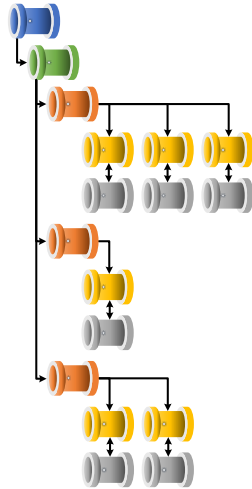




Creating a Metadata Driven Processing Framework (ProcFwk.com)

Using Azure Integration Pipelines



Paul Andrew | Technical Architect in Azure CoE



Mr Paul Andrew
Consulting Ltd



@MrPaulAndrew



In/MrPaulAndrew



MrPaulAndrew.com



c/MrPaulAndrew



<https://github.com/mrpaulandrew>

CommunityEvents

Demo code, content and slides from various community events.

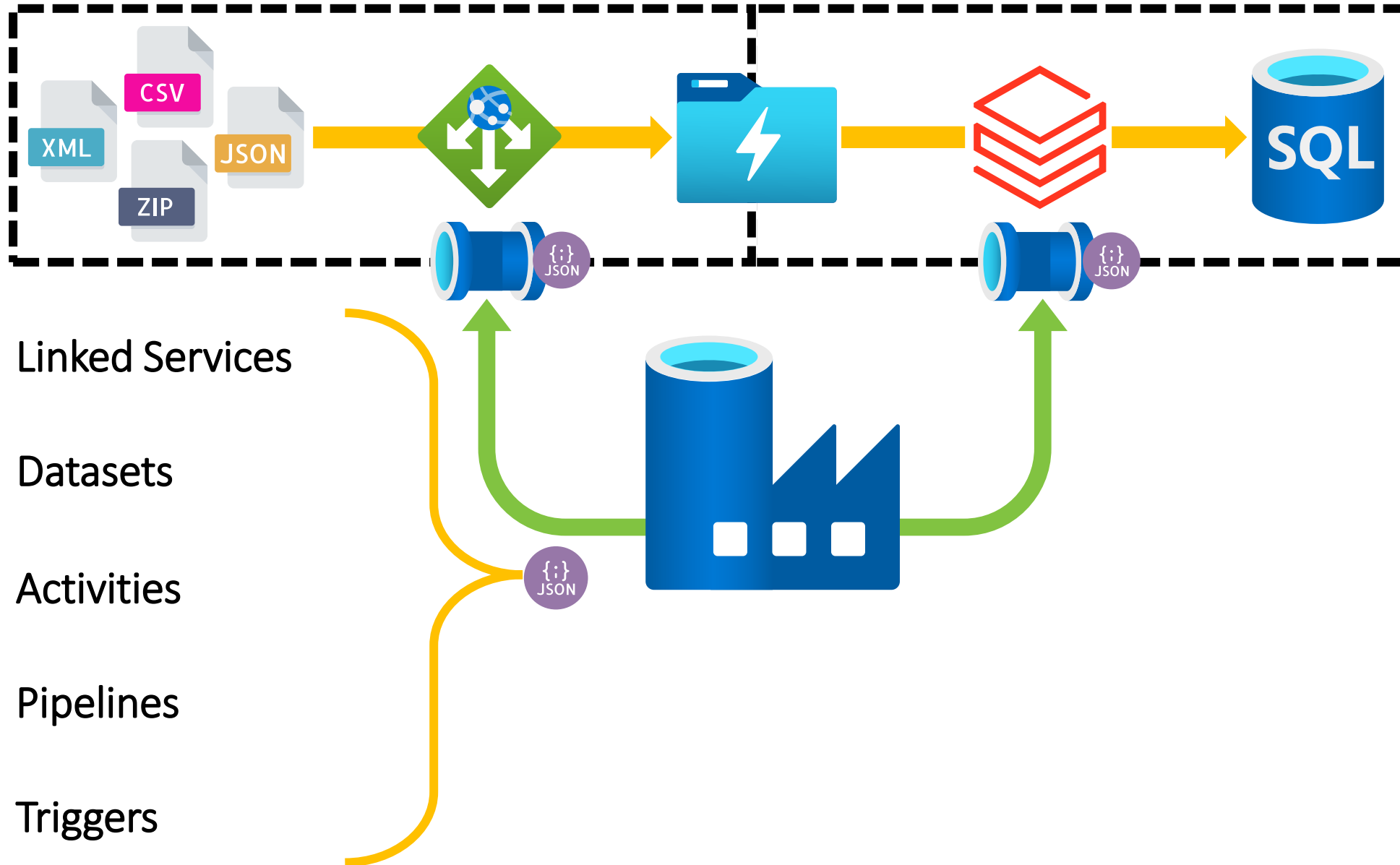
● C++

[{Event/Location}-{Month}-{Year}](#)

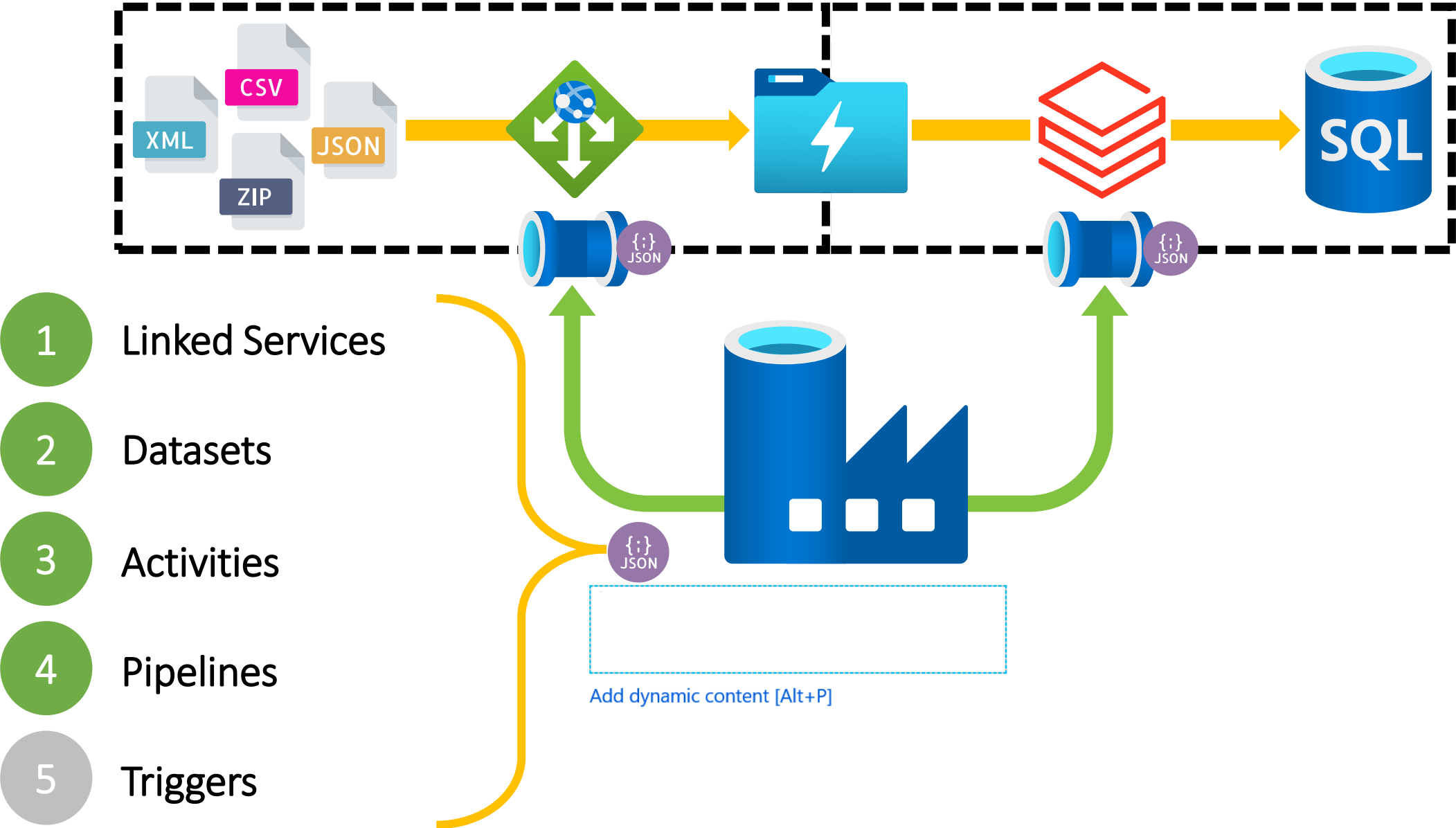
Data Factory – A Quick Overview

A large, solid blue shape that starts as a diagonal line from the bottom-left and extends towards the top-right, ending in a rounded corner. It occupies the right half of the slide.

