
DataFax Middleware - A Publishing Engine for DataFax System

Hanming Tu
Premier Research, A Division Of CRC Development, Ltd.
30 South 17th Street, Philadelphia, PA 19103
Hanming.Tu@Premier-Research.com
Tel: 215-282-5526; Fax: 215-282-5528
<http://www.premier-research.com/>

Content

- Our Issue and Path
- General Middleware Architecture
- DataFax Middleware Requirement
- DFM Configuration
- DFM Features
- Conclusions

Our Issue

- DataFax system is a very reliable and secure system for collecting data and generating reports within the system.
- Clients want to access their clinical data timely and flexibly.
- We create a data mart (DF_ORAbridge) for each client, generate reports and build a web portal for them.
- We repeat the process for each client and run the reports for different clients on different schedule. → Huge amount of WORK!

Our Path

- Our searching and struggling:
 - Resource Planning and Management for Clinical Data Management using DataFax, Oracle and Web server (DFUG 1999, Lake Louise)
 - Bridging DataFax with Oracle (DFUG 1999, Lake Louise)
 - DataFax Web Reporter (DFUG 2000, Whistler, BC)
 - Web Portal: From 'D' to 'E' (DFUG 2000, Whistler, BC)
 - Enable CDISC with DataFax (DFUG 2003, Duck Keys, FL)
- We need something between DataFax and clinical web portal – a middleware.

General Middleware Requirement

- Support applications: real-time or near real-time, collaborative
- Support underlying systems: Unix, Windows, etc.
- Configurable: deployment-time configurability and run-time reconfigurability
- Component architecture: reusability, dynamic extensibility, improved understandability, reduced development costs, and better support for long term system evolution.

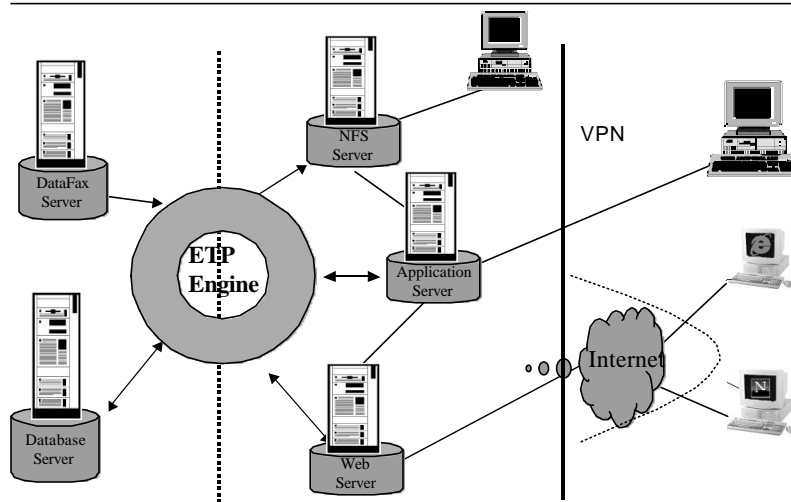
DataFax Middleware (DFM) Requirement

- Support DataFax and other data sources including Oracle
- Easy to configure
- Re-use existing components
- Flexible scheduling capability

