



**CloverDX**

# A Data Integration Tool Show Case

## Background

- As **IET evolves**, we are constantly faced with **opportunities and challenges** in equal measure as we better ourselves.
- Related or stand alone Data touches on **Students, Staff and Academics** in a number of **instances exist in silos** due to either platform protection or legal reasons.
- The question development teams face is on how to bring needed filtered components of this **data across various platform**.
- CloverDX - A go-to **Integration tool bringing solution(s)** to this most asked question in our sphere of operation.
- A viable ETL vehicle that deliver data in a simplified form / shape which many **end user integrators** revere.

# DISCLAIMER

Sam Irungu

No **Sales Rep** for CloverDX - 'Shape / form'

No credit awarded for marketing their product.

# INTRODUCTION

- **CloverDX** formerly known as **CloverETL**
- Robust platform - Help Automating data intensive workloads including transport and transformation of data between systems.
- Runs on Java Runtime Environment (JRE)

# CAPABILITIES



**Replacing Legacy  
Systems**



**Data Migration**



**Data Ingestion**



**Analytics /  
Business  
Intelligence**



**Risk &  
Compliance**



**Application  
Integration**



**Data Warehousing**



**DataOps**

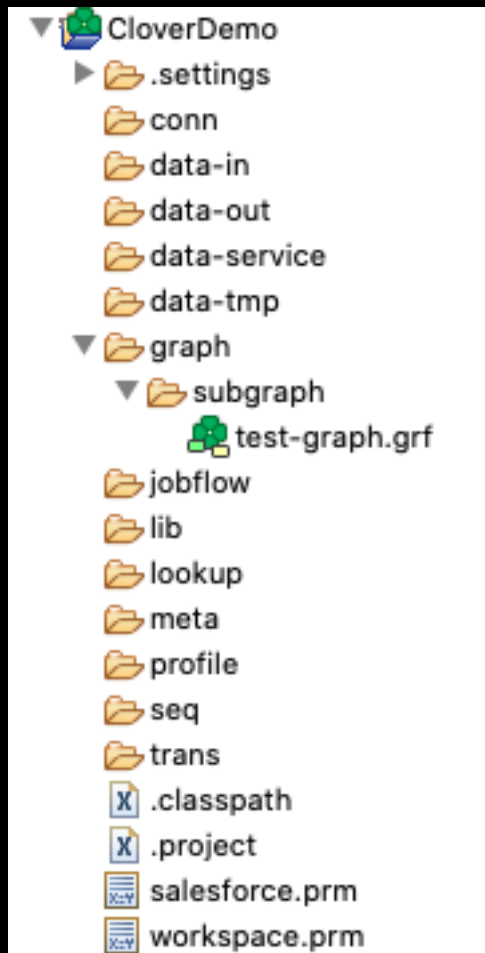
# COMPOSITION

- Is a **three components** product platform :
  - Designer
  - Server
  - Cluster

# 1. DESIGNER

- **CloverDX Designer** is an engine-based **standalone application** for creating and running Graphs.
- Its capability allows you to :
  - Connect to and communicate with **CloverDX Server**,
  - Create Projects
  - Graphs
  - Subgraphs
  - All others

# PROJECT



- Whole **project structure** is only on **your local computer**.
- All data and graphs reside locally in a project in a workspace, graphs runs locally with help of **CloverDX Runtime**.
- Local projects can be easily **versioned** with version system of your choice e.g. Git, Maven