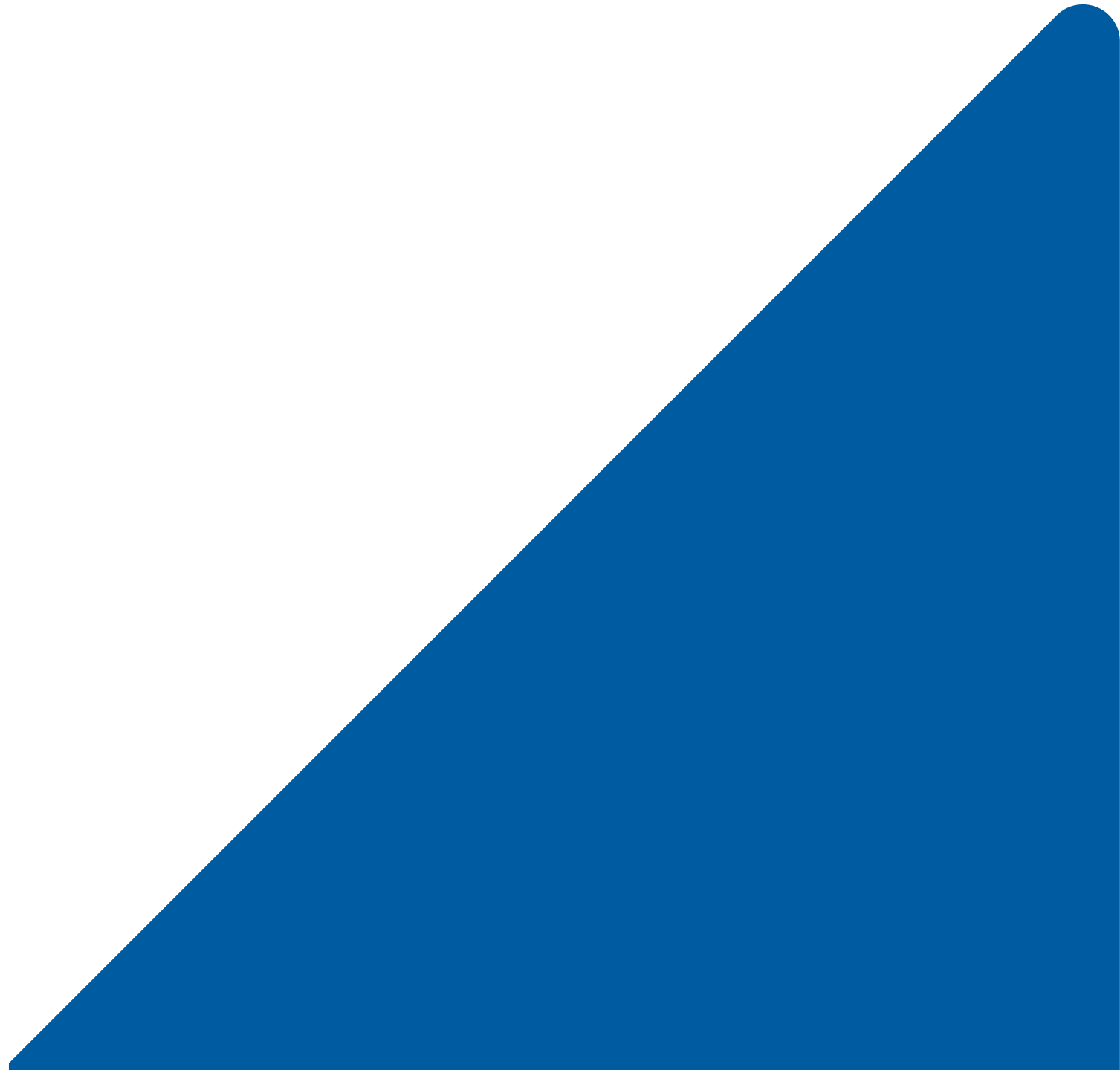
Data Factory A Quick Overview



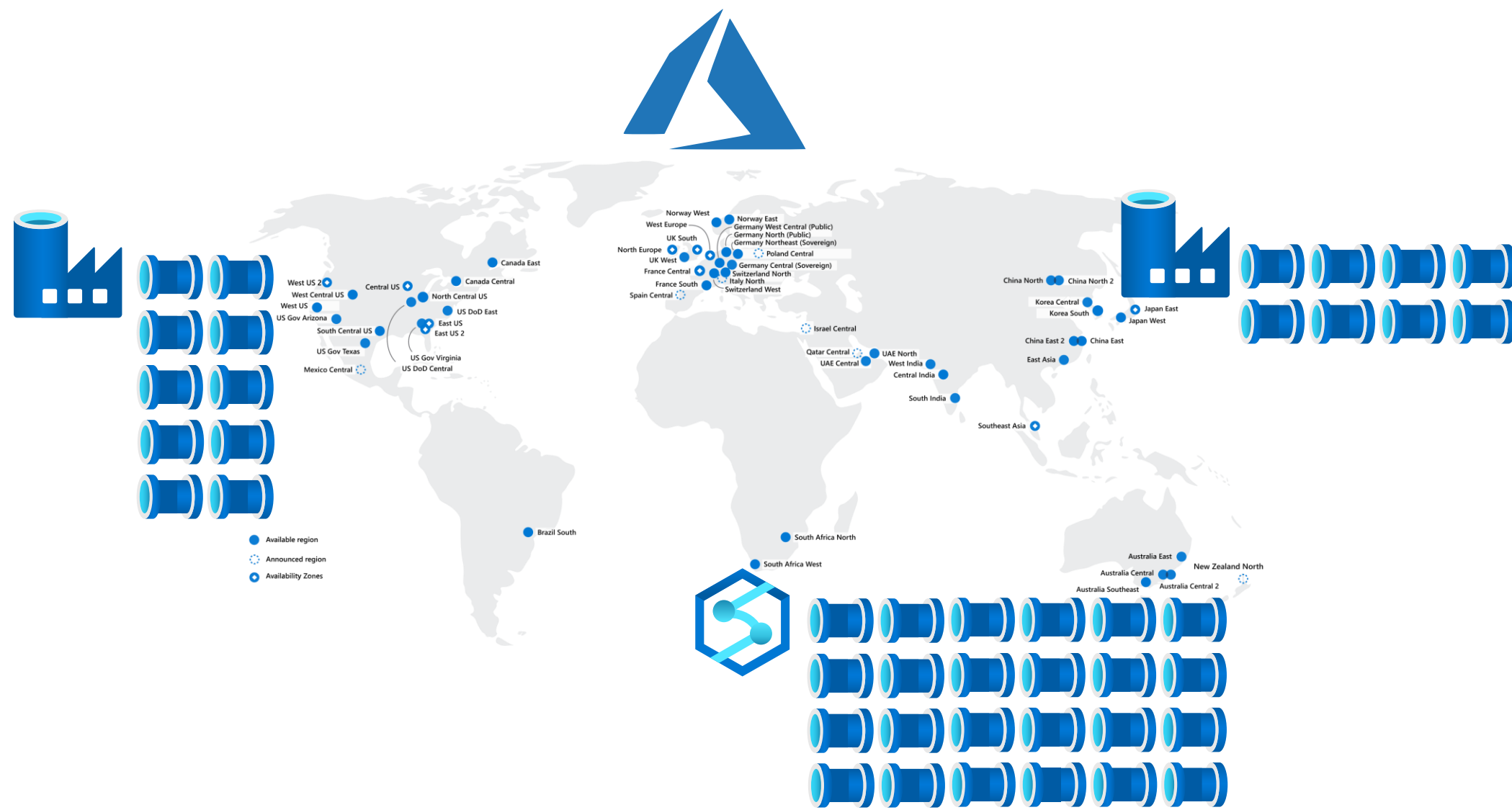
Data Factory Components – Add Dynamic Content



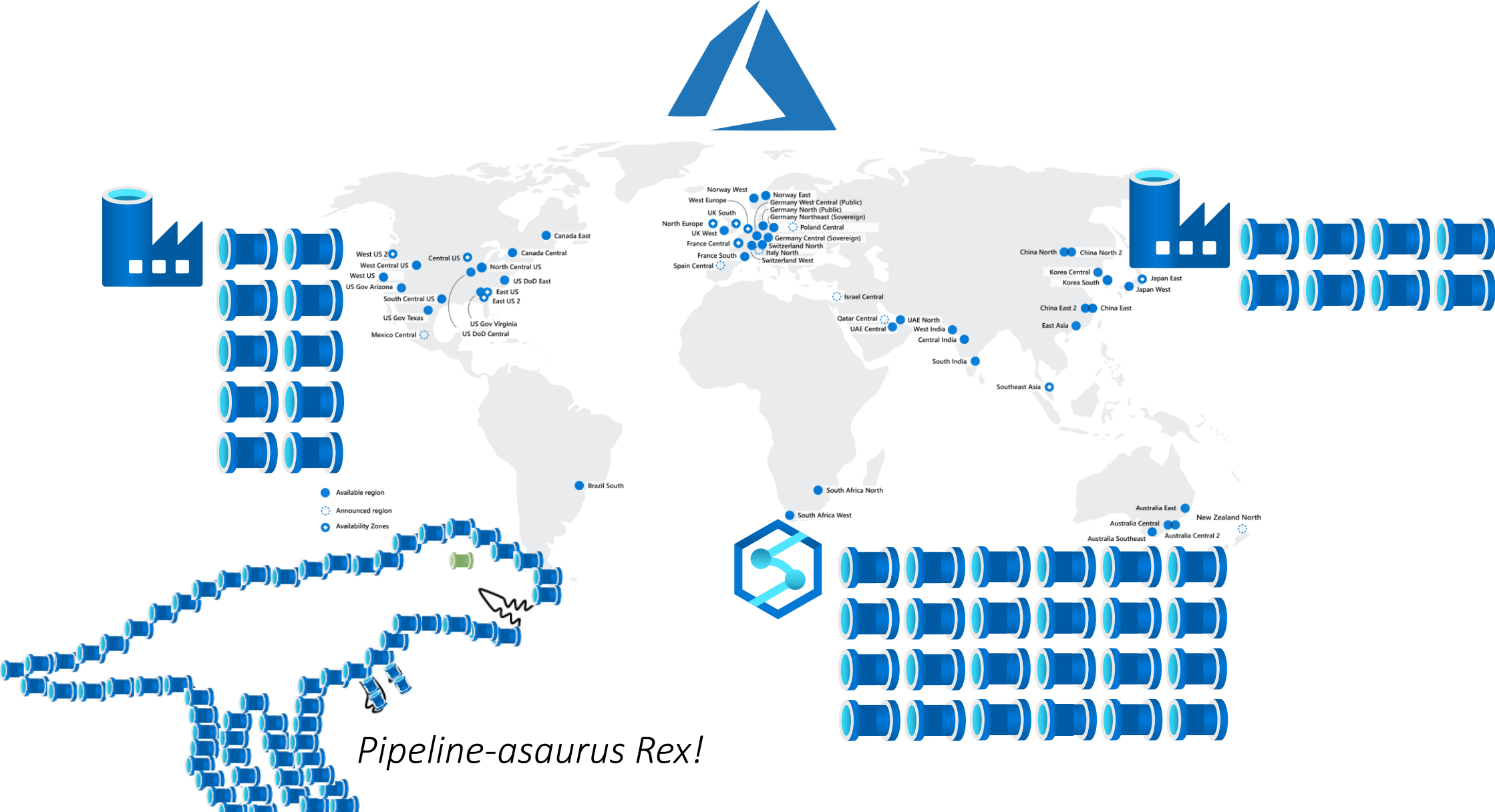
Problem



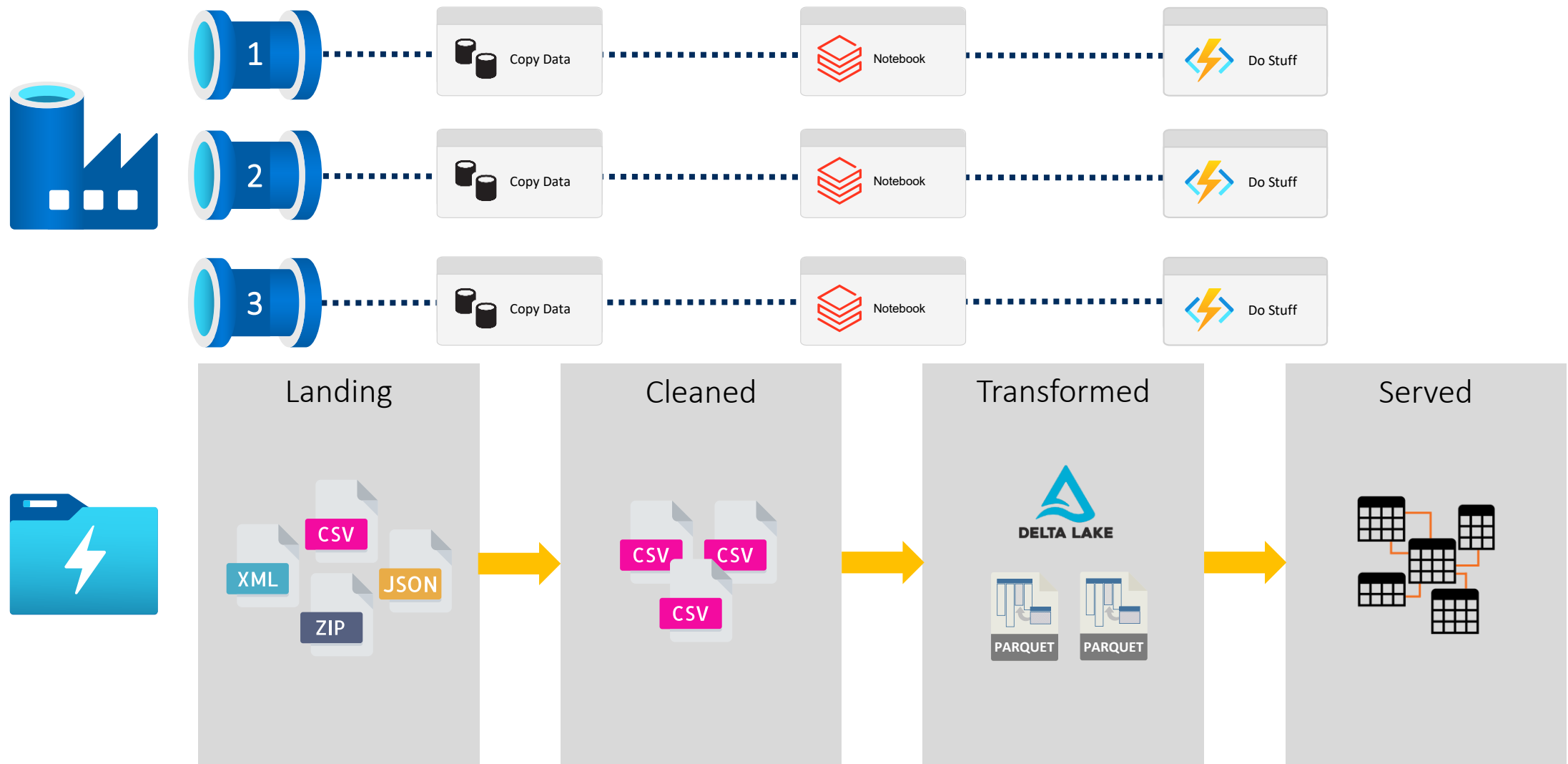
Problem: How should we structure our Integration Pipelines?



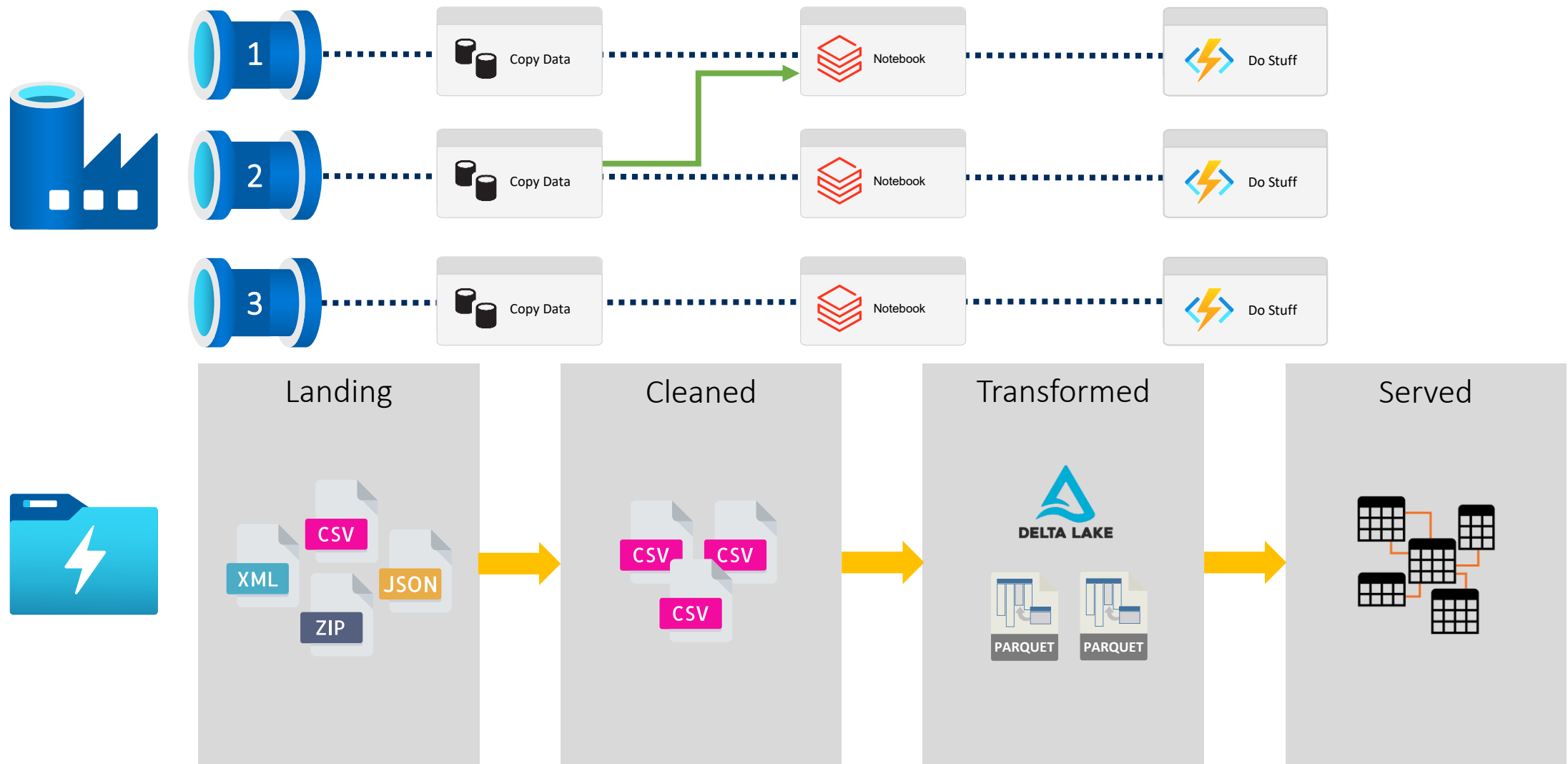
Problem: How should we structure our Integration Pipelines?