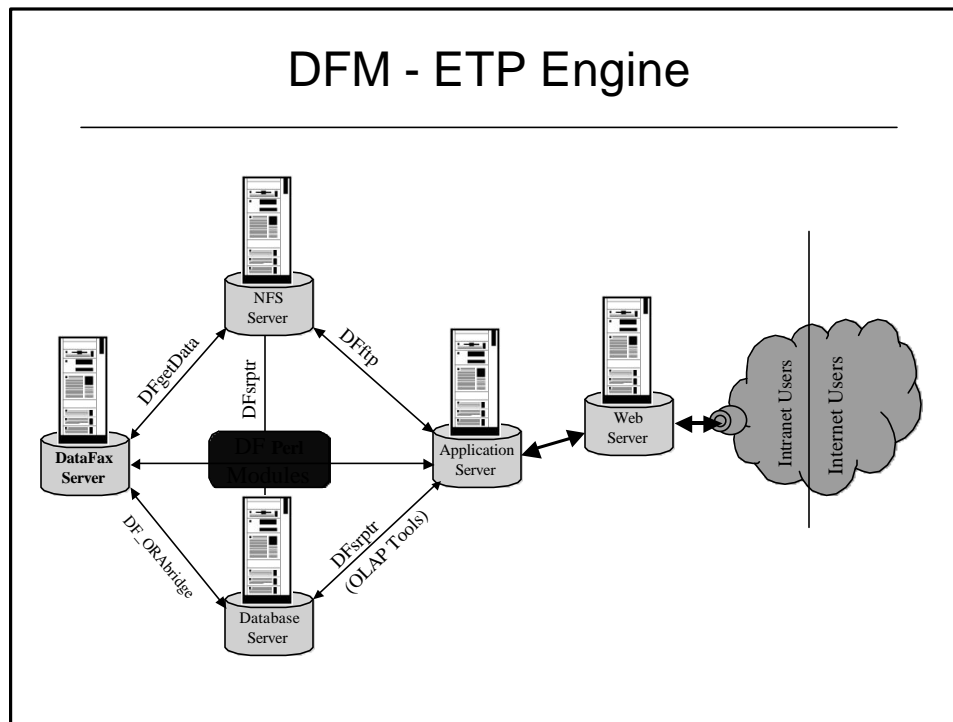
Our DFM - DFsrptr

- DFsrptr – DataFax standard reporter is a Unix shell application, and it ties all the DataFax components that we developed together to form an extracting, transforming, and publishing (ETP) engine.
- It acts as a middle tier in our clinical web portal architecture.
- It is a pushing engine, no outside pulling is allowed.

DFM - Architecture



A Three Tier Architecture



DFM - Features

- Configure for multiple studies;
- Publish DataFax reports (bypassing RDBMS) and new reports from data warehouse;
- Publish reports into multiple sites or based on subscription;
- Publish through different means such as ftp or email;
- Archive all the historical reports;
- Flexible scheduling capability to allow the reports to be published monthly, weekly, daily or the combinations within one study.

DFM - Usage

```
Usage: DFsrptr [options] [inputs]
options:
  -a <usr/local/bin> program directory
  -b </pub> remote data source directory
  -c <usr/pwd/@db> Oracle connection string
  -d </opt/datafax> DataFax directory at local host
  -f <addr user pass dir> file server ftp connection parameters
  -i <d|default|s|system|init_file> the initial file that defines the studies,
    status, schedule and parameters to be used for running this program.
    default will use DFsrptr.ini file in the program directory -
    system will use Protocol name file in a syswide directory
  -l <a|...|z|...> log file name modifier. It sets .
  -o <DFsas:SAS> output file names for DFsas job file and SAS job file.
  -p <user@host> remote ftp user password (de-supported)
  -s <schedule_type> d-daily; w-weekly; m-monthly
  -S <primary|secondary|all> datafax record select type
  -t Task type: exp|etp|etc
  -u <anonymous> remote ftp user name (de-supported)
  -v whether to echo message onto screen
  -w <addr user pass dir> web server ftp connection parameters

[inputs] a list of study numbers such as 1-3 or 23,45, etc.

30 7 * * 1-5 /opt/datafax/dba/bin/DFsrptr -v -i d -l elpd -t exp
0 6 * * 6 /opt/datafax/dba/bin/DFsrptr -v -i d -l elpw -t exp -s w
0 8 * * 1-7 /opt/datafax/dba/bin/DFsrptr -v -i d -l elpm -t exp -s m
```

Configuration – Options

- Option **-i**: specifies the input file location. It defaults to DFsrptr.ini.
- Option **-t**: specifies task type such as exp, etp, etc.
- Option **-s**: specifies which schedule to be run
- Option **-l**: specifies log file name

Configuration – initial file

- The initial file contains all the parameters for one or more studies. It has the following columns:
 - Study Number: DataFax study number
 - Status: Y – active, N – not run, V – Parameter indicator
 - Schedule: # Schedule: D – daily, W – Weekly, M - Monthly
 - Composite schedule:
 - W:0,1,5 - run on Sunday, Monday, and Friday
 - M:01,15,30 - run on 1st, 15th or 30th day in the month
 - Task: specifies task type, currently only ETP
 - Data source: specifies data repository or data warehouse such as usr/pwd@db for Oracle connection
 - Report name and its parameters: DataFax generic report or standard web reports
 - Archiving destination: ftp_server domain\ftp_user password ftp_directory
 - Publishing destination: web_server domain\ftp_user password web_directory

