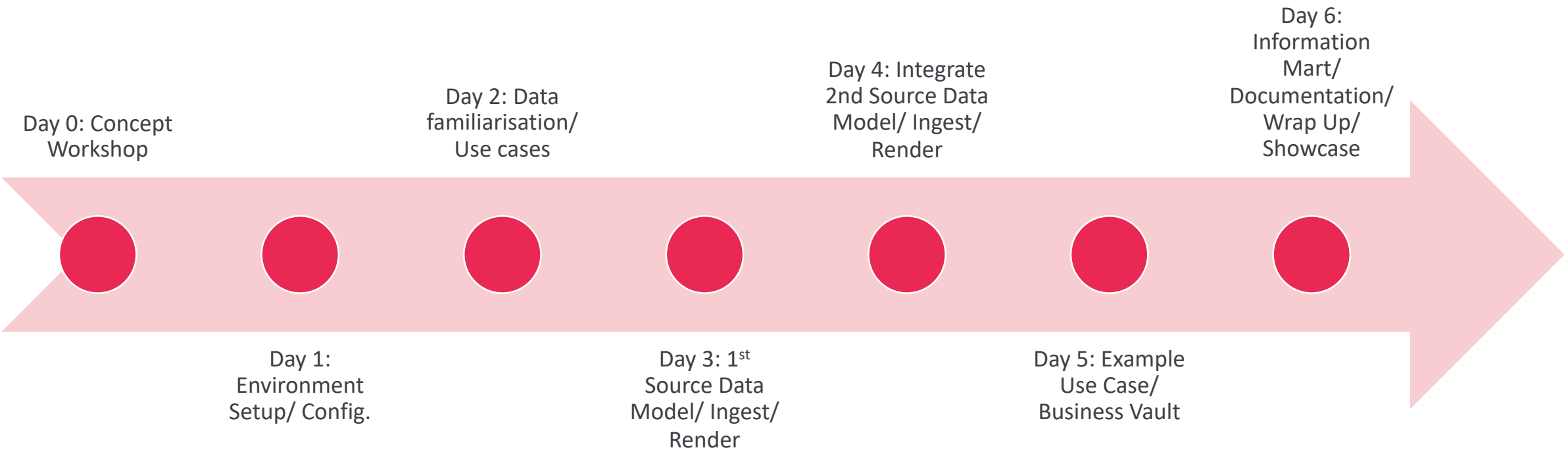
DATA VAULT 2.0 PROOF OF CONCEPT REPORT

Using IBM Information Server & Certus Assets

Originally completed for a south east Asian Central Bank.

CERTUS INSIGHT

PoC Timeline



CERTUS INSIGHT

INFORMATION MART & CONSUMPTION

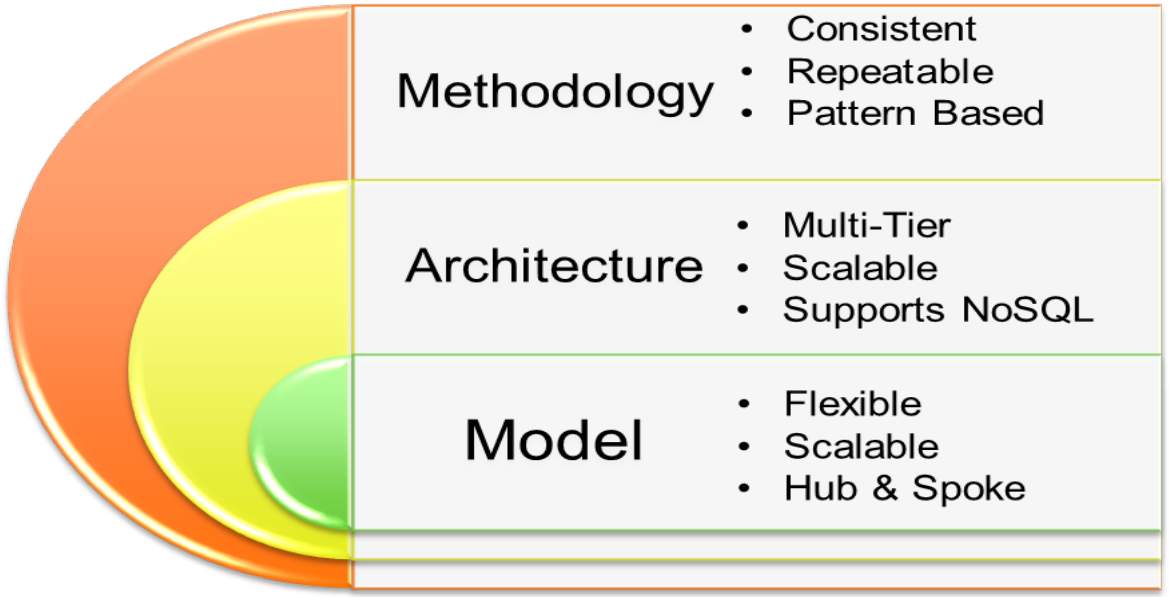
Data Vault 2.0 Definition:

*A **System of Information Management** containing the necessary components needed to accomplish **enterprise vision** in Data Warehousing and Information Delivery.*

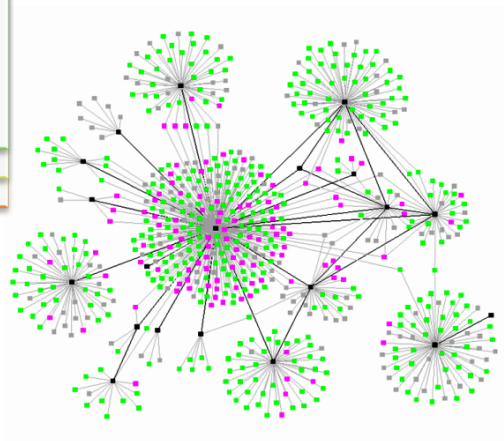
CERTUS INSIGHT

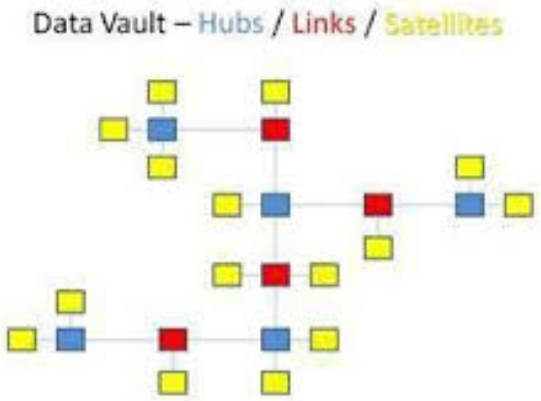
Data Vault 2.0:

… A system of Business Intelligence comprised of…



Hubs: Unique list of business keys
Links: Unique list of Relationships between keys
Satellites: Time Series of data records





