

BlueStar ETL

cedar

細達 鄭惟心

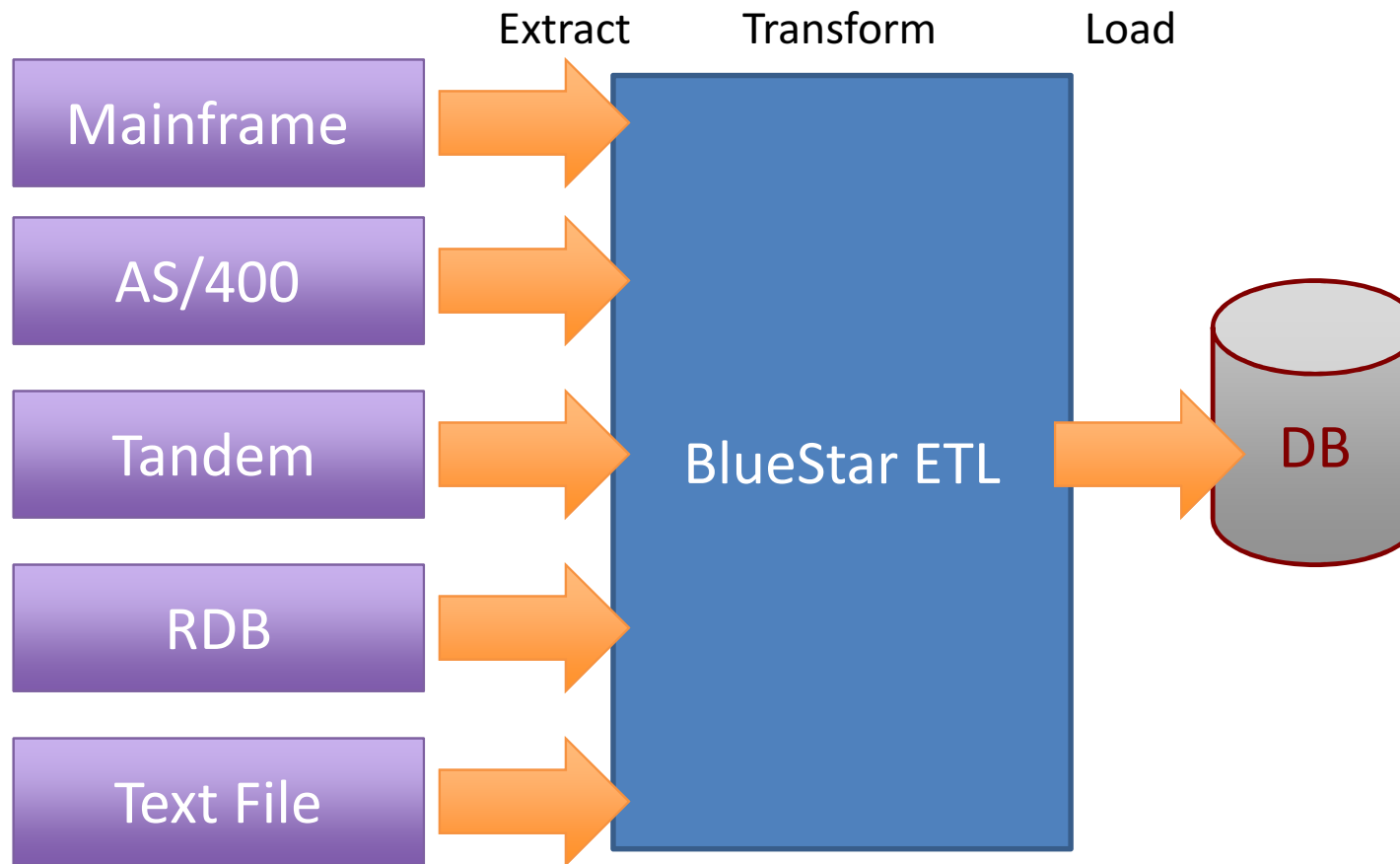
Agenda

- Data Transformation Issues
- What ETL is?
- Features Of BlueStar ETL
- Performance
- Case Studies

Data Transformation Issues

- Parsing
- Codepage Mapping
- Data Type Compatibility
- Language Dependencies
- System Dependencies
- Changes Of Business Models
- Code Versioning

What Is ETL?



