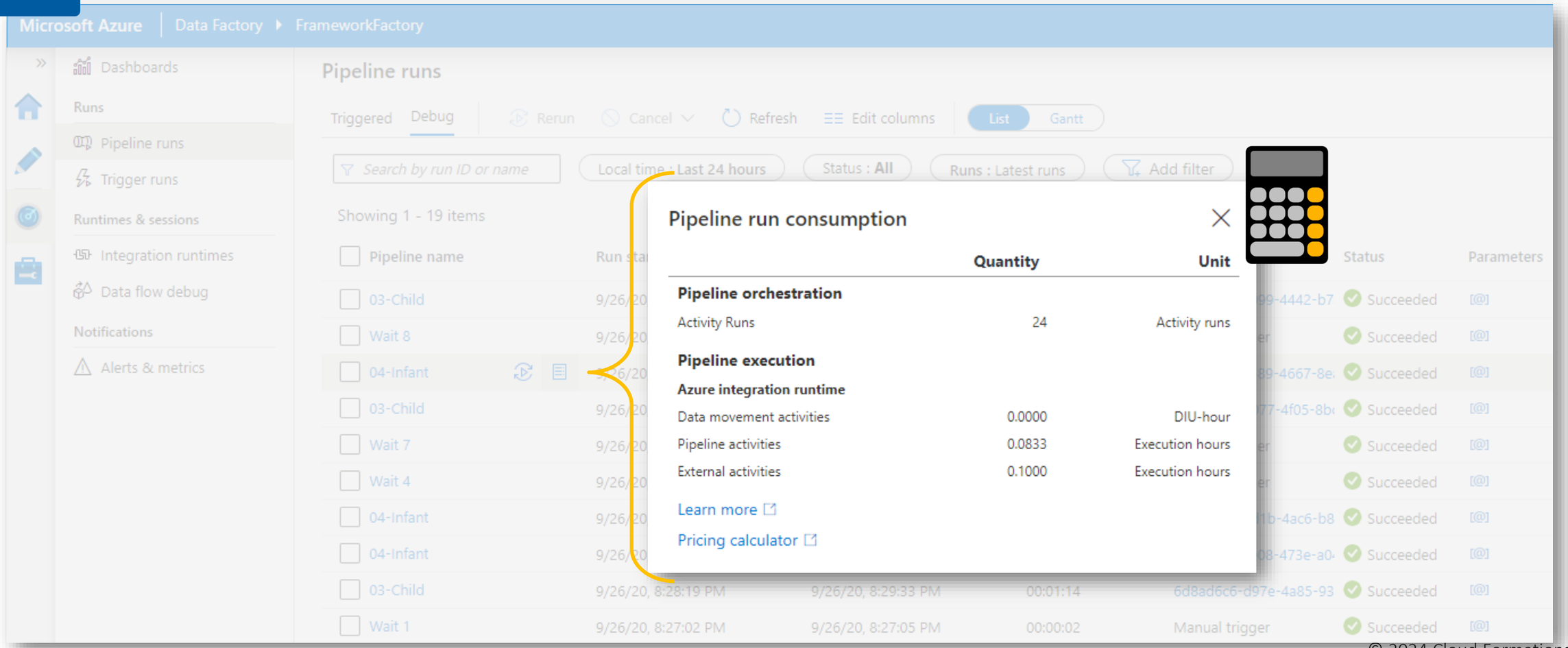


Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

How Much Does My Pipeline Cost?