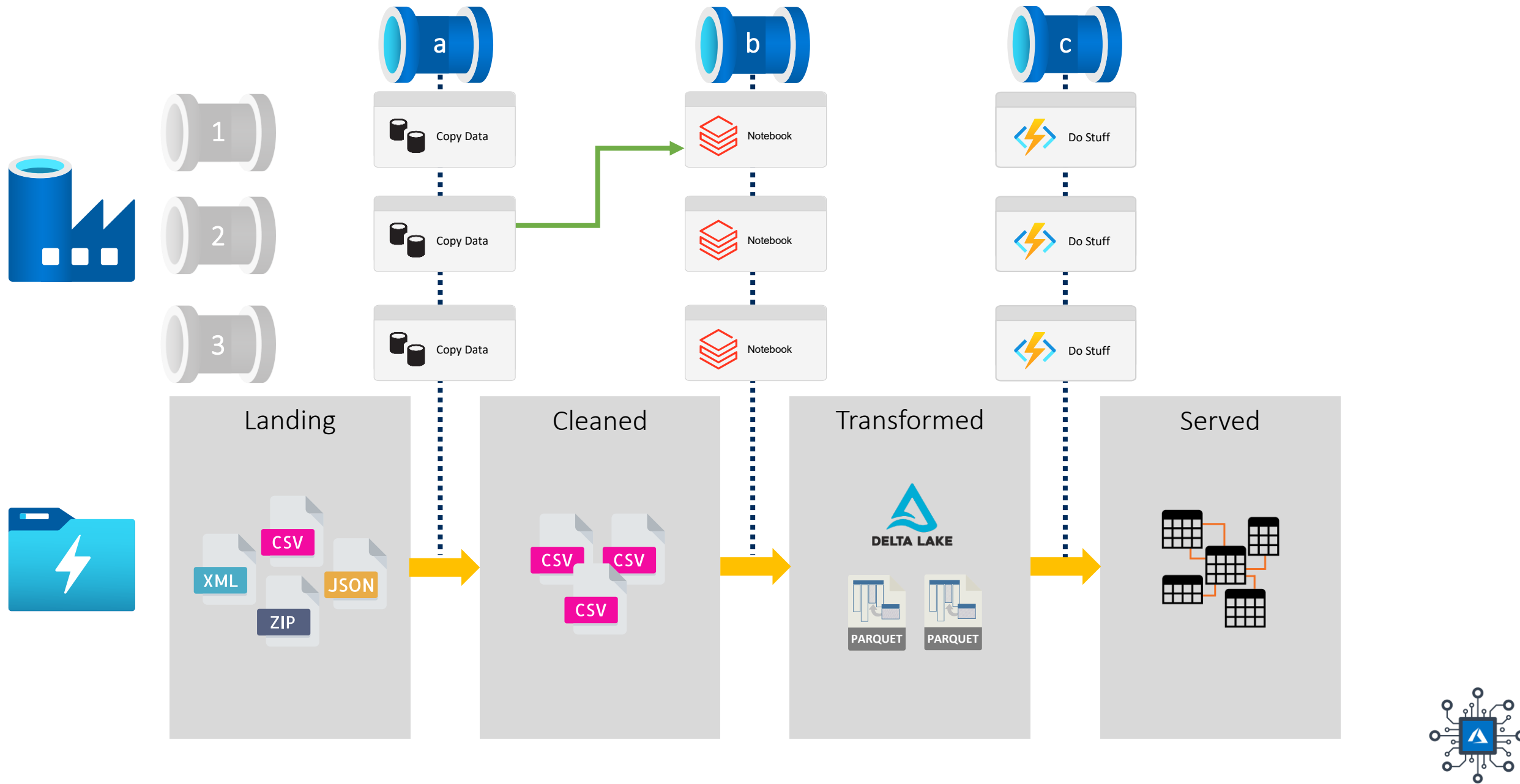
Problem



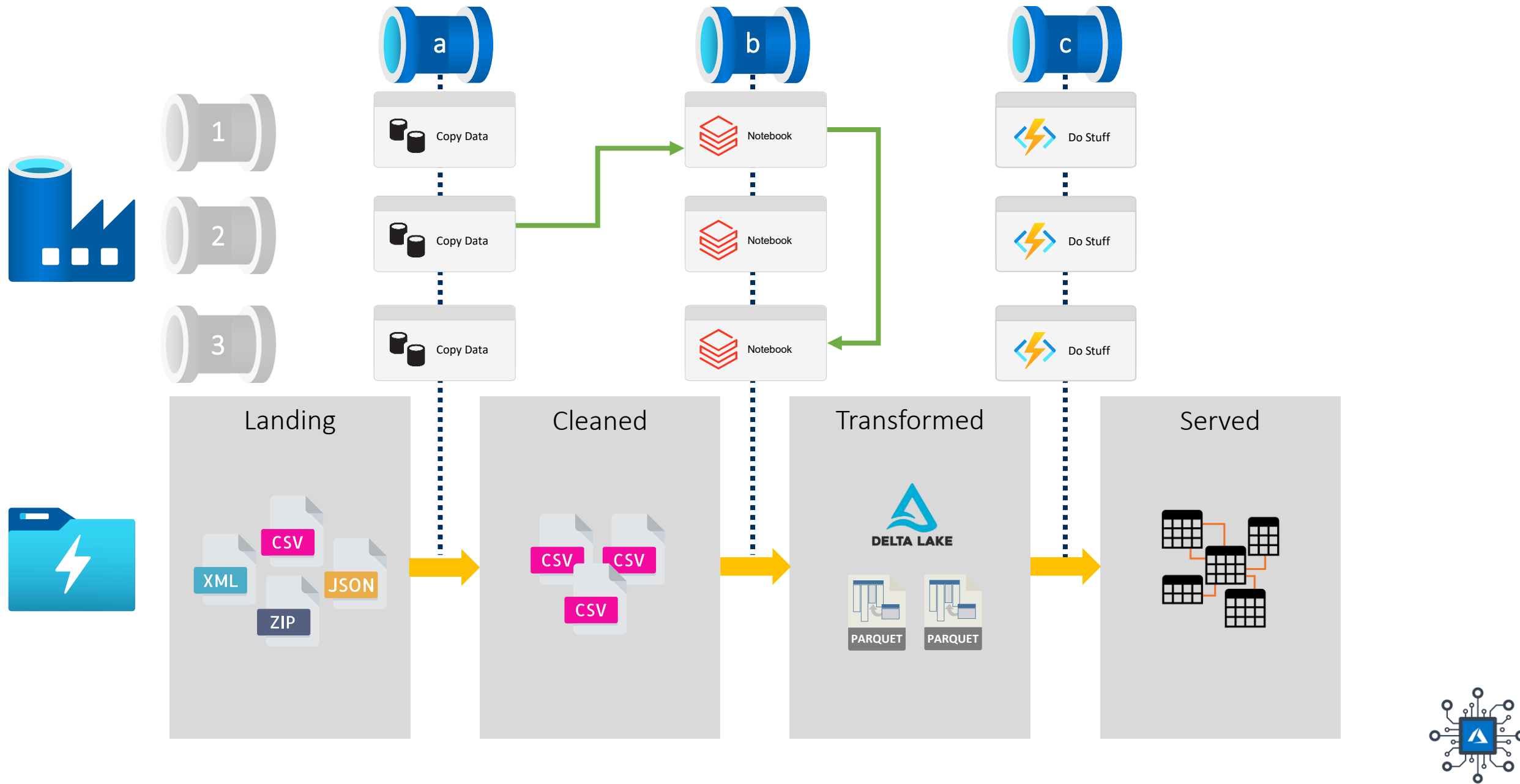
Problem



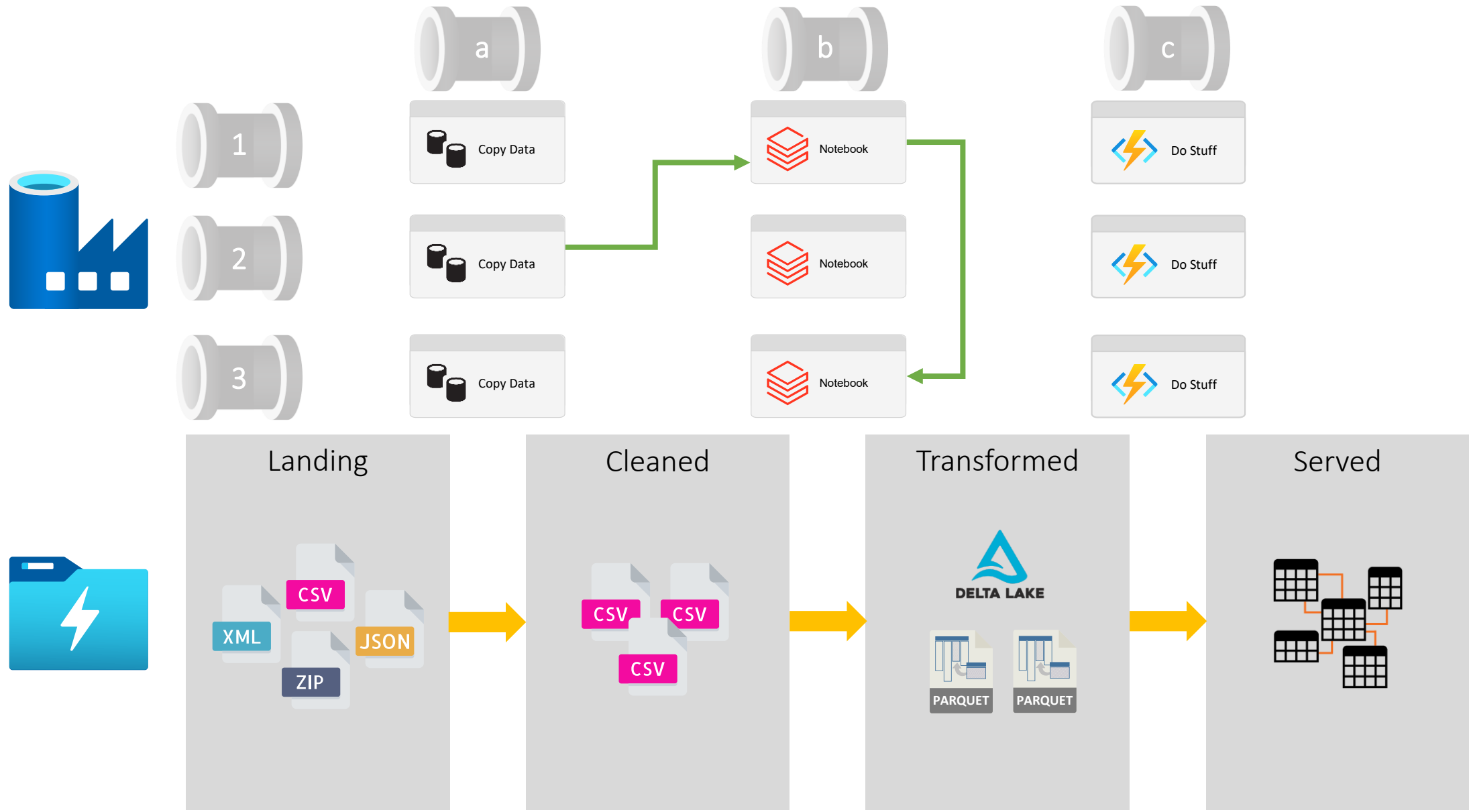
Problem




Problem

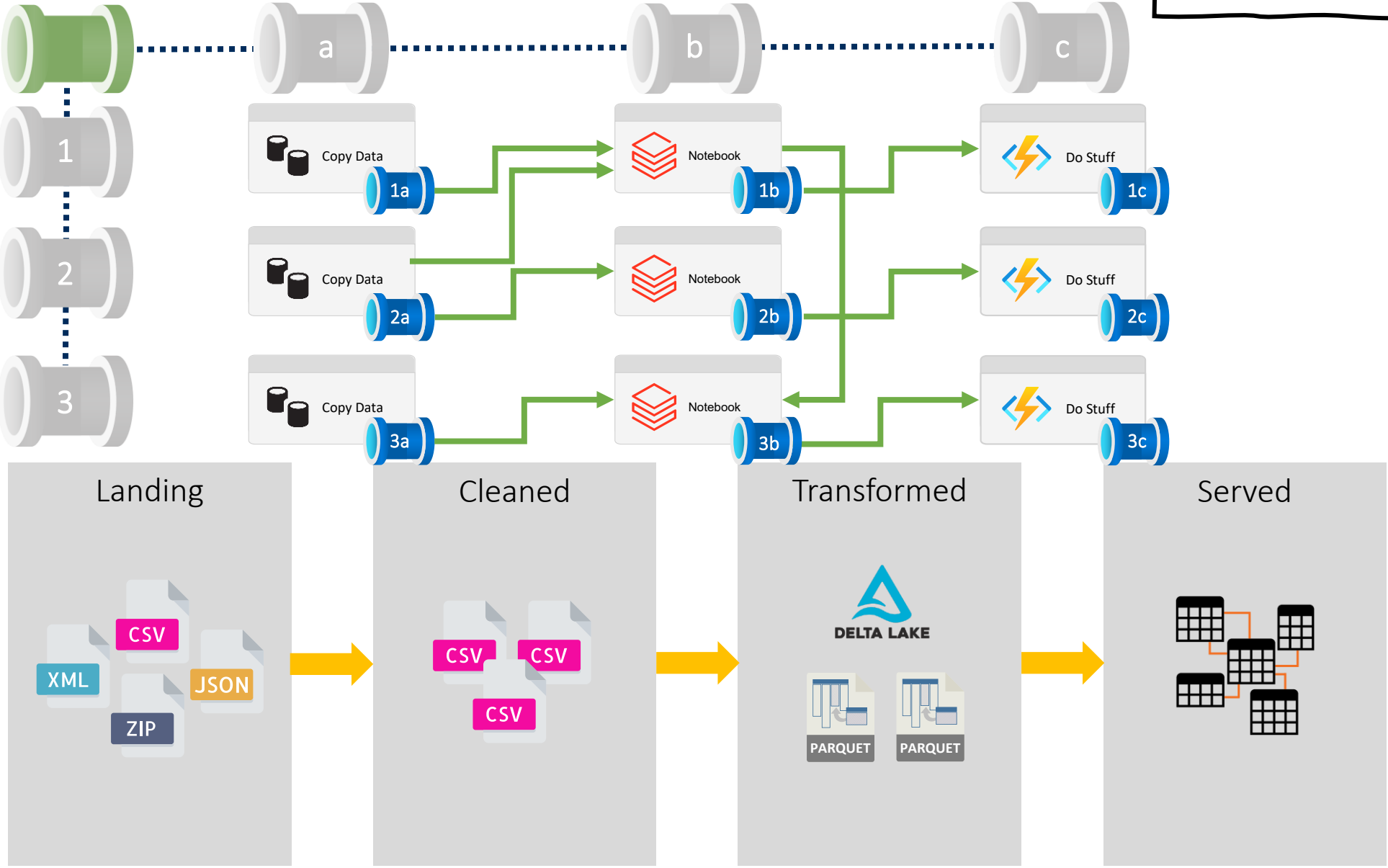
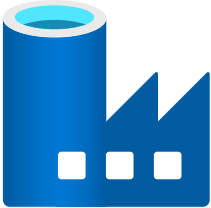


Problem

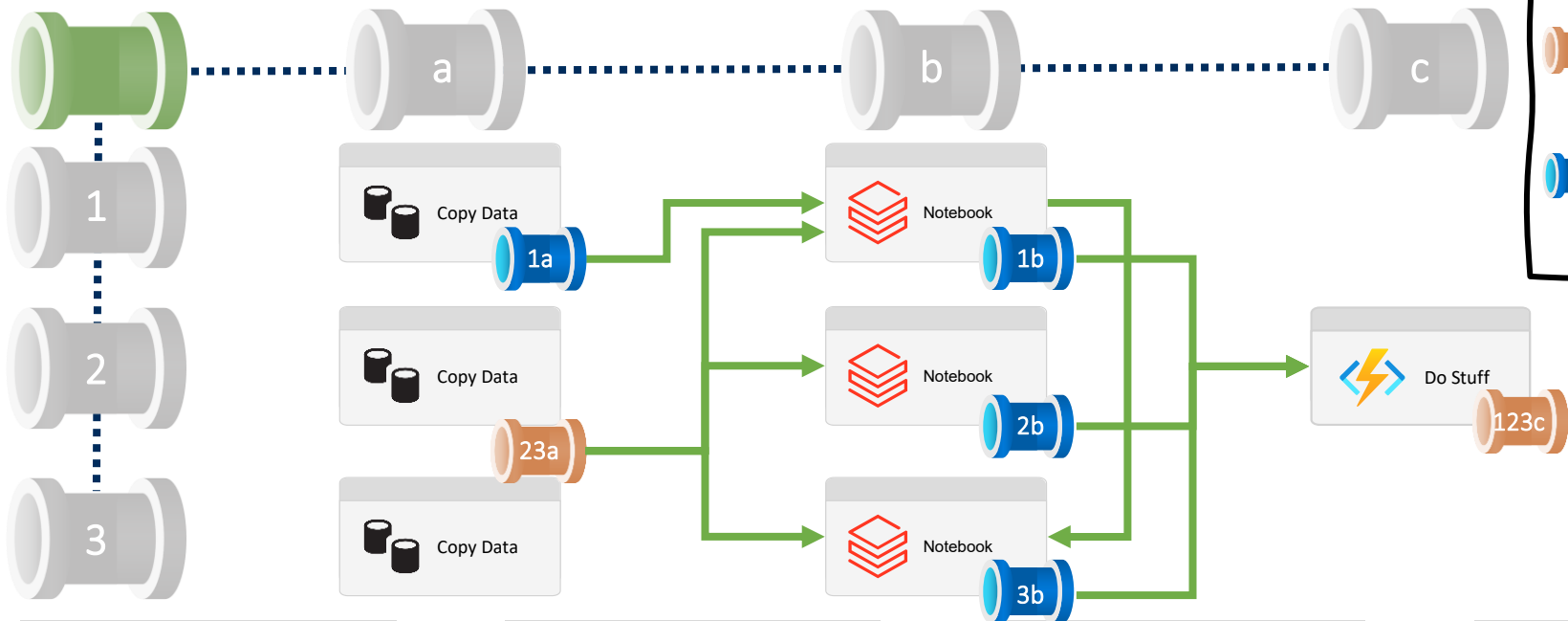
