Cloud Accelerate Factory (CAF)
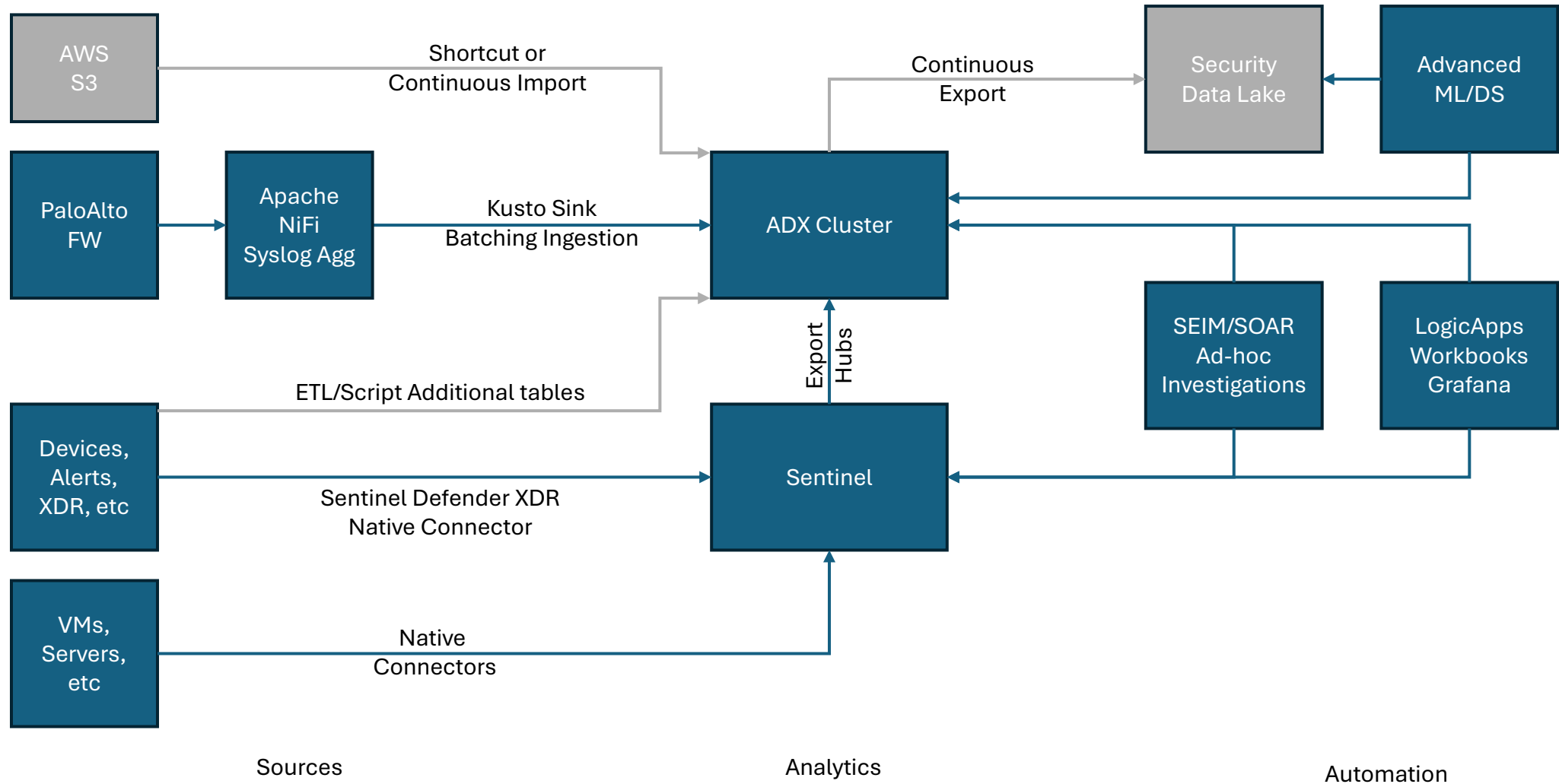# ILZ Fed Customer