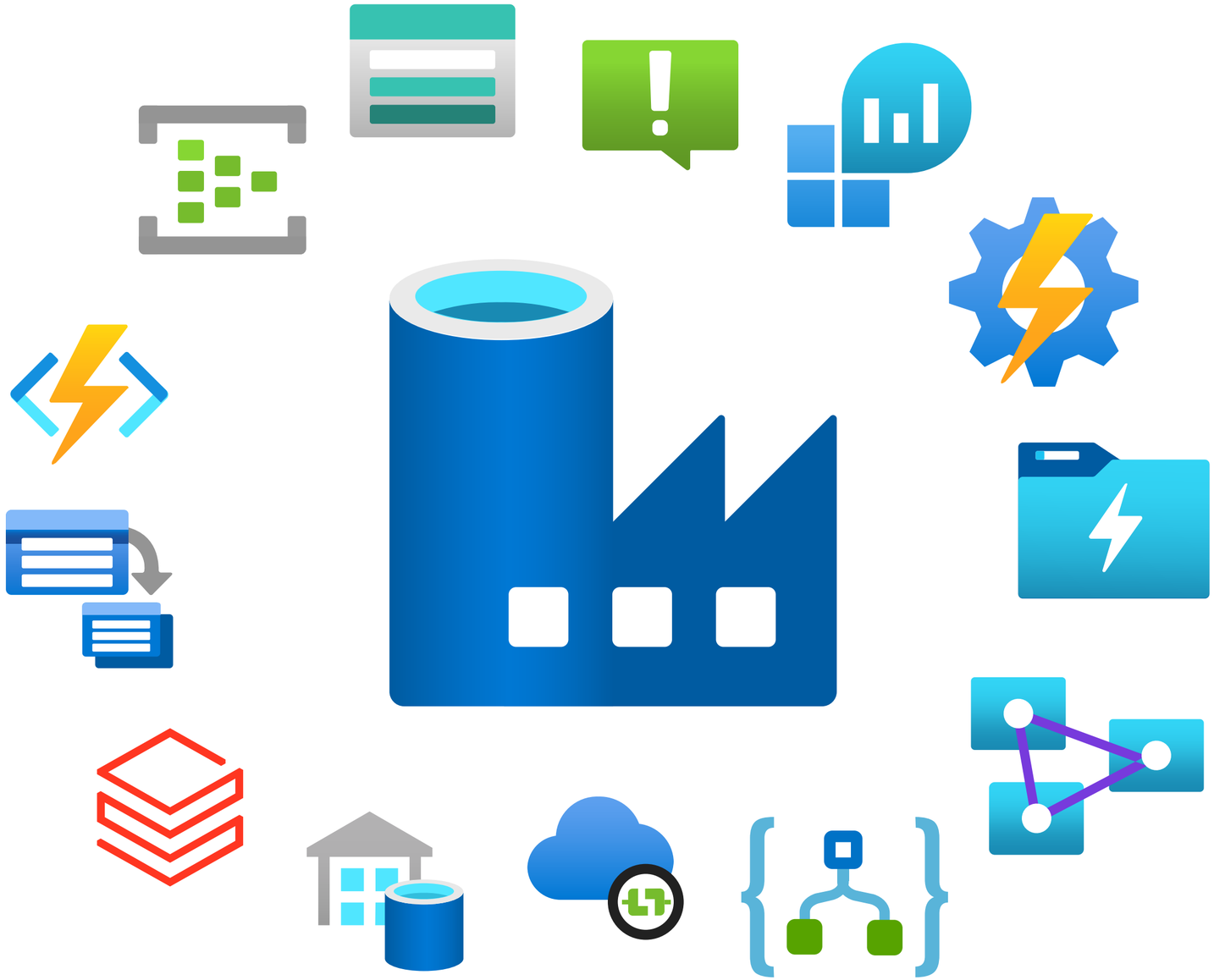




Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

How Much Does My Data Platform Solution Cost?



Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

CF.Cumulus Indicative Run Costs

Assuming the use of an Azure UK Region priced for GBP, with all CF.Cumulus compute processes running for 4 hours every day of the year, with locally redundant storage, the following Azure run costs **per month** can be assumed:

SQL Database
250GB of Metadata
£12.08

Data Factory
Azure IR Usage
£25.84

Functions
2 Node Dedicated App Service Plan
£44.30

Databricks
3 Node Cluster with 4 vCPUs & 16GB RAM
£152.92

Blob Storage
1TB of Log Data
£18.45

Data Lake
1TB Analytics Data
£38.82

Key Vault
50,000 Requests
£0.14

Monitoring
500MB Log & App Insights Data per Day
£45.61

£338.16*

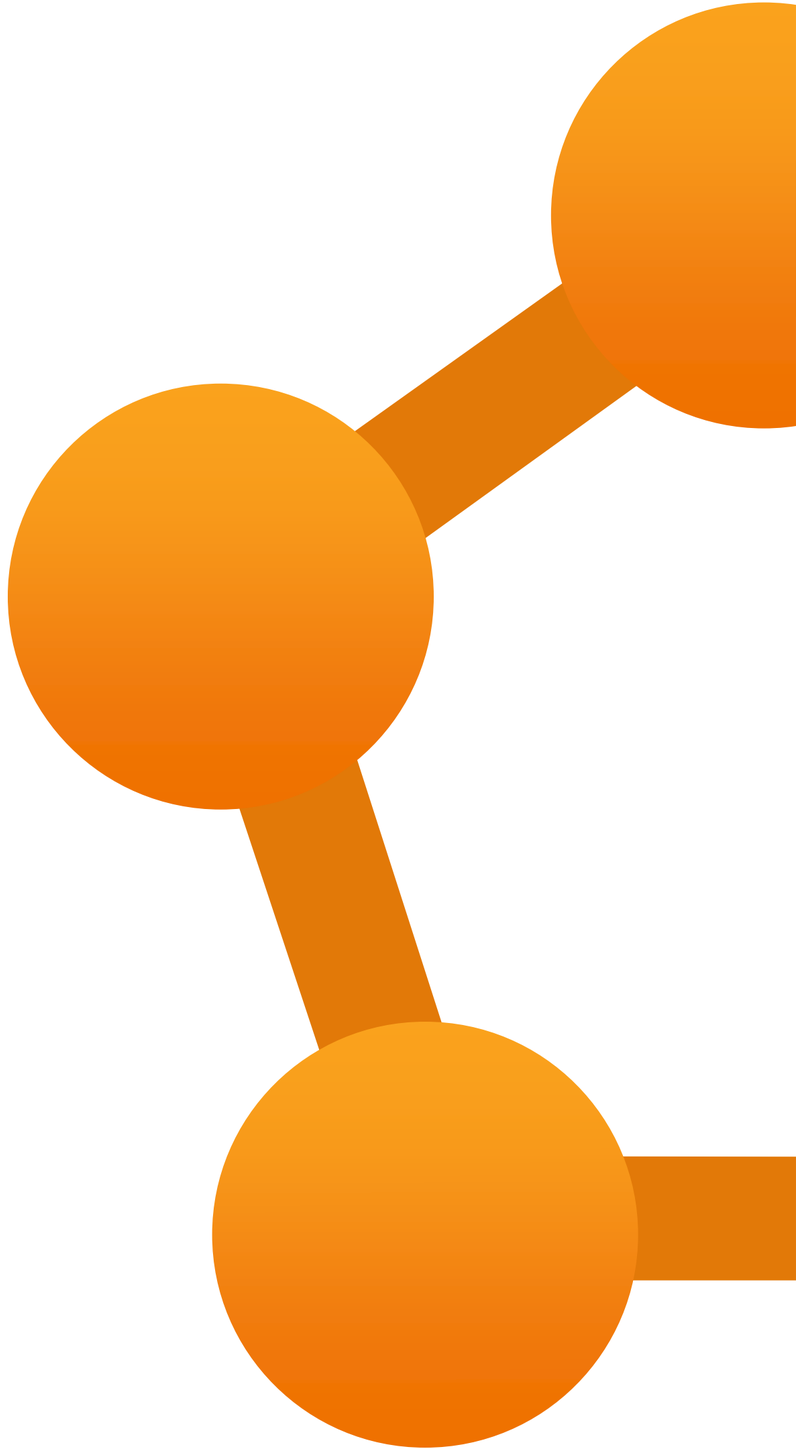
*With every customer implementation data volumes, table complexity and serving queries (to dashboards) may differ. All affecting the above run costs.