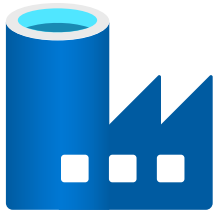


Problem

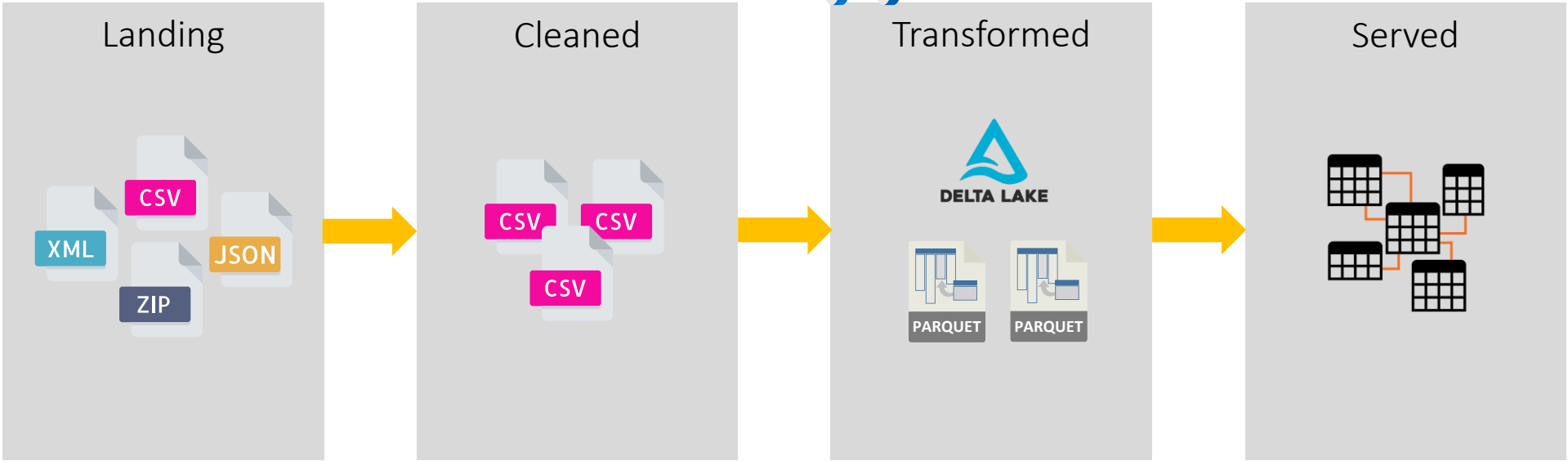
 Only 40 Activities per Pipeline.



Problem



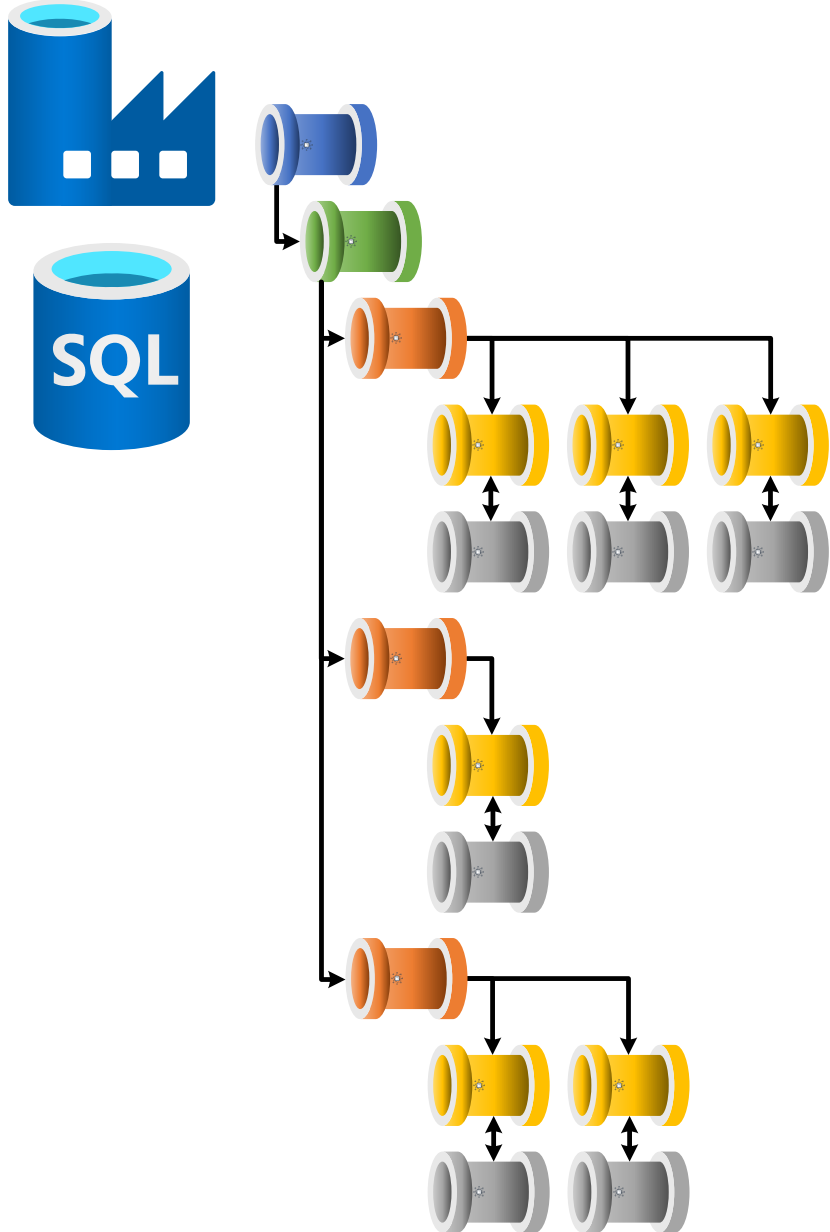
- Grandparent pipeline for all processing.
- Parent pipeline to consolidate work.
- Child pipelines for low level dependencies.