CERTUS INSIGHT

INGESTION & INTEGRATION

Initial source into new Raw Vault model

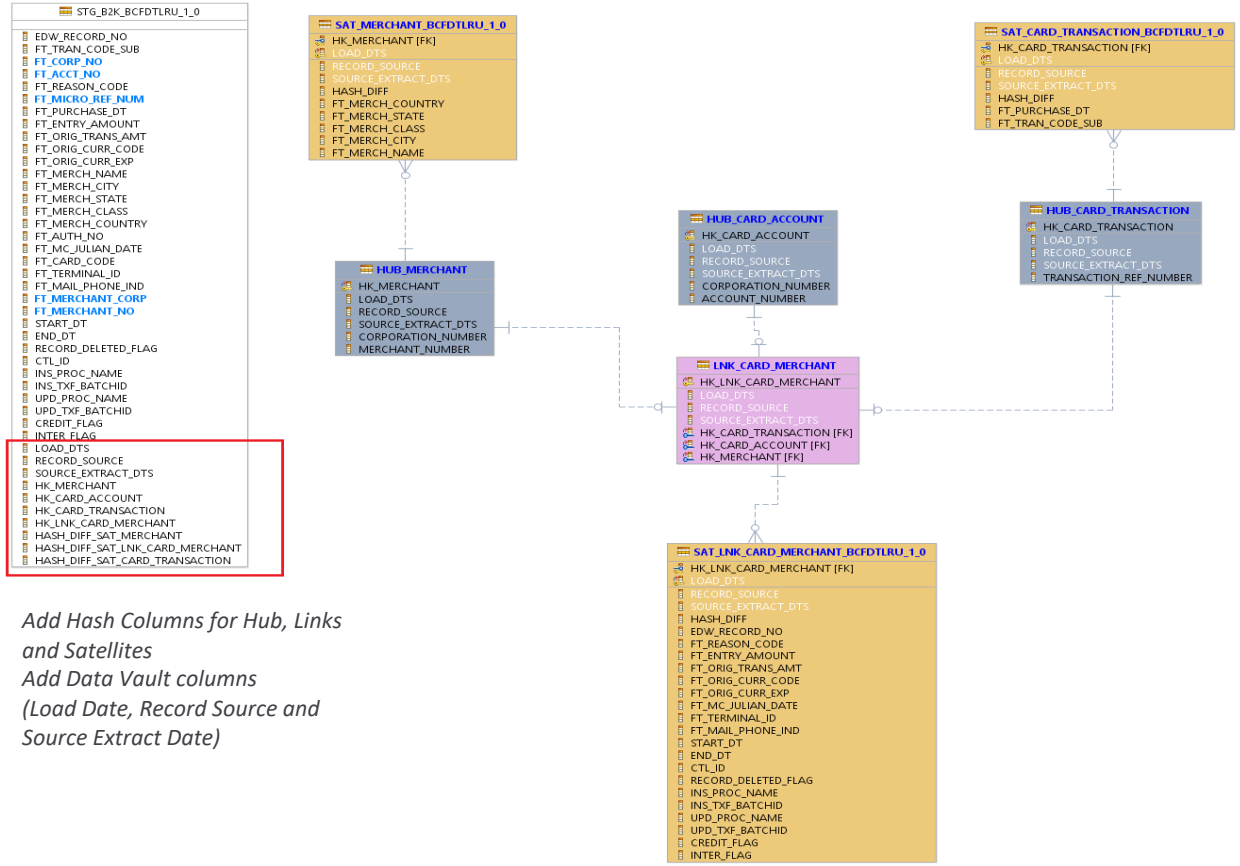
- One source table - native format from source
- Credit Card data
- Small volume

- Objective - source into Raw Vault
- Profile the data to identify the Business Keys (blue)
- Hubs & Links thus identified
- No Business Rules at this stage
- All column names as found
- No transformations at this stage
- As part of staging, we hash as we go
- Modeller not concerned with nature/ function of the data
- Not concerned with quality of data at this point

- Relationships become apparent due to the way the Raw Vault is laid out/ built. This helps downstream for consumption, even though this is Raw Vault

- Based on defined standards e.g.

- Naming conventions for all DV objects
- Hashing process and logic, e.g. Using CHAR and Casing



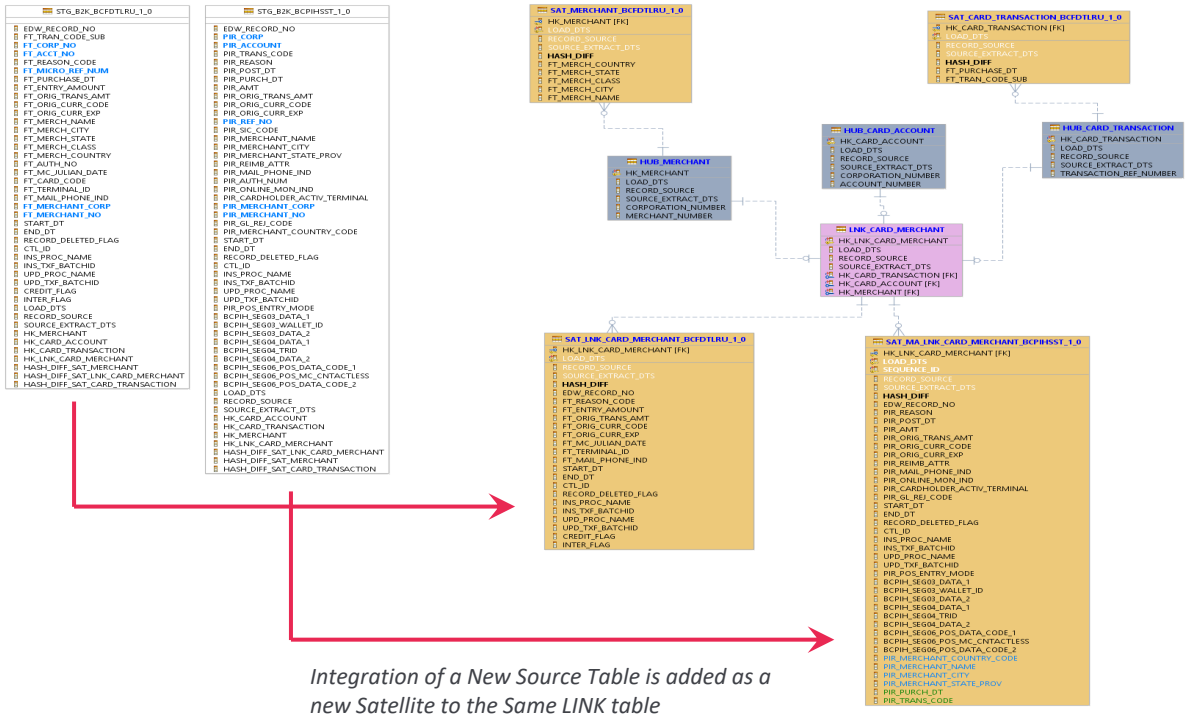
Add Hash Columns for Hub, Links and Satellites
Add Data Vault columns (Load Date, Record Source and Source Extract Date)

CERTUS INSIGHT

Integration of new source into existing Raw Vault

- Identify same business keys, same ‘unit of work’
- Profiling discovered that the data contains a different construct: Multi-Active data from multiple transactions for one unit of work on the same day
- Straight repeat of prev. source table, BUT diff construct dealt with via the separate satellite, which would have been added anyway
- Thus we have demonstrated easy and rapid integration of new source
- Process can be agile, with different sources being done in different sprints
- Or can be done in parallel with different teams
- Avoiding

- Regression Testing for shared structures
- Reengineering
- Parallel loading collisions
- Delays that might have come from any refactoring/ required



Integration of a New Source Table is added as a new Satellite to the Same LINK table

CERTUS INSIGHT

MAPPING &
CODE GENERATION

FastTrack Mapping

- Two levels of mapping
- Landing to Staging
- Then Staging to Raw Vault
- Every source follows these two level mappings
- Mapping in Fasttrack can be generated from the modelling exercise done for the RV load
- Two mappings for each source table
- Metadata driven, facilitated by the naming standards
- Consistency of approach aids reconciliation
- Fasttrack Discover utility allows rapid mapping, aided by the naming
- Hand coding would be significant effort

Mapping Assets

Name	Target	Created	Modified	Created By	Status	Percent Reviewed
DSDV_Demo						
Mapping Specifications						
MapSpec_B2K_BCFDTLRU_1_0_Staging	STG_B2K_BCF	6/16/20 11:47	6/19/20 12:08	isadmin	In Progress	0 %
MapSpec_B2K_BCFDTLRU_1_0_RawVault	LNK_CARD_ME	6/17/20 2:54	6/17/20 2:56	isadmin	In Progress	0 %
MapSpec_B2K_BCPIHSST_1_0_Staging	STG_B2K_BCPI	6/17/20 6:57	6/19/20 12:09	isadmin	In Progress	0 %
MapSpec_B2K_BCPIHSST_1_0_RawVault	LNK_CARD_ME	6/17/20 7:37	6/18/20 6:14	isadmin	In Progress	0 %
Mapping Components						
Mapping Compositions						