Azure Price Calculator workings available for review and adjustment: https://azure.com/e/30f8aeedfd564ba189d33f278845bb89

© 2024 Cloud Formations Ltd

Module 12 – Final Thoughts

Best Practices



Cloud Formations

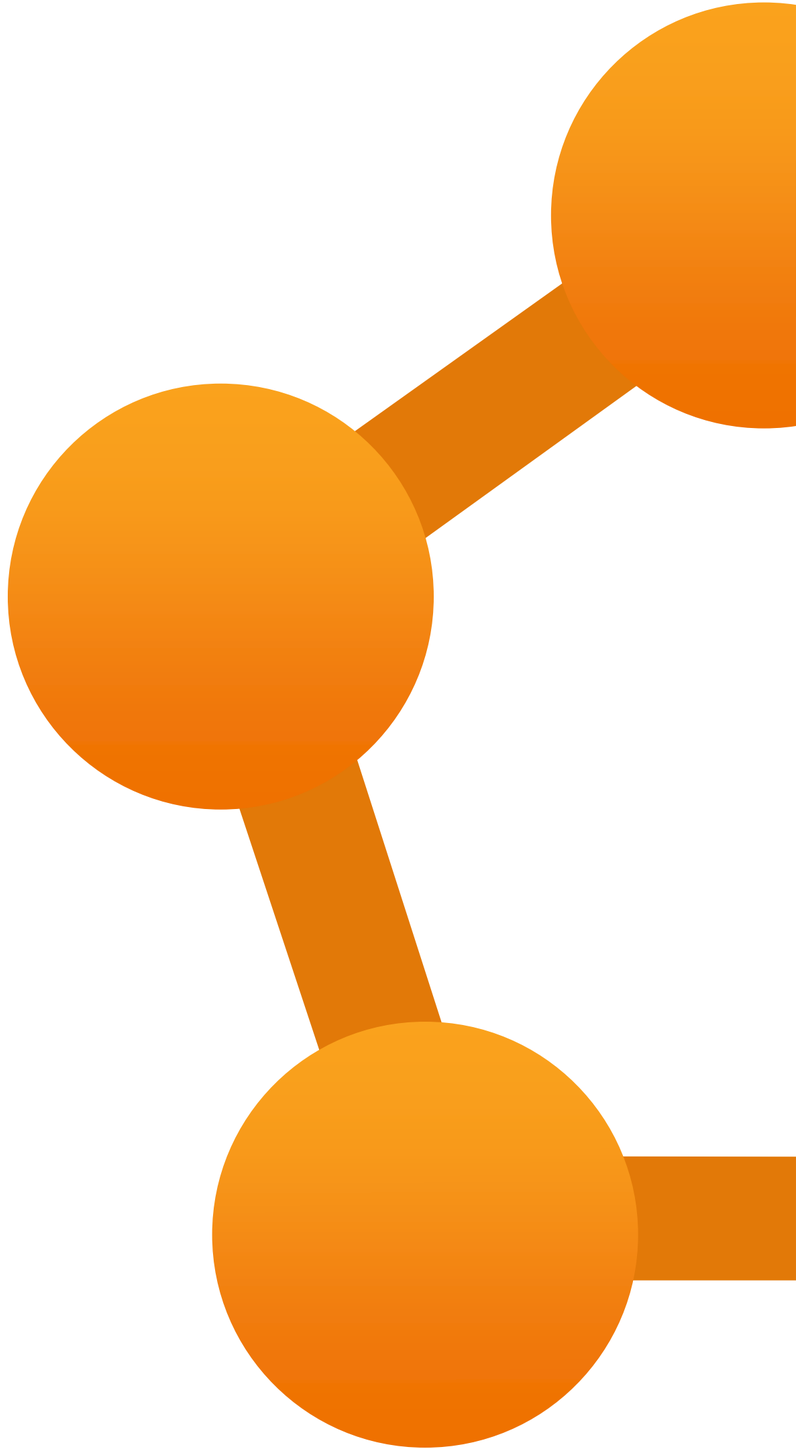
Key Points

- Environment Setup & Developer Debugging
- Deployments
- Automated Testing
- Naming Conventions
- Pipeline Hierarchies
- Pipeline & Activity Descriptions
- Factory Component Folders
- Linked Service Security via Azure Key Vault
- Dynamic Linked Services
- Generic Datasets
- Metadata Driven Processing
- Parallel Execution
- Hosted Integration Runtimes
- Azure Integration Runtimes
- Wider Platform Orchestration
- Custom Error Handler Paths
- Monitoring via Log Analytics
- Service Limitations
- Using Pipeline Templates
- Documentation