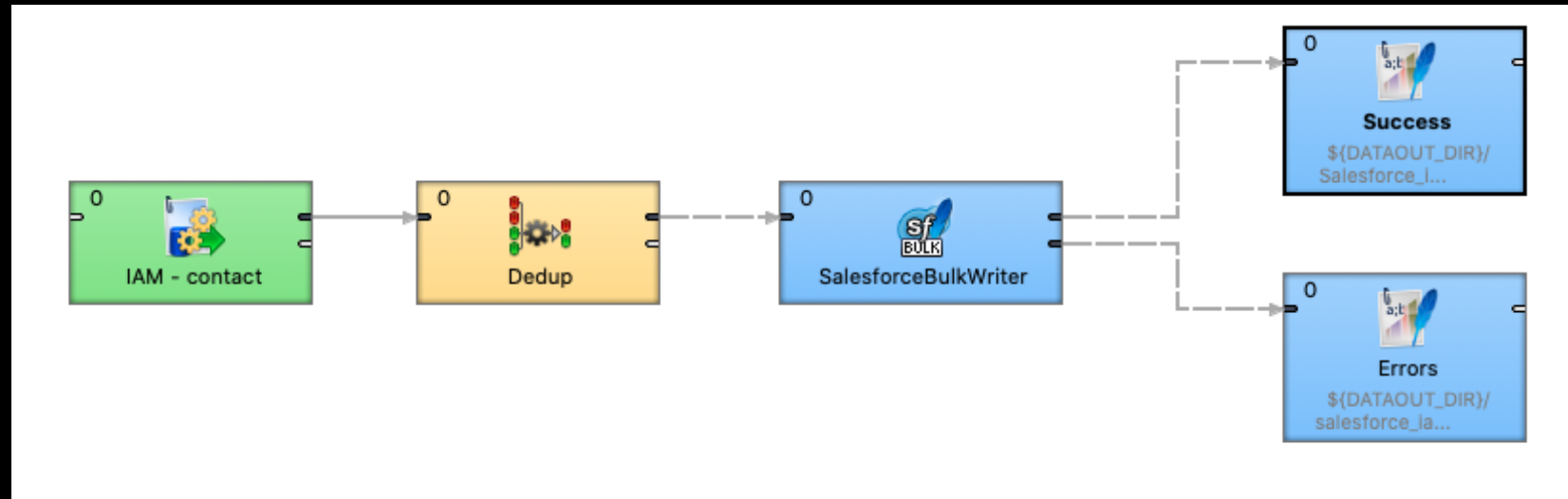
# GRAPH(S)

Smallest executable unit. Also known as a **WORKFLOW**

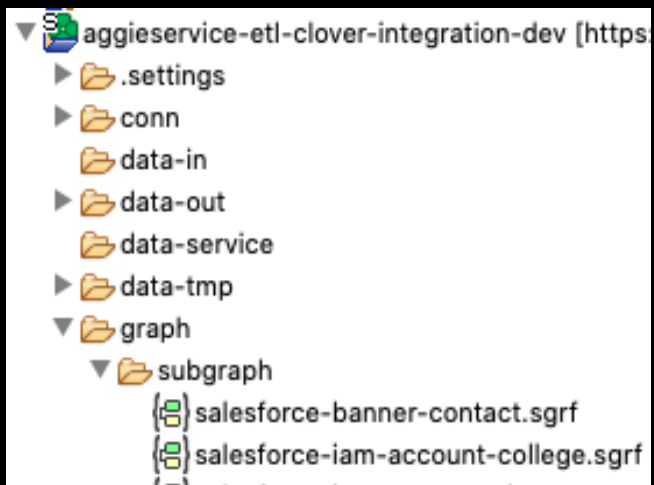
**Components (nodes)** are the most important **graph elements**.

They all serve to process data.

Most of them have ports through which they **can receive data** and/or **send the processed** data out.



# SUB-GRAPH



- **Subgraph** is a **user-defined reusable component** with logic implemented as graph instead of Java code.
- A **regular graph** and may use any graph elements (components, connections, lookups, sequences or parameters).
- Used to **Visually reduce the number of component** in complex graphs and highlight important processing logic
- Subgraphs **can be nested**; a subgraph definition may use other subgraphs.
- Stored in a separate file with **\*.sgrf extension**.
- In default project layout, a directory **\${PROJECT}/graph/subgraph** is created for storing subgraph files. You can reference this directory via the **\${SUBGRAPH\_DIR}** parameter.

