How to Optimize Data Integration in Manufacturing

By Fred Meyer
August 2005

Introduction

Data integration is a key enabler for supply chain modernization. Competition is driving increased integration of the supply chain, both due to direct cost pressures and to increased globalization. There can be little doubt that well implemented supply chain integration saves labor costs and improves inventory control. This has often been a complex and costly proposition. Standards, improved products and a common understanding of best practices for distributed systems are reducing the price of entry to supply chain automation. This will, in turn drive broader integration of manufacturing and distribution endpoints.

Past Practice

In the early days of integration, we had little idea of what the market would ultimately need. Standards were largely in the future. Simple tools like FTP led to integration via batch files. Our collective experience with Object-oriented programming led us to believe that sophisticated, object level interfaces between systems would be needed for integration, as they had been for human workflow. As integration became accepted and the Internet Bubble expanded, 'Big Vision' initiatives attempted to standardize entire domains of eCommerce. None of these approaches is entirely invalid, but all have drawbacks for mainstream integration, when compared with data integration.

- **Batch-oriented integration** - is a major stumbling block to today's business objectives. Demand-driven architectures require real time visibility of inventory and production status. Batch processes insert time lags which propagate through the supply chain, creating risks of excess or obsolete inventory, increasing carrying and liquidation costs. They are often full of bad data, but since they are simply data blobs from the infrastructure's perspective, no one knows.
- **Object layer integration** - while essential for human interaction is overly cumbersome and granular for data integration. Object layer integration is essential for human interaction, because workers carry much of the business process in their head. They need a fine-grained interface with which to manipulate underlying systems in a flexible way. Automated systems have much simpler interactions. Object integration is usually overkill between them.
- **'Big vision' standards** - have defined hundreds of messages, but customers typically use less than a dozen. The first wave of integration is behind us. Large corporate hubs and Tier 1 suppliers are well on the way to full integration. Global 2000 companies have effective, if sometimes cumbersome, links between backbone systems, but often cannot afford to integrate less strategic applications.

Global Presence, Distributed Data

M & A activity and internal consolidation are driving together disparate operations, systems and technologies. Outsourcing and globalization are distributing companies' data and

business processes. Supply chain automation is essential to remain competitive in this environment.

As we all try to drive down costs by reducing inventory and shortening lead times, old integration strategies are showing their weaknesses. We have learned enough about integration in the past ten years to make more practical and efficient choices. A side effect of this practical approach is that replacement of obsolete technology becomes cost-effective. Simplicity drives ROI.

Organizations that are able to effectively integrate across multiple tiers of the supply chain and to reach more internal endpoints will create competitive advantage for all participants in their sphere of influence. We must recognize though, that the nature of the challenge is different.

Today's Problem

We must now integrate companies with limited IT budgets and very limited IT skills, and reach less strategic, but more numerous internal systems. Tiers 2 through N are where the opportunity for increased efficiency and competitive advantage now lie. These smaller organizations cannot and will not implement 'big vision' standards. They need to rapidly implement just the parts of those standards required to communicate with other parts of their supply chain. A majority of internal systems have similar constraints.

A Range of Tools for Diverse Tasks

Pragmatic integration is driving segmentation. Large, complex, high-volume applications may still be built with very general and sophisticated software suites, but the 'one size fits all' approach is inappropriate for the vast majority of cases. The good news is that, though most vendors claim to do everything, there are already noticeable differences in the applicability of various types of integration software. Offerings are becoming more purpose-specific with time. Let's look at four examples.

Application Server/Platform

Application Server Platforms provide first class application construction tools, along with sufficiently good integration to connect those applications into the infrastructure, provided the rate of change of connections is not high. Claims to the contrary notwithstanding, they are not the best choice for 'pure' integration, as development environments are sophisticated and the relevant skills are scarce in most companies. 'Pure' integration focuses on interconnecting systems without building new applications.

These tools require experienced architects and developers, both for implementation and maintenance. A diverse supporting cast is required to integrate software with the OS, JVM, database, network, security and management systems. The same skill set is required for support and troubleshooting, so implementation in locations without IT resources is very difficult.

Enterprise Application Integration

EAI stacks are designed to plug together pre-existing applications at the BPM/Workflow level. Like App Servers, they require experienced architects, but there is a greatly reduced dependence on coders. Because EAI tools use configuration rather than coding, they respond more easily to change than Application Server platforms, but there is still a good amount of coding involved.

These tools are not appropriate for application development. A diverse supporting cast is still required to integrate the software with the OS, JVM, database, network, security and

management systems. Like Application Servers, EAI stacks generally need local IT staff for long-term support and maintenance.

Extract, Transform and Load

These products were originally intended for dealing with data sets (bulk load, data cleaning, duplicate removal, etc). ETL tools excel at very rapid database loads and data set operations. They are optimized for batch processes. Real time tools are not effective for dealing with data sets, since they view the world one transaction at a time.

Some of the ETL vendors have added real time extensions recently, but the underlying architectures are still batch mode, which limits their usefulness for event driven activities. ETL tools require coding. You must typically rerun an entire batch to recover after a failure. They have limited ability to validate and alert on individual transactions. As with App Servers and EAI, ETL requires a broad range of skills to implement and maintain.

Application Router

The Application Router is a new type of appliance specializing in data integration. The Application Router is ideal for data integration. The appliance approach eliminates 'per project' assembly of components. Appliances substantially reduce the skill set necessary for integration, because components such as the database, network and OS are pre-integrated and hidden from the user. This results in very short time-to-value. Because data is handled transactionally and in real time, the appliance can complement other types of integration, filling in where they are overly complex or not economically justifiable. Application Routers can be configured and managed remotely to eliminate on site IT staff.