mrpaulandrew.com/2019/12/18/best-practices-for-implementing-azure-data-factory/

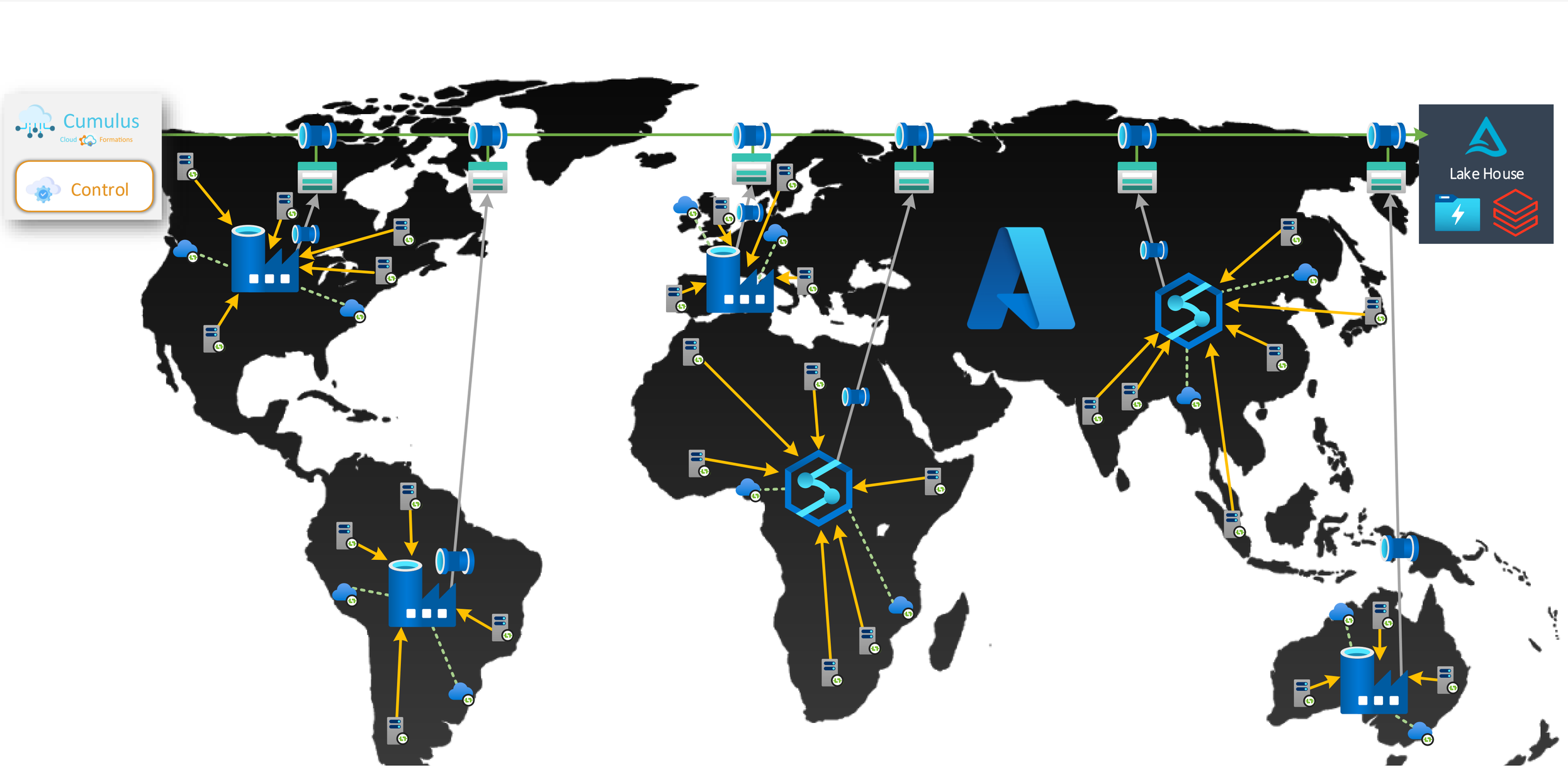
Module 12 – Final Thoughts

Conclusions



Cloud Formations

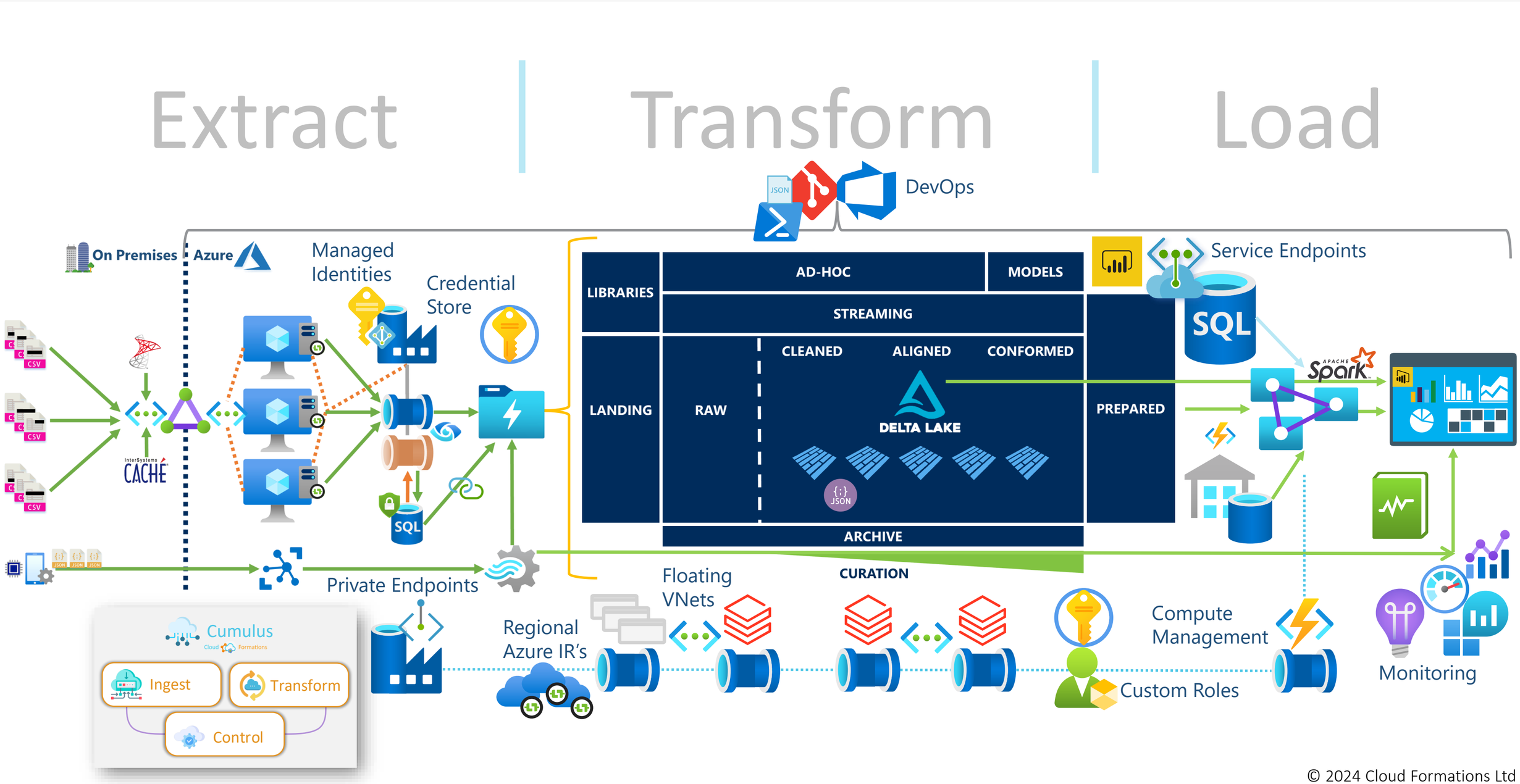
Exploring Data Factory at Enterprise Scale



Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

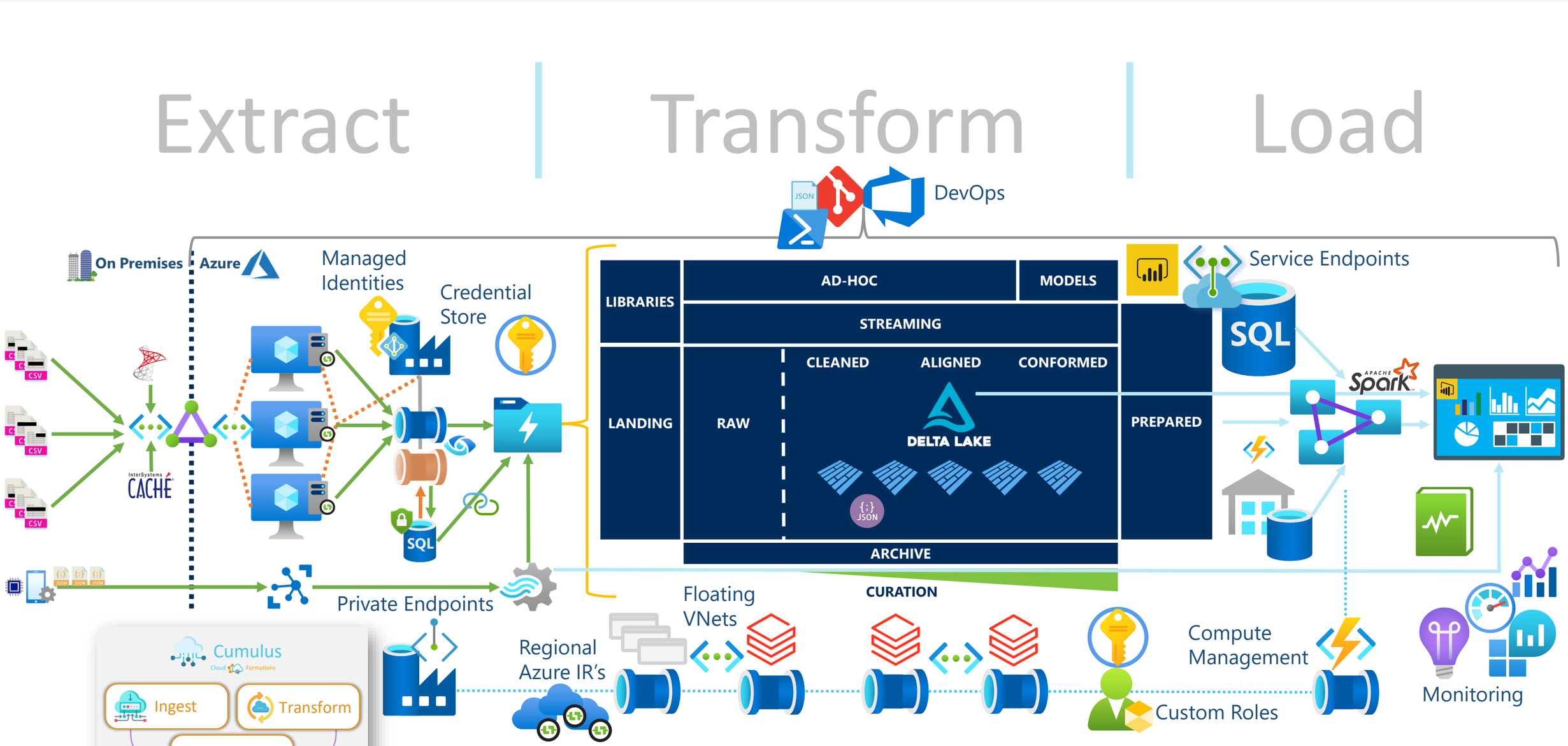
Overall Architecture



Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

Overall Architecture

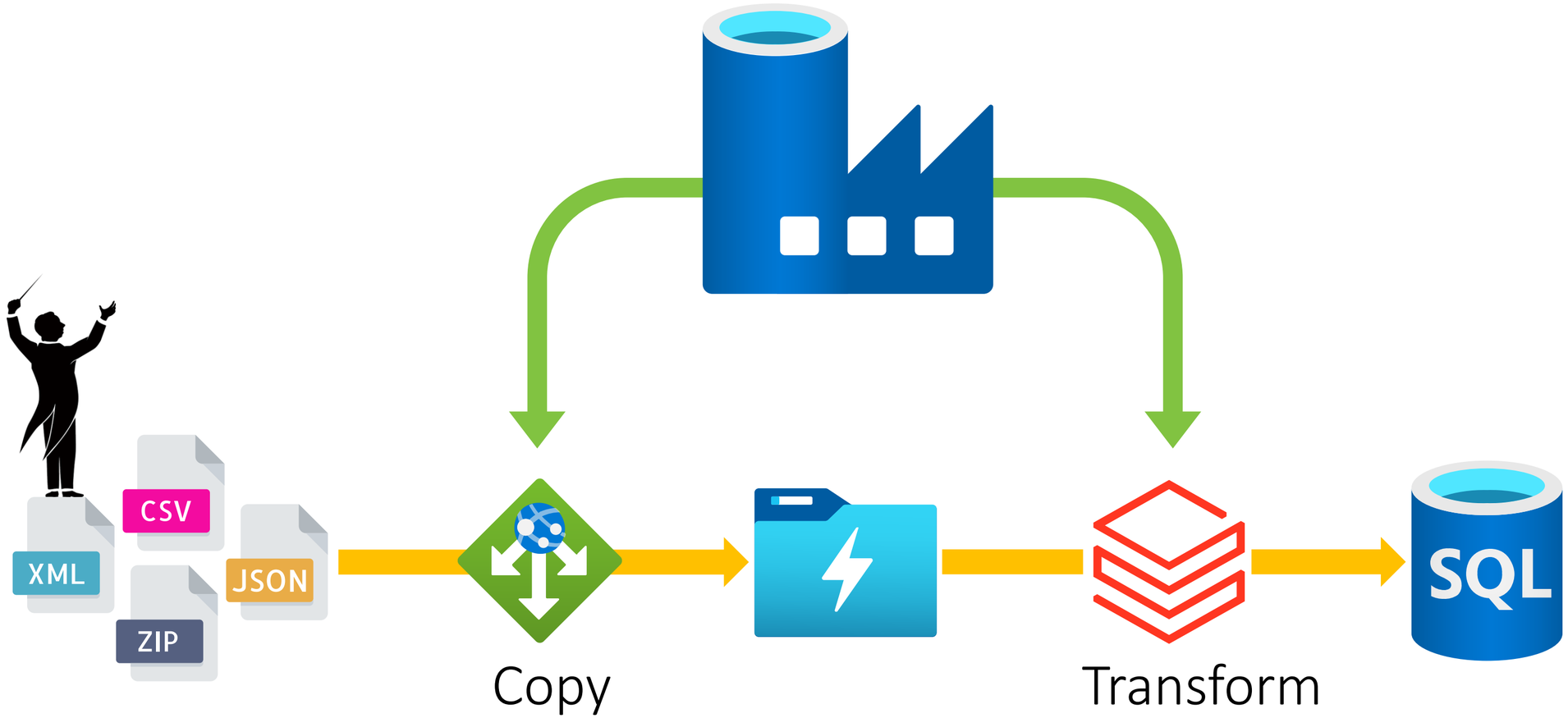