BlueStar ETL Features

- Support Multi-Data Source
 - IBM Mainframe(OS/390,MVS,VSE)
 - COBOL,COBOL II,COBOL 370...
 - PL/I
 - AS/400
 - COBOL,RPG - DDS
 - UNISYS
 - COBOL
 - Tandem
 - COBOL

BlueStar ETL Features

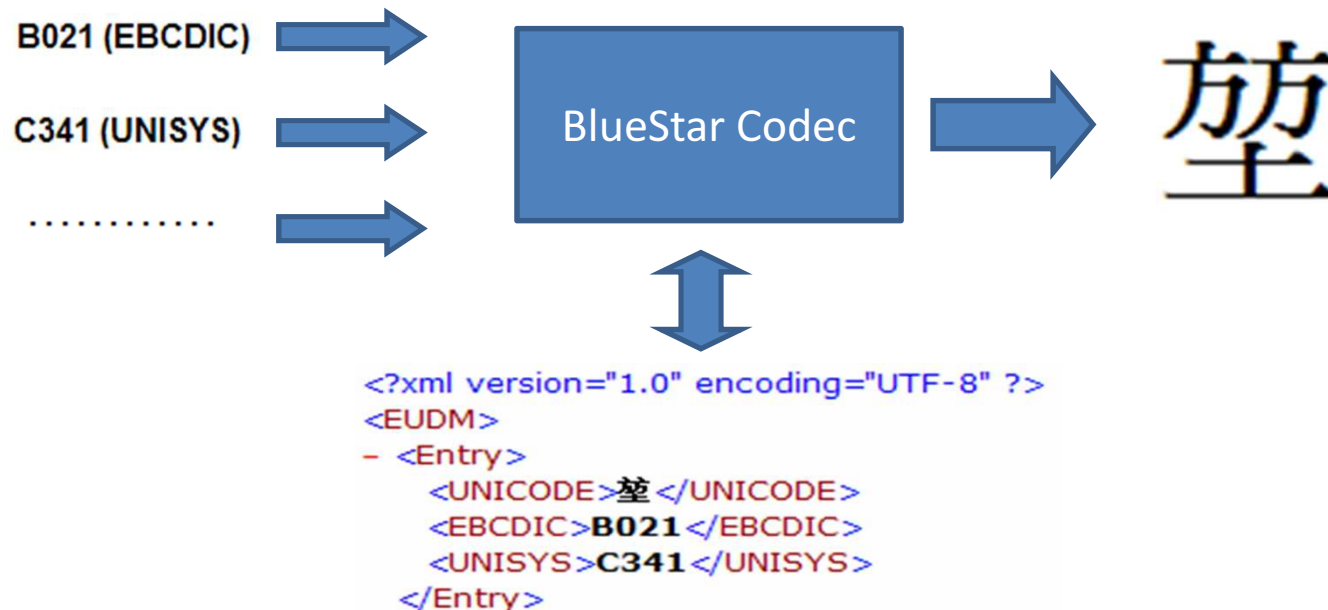
- Support All COBOL Data Type
 - Available
 - X (string)
 - 9 (Number)
 - S (Signed Numeric)
 - C (Composition)
 - C3 (Comp-3)
 - N (Native)
 - H (Hex)

BlueStar ETL Features

- Support Multi-Code Page
 - Available
 - ASCII
 - EBCDIC
 - UNISYS
 - CNS11643
 - UTF8.
 - Provide SDK for any new code page.

BlueStar ETL Features

- Native Unicode Support
 - Unicode Output
 - Support End User Defined Character – Map to Unicode.



BlueStar ETL Features

- Code Free: Auto Generate Message Schema

System : 4700

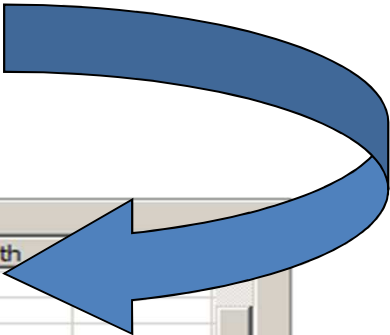
ID : P0

COPYBOOK: CPINPFI

Input

Page: I-2005/11/11

INPUT Length	Type	FIFO OFFSET OR LABEL	FIFO Length	Attr.	Chinese
21	C	FIFO_TX COMM HEAD 1	21	CL	Y
12	C	FIFO_TX COMM HEAD 2	12	CL	Y
6	N	FIPS_BRH_NO	6	CR	N
6	N	FIPS_ACT_NO			
1	P	FIPS_CDH_SRL_NO			
6	N	FIPS_TRF_IN_BRH_NO			
6	N	FIPS_TRF_IN_ACT_NO			
6	P	FIPS_TX_AMT			
8	C	FIPS_UNIT_NO			