CERTUS INSIGHT

Landing to Staging FastTrack Mapping

- Tagging identifies where the hashing needs to happen

	ID	Source Field	Target Field	Business Term	Transformation Rule Description
15	15	LND_B2K_BCFDTLRU_1_0.FT_MERCH_CLASS	STG_B2K_BCFDTLRU_1_0.FT_MERCH_CLASS		
16	16	LND_B2K_BCFDTLRU_1_0.FT_MERCH_COUNTRY	STG_B2K_BCFDTLRU_1_0.FT_MERCH_COUNTRY		
17	17	LND_B2K_BCFDTLRU_1_0.FT_AUTH_NO	STG_B2K_BCFDTLRU_1_0.FT_AUTH_NO		
18	18	LND_B2K_BCFDTLRU_1_0.FT_MC_JULIAN_DATE	STG_B2K_BCFDTLRU_1_0.FT_MC_JULIAN_DATE		
19	19	LND_B2K_BCFDTLRU_1_0.FT_CARD_CODE	STG_B2K_BCFDTLRU_1_0.FT_CARD_CODE		
20	20	LND_B2K_BCFDTLRU_1_0.FT_TERMINAL_ID	STG_B2K_BCFDTLRU_1_0.FT_TERMINAL_ID		
21	21	LND_B2K_BCFDTLRU_1_0.FT_MAIL_PHONE_IND	STG_B2K_BCFDTLRU_1_0.FT_MAIL_PHONE_IND		
22	22	LND_B2K_BCFDTLRU_1_0.FT_MERCHANT_CORP	STG_B2K_BCFDTLRU_1_0.FT_MERCHANT_CORP		
23	23	LND_B2K_BCFDTLRU_1_0.FT_MERCHANT_NO	STG_B2K_BCFDTLRU_1_0.FT_MERCHANT_NO		
24	24	LND_B2K_BCFDTLRU_1_0.START_DT	STG_B2K_BCFDTLRU_1_0.START_DT		
25	25	LND_B2K_BCFDTLRU_1_0.END_DT	STG_B2K_BCFDTLRU_1_0.END_DT		
26	26	LND_B2K_BCFDTLRU_1_0.RECORD_DELETED_FLAG	STG_B2K_BCFDTLRU_1_0.RECORD_DELETED_FLAG		
27	27	LND_B2K_BCFDTLRU_1_0.CTL_ID	STG_B2K_BCFDTLRU_1_0.CTL_ID		
28	28	LND_B2K_BCFDTLRU_1_0.INS_PROC_NAME	STG_B2K_BCFDTLRU_1_0.INS_PROC_NAME		
29	29	LND_B2K_BCFDTLRU_1_0.INS_TXF_BATCHID	STG_B2K_BCFDTLRU_1_0.INS_TXF_BATCHID		
30	30	LND_B2K_BCFDTLRU_1_0.UPD_PROC_NAME	STG_B2K_BCFDTLRU_1_0.UPD_PROC_NAME		
31	31	LND_B2K_BCFDTLRU_1_0.UPD_TXF_BATCHID	STG_B2K_BCFDTLRU_1_0.UPD_TXF_BATCHID		
32	32	LND_B2K_BCFDTLRU_1_0.CREDIT_FLAG	STG_B2K_BCFDTLRU_1_0.CREDIT_FLAG		
33	33	LND_B2K_BCFDTLRU_1_0.INTER_FLAG	STG_B2K_BCFDTLRU_1_0.INTER_FLAG		
34	34	LND_B2K_BCFDTLRU_1_0.LOAD_DTS	STG_B2K_BCFDTLRU_1_0.LOAD_DTS		
35	35	LND_B2K_BCFDTLRU_1_0.RECORD_SOURCE	STG_B2K_BCFDTLRU_1_0.RECORD_SOURCE		
36	36	LND_B2K_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS	STG_B2K_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS		
37	37	LND_B2K_BCFDTLRU_1_0.FT_CORP_NO, LND_B2K_BCFDTLRU_1_0.FT_ACCT_NO	STG_B2K_BCFDTLRU_1_0.HK_CARD_ACCOUNT		HubHash
38	38	LND_B2K_BCFDTLRU_1_0.FT_MICRO_REF_NUM	STG_B2K_BCFDTLRU_1_0.HK_CARD_TRANSACTION		HubHash
39	39	LND_B2K_BCFDTLRU_1_0.FT_MICRO_REF_NUM, LND_B2K_BCFDTLRU_1_0.FT_CORP_N	STG_B2K_BCFDTLRU_1_0.HK_LNK_CARD_MERCHANT		LinkHash
40	40	LND_B2K_BCFDTLRU_1_0.FT_MERCHANT_CORP, LND_B2K_BCFDTLRU_1_0.FT_MERCHA	STG_B2K_BCFDTLRU_1_0.HK_MERCHANT		HubHash
41	41	LND_B2K_BCFDTLRU_1_0.EDW_RECORD_NO, LND_B2K_BCFDTLRU_1_0.FT_REASON_C	STG_B2K_BCFDTLRU_1_0.HASH_DIFF_SAT_LNK_CARD_MERCHANT		HashDiff
42	42	LND_B2K_BCFDTLRU_1_0.FT_MERCH_COUNTRY, LND_B2K_BCFDTLRU_1_0.FT_MERCH_	STG_B2K_BCFDTLRU_1_0.HASH_DIFF_SAT_MERCHANT		HashDiff
43	43	LND_B2K_BCFDTLRU_1_0.FT_PURCHASE_DT, LND_B2K_BCFDTLRU_1_0.FT_TRAN_CODE	STG_B2K_BCFDTLRU_1_0.HASH_DIFF_SAT_CARD_TRANSACTION		HashDiff

CERTUS INSIGHT

Data Vault Mapping from Staging to Raw Vault

- Each standing component has its dependent objects mapped
- Very low error rate due to metadata driven nature

