



## ETL PERFORMANCE METRICS

SCOTT HEFFRON

---

---

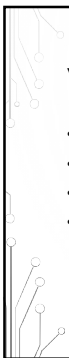
---

---

---


---

---



## WHO AM I

- Scott Heffron
- Title: Application Systems Analyst II
- Company: Verisk Health
- Have been working with data integration projects 25 years



---

---


---

---

---


---

---



## TODAY'S TOPIC

In architecting a data integration solution, the design decisions not only determine how successfully the solution meets functional requirements, but also how well the solution meets performance requirements. It will also show how to integrate it into legacy systems.



---

---

---

---

---

---

---

**AGENDA**

- Understanding Metrics
- Overview of System
- Information to collect
- Database Design
- Setting up Metrics
- Reports
- Case Studies
- Enhancements
- Questions and Answers

---

---

---

---

---

---

---

---

**UNDERSTANDING METRICS**

---

---

---

---

---

---

---

---

**REASON'S TO USE METRICS**

- Problem Analysis/Troubleshooting
  - When something is not how it should be.
- Process Improvement
  - Almost every system, process, or program will need adjustments at sometime. Metrics can show us the weak link
- Goal Setting
  - Specific goal setting processes, key performance indicators, milestones and validation

---

---

---

---

---

---

---

---

**REASON'S TO USE METRICS – CONTINUE**

- Trend Development
  - Understanding of processes over time
- Budgeting
  - Metrics can show improvements have been made. Can justify more work on the project.
- Validation

---

---

---

---

---

---

---

**CRITERIA OF A METRIC**

• Important / Significant	• Precise
• Descriptive of Process	• Timely
• Controllable	• Accessible
• Efficient	• Credible
• Understandable	

---

---

---

---

---

---

---

**DESCRIPTION OF CURRENT SYSTEM**

---

---

---

---

---

---

---

### PURPOSE OF CURRENT SYSTEM

Analyze provider charging patterns inside a carrier

---

---

---

---

---

---

---

---

### HISTORY OF THE INTEGRATION PROCESS

- Grab File
- Load Data from File(s) to Staging Table(s)
- Archive File
- Process Data into Appropriate Areas
- Process Rules and Analyze Data

---

---

---

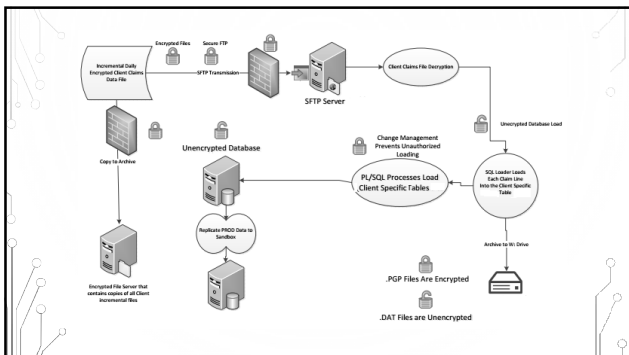
---

---

---

---

---



---

---

---

---

---


---

---

---

**STEPS IN PROCESS**

- 122 Steps in process (5 Levels)
  - 3 more levels deeper
- Combination
  - VBS script
  - SQLPlus/SQL Loader
  - PL/SQL
  - 3<sup>rd</sup> Party products