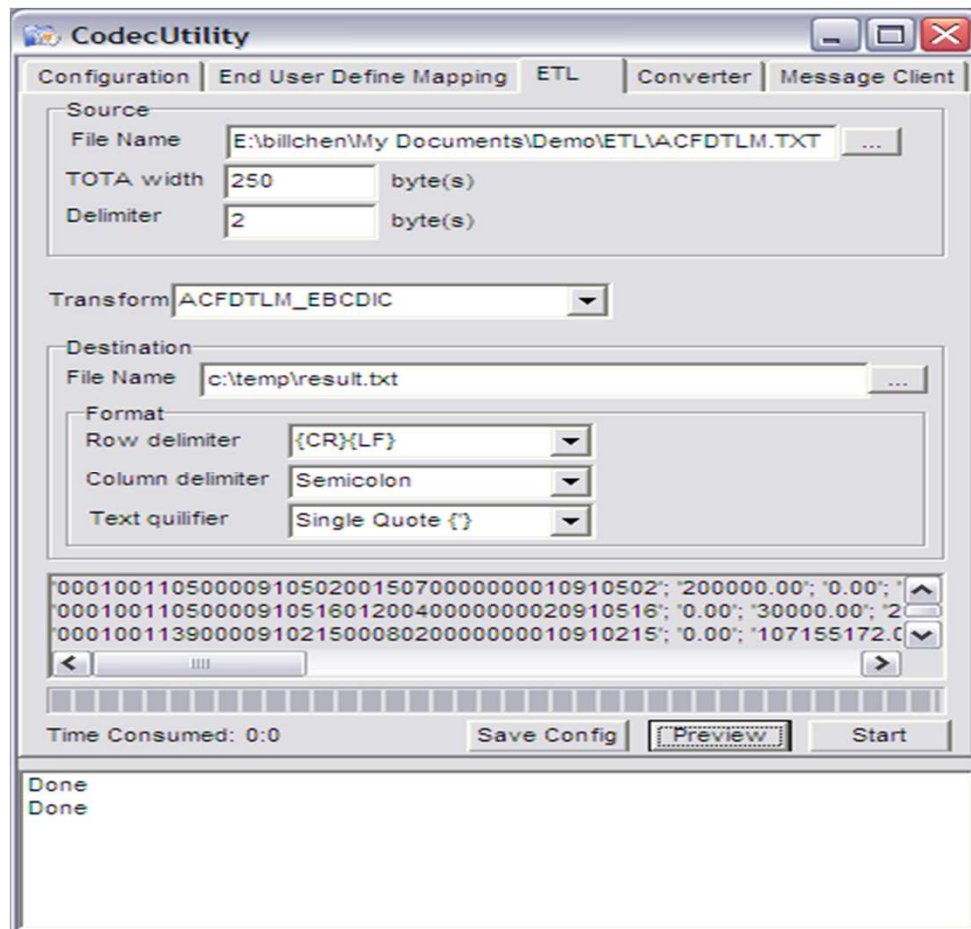


No	Name	Type	Length
3	FIFO_TX COMM HEAD 2	X (String)	6
4	FIPS_ACT_NO	9 (Posit...	6
5	FIPS_CDH_SRL_NO	C (Com...	1
6	FIPS_TRF_IN_BRH_NO	H (Hex ...	6
7	FIPS_TRF_IN_ACT_NO	C3 (Co...	6
8	FIPS_TX_AMT	X (String)	6
9	FIPS_UNIT_NO	X (String)	8
10	FIPS PIN 16	X (String)	14

Delete New Insert Edit Delete

BlueStar ETL Features

- Code Free: Transformation Tool



Performance

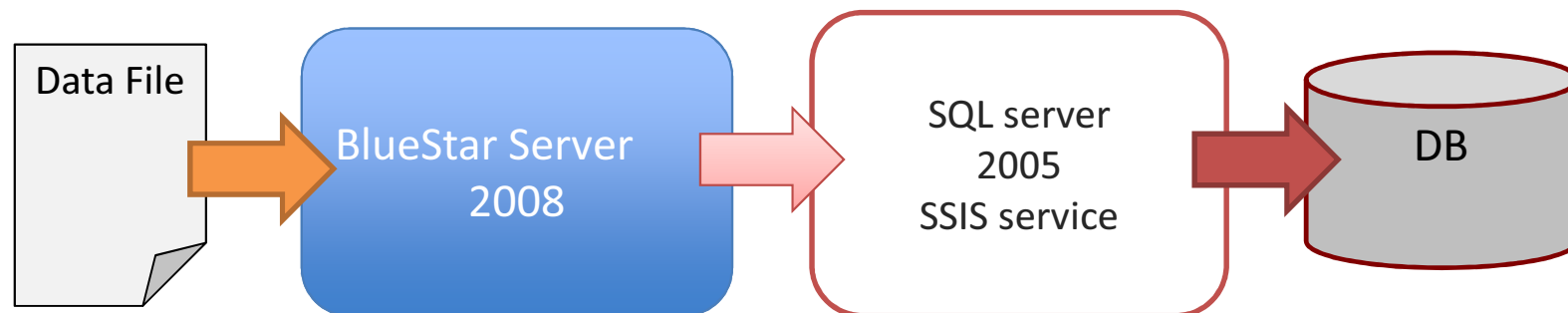
- Environment
 - Platform
 - Windows XP Professional
 - HW
 - CPU: Intel Pentium M 1.7GHz
 - RAM: 1GB
- Source Data
 - 5 Giga bytes
- Performance
 - Less than 8 minutes!