	ID	Source Field	Target Field
10	52	STG_B2K_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS	HUB_CARD_TRANSACTION.SOURCE_EXTRACT_DTS
11	53	STG_B2K_BCFDTLRU_1_0.FT_MICRO_REF_NUM	HUB_CARD_TRANSACTION.TRANSACTION_REF_NUMBER
12	54	STG_B2K_BCFDTLRU_1_0.HK_MERCHANT	HUB_MERCHANT.HK_MERCHANT
13	55	STG_B2K_BCFDTLRU_1_0.LOAD_DTS	HUB_MERCHANT.LOAD_DTS
14	56	STG_B2K_BCFDTLRU_1_0.RECORD_SOURCE	HUB_MERCHANT.RECORD_SOURCE
15	57	STG_B2K_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS	HUB_MERCHANT.SOURCE_EXTRACT_DTS
16	58	STG_B2K_BCFDTLRU_1_0.FT_MERCHANT_CORP	HUB_MERCHANT.CORPORATION_NUMBER
17	59	STG_B2K_BCFDTLRU_1_0.FT_MERCHANT_NO	HUB_MERCHANT.MERCHANT_NUMBER
18	60	STG_B2K_BCFDTLRU_1_0.HK_LNK_CARD_MERCHANT	LNK_CARD_MERCHANT.HK_LNK_CARD_MERCHANT
19	61	STG_B2K_BCFDTLRU_1_0.LOAD_DTS	LNK_CARD_MERCHANT.LOAD_DTS
20	62	STG_B2K_BCFDTLRU_1_0.RECORD_SOURCE	LNK_CARD_MERCHANT.RECORD_SOURCE
21	63	STG_B2K_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS	LNK_CARD_MERCHANT.SOURCE_EXTRACT_DTS
22	64	STG_B2K_BCFDTLRU_1_0.HK_CARD_TRANSACTION	LNK_CARD_MERCHANT.HK_CARD_TRANSACTION
23	65	STG_B2K_BCFDTLRU_1_0.HK_CARD_ACCOUNT	LNK_CARD_MERCHANT.HK_CARD_ACCOUNT
24	66	STG_B2K_BCFDTLRU_1_0.HK_MERCHANT	LNK_CARD_MERCHANT.HK_MERCHANT
25	67	STG_B2K_BCFDTLRU_1_0.HK_CARD_TRANSACTION	SAT_CARD_TRANSACTION_BCFDTLRU_1_0.HK_CARD_TRANSACTION
26	68	STG_B2K_BCFDTLRU_1_0.LOAD_DTS	SAT_CARD_TRANSACTION_BCFDTLRU_1_0.LOAD_DTS
27	69	STG_B2K_BCFDTLRU_1_0.RECORD_SOURCE	SAT_CARD_TRANSACTION_BCFDTLRU_1_0.RECORD_SOURCE
28	70	STG_B2K_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS	SAT_CARD_TRANSACTION_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS
29	71	STG_B2K_BCFDTLRU_1_0.HASH_DIFF_SAT_CARD_TRANSACTION	SAT_CARD_TRANSACTION_BCFDTLRU_1_0.HASH_DIFF
30	72	STG_B2K_BCFDTLRU_1_0.FT_PURCHASE_DT	SAT_CARD_TRANSACTION_BCFDTLRU_1_0.FT_PURCHASE_DT
31	73	STG_B2K_BCFDTLRU_1_0.FT_TRAN_CODE_SUB	SAT_CARD_TRANSACTION_BCFDTLRU_1_0.FT_TRAN_CODE_SUB
32	74	STG_B2K_BCFDTLRU_1_0.HK_MERCHANT	SAT_MERCHANT_BCFDTLRU_1_0.HK_MERCHANT
33	75	STG_B2K_BCFDTLRU_1_0.LOAD_DTS	SAT_MERCHANT_BCFDTLRU_1_0.LOAD_DTS
34	76	STG_B2K_BCFDTLRU_1_0.RECORD_SOURCE	SAT_MERCHANT_BCFDTLRU_1_0.RECORD_SOURCE
35	77	STG_B2K_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS	SAT_MERCHANT_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS
36	78	STG_B2K_BCFDTLRU_1_0.HASH_DIFF_SAT_MERCHANT	SAT_MERCHANT_BCFDTLRU_1_0.HASH_DIFF
37	79	STG_B2K_BCFDTLRU_1_0.FT_MERCH_CITY	SAT_MERCHANT_BCFDTLRU_1_0.FT_MERCH_CITY
38	80	STG_B2K_BCFDTLRU_1_0.FT_MERCH_CLASS	SAT_MERCHANT_BCFDTLRU_1_0.FT_MERCH_CLASS
39	81	STG_B2K_BCFDTLRU_1_0.FT_MERCH_COUNTRY	SAT_MERCHANT_BCFDTLRU_1_0.FT_MERCH_COUNTRY
40	82	STG_B2K_BCFDTLRU_1_0.FT_MERCH_NAME	SAT_MERCHANT_BCFDTLRU_1_0.FT_MERCH_NAME
41	83	STG_B2K_BCFDTLRU_1_0.FT_MERCH_STATE	SAT_MERCHANT_BCFDTLRU_1_0.FT_MERCH_STATE
42	84	STG_B2K_BCFDTLRU_1_0.HK_LNK_CARD_MERCHANT	SAT_LNK_CARD_MERCHANT_BCFDTLRU_1_0.HK_LNK_CARD_MERCHANT

CERTUS INSIGHT

Data Vault Mapping from Staging to Raw Vault

- Each standing component has its dependent objects mapped
- Very low error rate due to metadata driven nature

	ID	Source Field	Target Field
10	52	STG_B2K_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS	HUB_CARD_TRANSACTION.SOURCE_EXTRACT_DTS
11	53	STG_B2K_BCFDTLRU_1_0.FT_MICRO_REF_NUM	HUB_CARD_TRANSACTION.TRANSACTION_REF_NUMBER
12	54	STG_B2K_BCFDTLRU_1_0.HK_MERCHANT	HUB_MERCHANT.HK_MERCHANT
13	55	STG_B2K_BCFDTLRU_1_0.LOAD_DTS	HUB_MERCHANT.LOAD_DTS
14	56	STG_B2K_BCFDTLRU_1_0.RECORD_SOURCE	HUB_MERCHANT.RECORD_SOURCE
15	57	STG_B2K_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS	HUB_MERCHANT.SOURCE_EXTRACT_DTS
16	58	STG_B2K_BCFDTLRU_1_0.FT_MERCHANT_CORP	HUB_MERCHANT.CORPORATION_NUMBER
17	59	STG_B2K_BCFDTLRU_1_0.FT_MERCHANT_NO	HUB_MERCHANT.MERCHANT_NUMBER
18	60	STG_B2K_BCFDTLRU_1_0.HK_LNK_CARD_MERCHANT	LNK_CARD_MERCHANT.HK_LNK_CARD_MERCHANT
19	61	STG_B2K_BCFDTLRU_1_0.LOAD_DTS	LNK_CARD_MERCHANT.LOAD_DTS
20	62	STG_B2K_BCFDTLRU_1_0.RECORD_SOURCE	LNK_CARD_MERCHANT.RECORD_SOURCE
21	63	STG_B2K_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS	LNK_CARD_MERCHANT.SOURCE_EXTRACT_DTS
22	64	STG_B2K_BCFDTLRU_1_0.HK_CARD_TRANSACTION	LNK_CARD_MERCHANT.HK_CARD_TRANSACTION
23	65	STG_B2K_BCFDTLRU_1_0.HK_CARD_ACCOUNT	LNK_CARD_MERCHANT.HK_CARD_ACCOUNT
24	66	STG_B2K_BCFDTLRU_1_0.HK_MERCHANT	LNK_CARD_MERCHANT.HK_MERCHANT
25	67	STG_B2K_BCFDTLRU_1_0.HK_CARD_TRANSACTION	SAT_CARD_TRANSACTION_BCFDTLRU_1_0.HK_CARD_TRANSACTION
26	68	STG_B2K_BCFDTLRU_1_0.LOAD_DTS	SAT_CARD_TRANSACTION_BCFDTLRU_1_0.LOAD_DTS
27	69	STG_B2K_BCFDTLRU_1_0.RECORD_SOURCE	SAT_CARD_TRANSACTION_BCFDTLRU_1_0.RECORD_SOURCE
28	70	STG_B2K_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS	SAT_CARD_TRANSACTION_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS
29	71	STG_B2K_BCFDTLRU_1_0.HASH_DIFF_SAT_CARD_TRANSACTION	SAT_CARD_TRANSACTION_BCFDTLRU_1_0.HASH_DIFF
30	72	STG_B2K_BCFDTLRU_1_0.FT_PURCHASE_DT	SAT_CARD_TRANSACTION_BCFDTLRU_1_0.FT_PURCHASE_DT
31	73	STG_B2K_BCFDTLRU_1_0.FT_TRAN_CODE_SUB	SAT_CARD_TRANSACTION_BCFDTLRU_1_0.FT_TRAN_CODE_SUB
32	74	STG_B2K_BCFDTLRU_1_0.HK_MERCHANT	SAT_MERCHANT_BCFDTLRU_1_0.HK_MERCHANT
33	75	STG_B2K_BCFDTLRU_1_0.LOAD_DTS	SAT_MERCHANT_BCFDTLRU_1_0.LOAD_DTS
34	76	STG_B2K_BCFDTLRU_1_0.RECORD_SOURCE	SAT_MERCHANT_BCFDTLRU_1_0.RECORD_SOURCE
35	77	STG_B2K_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS	SAT_MERCHANT_BCFDTLRU_1_0.SOURCE_EXTRACT_DTS
36	78	STG_B2K_BCFDTLRU_1_0.HASH_DIFF_SAT_MERCHANT	SAT_MERCHANT_BCFDTLRU_1_0.HASH_DIFF
37	79	STG_B2K_BCFDTLRU_1_0.FT_MERCH_CITY	SAT_MERCHANT_BCFDTLRU_1_0.FT_MERCH_CITY
38	80	STG_B2K_BCFDTLRU_1_0.FT_MERCH_CLASS	SAT_MERCHANT_BCFDTLRU_1_0.FT_MERCH_CLASS
39	81	STG_B2K_BCFDTLRU_1_0.FT_MERCH_COUNTRY	SAT_MERCHANT_BCFDTLRU_1_0.FT_MERCH_COUNTRY
40	82	STG_B2K_BCFDTLRU_1_0.FT_MERCH_NAME	SAT_MERCHANT_BCFDTLRU_1_0.FT_MERCH_NAME
41	83	STG_B2K_BCFDTLRU_1_0.FT_MERCH_STATE	SAT_MERCHANT_BCFDTLRU_1_0.FT_MERCH_STATE
42	84	STG_B2K_BCFDTLRU_1_0.HK_LNK_CARD_MERCHANT	SAT_LNK_CARD_MERCHANT_BCFDTLRU_1_0.HK_LNK_CARD_MERCHANT

CERTUS INSIGHT

Data Vault Mapping from Staging to Raw Vault

- Mapping is then be used to generate the code to load data into the Raw Vault

- The Certus Code automation templates speed this process. Only Modelling and Mapping were required to get the DDL created, and data loaded (IDA was used to create the DDL).