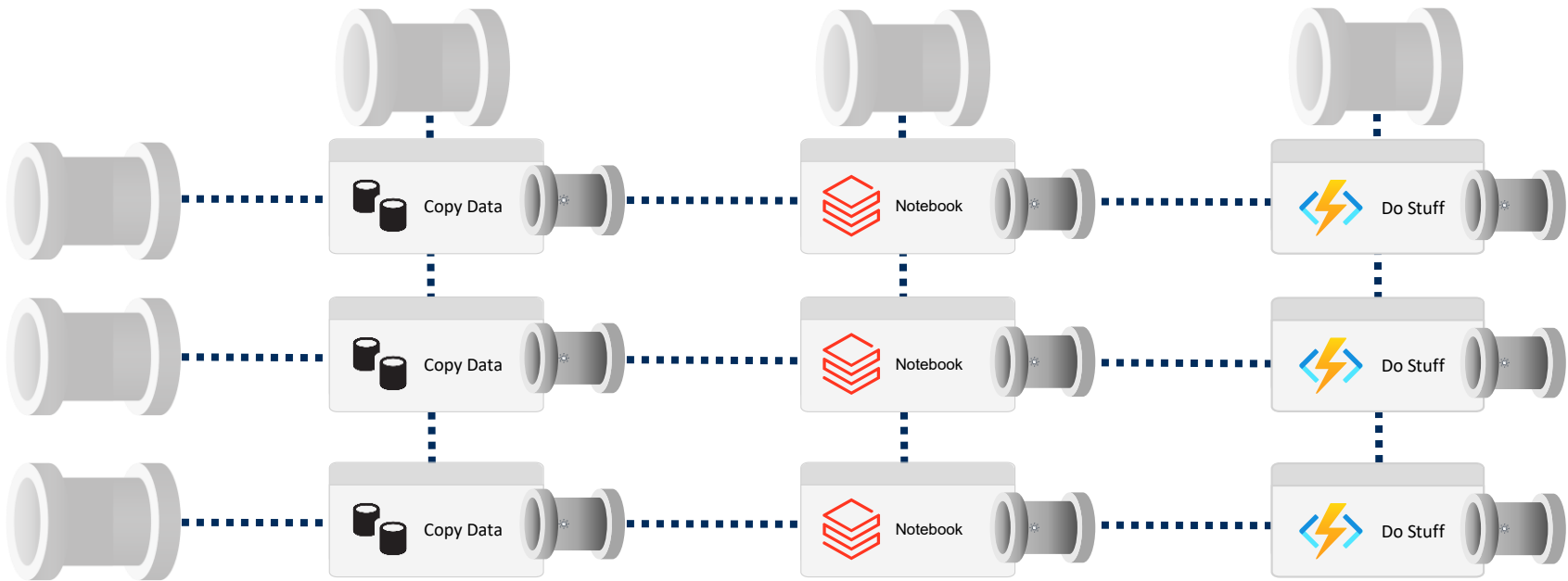
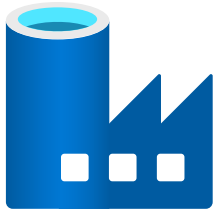
Solution



Solution: Use Metadata to Drive Data Factory Pipelines



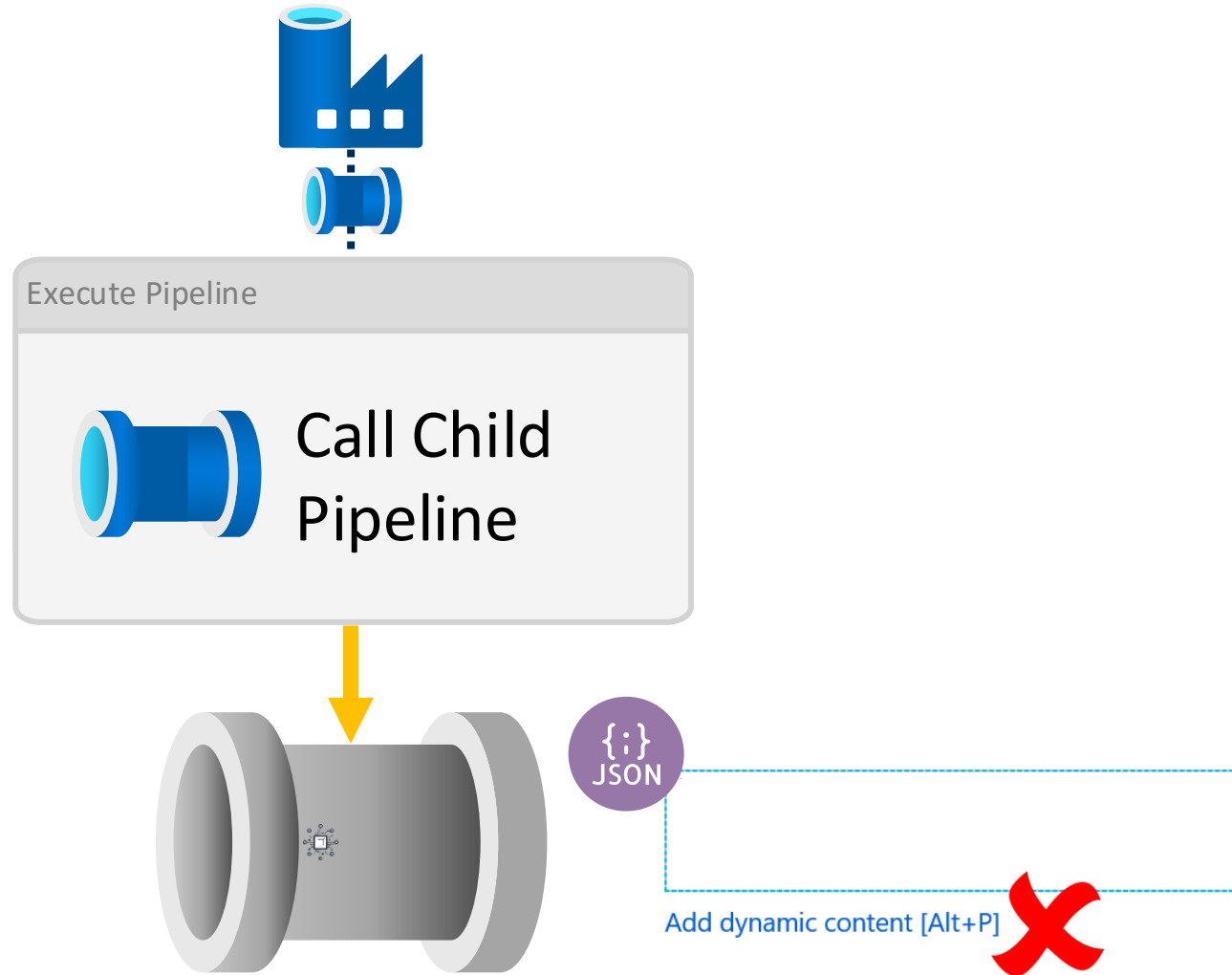
Solution



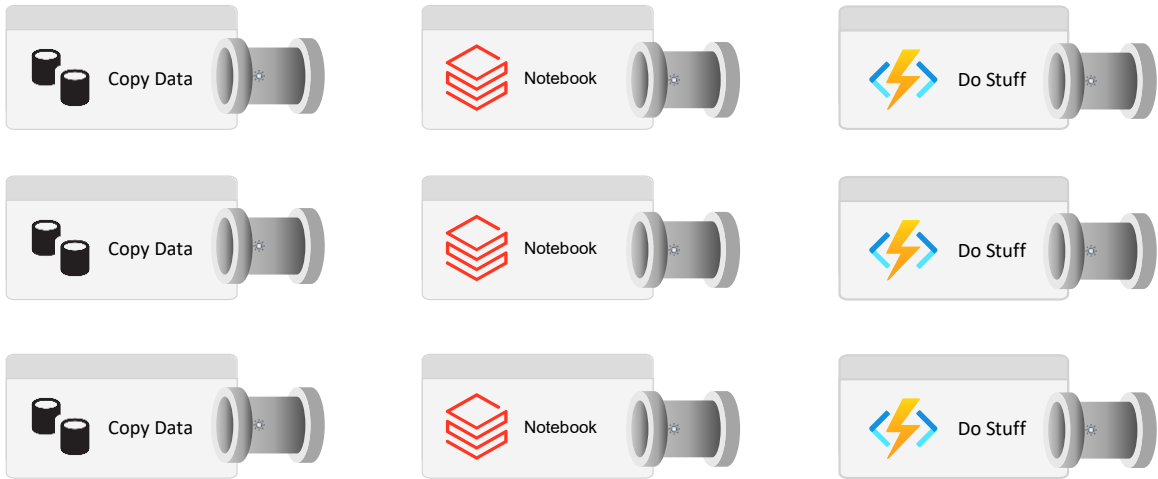
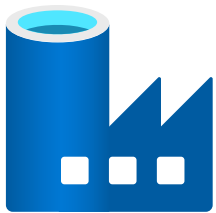
Stages	Pipelines
1	a
2	b
3	c
	d
	e
	f
	g
	h
	i

Stage	Pipeline
1	a
1	b
1	c
2	d
2	e
3	f
3	g
3	h
3	i

One More Problem



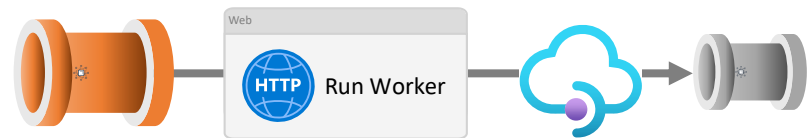
Calling Our Worker Pipelines



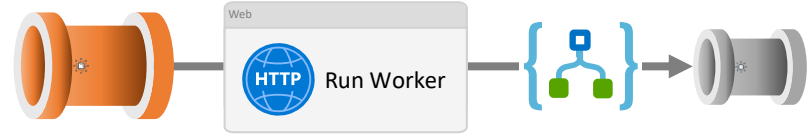
Stages	Pipelines
1	a
2	b
3	c
	d
	e
	f
	g
	h
	i

Stage	Pipeline
1	a
1	b
1	c
2	d
2	e
3	f
3	g
3	h
3	i

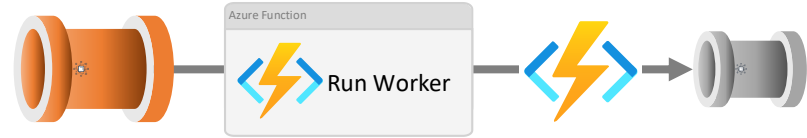
Option 1:



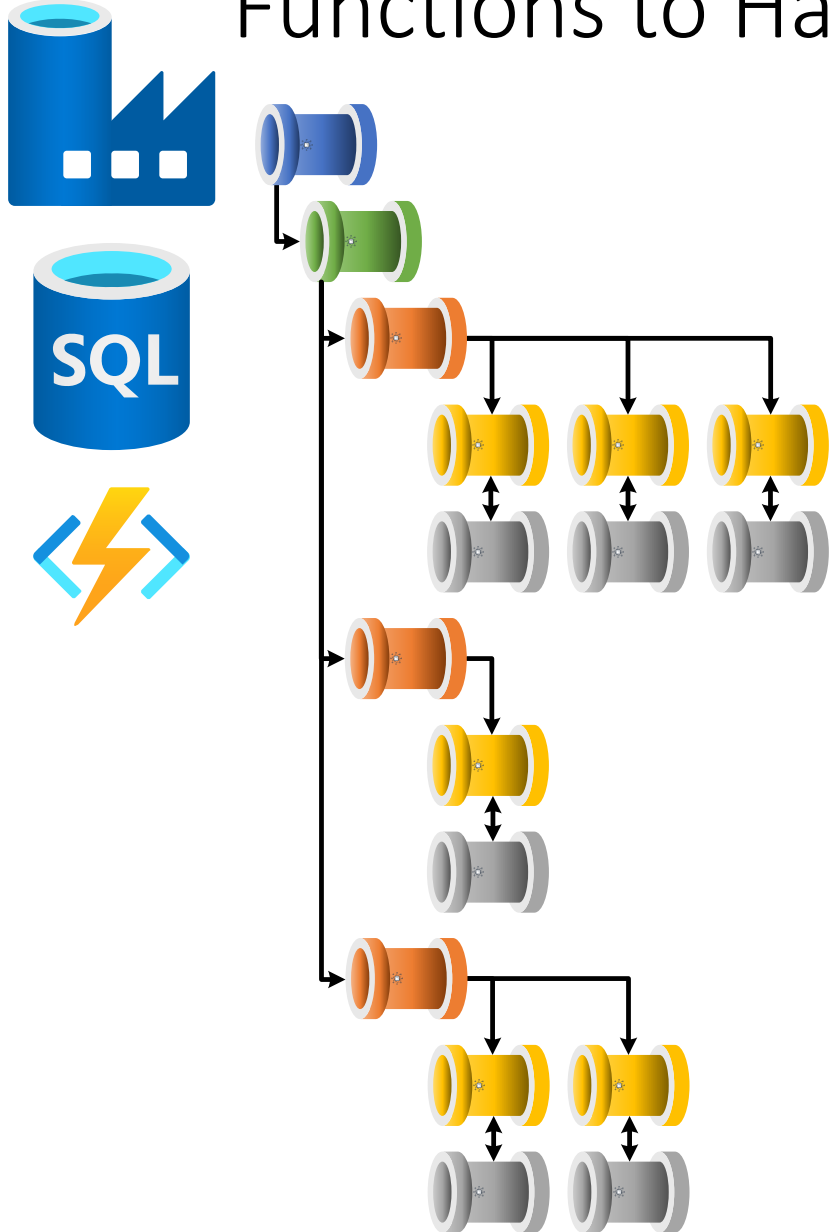
Option 2:



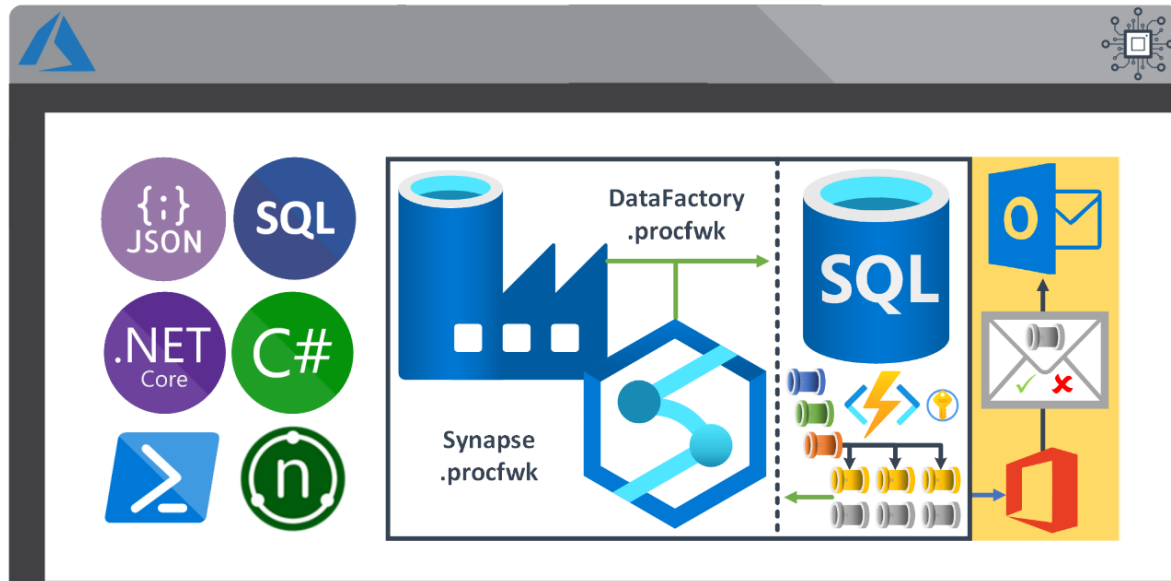
Option 3:



Solution: Use Metadata to Drive Data Factory Pipelines & Functions to Handle the Worker Execution



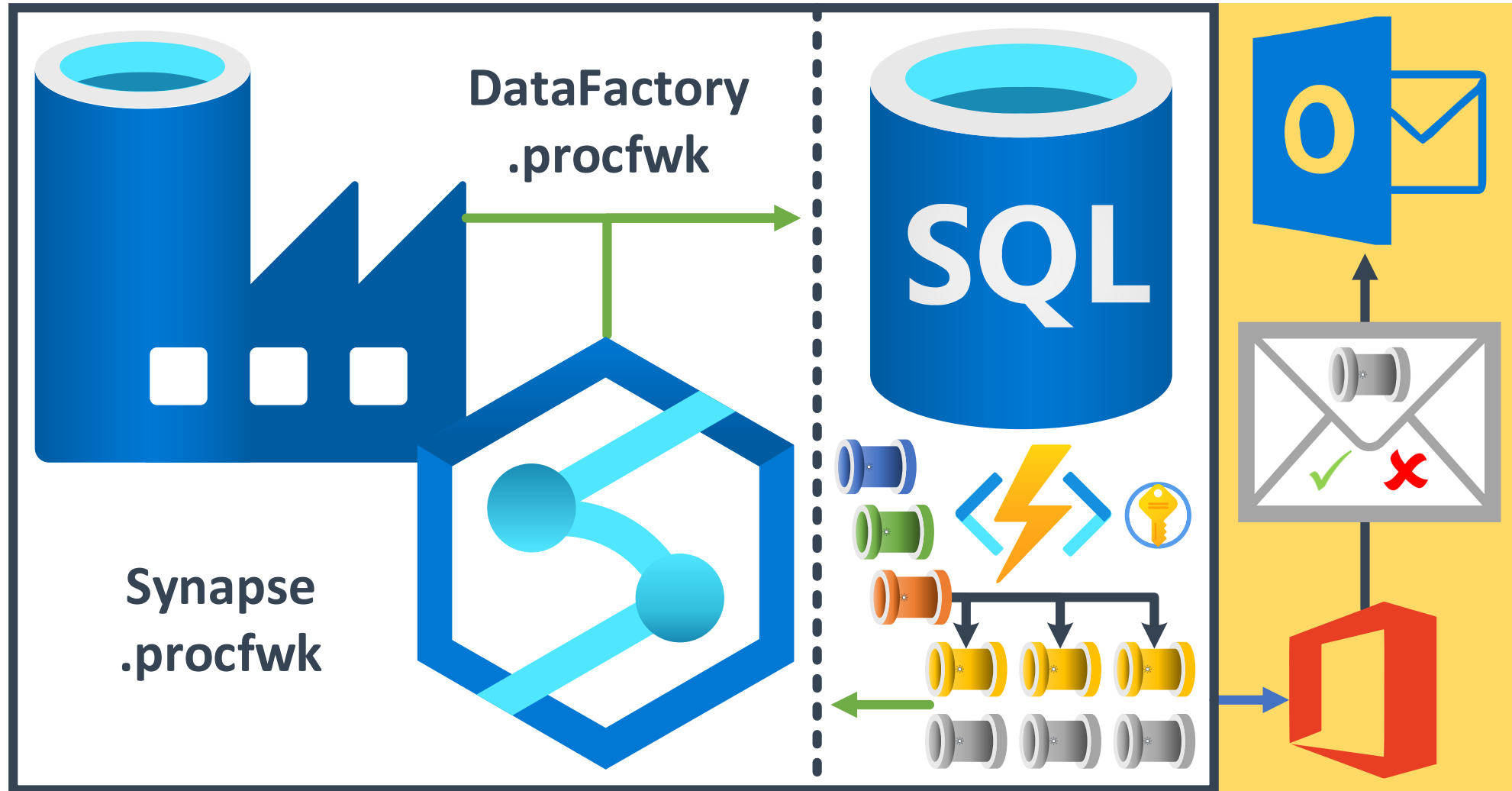
Introducing procfwk.com

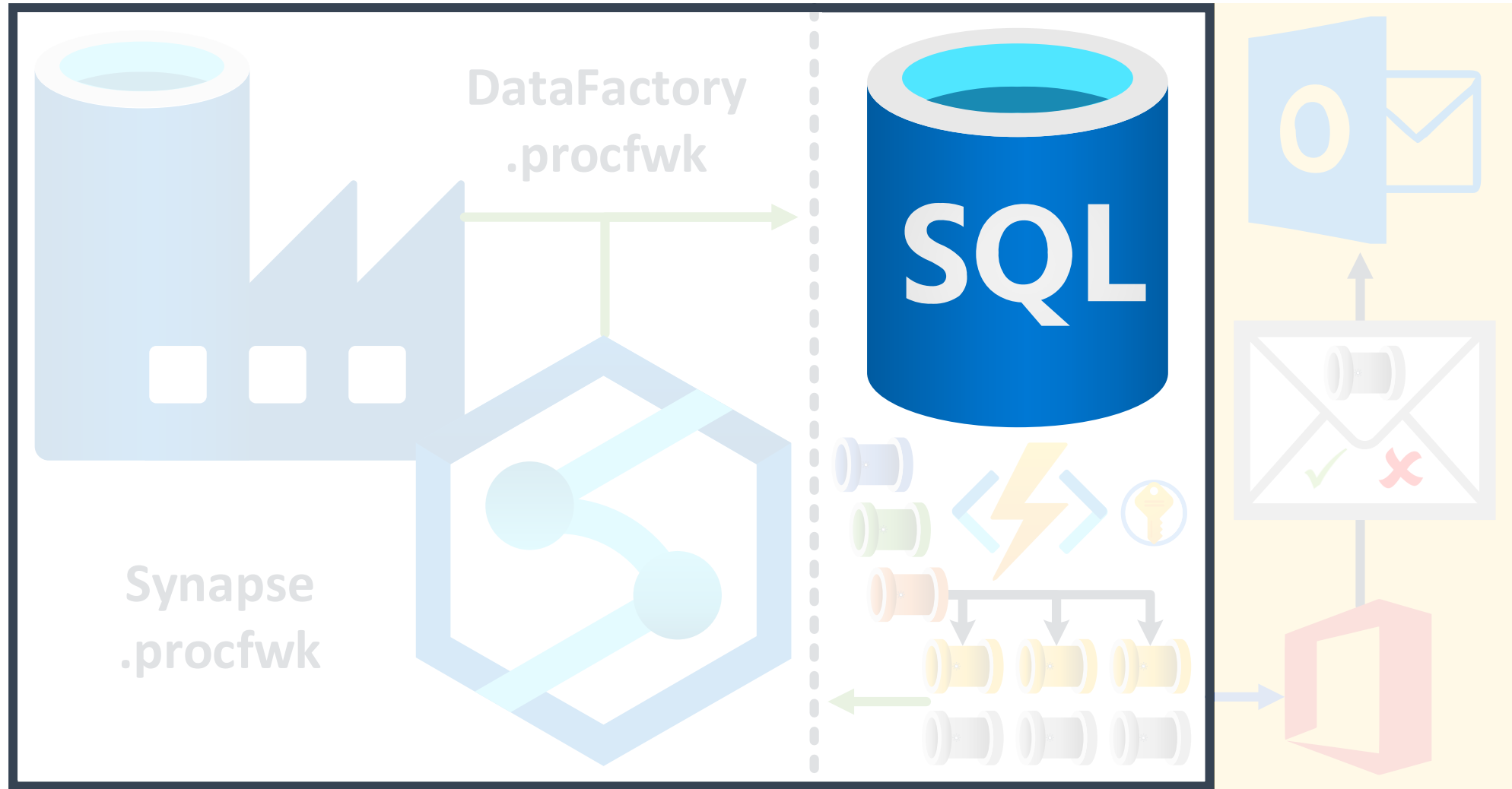
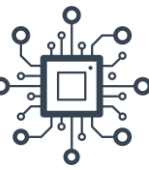




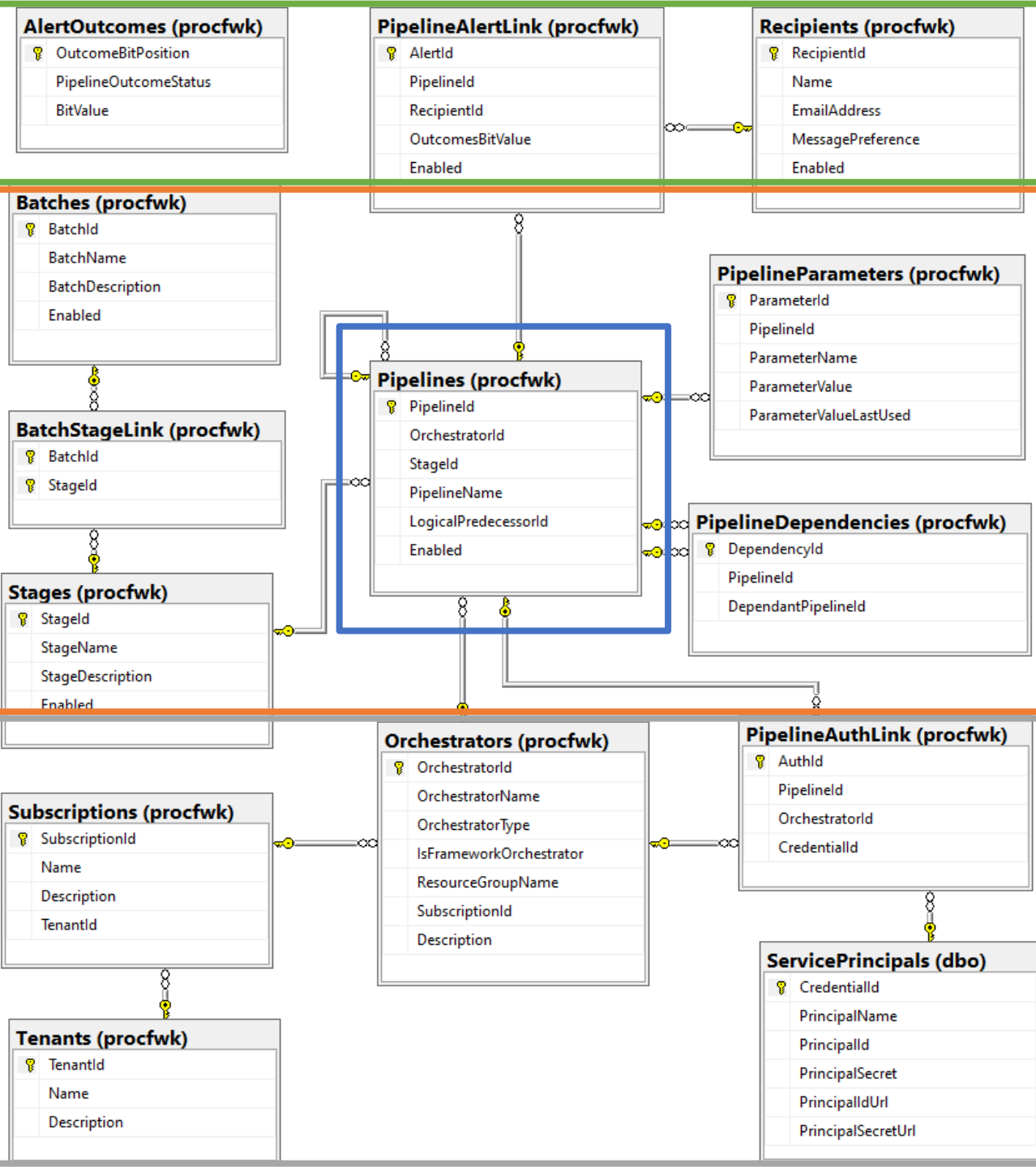
procfwk Features

- 00 Granular metadata control.
- 00 Metadata integrity checking.
- 00 Global properties.
- 00 Complete pipeline dependency chains.
- 00 Concurrent batch executions.
- 00 Execution restart-ability.
- 00 Parallel execution stages.
- 00 Full execution and error logs.
- 00 Operational dashboarding.
- 00 Low-cost orchestration.
- 00 Disconnection between framework and worker pipelines.
- 00 Cross Tenant/Subscription/Data Factory control flows.
- 00 Pipeline parameter support.
- 00 Simple troubleshooting.
- 00 Easy deployment.
- 00 Email alerting.
- 00 Automated testing.
- 00 Azure Key Vault integration.
- 00 Is pipeline already running checks.





