Automated Data Flow

Banks are now getting seriously into the implementation of ADF as stipulated by RBI. What does the implementation entail? Three senior bankers discuss the issues:

With a view to ensuring accuracy and integrity of data flow from the banking system to the Reserve Bank of India, the latter had set up a core group comprising experts from banks, the RBI, IDRBT and the IBA, which came out with an approach paper on automated data flow (a straight-through process) from the core banking solution (CBS) or other IT systems of commercial banks to the regulator. The paper outlines a cluster-based methodology to be adopted by banks (which classifies banks into various clusters based on technology and process dimensions used by them). The RBI wants the banks in the first phase to ensure a seamless flow of data from their transaction server to their management information system (MIS) server and automatically generate all returns from the MIS server, without any manual intervention. The regulator would then in the second phase introduce a system for the flow of data from the MIS server of banks in a straight through process.

RBI also wants the banks to achieve 'common end state' (the state of complete automation for submission of the returns without any manual intervention) independent of the current technologies being used by the banks and can be used by all banks irrespective of their current level of automation. This to some may appear to be a tall order and the banks may require high levels of transformation in processes and technology adopted. The approach paper has also made suggestions on the data architecture that may be applied. There are four key process elements identified - data acquisition, data integration and storage, data conversion and data submission.

RBI has subsequently decided to implement the project in a phased manner and banks were advised to submit a roadmap clearly indicating the returns which can be sourced directly from the banks' systems without manual intervention. RBI feels that ADF will significantly enhance the quality and timeliness of data, which in turn will benefit the entire sector through informed policy decisions and enactment of regulations.

How do bankers mandated with the task of implementing the system view it?

STANDARD PROCESS

Says J. Sridharan, head, MIS and Finance, IndusInd Bank: "RBI has suggested a standard process:
- Have a data warehouse (which RBI has called as CDR, or centralized data repository)
- Data from various source systems of the bank should feed into this CDR automatically, daily
- There should be validations to ensure correctness and completeness
- Based on such data, various RBI reports should be generated

However, RBI is yet to come out with a plan for seamless movement of the reports from banks to their own servers. Meanwhile, the multiple methods of feeding data to RBI will continue, there are still some returns that are sent in hard copy; some are to be keyed in RBI application directly; some are to be fed into small Foxpro/MS Access based applications provided by RBI and the output has to be e-mailed to them. A 'minimum' solution should look for an enterprise wide data warehouse (EDW) / data mart to collate all data points that are needed for the RBI reports, updated daily, and be able to generate reports in multiple formats that RBI has prescribed. But, since the data requirements are not cast in stone but tend to grow in scope as well as in complexity, a 'good' solution should look for maintaining all relevant data elements on a daily basis. For Banks, the entire process chain involving data validation and collation, storage and retrieval, processing reports, multiple reporting formats, and audit trail, etc, is complex and hence a 'bespoke' solution

development is not ideal."

Sridharan says most of the banks, including IndusInd Bank, have gone in for canned solutions that meet with the requirement. "Initially, when the market was nascent with no solution providers in sight, we too had started with a skeleton 'bespoke' development but abandoned once we found a credible partner," he adds.

He talks of two broad choices that are available to banks – (a) invest only in a limited solution to meet with RBI ADF requirements, suitable especially for banks which have a functioning data warehouse for business requirements and which do not need to revisit that investment, and (b) go in for a full-fledged data warehouse model, albeit with a staged implementation plan.

4 LAYERS
R.P. Marathe, GM, Planning, MIS & DWH, Bank of Baroda, talks about the 'common end state' approach for ADF which is the state of complete automation for submission of the returns by the banks to RBI without any manual intervention and is common across all banks. He explains the four distinct logical layers as data acquisition layer, which captures data from various source systems; data integration & storage layer, which extracts and integrates the data from source systems with maximum granularity; data conversion layer, which converts the data stored in the CDR to the prescribed formats using pre-defined business rules and data submission layer, which is a single transmission channel which ensures secure file upload mechanism in an STP mode

"As suggested by RBI, ADF can be implemented in two phases. In the first phase the bank would be required to ensure seamless flow of data from their transaction server to their MIS server and generate all returns from MIS server automatically without any manual intervention. In the second phase reserve bank will introduce a pull mechanism for flow of data from the MIS server of the banks in a straight through process (STP)," says Marathe, adding: "Based

J. Sridharan stresses that banks need to have internal teams experienced in managing large IT implementations

on the discussion with various banks, RBI has given the discretion to the banks to decide their own roadmap to complete phase I by December, 2012. In order to monitor the progress, RBI had advised the bank to automate minimum additional 15 and 25 returns during the quarter ended on December 2011 and March 2012 respectively."