Configuration - Parameters

The V in the status column indicates a parameter. The ETP has a system wide default parameters, which are defined in DataFax.ini file, but you can change them for each study for the following parameters:

```

dfsv - DataFax server (host name)
orcl - Oracle database connection string (usr/pwd@db)
fadr - ftp IP address or host name
fusr - ftp user name
fpas - ftp user's password
fdir - ftp destination directory
wadr - web server name or IP address
wusr - ftp user on web server
wpas - password for ftp user on web server
wdir - directory on web server

```

Configuration - Example

```

# define variables and reports for Study 253
253|V|dfsv|df-svr                               # defining study
253|V|orcl|usr/pwd@oradb                         # specific parameters
253|V|fadr|10.10.10.1                            #
253|V|fusr|crc\df_FTP_usr                       #
253|V|fpas|pwd4ftpusr                           #
253|V|fdir|/archive/destination/folder/S253     #
253|V|wadr|10.10.10.2                            #
253|V|wusr|dmz\df_WEB_usr                       #
253|V|wpas|pwd4webusr                           #
253|V|wdir|/web/content/folder/S253            #
253|N|D|exp|V|DF_Ptmissing:253 rpt|V           # this report will not run
253|Y|D|exp|V|DF_Wfcrfsperwk:253:-1:10 rpt|V   # DataFax report and its
253|Y|D|exp|V|DF_Wfcrfs:253:-u:all:-t:today rpt|V # parameter
253|Y|D|exp|V|DF_WFqcs:253 rpt|V              #
253|Y|W|exp|V|DF_qcsbyfield:253 rpt|V         #
253|Y|W|exp|V|DF_WFqcs_all:253 rpt|V          #
253|Y|D|exp|V|L20:1|V                          # std report for archiving
253|Y|W:5|exp|V|L20:3|V                        # and internal use
253|Y|M|exp|V|L20:4|V                          #
253|Y|W:3|exp|V|L20:31|V|V                    # std report for archiving
253|Y|W:3|exp|V|L20:32|V|V                    # and publishing

```

Standard Reports (I)

DataFax Generic Reports: DF_Ptmissing, DF_Wfcrfsperwk, DF_WFcrfs, DF_WFqcs, DF_WFqcs_all, DF_qcsbyfield, etc.

00 - CRF and QC Reports:

- 01 - CRF and QC Report *
- 02 - QC Aging Report *
- 03 - First QC By Center - Last 7 days *
- 04 - First QC By Center - All *
- 05 - Centers With More Than 25 Un-resolved QCs *
- 06 - CRF Status Report *
- 07 - QC Status Report *

20 - Productivity Reports

- 21 - Monthly DataFax User Productivity Report *
- 22 - Weekly DataFax User Productivity Report *
- 23 - Daily DataFax User Productivity Report *
- 24 - Simplified Monthly DataFax User Productivity Report *
- 25 - Simplified Weekly DataFax User Productivity Report *
- 26 - Simplified Daily DataFax User Productivity Report *

Standard Reports (II)

- 30 - Subject Enrollment Reports
 - 31 - Subject Enrollment And Visit Status By Site **
 - 32 - Subject Enrollment By Time **
 - 33 - Subject Enrollment And Visit Log **
 - 34 - Subject Enrollment Comment List **
 - 35 - Subject Exception Report **
 - 36 - Subject Termination Reason Counts By Site **
 - 37 - Subject Termination Reasons **
 - 38 - Subjects Who Have Prematurely Discontinued **
- 40 - AE Reports
 - 41 - Adverse Event List **
 - 42 - Adverse Event Severity By Site **
 - 43 - Adverse Event Relation By Site **
 - 44 - Adverse Event Action Taken By Site **
 - 45 - Adverse Event Outcome By Site **

Standard Reports (III)