Case Studies

- HCT
 - Data integration from mobile device to centralized database
- CTCB
 - Notification platform
- To Achieve Process Automation
 - File Trigger Service

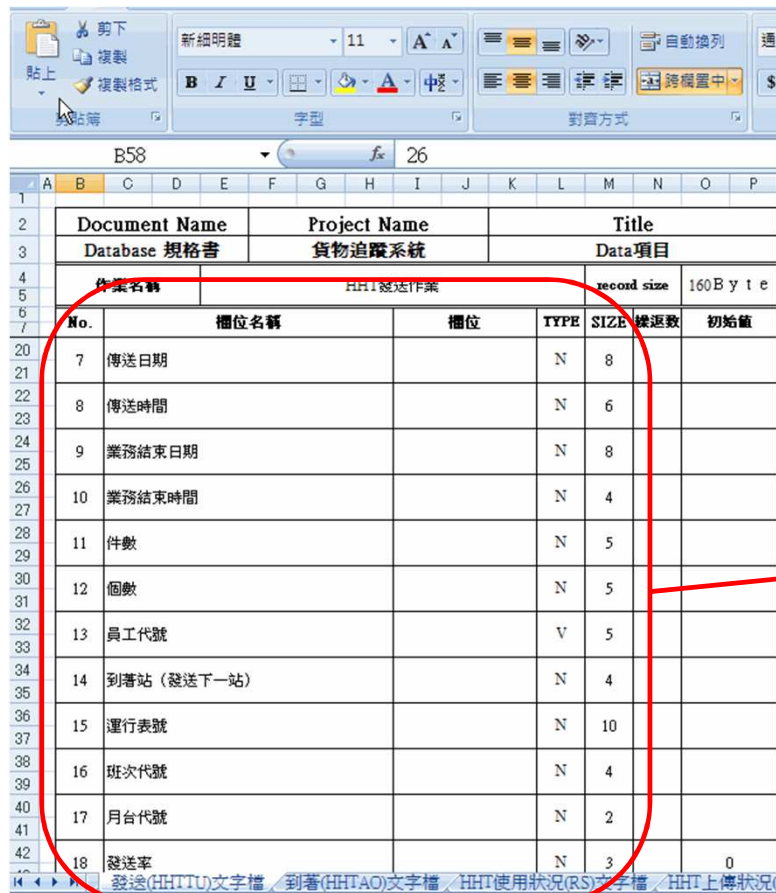
Case Study: HCT

- Transporting ASCII format data files from mobile device to database server
- Process Flow