He says currently, banks are more focused on implementing phase I, that is, automated generation of returns from centralized data base without any manual intervention. Different banks are adopting different approaches to achieve this and a few banks are able to automate few returns from their existing infrastructure/ systems available while others are looking for holistic approach and searching for comprehensive solutions.

N.V.L. Narasimham, AVP, Strategic Projects & Corporate Technology Group, Dhanlaxmi Bank, feels that as a first step, banks would need to adopt a combination of data acquisition (layer 1) and data integration and storage (layer 2) in order to create a CDR which would

need to be used as a common source for generating the various reports that are to be submitted to the RBI. "As a second step, the RBI would introduce a mechanism by which data from individual banks' CDR would flow seamlessly to the RBI, through STP. Depending upon their individual process and technology maturity, banks can adopt strategies that align to establishment of the four layers," he adds.

MATURITY LEVELS
Marathe says banks should have the centralized database for all applications and the system to integrate/ pool the data from all sources in the single database for automated generation of returns without manual interventions. "Further there is need to document processes for data capture and data collation. Guidelines or checklist for maintaining data quality, audit and report submission should also be in place. Documented process in case of error (root cause analysis, audit trail for tracing back) is also important. Further proper IT security and network is also required for STP of data as envisaged in phase II," he adds.

He says Bank of Baroda is in the process of implementation of an EDW and at present an operational data store is in place through which more than 100 reports are being generated for various departments and RBI. The bank has also put in place a process of creating of various required documents during implementation phase. Proper network and IT security have also been implemented.

Sridharan says given the situation that the ADF solution is more likely to be implemented as standardized IT solution developed by a vendor, banks need to have an internal team experienced in managing the implementation of large IT projects and then running them thereafter. This in itself calls for higher levels of process and technology maturity.

He elaborates: "Seamless flow of data into RBI's systems is still a vision. Right now, banks are trying to ensure that the collated data is readily available in a single

place, in an electronic form, so that when needed they can be pushed to RBI."

He adds that for IndusInd Bank, technology is a key business enabler and not merely a business support system. The bank already has a number of complex and sophisticated IT systems in use. For ADF, it has found a technology partner, iCreate, which has an appropriate solution. "And to leverage our return on this investment, we have embarked on a wider EDW which can support business intelligence and analytics, besides RBI's ADF requirements," says he.

Narasimham of Dhanlaxmi Bank says RBI has advised banks to conduct a self–assessment based on their technology and process maturity profiles and accordingly place themselves in one of the six clusters established for this purpose. While cluster 1 represents the highest maturity level, cluster 6 represents the lowest with the intervening clusters falling in between.

He says across the Indian banking industry, some 223 individual reports are submitted to the RBI at present by the banks. "For the purpose of ADF, these returns need to be classified into five different groups: Simple, Medium, Complex - I, Complex - II and Others."

Based on the self-assessment done by individual banks, they may follow a varied set of timeframes for implementing the 5 individual report groups, ranging from under six months to about 24 months.

TWO APPROACHES

He also points out that the approach paper has specified two distinct approaches: A standard (or overall) approach that needs to be followed by all the banks that assumes that the implementing bank has little or no automation for submitting the returns and therefore the work has to start from scratch to build the 4 distinct layers, one by one; and a variations approach that is applicable in those cases where the bank concerned has a reasonable degree of automation for the returns submission process. This approach calls for customization of the bank's current automation set up to complete the end to end automation process.

RP Marathe points out that currently, banks are more focused on implementing automated generation of returns from centralized data base without any manual intervention

"At present, we at Dhanlaxmi Bank submits to the RBI a set of 173 reports across various business functions and frequencies. An internal assessment based on the small number of systems and processes in practice reveals that the bank would take about 12 -15 months to implement the ADF," says Narasimham.

