Workflow – Load network event

Document describe idea of how to load of files comes from firewall in to DHW.

firewall-blocked-yyyy-mm-dd.csv

Stage Load

SSIS Package

Data Cleansing

Stage Tables



DWH Load

SSIS Package SQL Task

DWH Tables



Presentation

Reporting

1. Stage files path

Manual download from S3 source - >

N:\Source\_csv\firewall-blocked-yyyy-mm-dd.csv

Files are in CSV format and will be loaded using SSIS functionality.

Data from stage will load in fact table FAC\_firewall\_blocked

1. Source file format –

The data is in CSV, where each record has the following fields:

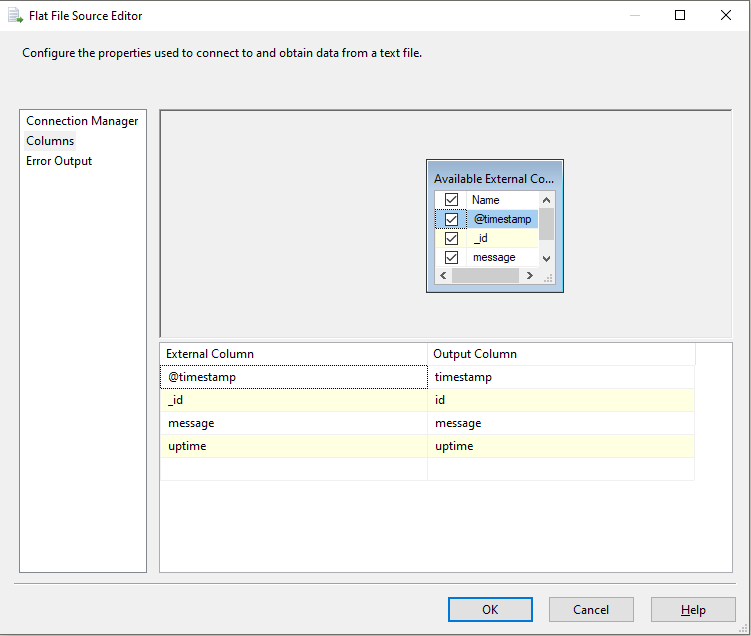
- timestamp: UTC timestamp of the log message

- \_id: unique id of the log message in the system that tracks log messages

- message: the raw log message from the firewall service

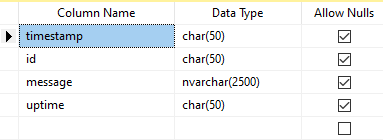
- uptime: the uptime of the firewall service

SSIS source description:



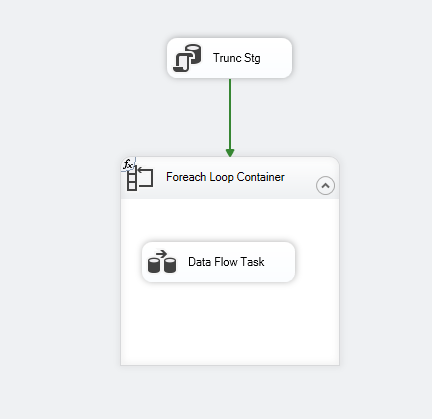
1. Target - Stage table

firewall\_blocked

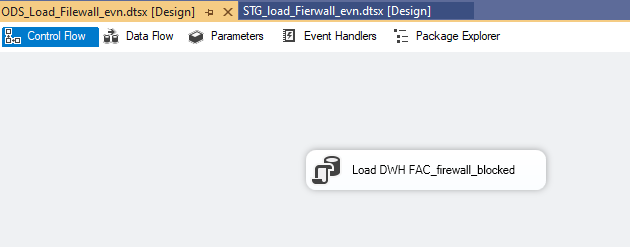


1. Files load
   1. First step truncate stage table - firewall\_blocked

Loading utilise SSIS, For Each Container

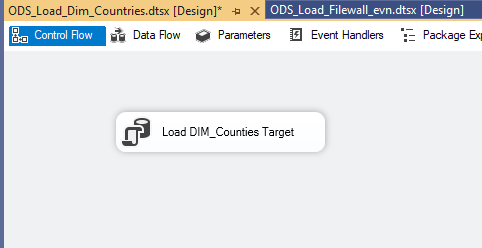


1. DWH transformation
   1. Table - FAC\_firewall\_blocked



Logic of data transformation is in the script component

|  |  |  |
| --- | --- | --- |
| Input filed | Transformation | Target Fact |
| Id | - | id |
| timestamp | - | timestamp\_org |
| timestamp | CAST( replace([timestamp],',', '') as datetime) as TimeStamp , CAST( Left([timestamp] ,10) as date) | TimeStamp |
| timestamp | CAST( replace([timestamp],',', '') as datetime) as TimeStamp , CAST( Left([timestamp] ,10) as date) | TheDate |
| message | - | message |
| message | RTRIM( SUBSTRING (message , CHARINDEX('SRC=', message )+4, ( CHARINDEX('DST', message ) - CHARINDEX('SRC=', message )-4 ) )) | message\_ip |
| message | REPLACE( RTRIM( SUBSTRING (message , CHARINDEX('SRC=', message )+4, ( CHARINDEX('DST', message ) - CHARINDEX('SRC=', message )-4 ) )) , '.', '') | ip\_number |
| message | RTRIM( SUBSTRING (message , CHARINDEX('MAC=', message )+4, ( CHARINDEX('SRC=', message ) - CHARINDEX('MAC=', message )-4 ) )) | MAC |
| message | RTRIM( SUBSTRING (message , CHARINDEX('PROTO=', message )+6,  iif( ( CHARINDEX('SPT=', message ) - CHARINDEX('PROTO=', message )-6 )<0, CHARINDEX('SPT=', message ), ( CHARINDEX('SPT=', message ) - CHARINDEX('PROTO=', message )-6 ) )  )  ) as protocol | protocol |
| message | RTRIM( SUBSTRING (message , CHARINDEX('DPT=', message ) +4 ,  iif( ( CHARINDEX('WINDOW=', message ) - CHARINDEX('DPT=', message )-4 )< 0 , CHARINDEX('DPT=', message ) +4 , ( CHARINDEX('WINDOW=', message ) - CHARINDEX('DPT=', message )-4 )))) | DPT |
| message | - | CC |
| - | getDate() | ETL\_Insert |

* 1. Table - DIM\_Counties 

|  |  |  |
| --- | --- | --- |
| Input filed | Transformation | Target Fact |
| cc | - | Cc |
| country |  | country |