---

---

---

---

---

---

---

---

**DESCRIPTION OF NEW PROJECT**

---

---

---

---

---

---

---

---

**PURPOSE OF PROJECT**

Ability to analyze provider charging patterns across multiple carriers

---

---

---

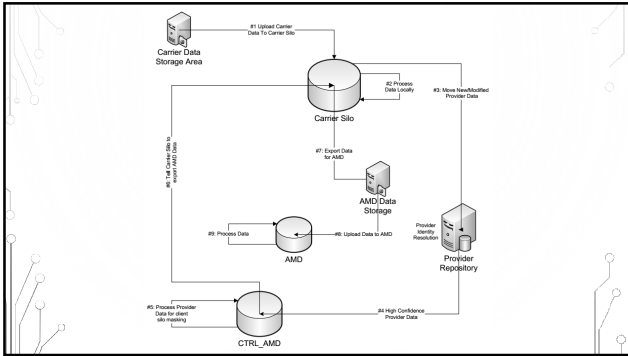
---

---

---

---

---



---

---

---

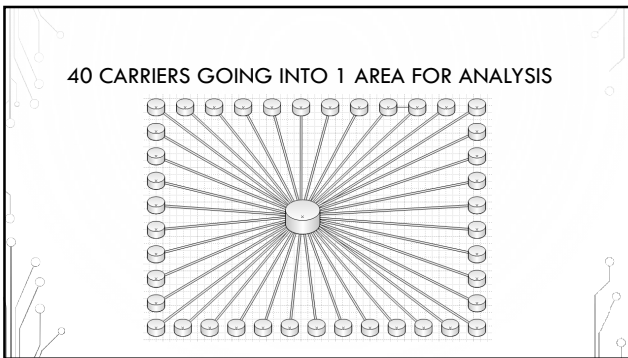
---

---

---

---

---



---

---

---

---

---

---

---

---

**OPERATIONAL QUESTION(S)**

- How long does each unit of work take
- How many records are processed within that unit of work
- How big of impact is there on the system over time
- How big is the impact as more carriers are added
- Which unit of work is the current process on right now
- How does each unit of work perform over time

---

---

---

---

---

---

---

---

### INFORMATION NEEDED TO COLLECT

- Schema Name
- Process Name (Unit of Work)
- Record Count
- Start Time
- End Time

---

---

---

---

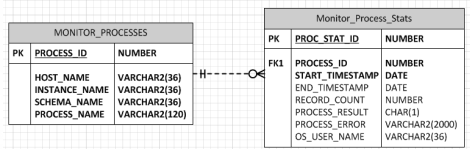
---

---

---

---

### DATABASE DESIGN



---

---

---

---

---

---

---

---

### SETTING UP TO USE METRICS

---

---

---

---

---

---

---

---

### SETTING UP PERMISSIONS

```
clear screen
set serveroutput on;

-- =====
-- Permissions for DBA_Monitoring Area Needed (Begin)
-- =====
ALTER SESSION SET CURRENT_SCHEMA=DBA_MONITOR;
GRANT EXECUTE on MONITOR_STATS to <Client Code>;
GRANT EXECUTE on MONITOR_STATS to <Client Code>_Role;
-- =====
-- Permissions for DBA_Monitoring Area Needed (End)
-- =====
```

---

---

---

---

---

---

---

---

### SYSTEM\_SETTINGS TABLE

SETTING_NAME	SYSTEM_VALUE	SUB_SYSTEM	NOTES
Run_DBA_Monitor_Stats	T	Monitor Process Stats	Setting value of 'T' will allow DBA_MONITOR.SUBMIT...

System\_Settings\_CFG\_Historical\_Export.sql

---

---

---

---

---

---

---

---

### CODE EXAMPLE

(METRICS\_CODE.SQL)

---

---

---

---

---

---

---

---



CAPTURE ROWS TOUCHED

```
n_Metric_1_RC := n_Metric_1_RC + SQL%RowCount;
```

Note: Do before "COMMIT"

---

---

---

---

---

---

---

---

ACTUAL CODE EXAMPLE

(P\_PROCESS\_BATCH.SQL)

---

---

---

---

---

---

---

---

REPORTS

- General Report
- Client Step Specific
- Client Silo Job Status
- File Duration
- Step Trend Analysis
- AMD Client Silo Status
- Client Silo File Processing
- Client Silo Not Proc Records

---

---

---

---

---

---

---

---

### GENERAL PERFORMANCE REPORT

- Used for troubleshooting
- Used to see what current step the process is on for the target schema
- Validation of steps being completed