Real-Time Intelligence Nomination

Partner: Cloud Fit Software

High-level Architecture Diagram
Cyber Workflow w/ available features in Gov Cloud

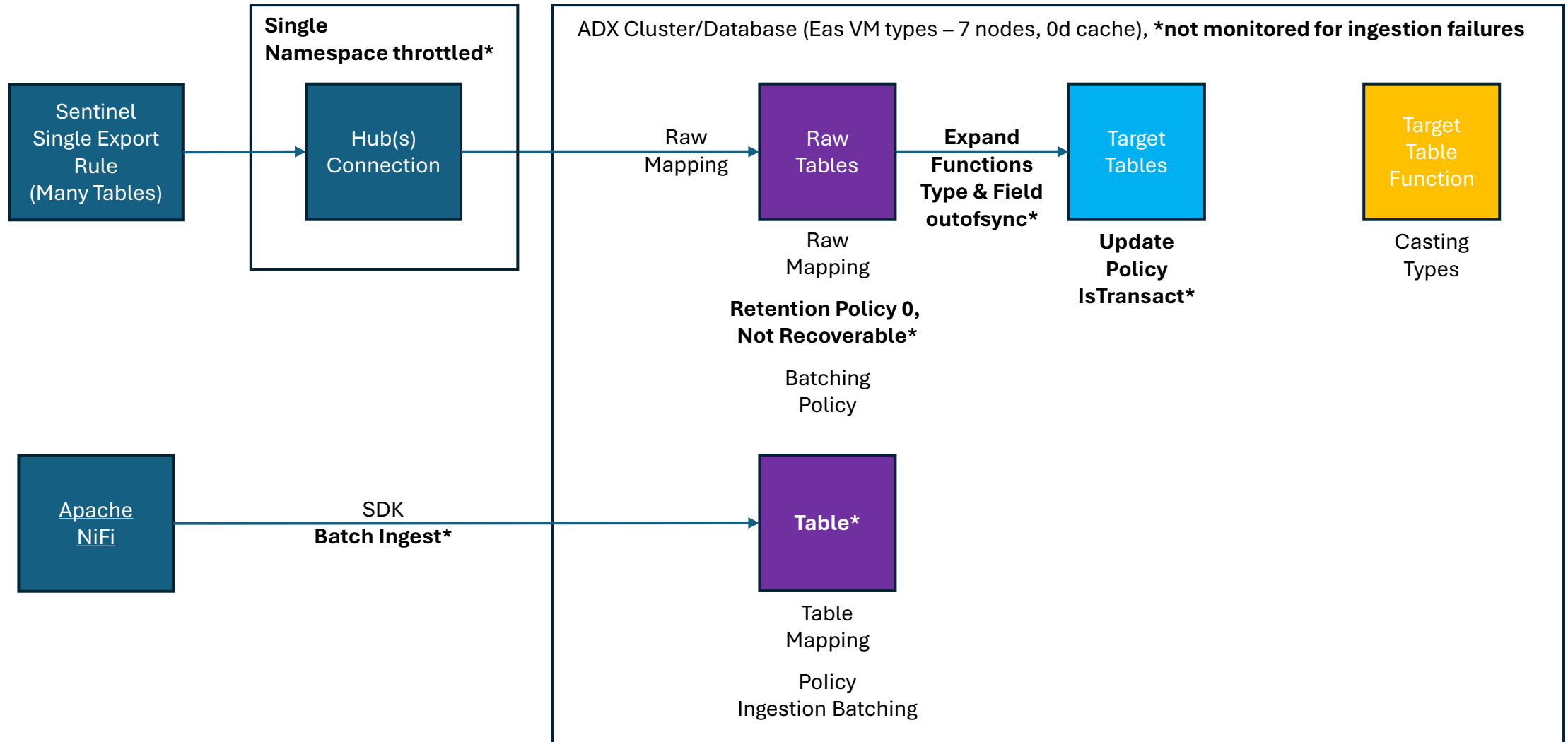Sources   Analytics   Automation

Sentinel + ADX
Considerations

- Limit 10 export rules per LAW workspace (10-20 tables per rule), one hubs per table, many hubs per namespace. Reference [Sentinel Tools ADX guide](#).