Q: Should we build our data platform solution like this?... A: It depends!

Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

What is Azure Data Factory (ADF)?

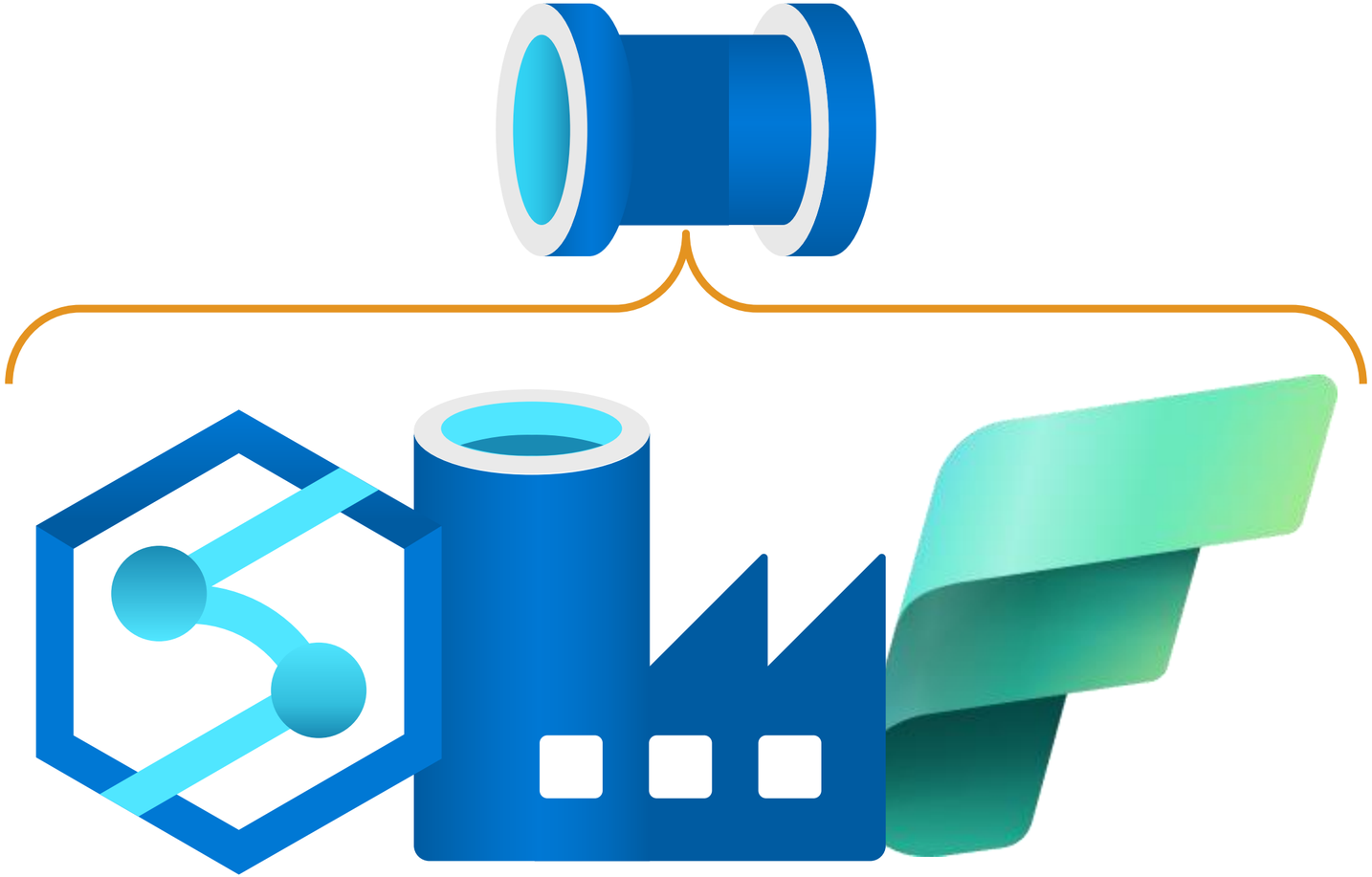


Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

What are Integration Pipelines?