- This is all that is needed to ingest the data
- Repeatability
- Automation
- Trainability

Mapping Assets

Name	Target	Created	Modified	Created By	Status	Percent Reviewed
DSDV_Demo						
Mapping Specifications						
MapSpec_B2K_BCFDTLRU_1_0_Staging	STG_B2K_BCF	6/16/20 11:47	6/19/20 12:08	isadmin	In Progress	0 %
MapSpec_B2K_BCFDTLRU_1_0_RawVault	LNK_CARD_ME	6/17/20 2:54	6/17/20 2:56	isadmin	In Progress	0 %
MapSpec_B2K_BCPIHSST_1_0_Staging	STG_B2K_BCPI	6/17/20 6:57	6/19/20 12:09	isadmin	In Progress	0 %
MapSpec_B2K_BCPIHSST_1_0_RawVault	LNK_CARD_ME	6/17/20 7:37	6/18/20 6:14	isadmin	In Progress	0 %
Mapping Components						
Mapping Compositions						

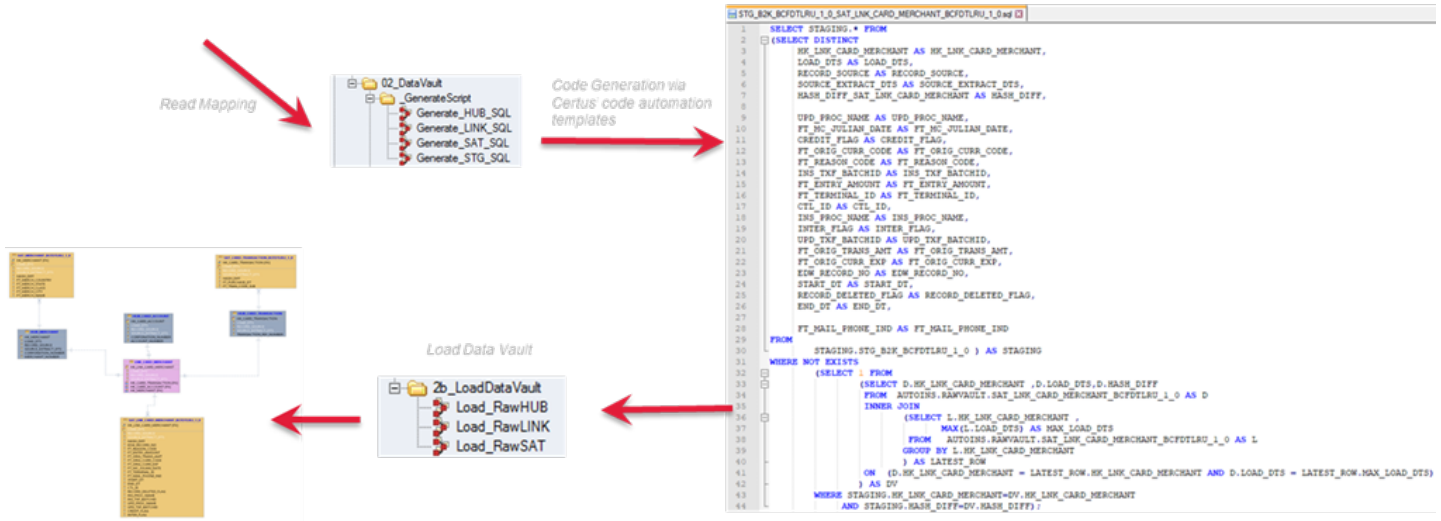


Illustration of the code structure

- The code is similar across the solution end to end
- Machine generated for consistency and reliability
- Hand-coding would introduce significant effort

STG_B2K_BCFDTLRU_1_0_SAT_LNK_CARD_MERCHANT_BCFDTLRU_1_0.sql

```
SELECT STAGING.* FROM
(SELECT DISTINCT
      HK_LNK_CARD_MERCHANT AS HK_LNK_CARD_MERCHANT,
      LOAD_DTS AS LOAD_DTS,
      RECORD_SOURCE AS RECORD_SOURCE,
      SOURCE_EXTRACT_DTS AS SOURCE_EXTRACT_DTS,
      HASH_DIFF_SAT_LNK_CARD_MERCHANT AS HASH_DIFF,

      UPD_PROC_NAME AS UPD_PROC_NAME,
      FT_MC_JULIAN_DATE AS FT_MC_JULIAN_DATE,
      CREDIT_FLAG AS CREDIT_FLAG,
      FT_ORIG_CURR_CODE AS FT_ORIG_CURR_CODE,
      FT_REASON_CODE AS FT_REASON_CODE,
      INS_TXF_BATCHID AS INS_TXF_BATCHID,
      FT_ENTRY_AMOUNT AS FT_ENTRY_AMOUNT,
      FT_TERMINAL_ID AS FT_TERMINAL_ID,
      CTL_ID AS CTL_ID,
      INS_PROC_NAME AS INS_PROC_NAME,
      INTER_FLAG AS INTER_FLAG,
      UPD_TXF_BATCHID AS UPD_TXF_BATCHID,
      FT_ORIG_TRANS_AMT AS FT_ORIG_TRANS_AMT,
      FT_ORIG_CURR_EXP AS FT_ORIG_CURR_EXP,
      EDW_RECORD_NO AS EDW_RECORD_NO,
      START_DT AS START_DT,
      RECORD_DELETED_FLAG AS RECORD_DELETED_FLAG,
      END_DT AS END_DT,

      FT_MAIL_PHONE_IND AS FT_MAIL_PHONE_IND
FROM
         STAGING.STG_B2K_BCFDTLRU_1_0 ) AS STAGING
 WHERE NOT EXISTS
         (SELECT 1 FROM
                (SELECT D.HK_LNK_CARD_MERCHANT ,D.LOAD_DTS,D.HASH_DIFF
                 FROM  AUTOINS.RAWVAULT.SAT_LNK_CARD_MERCHANT_BCFDTLRU_1_0 AS D
                 INNER JOIN
                        (SELECT L.HK_LNK_CARD_MERCHANT ,
                              MAX(L.LOAD_DTS) AS MAX_LOAD_DTS
                         FROM   AUTOINS.RAWVAULT.SAT_LNK_CARD_MERCHANT_BCFDTLRU_1_0 AS L
                        GROUP BY L.HK_LNK_CARD_MERCHANT
                        ) AS LATEST_ROW
                 ON  (D.HK_LNK_CARD_MERCHANT = LATEST_ROW.HK_LNK_CARD_MERCHANT AND D.LOAD_DTS = LATEST_ROW.MAX_LOAD_DTS)
                ) AS DV
         WHERE STAGING.HK_LNK_CARD_MERCHANT=DV.HK_LNK_CARD_MERCHANT
               AND STAGING.HASH_DIFF=DV.HASH_DIFF);
```