Perf\_Metrics\_General\_Report.sql

---

---

---

---

---

---

---

---

### GENERAL REPORT

RECID	...	PROCESS_NAME	RECORD_COUNT	START_TIME	END_TIME	DURATION_MN
3079430	AMD	P_Pepsi: P_GetMainType	(null)	Jul-29 15:53:25	(null)	(null)
3079429	AMD	P_Pepsi: P_GetStatistics		0 Jul-29 15:53:24	Jul-29 15:53:26	0
3079428	AMD	P_Pepsi: P_GetFlag		0 Jul-29 15:53:25	Jul-29 15:53:26	0
3079427	AMD	P_Pepsi: P_SetBarrierFlag_N		0 Jul-29 15:53:22	Jul-29 15:53:25	0
3079426	AMD	P_prep: Pep Code 015		0 Jul-29 15:53:22	Jul-29 15:53:22	0
3079425	AMD	P_prep: Pep Code 197		0 Jul-29 15:53:22	Jul-29 15:53:22	0
3079424	AMD	P_prep: Pep Code 104		0 Jul-29 15:53:21	Jul-29 15:53:22	0
3079423	AMD	P_prep: Pep Code 102		0 Jul-29 15:53:21	Jul-29 15:53:21	0
3079422	AMD	P_prep: Pep Code 074		0 Jul-29 15:53:21	Jul-29 15:53:21	0
3079421	AMD	P_prep: Pep Code 073		0 Jul-29 15:53:21	Jul-29 15:53:21	0
3079420	AMD	P_prep: Pep Code 072		0 Jul-29 15:53:21	Jul-29 15:53:21	0

Current Time: 4:02 PM / 16:02

---

---

---

---

---

---

---

---

### CLIENT SPECIFIC STEP REPORT

- Used for troubleshooting
- Trend Development
- Validation of steps being completed

Perf\_Metrics\_Specific\_Client\_and\_Task.sql

---

---

---

---

---

---

---

---

### CLIENT STEP SPECIFIC

Client #1

PROCESS_NAME	RECORD_COUNT	START_TIME	END_TIME	DURATION_MIN
P_Pepsi: P_Sec1statype	50004	07/29/2013 06:42:41 PM	07/29/2013 06:27:17 PM	106
P_Pepsi: P_Sec1statype	50003	07/29/2013 06:57:47 PM	07/29/2013 09:12:14 PM	134
P_Pepsi: P_Sec1statype	20770	07/29/2013 09:04:39 PM	07/29/2013 11:07:03 PM	92

Client #2

P_Pepsi: P_Sec1statype	32976	07/04/2013 09:18:44 PM	07/04/2013 09:18:55 PM	0
P_Pepsi: P_Sec1statype	27976	07/11/2013 09:20:42 PM	07/11/2013 09:20:52 PM	0
P_Pepsi: P_Sec1statype	32299	07/18/2013 09:29:10 PM	07/18/2013 09:29:21 PM	0
P_Pepsi: P_Sec1statype	33621	07/25/2013 09:22:30 PM	07/25/2013 09:22:42 PM	0

---

---

---

---

---

---

---

---

SHOW  
WEEKLY STATUS REPORTS

---

---

---

---

---

---

---

---

- ### METRIC CASE STUDIES
- Finding problem areas
    - Process stops, we receive an error email that says the highest level procedure failed.
    - Implementation – check to see which step is slow
  - Operational Issues
    - One month after project went live, Database started having issues.
    - We can find where processes are increasing in time to finish
    - Can see that a procedure was running twice
  - Validation
    - Shows process is working as requested

---

---

---

---

---

---

---

---

**HOW DO WE KNOW IT WORKS**

- Finding problem areas
  - Troubleshoot where the process is running right now and how much time has been running
  - See processes over time.
- Someone else captures metrics
  - A manager noticed increased timings in a certain area. Asked another developer to add metrics to see what was happening. They went to the wiki and used the documentation.
- Going to general areas

---

---

---

---

---

---

---

---

**ENHANCEMENTS**

- Client Code
  - Found that one schema can contain multiple client's
- Error Information
  - Want error that is associated with unit of work, not just logged.
  - Catching
- Set Metric Processing Level
- Parent Calling Processes
  - Better Analysis
  - Better Reports

---

---

---

---

---

---

---

---

**QUESTIONS AND ANSWERS**

---

---

---

---

---

---

---

---

**CONTACT INFORMATION**

- Scott Heffron
- Email: [Sheffron@VeriskHealth.com](mailto:Sheffron@VeriskHealth.com) or [Scott.Heffron@CTR-SQL.com](mailto:Scott.Heffron@CTR-SQL.com)
- Linked In: <http://www.linkedin.com/in/ScottPHeffron>
- Twitter: @CTR\_SQL

---

---

---

---

---

---

---