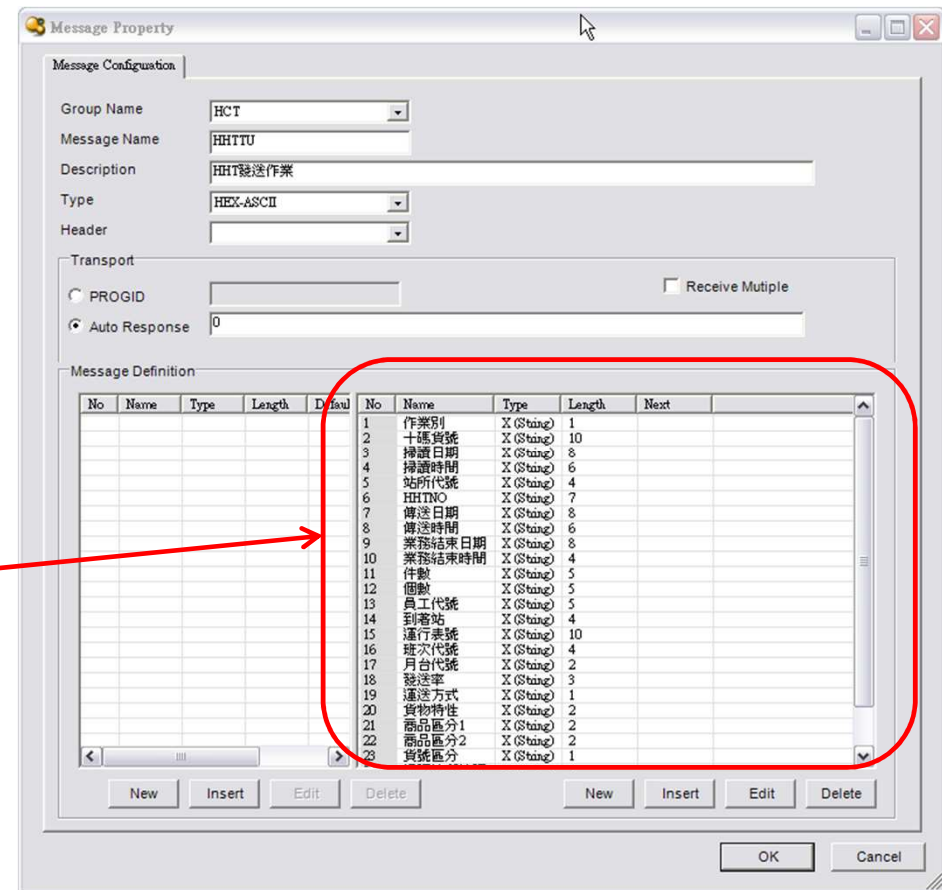


Quickly Build Up Service

- Message schema definition:



Document Name		Project Name		Title	
Database 規格書		貨物追蹤系統		Data項目	
作業名稱		HHT1發送作業		record size	160 Byte
No.	欄位名稱	欄位	TYPE	SIZE	樣本數
7	傳送日期		N	8	
8	傳送時間		N	6	
9	業務結束日期		N	8	
10	業務結束時間		N	4	
11	件數		N	5	
12	個數		N	5	
13	員工代號		V	5	
14	到站站 (發送下一站)		N	4	
15	運行表號		N	10	
16	班次代號		N	4	
17	月台代號		N	2	
18	發送率		N	3	0



Message Property

Message Configuration

Group Name: HCT

Message Name: HHTTU

Description: HHT發送作業

Type: HEX-ASCII

Header:

Transport

☐ PROGID

☒ Auto Response: 0

☐ Receive Multiple

Message Definition

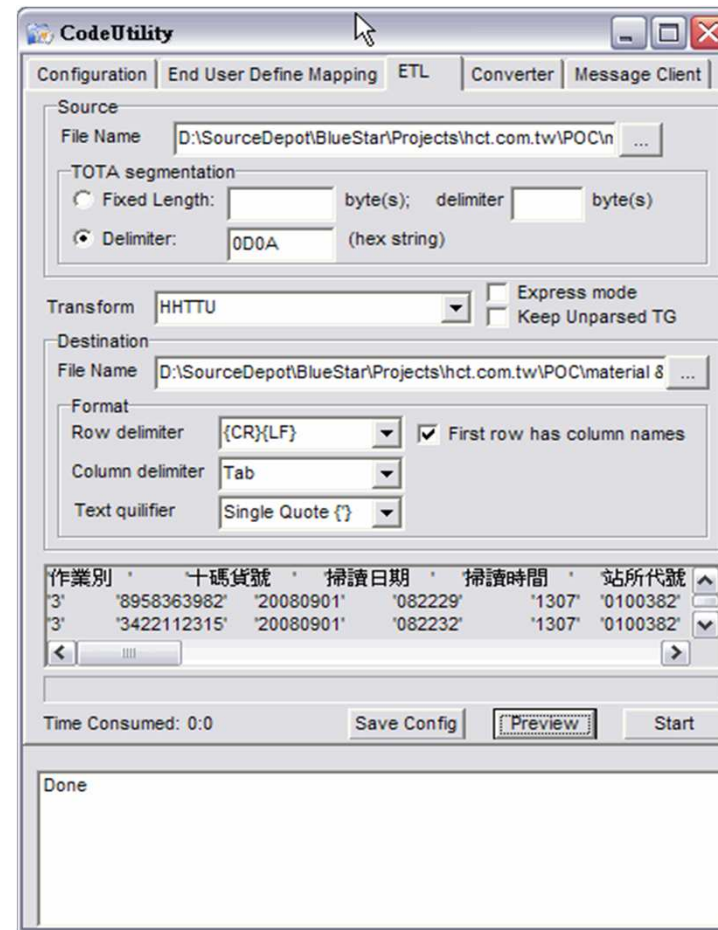
No	Name	Type	Length	Next
1	作業別	X (String)	1	
2	十碼貨號	X (String)	10	
3	掃描日期	X (String)	8	
4	掃描時間	X (String)	6	
5	站所代號	X (String)	4	
6	HHTMO	X (String)	7	
7	傳送日期	X (String)	8	
8	傳送時間	X (String)	6	
9	業務結束日期	X (String)	8	
10	業務結束時間	X (String)	4	
11	件數	X (String)	5	
12	個數	X (String)	5	
13	員工代號	X (String)	5	
14	到站站	X (String)	4	
15	運行表號	X (String)	10	
16	班次代號	X (String)	4	
17	月台代號	X (String)	2	
18	發送率	X (String)	3	
19	運送方式	X (String)	1	
20	貨物特性	X (String)	2	
21	商品區分1	X (String)	2	
22	商品區分2	X (String)	2	
23	貨號區分	X (String)	1	