Compare Row HashDiff and only load where row has changed

Compare only with Latest Row from RawVault

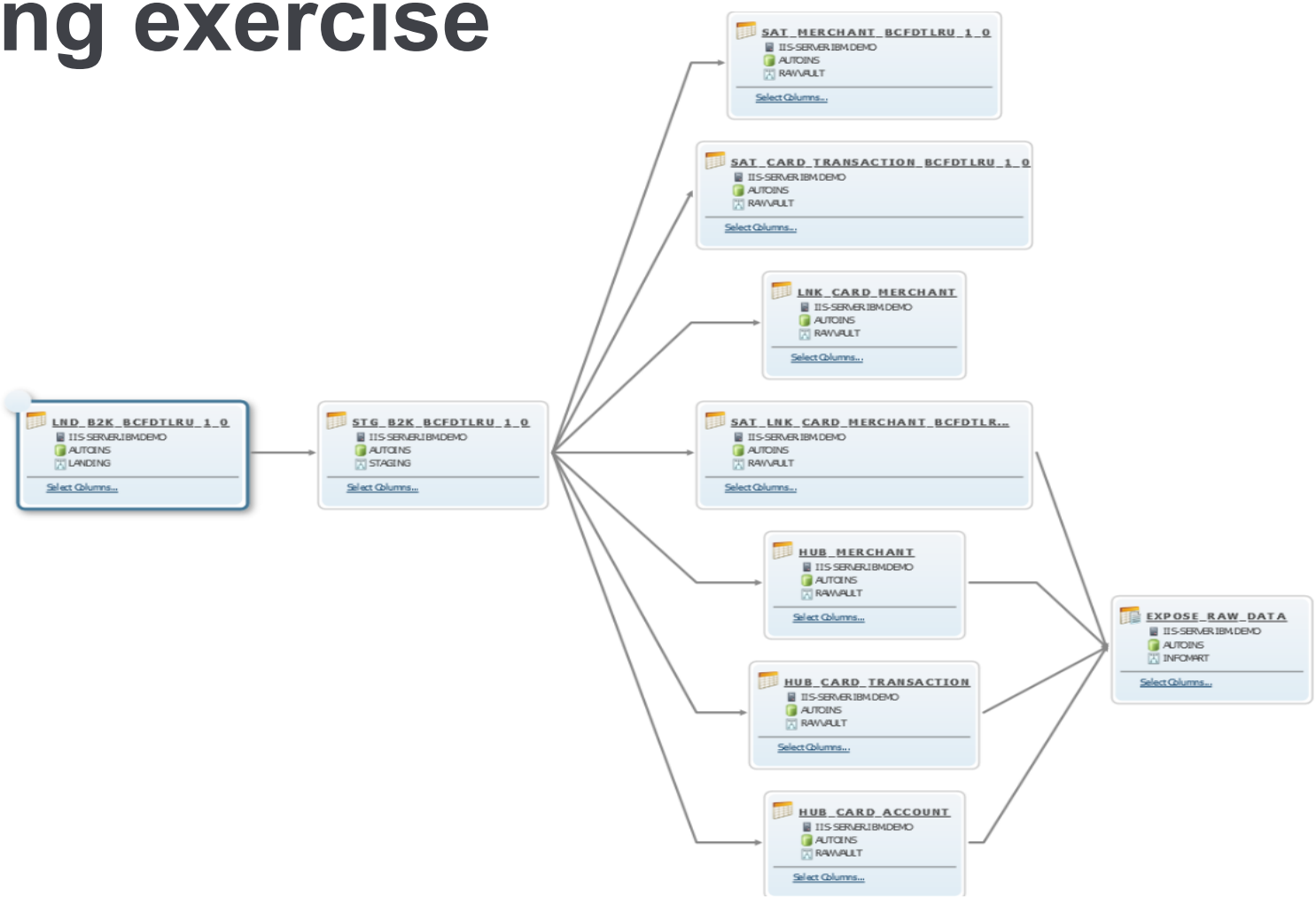
Code Structure is same across Data Vault objects, thus making the process repeatable and easy to understand

CERTUS INSIGHT

LINEAGE

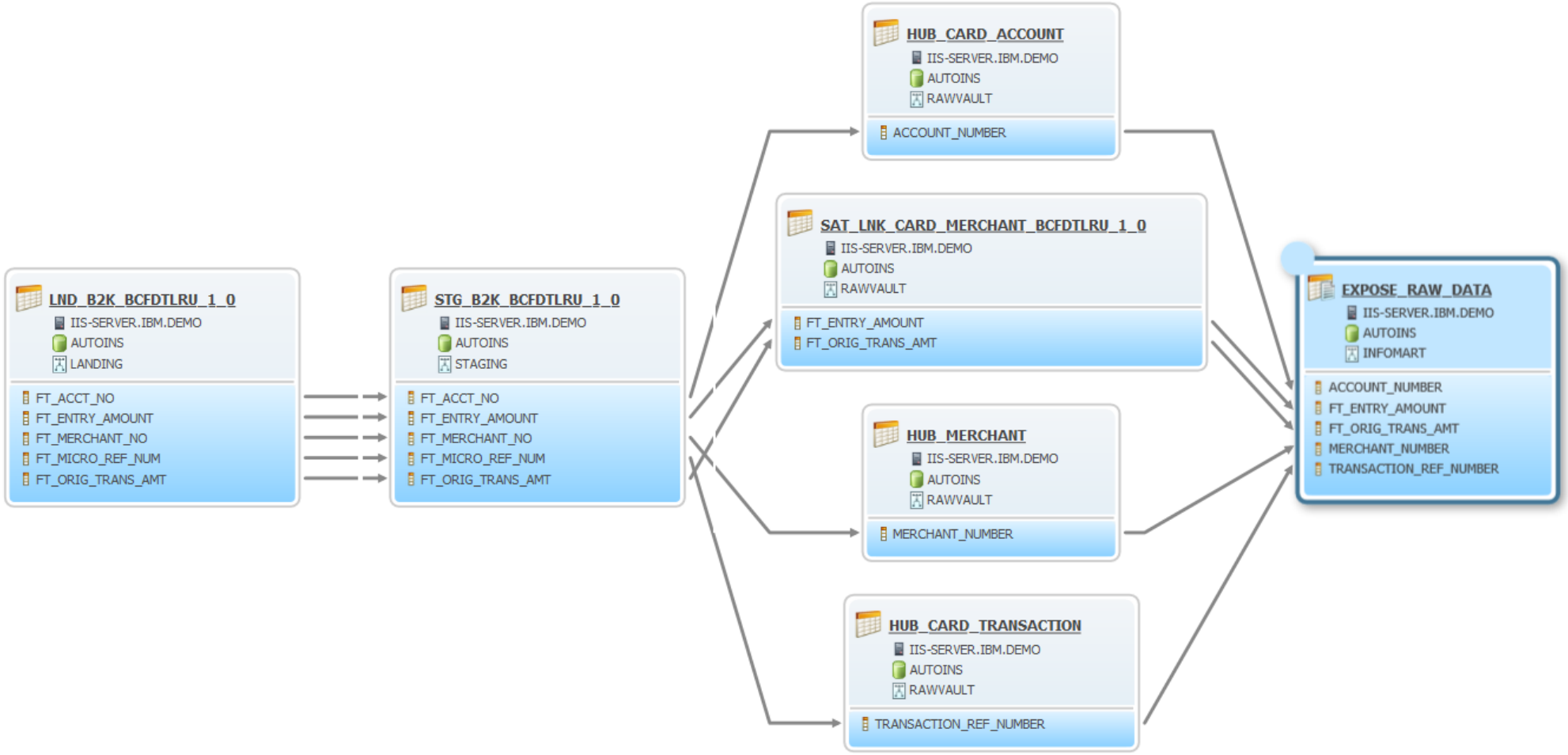
Lineage is automatically generated as part of mapping exercise

- No additional work is required to generate lineage in IGC
- The Data Vault mapping facilitates visibility and auditability as structures are known and consistent.