DATA IN ELECTRONIC FORM

Sridharan says availability of valid data in electronic form is the pre-requisite for a data warehouse or a data mart and from this perspective, banks having dedicated platforms for specialized businesses such as treasury, forex, trade finance, retail loans, credit cards, cash management etc, besides CBS, will have a definite edge. Says he: "A good solution needs to have three building blocks: (i) strong technology tool for ETL, or Extract, Transform and Load, that can collate data from multiple source systems, (ii) comprehensive database design with strong metadata definition, which is

more of a solution design element than technology and in terms of database, it can be any standard one such as Oracle, Microsoft, SQL, etc, and (iii) robust report writing / generating tool which is essential to analyze the data and prepare RBI returns."

He points out that RBI has mandated that the preparation of data and reports based on the data within the bank should be a STP but, it is yet to come out with a clear mandate on a STP transmission of data from the banks to RBI. "RBI's vision is to adopt XBRL technology but this calls for a clear definition of taxonomy and RBI has already initiated the process and this may, to a large extent, get completed this fiscal," he adds.

Sridharan also says that at IndusInd Bank, to the largest extent, feeding data to the warehouse from various source systems is fully automated, the system design provides for data validation and cross-tabbing and there are system screens for accepting corrections to the data wherever needed, some of the RBI reports call for qualitative inputs and as such manual intervention in such cases is inevitable.

Narasimham also says that the solution depends on the implementation of a CDR, which in itself is a collection of data marts across several dimensions such as customers, accounts, transactions, etc. Thus, technologies used in building OLAP systems such as a data ware house are an apt choice.

"Irrespective of the technology platform and the technology partner chosen by a bank for implementing the ADF, the most important activity that needs to be undertaken is the establishment of the CDR," says he. "A relatively simple way of building the CDR is to first set up a common reception (or stage) area that is filled with data from one or more source systems. Thus, irrespective of the data platform of the source systems, the combined data resides in one specific database. Thereafter, data in the reception area can be used to build the CDR, as per requirement."

He says in Dhanlaxmi Bank, this approach is in place. The reception area

is built on an Oracle database, and is programmed to contain EOD data of the various source systems. From the reception area, data is subjected to an ETL, which populates the CDR with the required information. The actual RBI reporting application consumes the data from the CDR, thereafter.

As for the seamless transfer of data to and from RBI is concerned, banks would need to undertake the first steps to convert their respective reporting data into standard transmission forms such as XBRL, XML, etc, he says. This can be achieved by implementing conversion tools that allow transformations of information from one type of file to another.

MANUAL REPORTING

Narasimham also states that it is evident that in the present state, banks do follow manual reporting to a large extent. Implementation of ADF should ease many of the problems faced by individual banks, especially those relating to the availability of data itself, followed by the delays in reporting.

"That manual intervention will become a thing of the past is an understatement. Even supposing banks achieve 100% automation of the submission, it should be noted that the entire program is dependent on several individual functions. While the list of functions would be quite lengthy, some of them are: maintenance of data in the source systems, which relies heavily on data entry by usergroups, successful execution of the ETL procedures involved, at various stages, compilation and generation of the reports and successful transmission to the RBI, data quality checks at various points in the program and technology checks at various points in the program to monitor scheduling of jobs, successful execution of the scripts, etc," says Narasimham.

Marathe says technologies are required to support (i) data acquisition from source data in electronic form, (ii) developing a CDR, (iii) reporting tool for generation of reports and (iv) utility for submission/ transmission of data.

NVL Narasimham points out that RBI has advised banks to conduct a self–assessment based on their technology and process maturity profiles

The role of STP is associated with the data submission layer and it is a single transmission channel which ensures secure data transmission/ file upload mechanism, says he.

"RBI envisages that data submission from the banks can be triggered automatically using a scheduler. Data submission can be designed to support either 'push' or 'pull' mechanisms. If the data submission is initiated by the bank then it is considered to be a 'push'. In a 'pull' mechanism the submission process is initiated by the RBI," explains Marathe.

Marathe points out that RBI has prescribed formats and thus data requirement is prescribed. The banks can use any technology to capture data from sources. And thereafter the data has to be stored in one place. In data conversion process, the data to be converted in the specified format of the RBI like XBRL, XML or Excel based files. "This is the common end approach for ADF. The final submission can be made through web base submission or by uploading the data file."