## 2. SERVER

- CloverDX Server application for running, scheduling and monitoring Graphs.
  - Centralized Job Management
  - Integration into enterprise workflows
  - Multi-User environment (License)
  - Parallel Execution of graphs
  - Tracking of Execution of graphs
  - Scheduling Tasks
  - Clustering and distributed execution of graphs

- Monitoring
  - Status History
  - Server Logs
  - Execution History
  - Task History
- Schedules
- Event Listeners
- Sandboxes
- Configuration
  - Security
- Help

RESOURCE UTILIZATION



WORKER

Status ● **RUNNING**

Uptime 45 days 4 hrs

Name [worker0@node01]

Ports ⓘ 10500, 10501, 10502, -, -

Running jobs 0

SYSTEM

Current time 2021-04-12 20:36:53 PDT

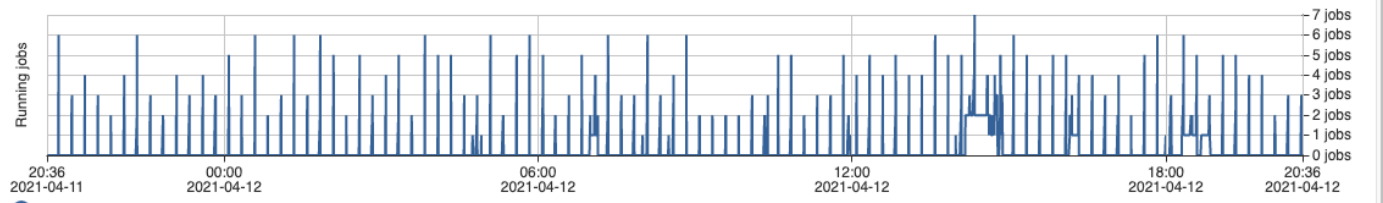
Java timezone America/Los\_Angeles

UTC time 2021-04-13 03:36:53 GMT

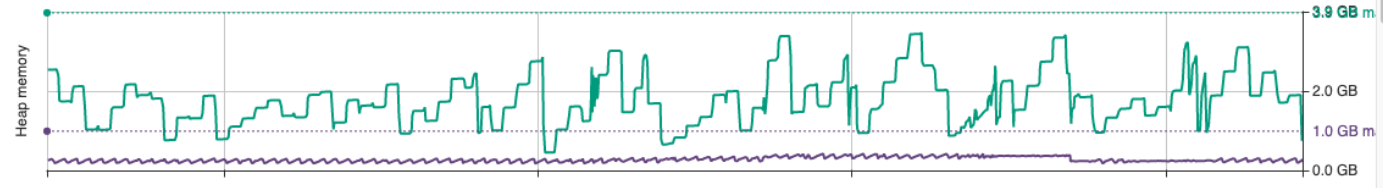
Uptime 45 days 4 hrs

Clock drift ⓘ 84 ms

PERFORMANCE



Running jobs



Refresh Actions

STATUS HISTORY (LAST 10)

READY	updated by node01	2021-02-26 14:36:53
STARTING	updated by node01	2021-02-26 14:36:53
STOPPED	updated by node01	2021-02-26 14:36:53
STOPPED	updated by node01	2021-02-26 14:36:53
READY	updated by node01	2021-02-19 15:57:12
STARTING	updated by node01	2021-02-19 15:57:21
STOPPED	updated by node01	2021-02-19 15:57:12
STOPPED	updated by node01	2021-02-19 15:57:12
READY	updated by node01	2021-02-16 21:16:31
STARTING	updated by node01	2021-02-16 21:16:12

- Start worker
- Finish jobs & restart worker
- Restart worker now
- Go to worker logs
- Show worker command line
- Suspend server core
- Force suspend server core
- Show dismissed warnings
- Show details
- Stop Refreshing

## 3. CLUSTER

- Consist of **several instances** of **CloverDX Servers** running parallel on different **nodes**.
- Provides the functionality of **CloverDX Server** plus **High Availability and Scalability** resulting from distributed execution environment.

# DRAWBACKS

- License quite pricy per head
- Need Designer and server instance

- References

- <https://www.cloverdx.com/>
- <https://doc.cloverdx.com/latest/designer/cloverdx-server.html>

Q & A