CERTUS INSIGHT

- Mapping lineage will include individual column level also

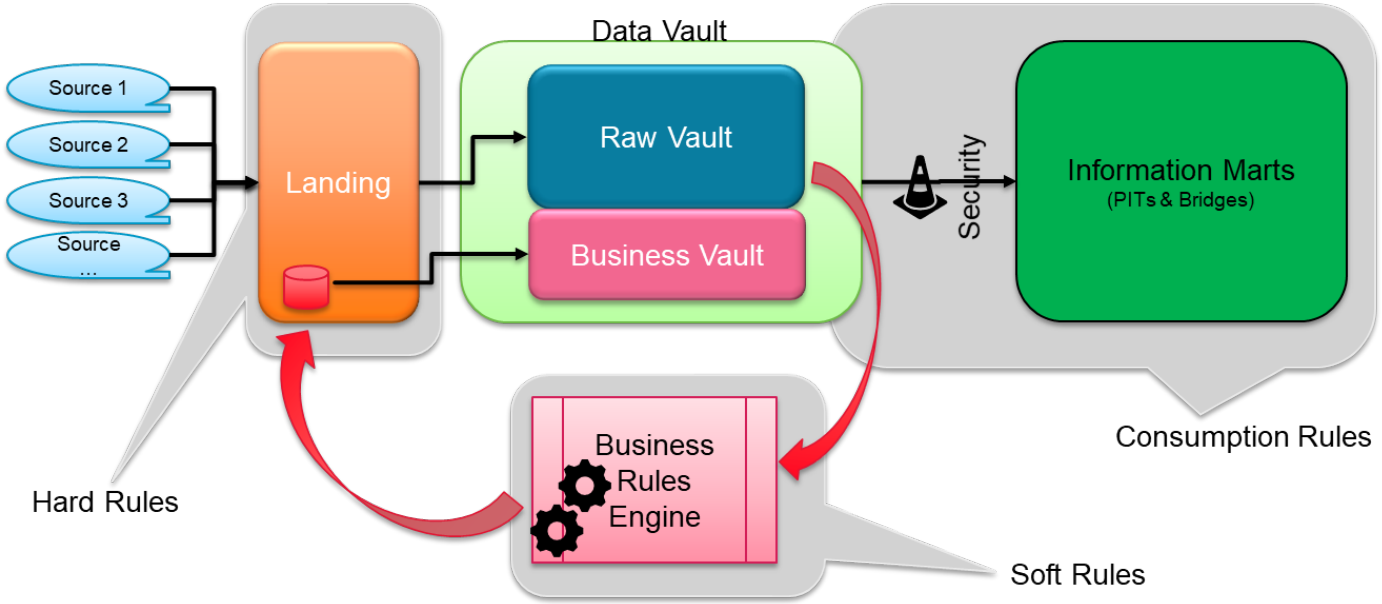


CERTUS INSIGHT

BUSINESS VAULT

Architecture…

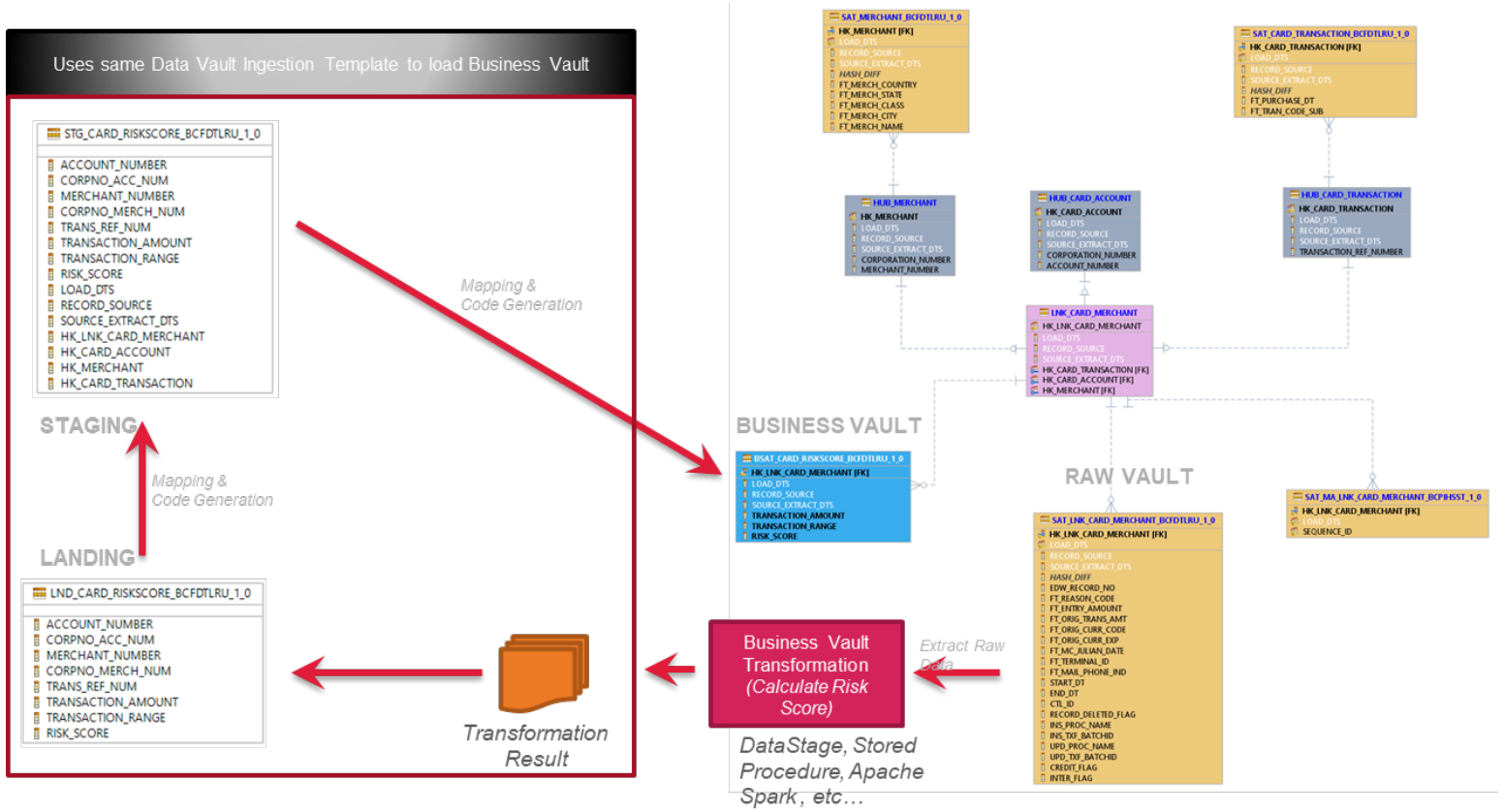
- IN Data Vault 2.0 the outcomes for business rules are physical. It will be the first time ever that the data comes to rest.
- Business Vault data is historized and auditable like all other information item in the Data Vault.
- The use of Business Vault removes all dependency from Information marts
- Information marts become completely disposable.
- First time ever, managed self service BI is possible.
- Business Rule engines are specifically separated (de-coupled) from the Data Vault platform.
- The Data Vault platform can be upgraded without impact to Business Rule engines and vice versa



CERTUS INSIGHT

Business Vault Transformation

- This slide slide shows the actual Business Vault applied within the POC Model (blue table), calculating a risk score for that transaction
- This Business Vault approach provides a repeatable process to create and populate consistent Data Vault objects
- This table could be a view, however physicalising it allows capture of the change over time