New Insert Edit Delete

OK Cancel

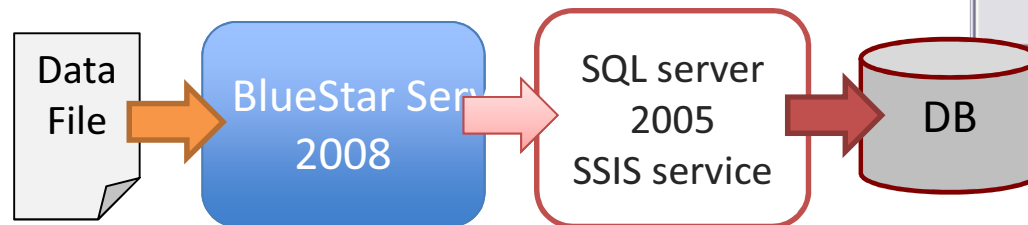
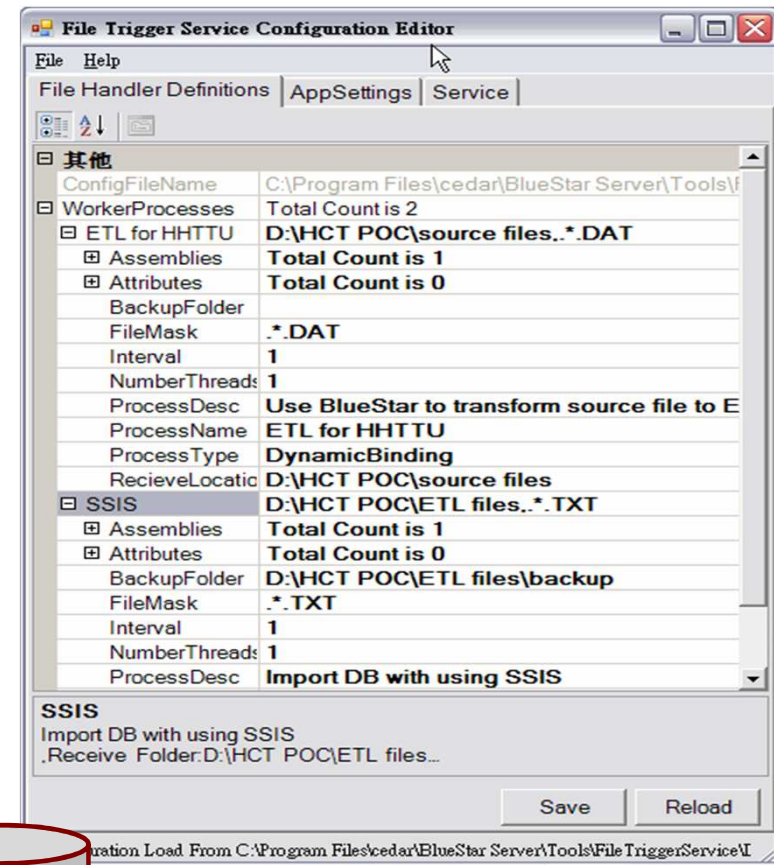
Quickly Build Up Service