Answer: A data integration pipeline is a logical package of work that can be used move, transform and orchestrate data tasks from and to various compute and storage resources.



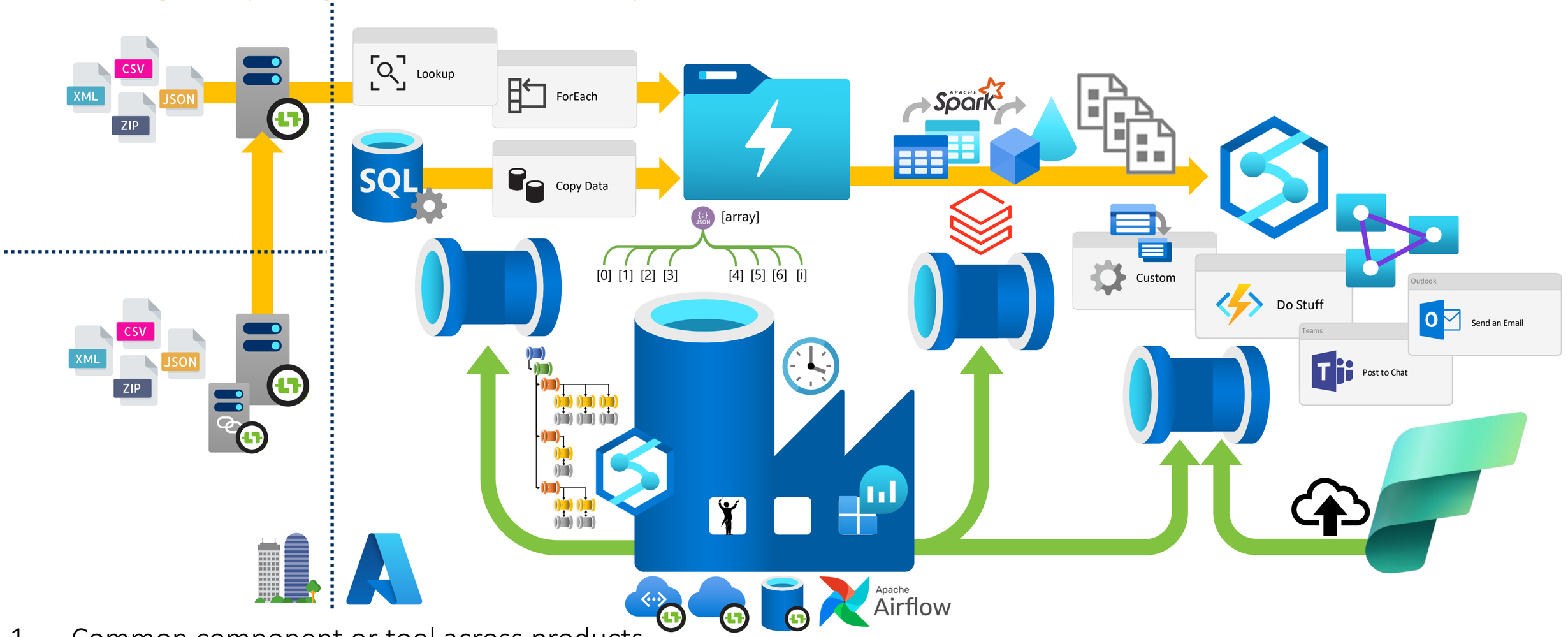
In the context of Azure Data Factory, Azure Synapse Analytics, and Microsoft Fabric, integration pipelines represent a common component or tool across products for engineering teams to utilise in the delivery of a data platform solution.

Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

What are Azure ~~Data Factory~~ Integration Pipelines?

Considering everything we've covered today.



1. Common component or tool across products.
2. The ability to orchestrate our Control Flow operations – with scale out Activities.
3. The ability to orchestrate our Data Flow transformations – using cloud native services.
4. The scheduler of solutions – using a variety of Pipeline Triggers and dynamic frameworks.

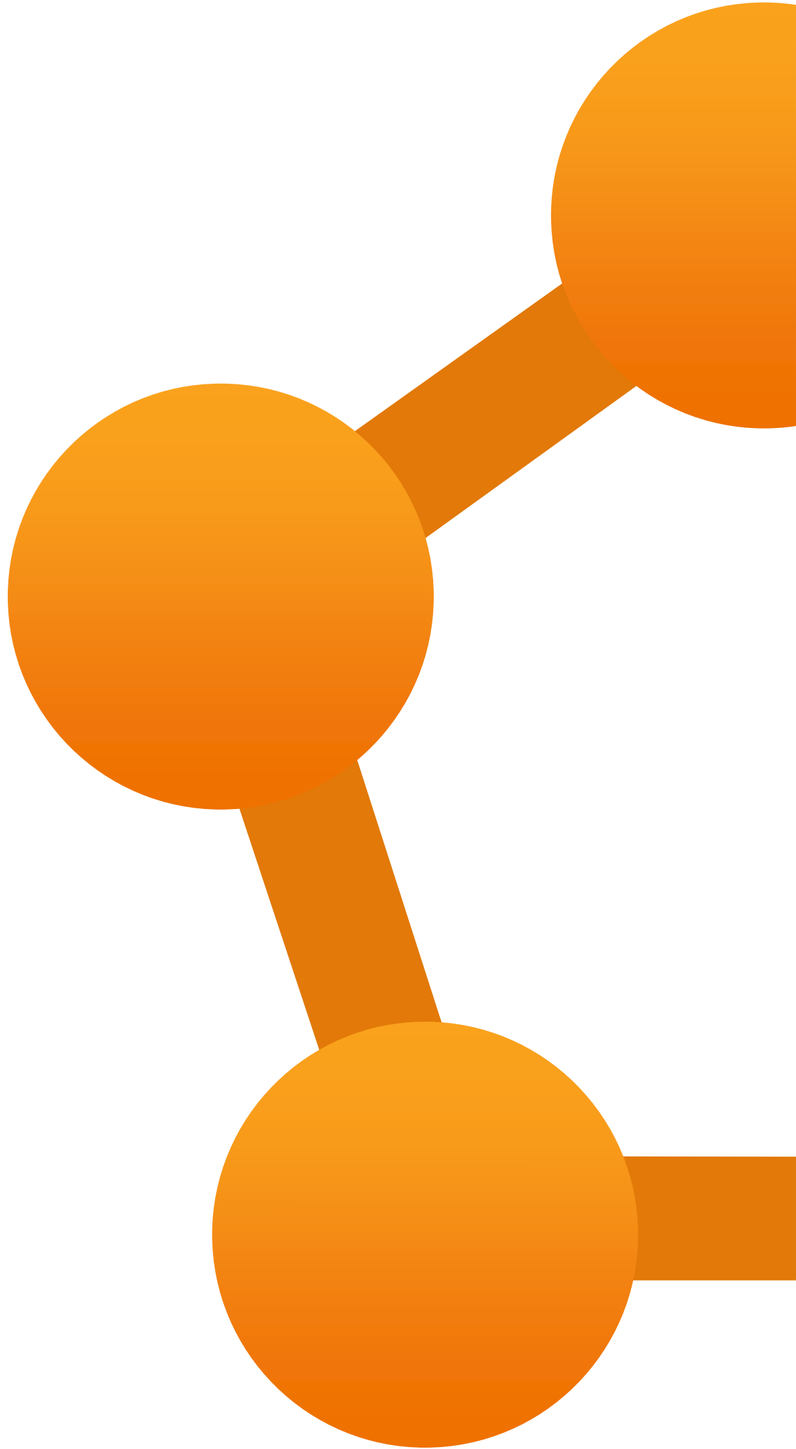
Cloud Formations - Knowledge Transfer & Training