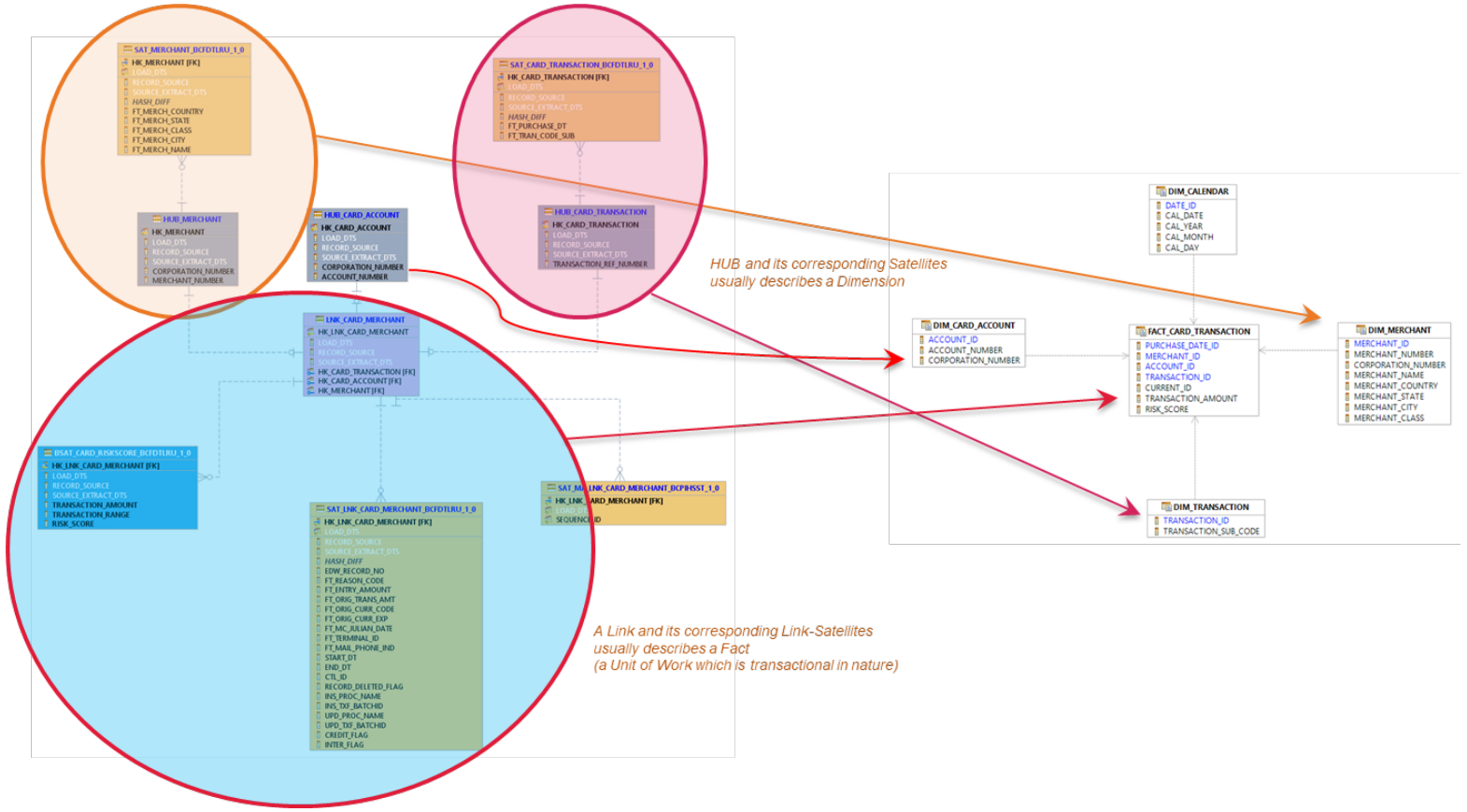
CERTUS INSIGHT

INFORMATION MART & CONSUMPTION

Information Mart

- Information Marts are constructed from Raw Vault and Business Vault objects as needed.
- The Information Mart is a star schema, supporting most BI tool and to create a data set that a user wold expect to see.
- Information Marts are typically implemented as views.
- Can be physicalised if needed, e.g. For performance
- Highly iterative and repeatable
- As Business Rules are de-coupled from Raw Vault we gain the benefit of independent processes for things like maintenance & testing.

  - For example we can avoid regression testing of unrelated data sets

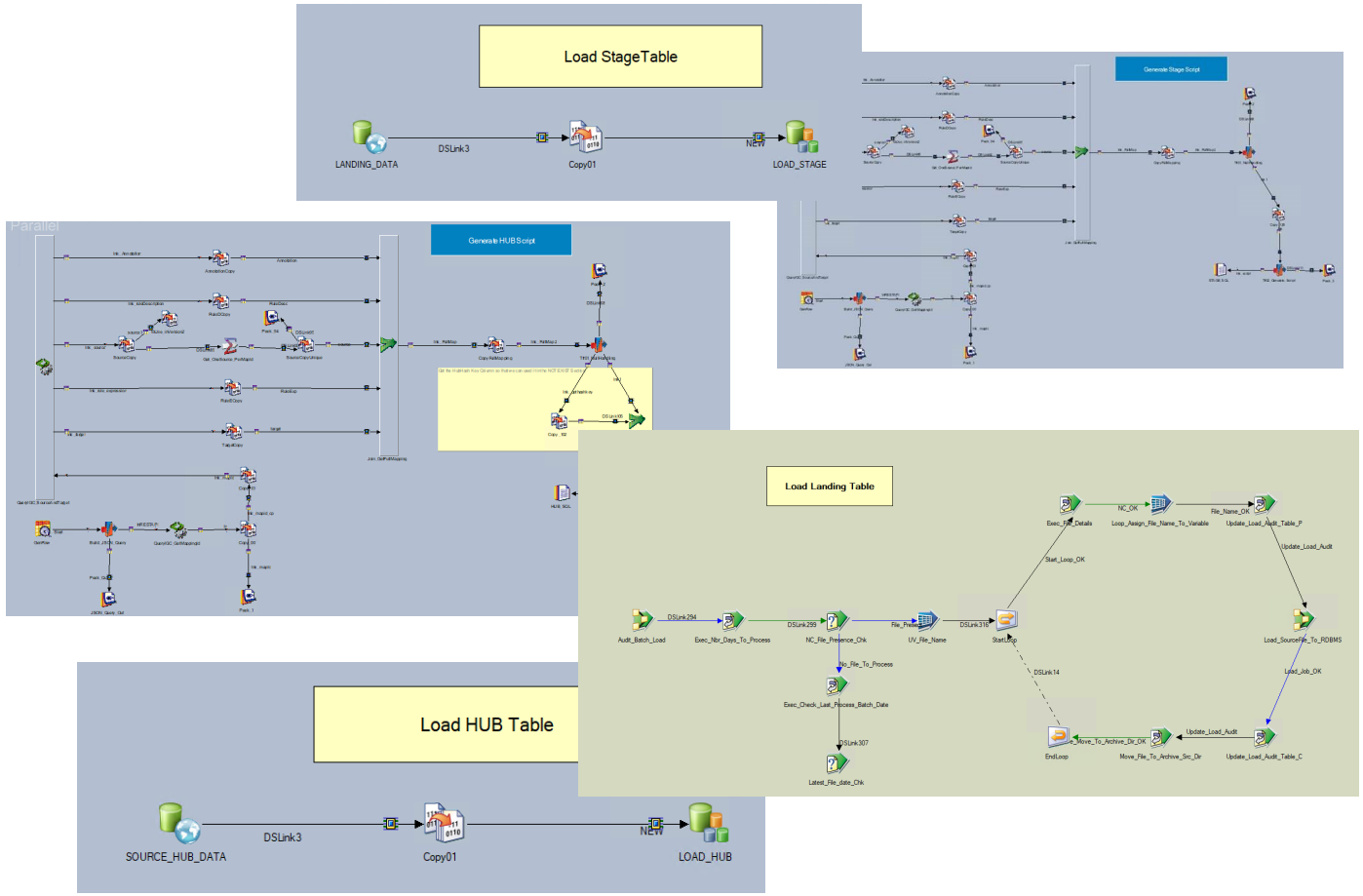


CERTUS INSIGHT

CERTUS DATA VAULT 2.0 ACCELERATORS FOR IBM INFOSHERE

Data Vault 2.0 Accelerators

- Certus has designed and built a set of Data Vault 2.0 accelerators for the IBM InfoSphere suite.
- The accelerators cover the processed for modelling and mapping, Data Vault 2.0 code automation templates for DataStage and them management of metadata for lineage and impact analysis.
- These accelerators assist customers to rapidly adopt the Data Vault 2.0 methodology and ensure the leverage the standards to their full.



CERTUS INSIGHT

THANK YOU

CERTUS
INSIGHT