Framework Database



Configuration & Behaviour

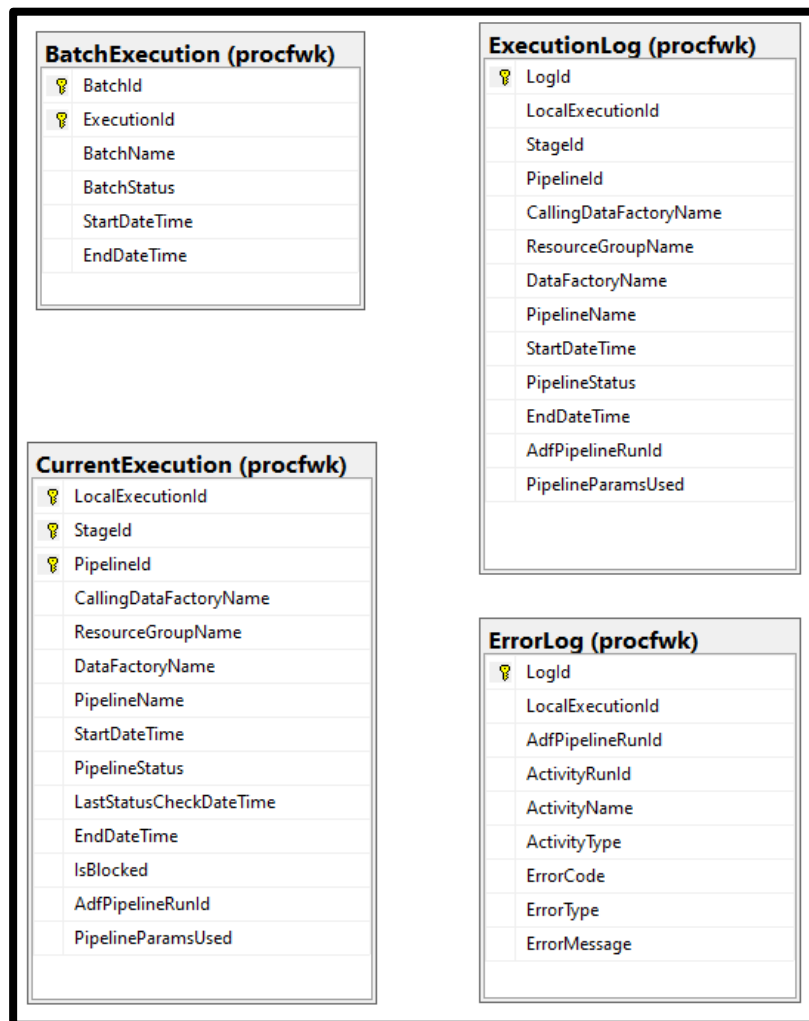
Core Metadata

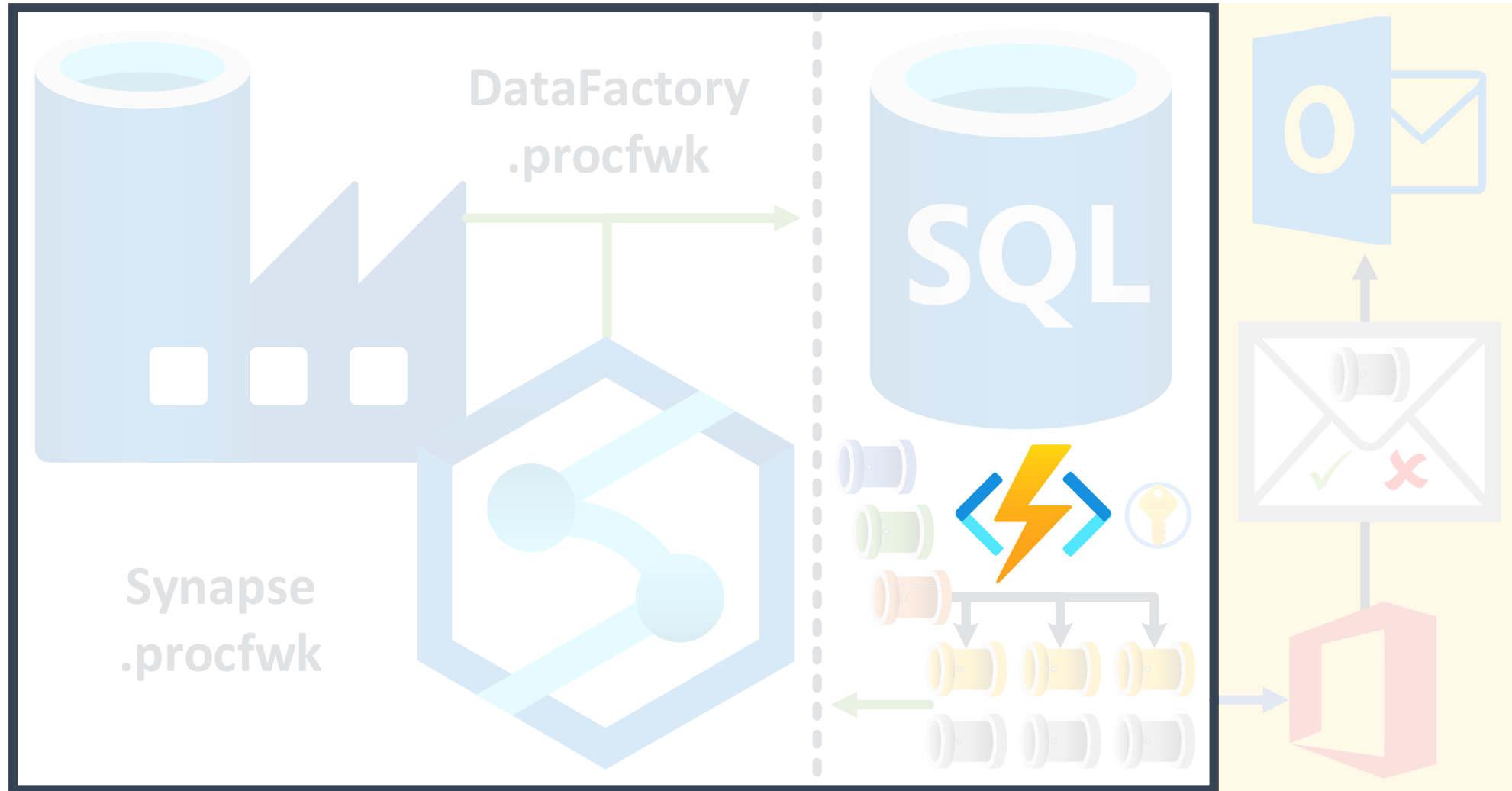
Execution Handling

Location & Authentication

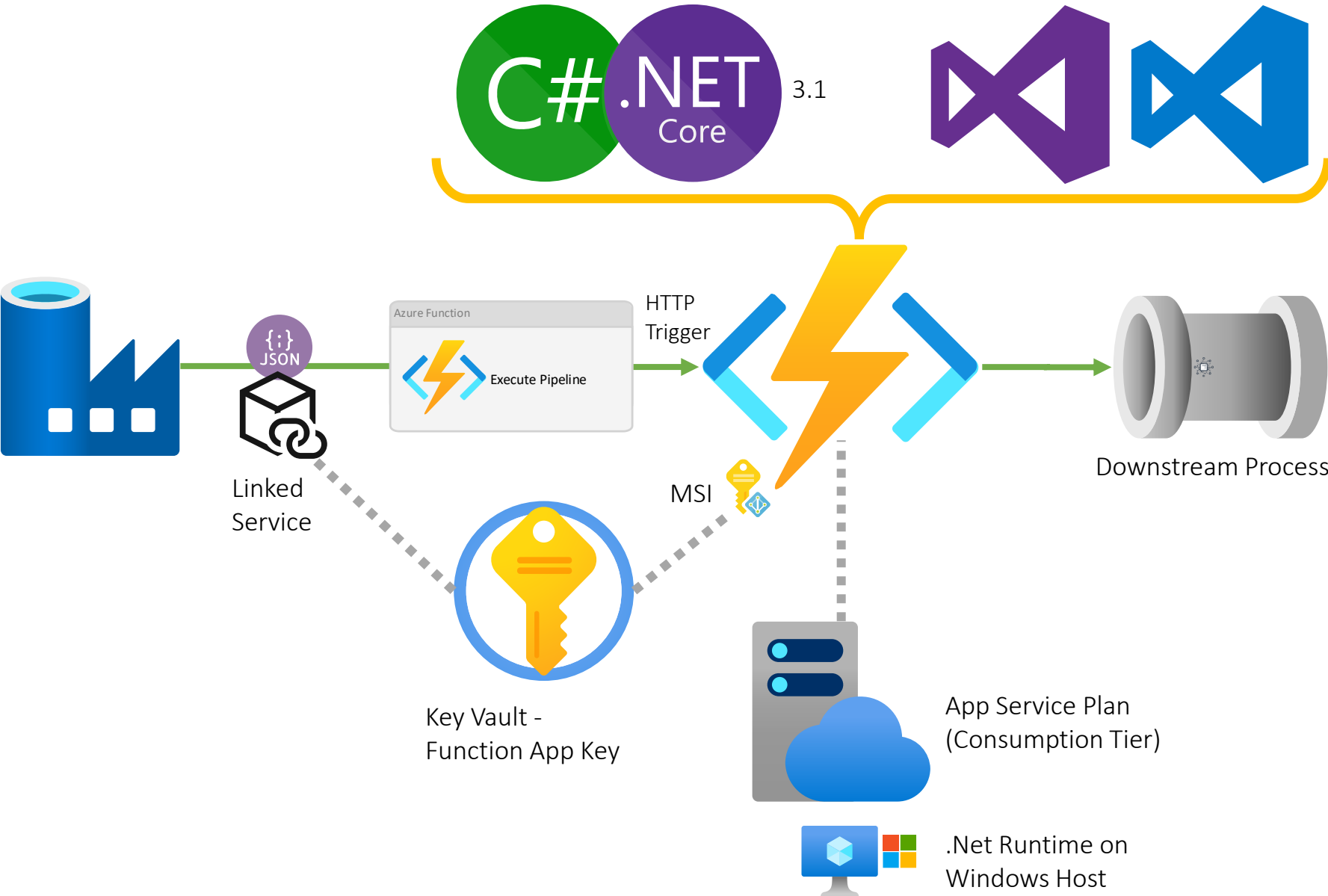
Email Alerting

Runtime & Logging



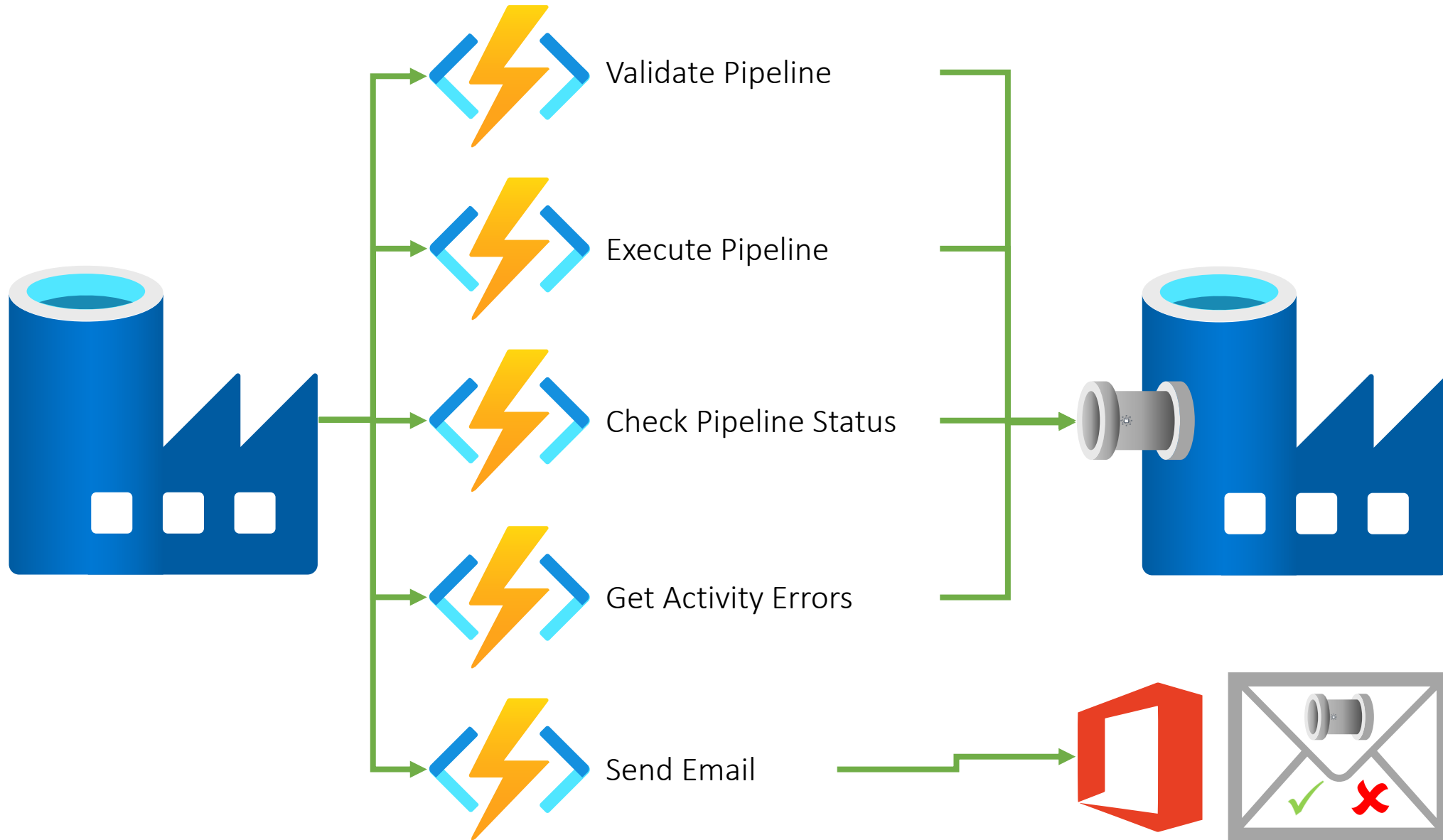


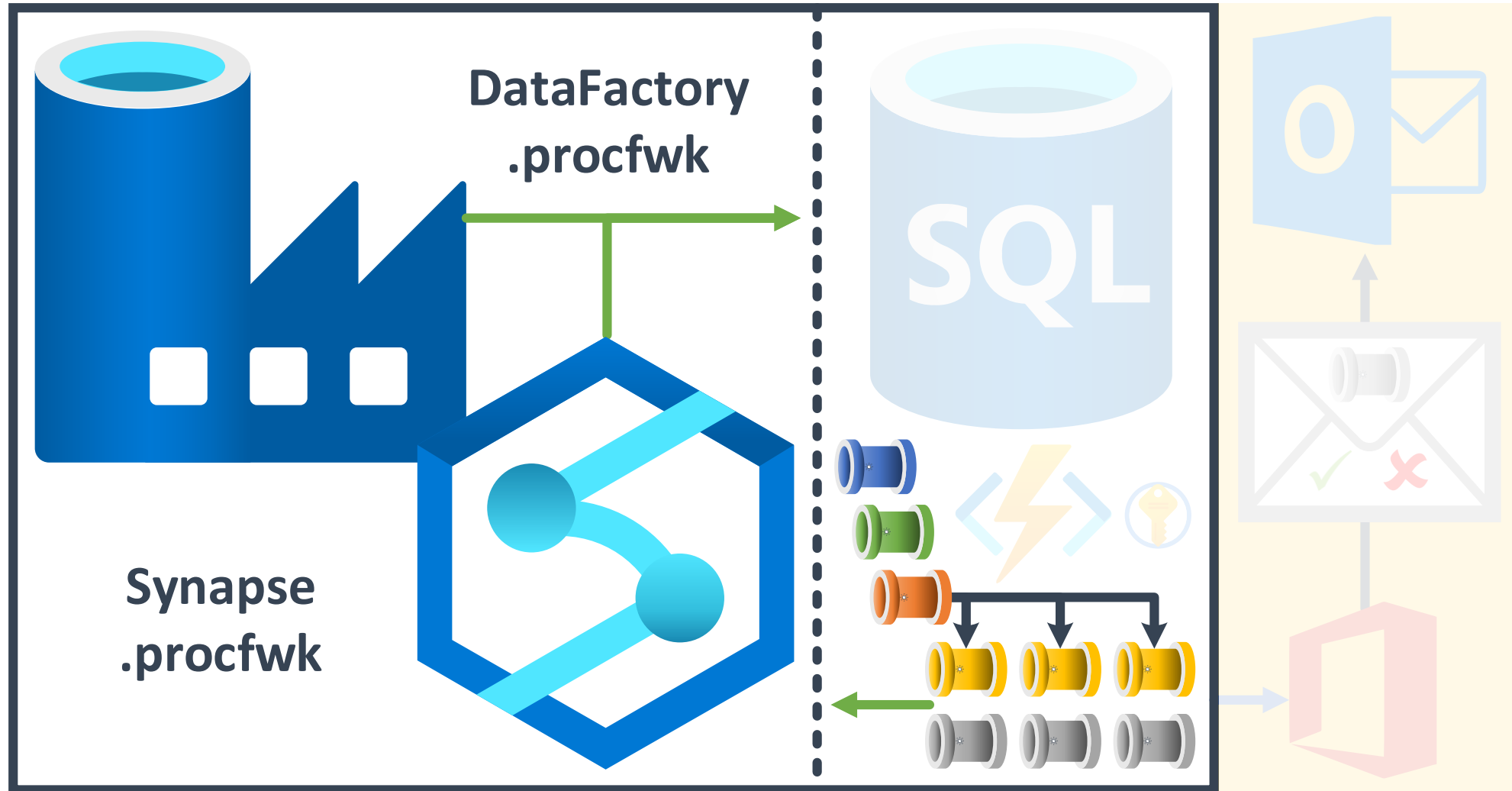
Functions Creation & Configuration



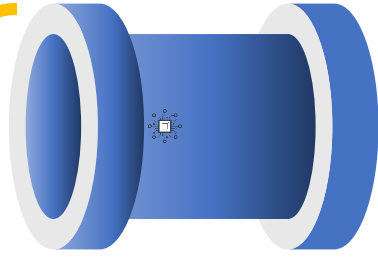


procfwk Functions



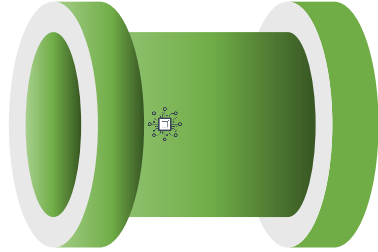


Framework Pipeline Hierarchy



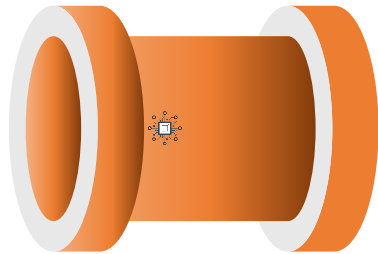
- Grandparent

Role: Optional level platform setup, for example, scale up/out compute services ready for the framework to run.



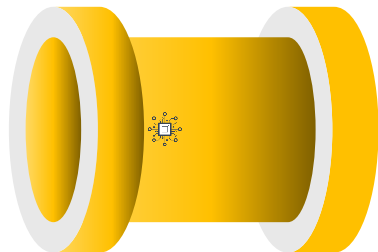
- Parent

Role: Execution run wrapper for batches and execution stage iterator.



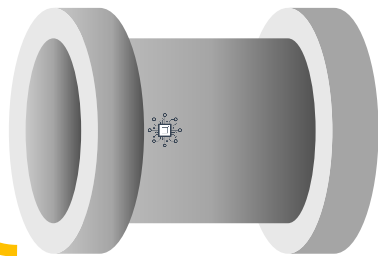
- Child

Role: Scale out triggering of worker pipelines within the execution stage(s).