- Testing Your Definition:
 - Codec Utility



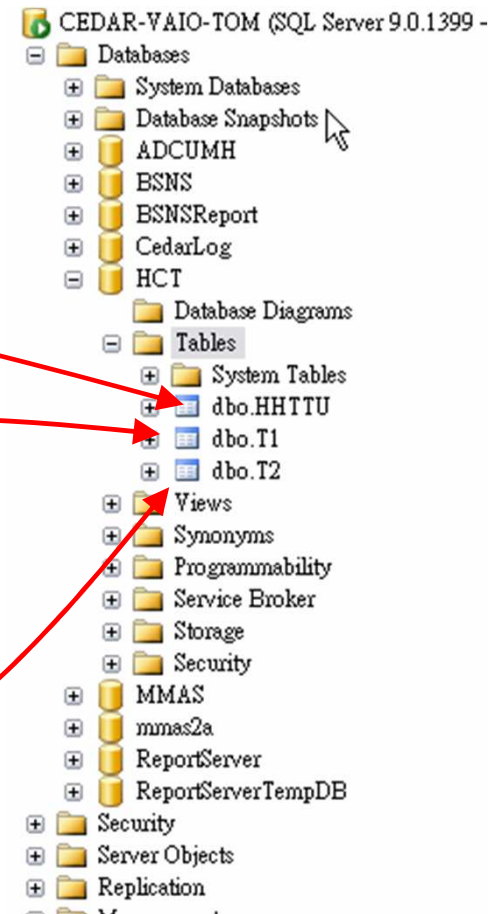
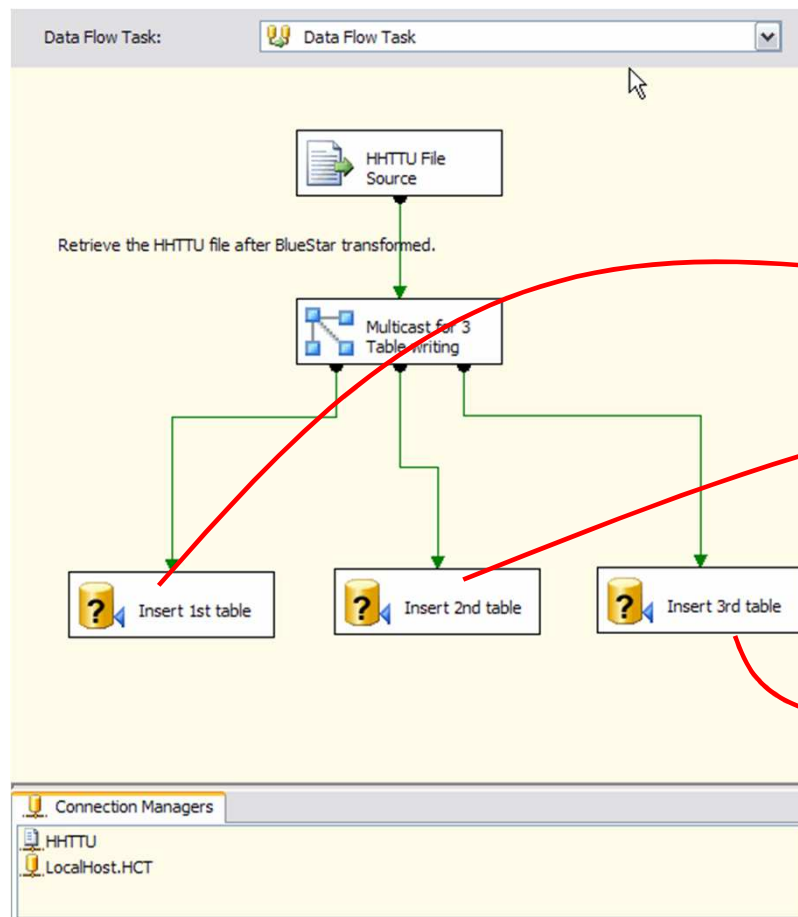
Quickly Build Up Service

- Defining Monitoring Mechanism:
 - File Trigger Service



Quickly Build Up Service

- SSIS DataFlow Designer:



Performance Test Result:

- Testing platform:
Hardware :
CPU : Dual Core 1.6G
RAM : 1GB
HD : SATA 5400 rpm
- Software :
OS : Windows Vista Ultimate
ETL : BlueStar 2008 (.net framework 3.5)
DB : SQL server 2005

Performance Result:

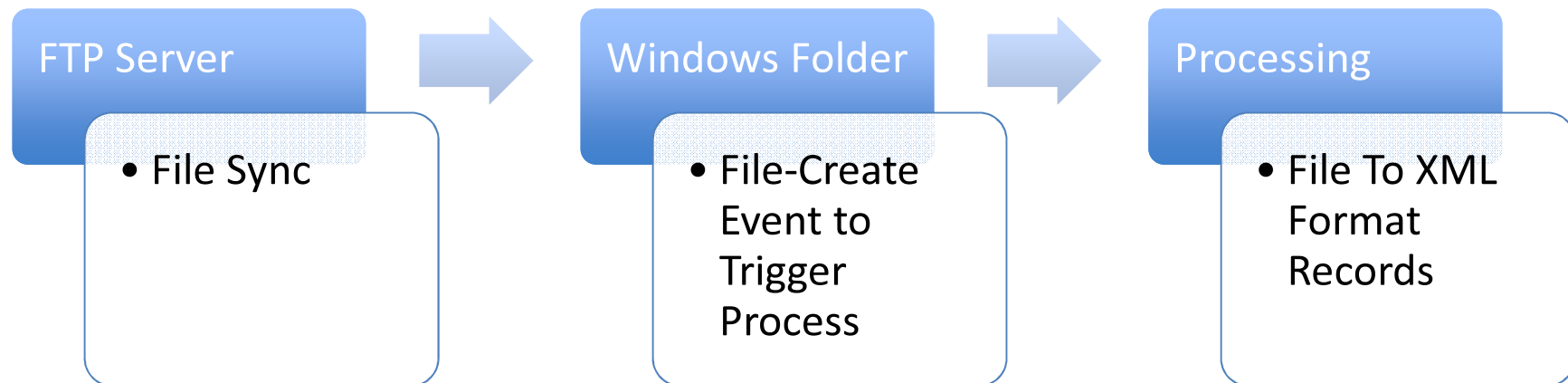
- ETL:
182,566 HHTTU data file transformation:
Avg. Cost **25.8 seconds**
- SSIS DB import:
182,566 * 3 tables = 547,698 data rows :
Cost about **1 minute**
(Highly depend on DB hardware power)