- Sentinel Export API payloads differs Sentinel table schemas as some columns are internal for Sentinel-only billing.

- Sentinel custom tables are not supported by Continuous Export feature.

- Use premium hubs and **monitor** for throttling.

- Scale by adding partitions, PUs, namespaces and/or hubs.

- Sentinel data not intended to ever be read/queried should be exported to blob storage to lower costs, instead of hubs & ADX.

- Land data raw then use update policy to flatten. If you choose to flatten on ingestion is simpler but use DropMappedFields mapping option to catch full record. Objective is to have ability to revert to the full-record.

- Minimize caching for raw layer. Maximize caching for silver & gold layers. ie. Target tables or materialized-views used for queries/reporting.

- Increase retention policy.

- Use Las vm-sku types if available in your region. Scale by adding nodes.

- Enable diagnostics and Optimized Auto-Scale.

- Monitor ADX using system views, Insights Blade and Azure Advisor.

ADX Dataflow (Before)
Cyber Workflow w/ available features in Gov Cloud

Single Namespace throttled*

Sentinel Single Export Rule (Many Tables)

Hub(s) Connection

ADX Cluster/Database (Eas VM types – 7 nodes, 0d cache), ***not monitored for ingestion failures**

Raw Mapping

Raw Tables

Raw Mapping

Retention Policy 0, Not Recoverable*

Batching Policy

Expand Functions Type & Field outofsync*

Target Tables

Update Policy IsTransact*

Target Table Function

Casting Types

Apache NiFi

SDK
Batch Ingest*

Table*

Table Mapping

Policy Ingestion Batching

ADX Dataflow (After)
Cyber Workflow w/ available features in Gov Cloud

Multiple Namespace 1:20 (Premium 4+ PUs)

ADX Cluster Database (Las VM types, **Enable Optimized Auto-Scale**, Enable Diags, Enable Streaming, 365d cache, 913d retention)

Sentinel Export Rules (limit 10 per wks)

Hub(s) Connection (8+ part.)

Monitor for throttling

Raw Mapping

Raw Tables

Raw Mapping

Caching Policy 1d
Retention 7d & recoverable

Batching Policy 1m, 1GB

Expand Functions All fields & types

Target Tables

Update Policy

Caching Policy 365d
Retention 913-3650d

Mat-views

Functions

Logstash (Cluster)

SDK Streaming Ingest

Table

Table Mapping

Expand Function

Target Tables

Update Policy

Caching Policy 365d
Retention 913-3650d

Bronze layer

Silver layer

Gold layer

Presidential Mandate: 12 months hot, 30 months cold. Guidance for Implementing M-21-31: Improving the Federal Government's Investigative and Remediation Capabilities, M-21-31

Validations

- Get count for both systems (Sentinel vs ADX) where ingestion_time()>=ago(10m). This should match...very very closely in reference to the particular exported table. ie.

rawtable ... targettable | where ingestion_time() >=(10m) | count

- PS > LAW table 50 cols > get schema > generates expand function dynamically. > generates the target dynamically.

- LAW Export API > outputs payload schema > raw > expect the expand function > flatten all fields to target.

- Raw Payload if 43 > all 43 fields expanded to target. (Get Data UI preview should show them populated for a recent raw json, therefore target should show them populated also)

- Check table(s) retention, cache, update policy is(not)transact, hub partitions, namespace is prem/stand – sku, PUs, etc, check for requests, **throttling**.

- Check adx metric insights blade – no failures and using system view .show ingestion failures | where FailedOn >=ago(10m)

- Scripts located [here](#).