Application Routers are not suitable for processes with intensive Business Process Management or manual workflow requirements. Human workflow requires object level interfaces and extensive customization, which are antithetical to the appliance approach. The Application Router is also not suitable as a platform for custom application development.

Data Integration in Manufacturing

Many integration problems in manufacturing revolve around supply chain and logistics issues. These problems are often data driven: Inventory management, fulfillment, WIP tracking, etc. Simpler endpoints tend to be most amenable to data integration. This is more:

WIP Tracking

WIP Tracking is a conceptually simple problem that can be extremely challenging in practice. At the highest level, it is about synchronizing demand, production and forecast data for each product and major subassembly. Common problems include lack of common part numbers, data that is distributed across departments and geographies and incompatible data formats. Existing systems often use batch transfer for data synchronization. Batch transfer prevents real time updates. Worse, batch data is often incomplete and of bad quality. It is impossible to monitor, manage or validate data while it is in a batch file.

What's needed is to create a virtual database with valid, complete real time information. The barriers are economic justification and lack of IT staff at the endpoints. The Application Router, sometimes in combination with EAI at the hub, offers more reach and visibility than can be economically obtained with a software solution alone.

Procurement Planning

A large consumer goods manufacturer has worldwide subsidiaries. The nature of their product permits them to order raw materials through a central system for better prices. The

Application Server Platform chosen for their hub required too much investment and too high a level of skill to reach many end markets. A combination of App Server technology with Application Routers preserves flexibility at the hub while extending visibility to the edges of the company.

Asymmetric B2B

As mentioned earlier, much of the next wave of integration growth will come in the lower tiers of the supply chain. This is being driven by the large 'hub' customers, and by the desire of tier 2 and 3 firms to achieve the same cost reductions and flexibility being enjoyed by the hubs. These vendors do not, however, have budget or skills to implement the massive EDI, RosettaNet or UCCNet solutions used by the hubs. They typically need less than a dozen types of standard transactions. They often need transactions from more than one of the standards: Mixtures of EDI and RosettaNet, for instance, are common in the high tech manufacturing sector. The situation is one of fat hubs, thin spokes. We refer to this as Asymmetric B2B.

Companies facing this challenge have a real problem on their hands. Conventional approaches require one integration component to connect to internal ERP systems, one for RosettaNet (for instance), and another for EDI. This is clearly out of the question for automation of a dozen transaction types. A hybrid model with heavyweight systems at the hub, and lightweight, data-oriented solutions at the spokes is the best solution.

Cut to the Chase

High end BPM and Workflow tools are useful and appropriate for implementing business processes. For data integration, however, these tools are overkill. A data-centric view of the world can help connect more of the supply chain, more rapidly:

- Data integration allows rapid connection to legacy and proprietary systems – Going directly to the data store in these cases may be simpler and faster than writing custom code against an API.
- Data integration eliminates adapters and most of the changes to endpoint systems – Adapters are one part of an object layer integration which does not scale. No company has working adapters to every version of every commercial software package. Custom systems always require custom adapters. Data interfaces are much more standardized than are API's. As XML gateways proliferate, the case for data integration becomes even more compelling.

  Since adapters at the application can be eliminated with data integration, the application changes and cost involved in hosting an adapter are also eliminated.
- If properly architected, data integration reduces the complexity of service-based integration without sacrificing flexibility. The idea here is to nominate one system as the master system for each data type (e.g., customer data), then synchronize all systems to this master via data integration. The SOA layer can then be built across only the master systems, reducing cost and complexity. Data integration and object integration complement one another again in this example.
- Data appliances such as the Application Router result in dramatic reduction in time to value. Recently, a supply chain customer reduced implementation time from two weeks with a leading EAI product to 1.5 days using data integration.
- Data appliances eliminate the need for IT staff at remote endpoints. Remote management, monitoring and configuration allow centralized teams to manage distributed appliance networks, enabling companies to reach and integrate endpoints that have not previously been cost-effective.

- Data appliances have the potential to increase the number of integrated endpoints by an order of magnitude. This is the critical factor if BI tools and analytics are to achieve their full promise.

Picking the Right Tool

As the integration market grows in sophistication, a range of purpose-specific tools based on standards is appearing. This provides the user with an opportunity to obtain better project predictability, more business benefits and reduced TCO by selecting a tool with the proper range of functions and the best price/performance ratio for a particular task. Data integration will play a growing role because it is simple and economical. Proper tool selection, along with well-designed integration architecture is essential for initial project success and control of TCO.

About the Author

Fred Meyer is the President and CEO for Cast Iron Systems and a recognized expert in the integration industry. Before coming to Cast Iron, Fred spent seven years at TIBCO building it from an early concept to the leading Enterprise Application Integration (EAI) vendor. Fred's most recent role was Chief Strategy Officer and prior to that, he served as TIBCO's Chief Marketing Officer for three years pioneering the concept of Service-Oriented Architectures. Fred also managed the Advanced Technology team which prototyped market leading products including the TIBCO Message Broker and TIBCO's first Business Process Management (BPM) tool, TIB/IntegrationManager. Prior to joining TIBCO, Fred spent 10 years developing real time manufacturing management solutions for clients in the automotive, process and pharmaceutical industries. Fred has a BS degree in Chemical Engineering from the University of Minnesota, Minneapolis.

Contact Us

To learn more about the Cast Iron Systems Application Router, please call us at 650.230.0658 or visit us online at www.castironsys.com