Case Study: CTCB

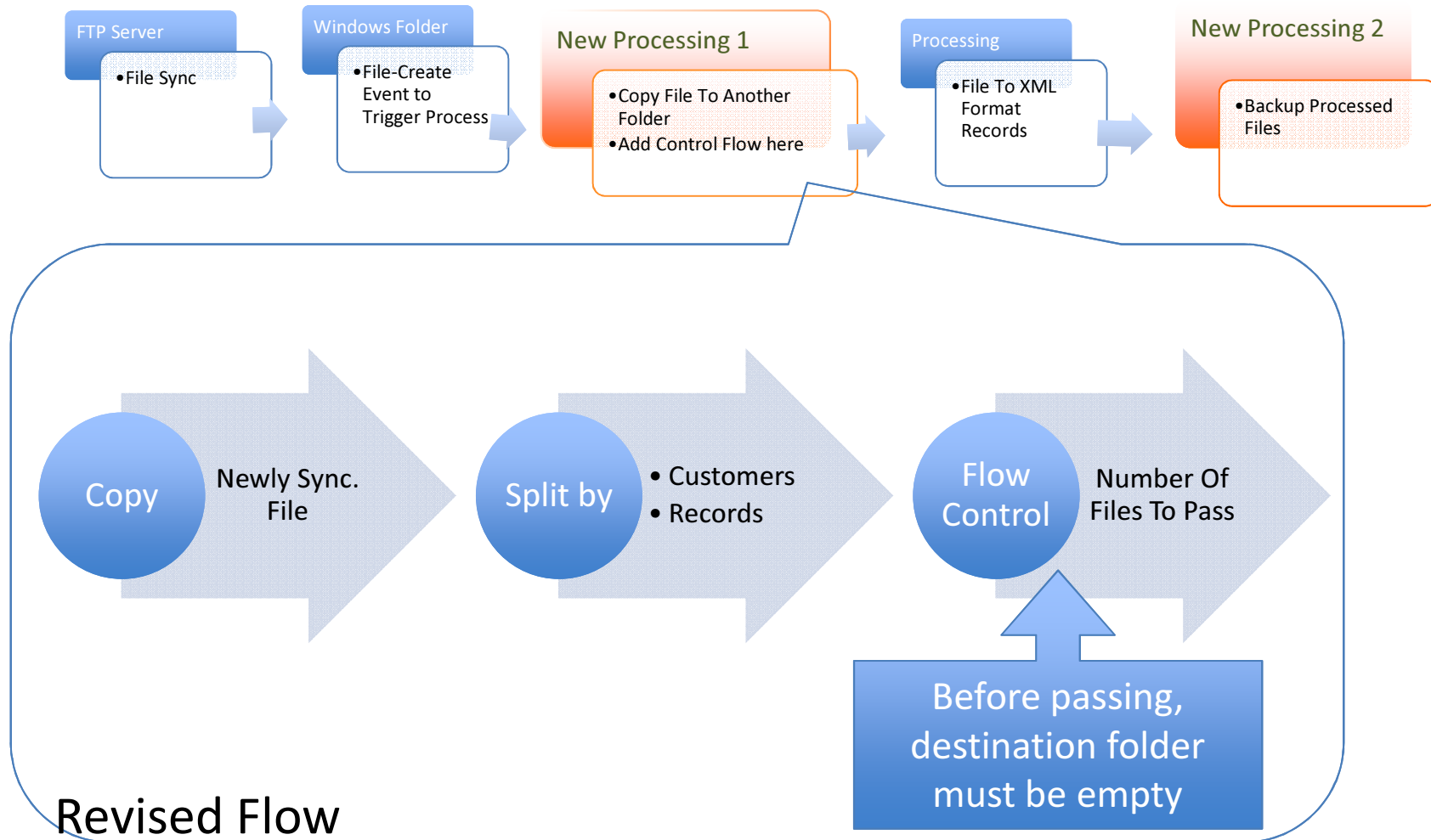
- Resolve Issues
 - Resource Consumption
 - Online Transaction vs. Report Generation
- Solution:
 - Refine Workflow (BlueStar File Trigger Service)

Case Study: CTCB

Original Process Flow



Case Study: CTCB



Conclusion

- Data Transformation via BlueStar
 - Support Multiple Data Sources
 - Support Multiple Codepages
 - Code Free ETL
 - High Performance
 - Workflow

Thank You

- Q&A