- Infant

Role: Worker validator, executor, monitor and reporting of the outcome for the single worker pipeline.



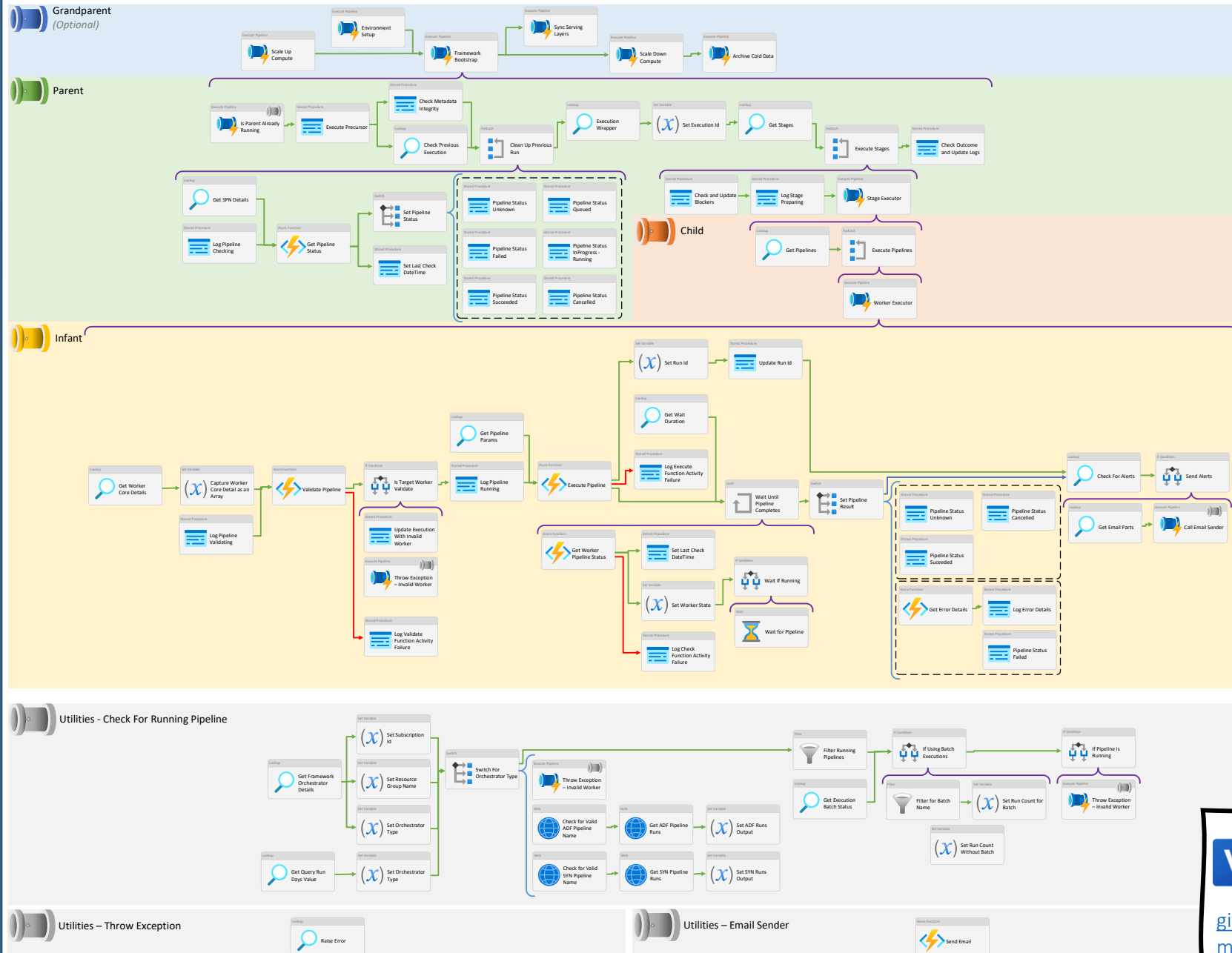
- Worker

Role: Anything specific to the process needing to be performed.



Processing Framework - Activity Chain

Orchestration Framework



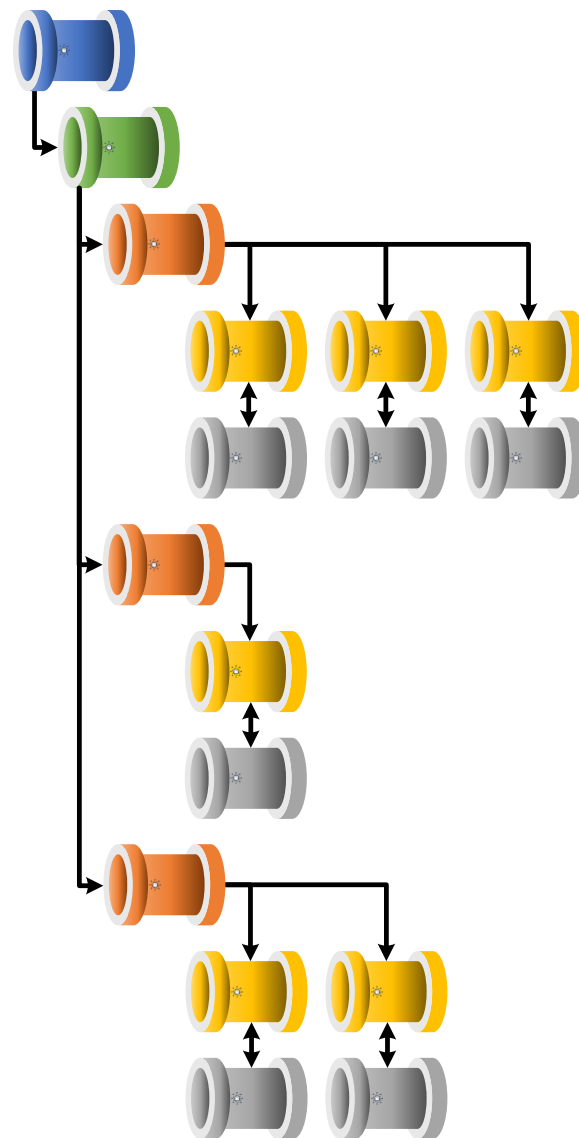
Worker Pipelines

- Worker 1 - Extract
- Worker 2 - Clean
- Worker 3 - Transform
- Worker 4 - Load
- Worker 5 - Serve
- Worker n -



Go to Visio file in
GitHub:

github.com/mrpaulandrew/procfwk/blob/master/Images





Thank you for listening...

Paul Andrew



Mr Paul Andrew
Consulting Ltd

Blog: mrpaulandrew.com
YouTube: [c/mrpaulandrew](https://www.youtube.com/c/mrpaulandrew)
Email: paul@mrpaulandrew.com

Twitter: [@mrpaulandrew](https://twitter.com/mrpaulandrew)
LinkedIn: [In/mrpaulandrew](https://www.linkedin.com/company/mrpaulandrew)

GitHub: github.com/mrpaulandrew

Contact Details



sqlbits
Session Feedback