Sridharan says the RBI has defined

Neorithm's OLAP-based solution

Neorithm is a service provider to the banks which in spite of having CBS in place find that critical business decisions are based on strategic data. The company has developed i-NavigatorSuite Business Analysis and Reporting Solution which is a centralized web-based solution, using Online Analytical Processing (OLAP) which is designed for analysis and decision support by providing unlimited views of various business dimensions. The value in an OLAP database is such that many complex calculations and predefined queries are pre-processed and results are stored.

iCreate's Biz$core facilitates reporting

iCreate Software's Biz$core is an integrated suite of business intelligence, analytics and performance management solutions tailor-made for banks. Biz$core's ETL layer with pre-configured integrations with source banking systems extracts data from multiple systems and loads into a CDR which is engineered for the automated generation of RBI returns through pre-built reporting templates. A fully integrated metadata framework enables auditability of the data lineage from source to report.

Ramco has automated solution

Ramco Banking Analytics is an ADF solution designed to help banks achieve complete automation of regulatory reporting. Not only is it a pre-built product, it also follows a data model for ESS/DSS/MIS purpose rather than a report-centric approach, ensuring in the process that a bank can not only take out all the reports required for ADF compliance but also use the data model for performance management at various hierarchy levels, capable of extending to accommodate further changes and overcome an ever-changing regulatory environment.

the report format. It is only a question of automating data collation and report generation at the banks' end and pushing the data directly to RBI. "From RBI's perspective, the data points for reports are well defined; the choice of technology and process at banks' end is immaterial as long as the requisite data points can reach RBI. Hence, this is not an issue at all. The most important benefit for RBI is availability of information in a standardized format, in electronic form which can be readily used for downstream analysis," he elaborates further.

"There will be tremendous productivity gains at RBI. At present, lot of resources – human, technology, processing – are used in RBI in collating the data submitted by banks in various forms and formats. These resources will now be freed and RBI will have a near 'real time' banking statistics," says he.

SECURITY ISSUES

Implementation of such a large program such as this is replete with security issues. Narasimham has a list of key areas which banks should ideally focus on: (i) Database Security. It is essential that the databases participating in the program are completely secured with highly restricted access, and follow information security policies and procedures of the respective institution. In fact, it would be prudent to allow a 'read only' grant, even to those ETL objects which extract data stored in the source

systems. This will prevent accidental writes even by computer programs on such source databases. (ii) Application Security. Access to the application used in the ADF program implementation is the next level where access needs to be provided on a 'need to access / operate' basis. Definitely, a maker-checker approach should be adopted such that all reports generated on the application carry the stamp of the person generating the same as well as the stamp of the person verifying the report. In addition to following this concept, it is also important that access is governed and aligned to individual business functions. (iii) Network and Transmission Security. Networks on which the ADF platform is implemented, are vulnerable to breaches. Therefore, banks would need to ensure that both intranet and internet platforms are protected with firewalls, security policies and other security procedures that ensure breaches are contained and managed so that transmission of data within the bank and with the RBI is done in a secured and encrypted manner. (iv) Report Manipulation and Copy. One of the easy ways to Access data is to download files from the applications. Once the information (either a report or spreadsheet) is available on the local workstation, there is no control on the usage (or) manipulation of such data. In order to circumvent this problem, it is recommended that report download happens only on 'read only' format, the simplest of which is the PDF. It is also

in the interest of the banks to disable burning of discs (DVD / CD) on individual workstations as well as disable USB ports.

NEXT STEP

Sridharan maintains that RBI is yet to come out with a detailed plan for seamless flow of data from banks to their site, but when it happens, banks may have to evaluate the security issues. "Knowing how RBI operates, banks generally do not have a concern here," says he, adding that in spite of this, "The entire set up should be within the control of the bank and those banks which are going to put together a CDR / data warehouse for the first time need to evaluate their process and IT controls to ensure data security, as the CDR / DW will have comprehensive data on the entire bank."

For Marathe, the focus areas should be network security (data must be transmitted using a secure, encrypted format), application security (limit access from centralized data depository to the authorized users only), data security (to control and monitor access, update rights and for deletion of data held within the database) and workstation security (use anti-virus protection, screen saver passwords, policies around installation and personal software and access).

While RBI came with the process requiring banks to follow specific exercise in filing returns, it's concern about the decreasing profitability of the PSU banks on account of prevailing scope for errors and manipulations, that really promoted it to go for ADF. Now, when one looks at the banks' efforts in complying with ADF, it has become more of a IT issue that an operational or business issue. Compliance and practices that go with it should be a core business function of the banks and as the final phase of ADF comes into being, banks should take it up with all seriousness. There is a view that banks do not do much to gain adherence to the ADF program and only a high level of participation by the senior management will help identify the gaps in order to take cohesive corrective action.

mohan@bankingfrontiers.com