- 50 - Projected Payment Reports
 - 51 - Weekly Actual And Projected Payment ***
 - 52 - Weekly Accumulative Actual And Projected Payment ***
 - 53 - Monthly Actual And Projected Payment ***
 - 54 - Monthly Accumulative Actual And Projected Payment ***
 - 55 - Projected Pass-through CRA Travel Cost ***
 - 56 - Monitor Monthly Calendar Report ***
 - 57 - Monitor Weekly Calendar Report ***
- 60 - Projected CRF Reports
 - 61 - Weekly Actual And Projected CRFs ***
 - 62 - Weekly Accumulative Actual And Projected CRFs ***
 - 63 - Monthly Actual And Projected CRFs ***
 - 64 - Monthly Accumulative Actual And Projected CRFs ***
 - 65 - Daily Actual And Projected CRFs and QCs ***

Standard Reports (IV)

70 - Projected Subject Reports

71 - Weekly Actual And Projected Subjects ***

80 - Drug Accountability Reports

81 - Drug Accountability List **

82 - Drug Dose Level By Site **

90 - Miscellaneous Reports

91 - Drug Accountability Report **

Report Publishing Steps

- **Report creation and publishing steps:**
 - Get client requirement
 - Determine data source and logics
 - Create Oracle views
 - Add the report number to the configuration file
 - Link the report onto the web portal
- **Since most reports are standardized, we can just re-use the template with data from the new study.**
- **Information needed for configuration:**
 - Schedule: how often do you want the report to be run
 - Destinations: archiving and publishing

Standard Report Tool

We developed a web tool for running DataFax reports and developed a similar tool for viewing standard reports online as well.

DataFax Specialized Reports

Select studies	001 - Mallin 1092-95-01	acc.	
Select Reports	01 - CRF and QC report		1
Select Options	<input type="checkbox"/> Create ASCII and Excel files. <input type="checkbox"/> Real time <input type="checkbox"/> Userbase mode <input checked="" type="checkbox"/> Format output <input type="checkbox"/> Add InPage Links <input type="checkbox"/> Add Sub-visual <input type="checkbox"/> Add Cell Links		
Time Range	Start: (YYYYMMDD.HH:MM:SS)	End: 20040206 005123	(YYYYMMDD.HH:MM:SS)
Actions	<input type="button" value="Reset"/> <input type="button" value="Submit"/> <input type="button" value="Help"/> <input type="button" value="Refresh"/>		