© 2024 Cloud Formations Ltd

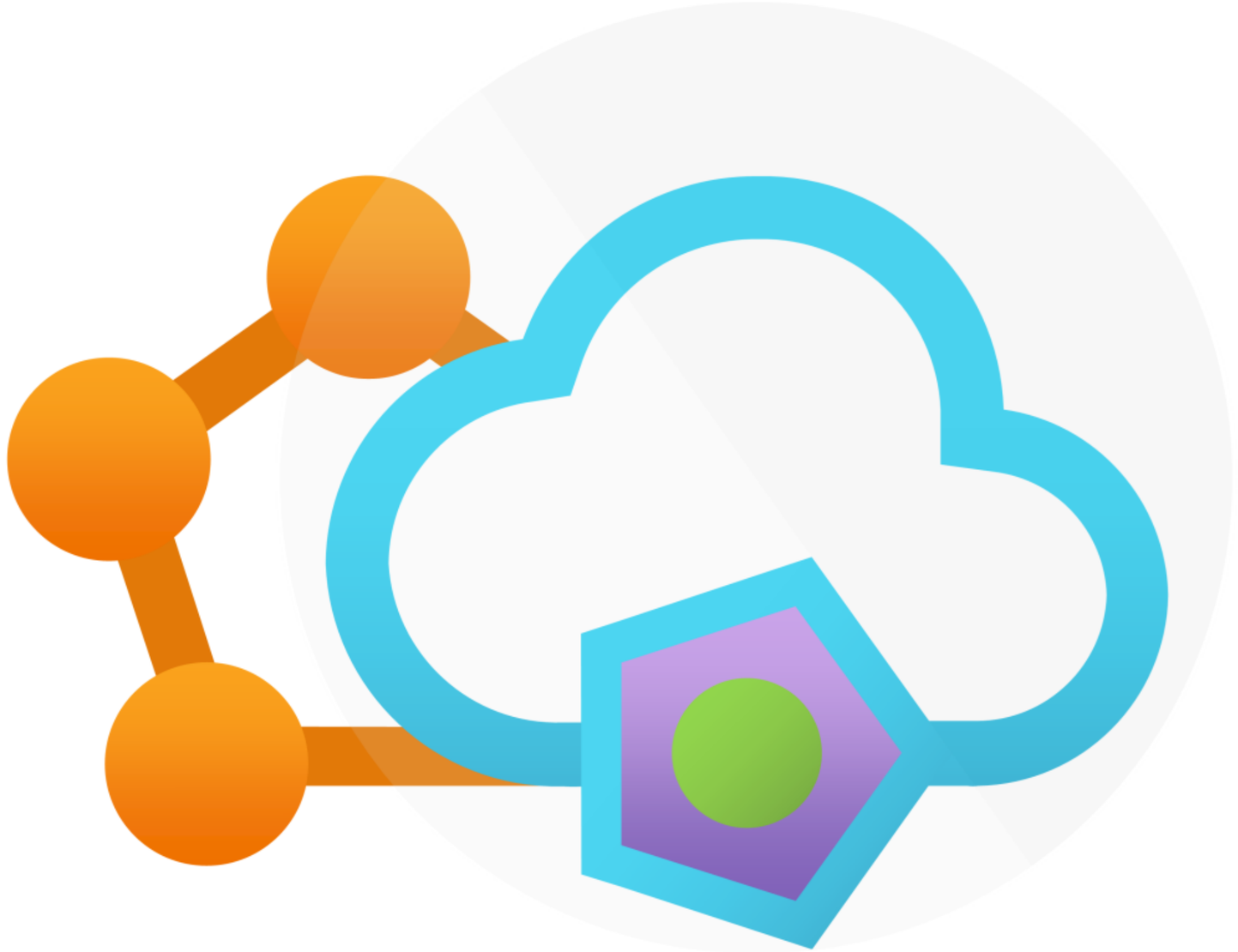
Module 12

Final Thoughts ✓

Any questions?



Cloud Formations



Thank You

paul.andrew@cloudformtions.org

In/mrpaulandrew

@mrpaulandrew

https://cloudformations.org

contactus@cloudformations.org

In/CloudFormations

@CloudFormsLtd

CloudFormationsLtd

Cloud Formations