Example of Archived Reports

<ul style="list-style-type: none"> [-] Audit [-] EditCheck [-] Extract [-] Listings [-] Qc [-] Reports <ul style="list-style-type: none"> [-] Y2002 [-] Y2003 [-] Y2004 <ul style="list-style-type: none"> [-] 01 [-] 02 <ul style="list-style-type: none"> [-] 02Mon [-] 03Tue [-] 04Wed [-] 05Thu [-] 06Fri [-] Stat [-] Tables [-] Pdf [-] Programs [-] Script 	<table border="0"> <tr><td></td><td>S112_L2O_01.csv</td><td>4 KB</td><td>Microsoft Excel Com...</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_01.htm</td><td>44 KB</td><td>HTML Document</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_02.csv</td><td>2 KB</td><td>Microsoft Excel Com...</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_02.htm</td><td>29 KB</td><td>HTML Document</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_03.csv</td><td>1 KB</td><td>Microsoft Excel Com...</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_03.htm</td><td>18 KB</td><td>HTML Document</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_04.csv</td><td>9 KB</td><td>Microsoft Excel Com...</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_04.htm</td><td>48 KB</td><td>HTML Document</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_05.csv</td><td>1 KB</td><td>Microsoft Excel Com...</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_05.htm</td><td>22 KB</td><td>HTML Document</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_06.csv</td><td>1 KB</td><td>Microsoft Excel Com...</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_06.htm</td><td>20 KB</td><td>HTML Document</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_31.csv</td><td>2 KB</td><td>Microsoft Excel Com...</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_31.htm</td><td>32 KB</td><td>HTML Document</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_32.csv</td><td>3 KB</td><td>Microsoft Excel Com...</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_32.htm</td><td>41 KB</td><td>HTML Document</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_33.csv</td><td>21 KB</td><td>Microsoft Excel Com...</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_33.htm</td><td>140 KB</td><td>HTML Document</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_34.htm</td><td>20 KB</td><td>HTML Document</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_38.csv</td><td>12 KB</td><td>Microsoft Excel Com...</td><td>2/6/2004 7:49 AM</td></tr> <tr><td></td><td>S112_L2O_38.htm</td><td>56 KB</td><td>HTML Document</td><td>2/6/2004 7:49 AM</td></tr> </table>		S112_L2O_01.csv	4 KB	Microsoft Excel Com...	2/6/2004 7:49 AM		S112_L2O_01.htm	44 KB	HTML Document	2/6/2004 7:49 AM		S112_L2O_02.csv	2 KB	Microsoft Excel Com...	2/6/2004 7:49 AM		S112_L2O_02.htm	29 KB	HTML Document	2/6/2004 7:49 AM		S112_L2O_03.csv	1 KB	Microsoft Excel Com...	2/6/2004 7:49 AM		S112_L2O_03.htm	18 KB	HTML Document	2/6/2004 7:49 AM		S112_L2O_04.csv	9 KB	Microsoft Excel Com...	2/6/2004 7:49 AM		S112_L2O_04.htm	48 KB	HTML Document	2/6/2004 7:49 AM		S112_L2O_05.csv	1 KB	Microsoft Excel Com...	2/6/2004 7:49 AM		S112_L2O_05.htm	22 KB	HTML Document	2/6/2004 7:49 AM		S112_L2O_06.csv	1 KB	Microsoft Excel Com...	2/6/2004 7:49 AM		S112_L2O_06.htm	20 KB	HTML Document	2/6/2004 7:49 AM		S112_L2O_31.csv	2 KB	Microsoft Excel Com...	2/6/2004 7:49 AM		S112_L2O_31.htm	32 KB	HTML Document	2/6/2004 7:49 AM		S112_L2O_32.csv	3 KB	Microsoft Excel Com...	2/6/2004 7:49 AM		S112_L2O_32.htm	41 KB	HTML Document	2/6/2004 7:49 AM		S112_L2O_33.csv	21 KB	Microsoft Excel Com...	2/6/2004 7:49 AM		S112_L2O_33.htm	140 KB	HTML Document	2/6/2004 7:49 AM		S112_L2O_34.htm	20 KB	HTML Document	2/6/2004 7:49 AM		S112_L2O_38.csv	12 KB	Microsoft Excel Com...	2/6/2004 7:49 AM		S112_L2O_38.htm	56 KB	HTML Document	2/6/2004 7:49 AM
	S112_L2O_01.csv	4 KB	Microsoft Excel Com...	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_01.htm	44 KB	HTML Document	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_02.csv	2 KB	Microsoft Excel Com...	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_02.htm	29 KB	HTML Document	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_03.csv	1 KB	Microsoft Excel Com...	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_03.htm	18 KB	HTML Document	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_04.csv	9 KB	Microsoft Excel Com...	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_04.htm	48 KB	HTML Document	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_05.csv	1 KB	Microsoft Excel Com...	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_05.htm	22 KB	HTML Document	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_06.csv	1 KB	Microsoft Excel Com...	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_06.htm	20 KB	HTML Document	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_31.csv	2 KB	Microsoft Excel Com...	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_31.htm	32 KB	HTML Document	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_32.csv	3 KB	Microsoft Excel Com...	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_32.htm	41 KB	HTML Document	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_33.csv	21 KB	Microsoft Excel Com...	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_33.htm	140 KB	HTML Document	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_34.htm	20 KB	HTML Document	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_38.csv	12 KB	Microsoft Excel Com...	2/6/2004 7:49 AM																																																																																																						
	S112_L2O_38.htm	56 KB	HTML Document	2/6/2004 7:49 AM																																																																																																						

Example of Published Reports

- Click [here](#) for a list of published reports.

Conclusion

- The DFM provides not only a platform for integrating DataFax, Oracle and web portal but also a framework for DataFax federation and collaboration with other systems.
- The component approach makes the development easy to manage while keeping the system flexible.
- Limitation: The current version is in Korn shell and it is difficult to port to other environment.
- Future development: We are planning to re-write it in Perl so that it can integrate with other systems such as EDC.

Reference

- **The design of a configurable and reconfigurable middleware platform:**
http://imagbean.ffclrp.usp.br/~adriano/bibliography/doc/middleware/DistribComp-design_configurable_middleware-coulson2002.pdf
- **Middleware Architecture Committee for Education (MACE):**
<http://middleware.internet2.edu/MACE/>
- **Middleware Architecture URLs:**
http://www.iturls.com/English/TechHotspot/TH_43.asp