

ADAPTING CIVILIAN/COMMERCIAL MODELS OF
DATA WAREHOUSING TECHNOLOGY TO THE MILITARY
By Jim Wakefield

In 1859, when Charles Darwin published "On the Origin of Species," evolutionary leaps were measured in millions of years. In today's technology world, similar leaps are counted in decades or less.

Twenty to thirty years ago, On Line Transactional Processing or OLTP (defined as the ability to store and retrieve information within a subject area via short tactical queries) was "as good as it gets" back then. Decision Support Systems (a.k.a. Data Warehouses) began to appear in the mid 1980's and then truly blossomed within the last ten years. Data warehouses took information management to the next level, allowing users to query vast amounts of data from multiple sources and subject areas. The use of analytics or data mining helped advance the technology even further by predicting trends based on past behavior patterns.

The "active data warehouse" is the next evolutionary advance for data warehousing and is substantiated by the growing number of sites already venturing into new, active directions. The origin of the active data warehouse is defined as the timely, integrated store of detailed data available for analytic business decision making. It is only from that source that the additional traits needed by the active data warehouse can evolve. These new active traits are supplemental to traditional data warehouse functionality. For example, the work mix in the database still includes complex decision support queries, but expands to take on short, tactical queries, background data feeds, and possibly event-driven updates (triggers) all at the same time. Data volumes and user concurrency levels may explode upward beyond anyone's expectation. Restraints may need to be placed on the longer, analytical queries to guarantee tactical work throughput. While accessing the detailed data directly remains an important opportunity for analytical work, tactical work may thrive on shortcuts and summaries. And for both strategic and tactical decisions to be useful to the business, today's data, this hour's data, even this minute's data must be at hand.

Active Data Warehousing Case Study

Some of the toughest decision support opportunities are those that take place directly in the field. One such example is at the break bulks located across North America in a large overnight package shipping company. One purpose of the break bulk is to be a hub for shuffling packages between trucks to optimize the

cost of package transportation to and from points all over the country. At any moment, there may be hundreds of trucks destined for the break bulk and dozens of trucks currently parked at the break bulk preparing for the next leg of their trip. A truck may contain hundreds or thousands of packages.

A critical decision that gets made thousands of times per day is when to release a truck for the next leg of its journey. If you release truck A, which is waiting for a package contained on Truck B before truck B arrives, you could cause a missed package connection. The package is likely to fail its service level target if it misses the connection. On the other hand, if you hold Truck A too long, other packages on truck A are at risk of not meeting their service levels.

Assisting with this complex decision is precisely where the active data warehouse comes into play. To optimize the decision, you must evaluate service level implications for every package on truck A, as well as those that haven't yet been put on the truck. You must also consider the estimated time of arrival for truck B. Moreover, you must analyze aggregated data related to service level performance for year-to-date package shipments of each ship-to and ship-from customer so that you can understand the service level implications of the individual package shipments at hand. And then, you must relate these to the big picture of service level compliance for the year so far. It may be acceptable to allow a package to miss its service level so that another can make its target delivery date when a package has never been delivered late previously (versus a scenario where the last ten packages have been late, and the customer is at risk of attrition because of poor service level compliance). Clearly, very sophisticated analysis is required to optimize this decision when dealing with hundreds of trucks and thousands of packages per truck.

Moreover, the highest paid person in the break bulk is going to make an inquiry to the active data warehouse and expect an answer back in a few seconds to drive the business decision. And, of course, the decision must be made using data that is upto-date. It won't be a very effective decision if the data is a day — or even an hour — stale in this decision making environment.

Furthermore, the break bulk operates 24 hours per day almost every day of the year, so the active data warehouse must provide near 100% availability. An organization's overall strategy will be influenced by the traditional use of decision support against the corporate data warehouse.

Other Commercial Examples

In today's competitive business environment, every successful business is data driven. Information drives decision making based on fact rather than instinct or haphazard guessing.

Telecommunication companies make use of data warehousing for everything from network performance modeling to Customer Relationship Modeling (CRM), which is analyzing customer's behavior and buying patterns to determine who their truly profitable customers are. Airlines rely on data warehousing to track aircraft maintenance actions, asset inventory, analyze lost luggage patterns and marketing campaign effectiveness. Railroads depend on data warehouses to track freight shipments, asset management and determine "dwell time" or the amount of time that assets are idle and otherwise unproductive. Banking and insurance organizations rely on data warehousing for financial applications, CRM and fraud detection.

Military Uses For Data Warehousing

While the business of today's military is vastly different from civilian/commercial counterparts, many of the commercial data warehousing applications readily apply with very little modification. Tracking logistics and assets, planning maintenance, determining asset replacement cycles, preventing fraud, predicting personnel turnover, financial applications and intelligence operations (civilians call it CRM) are some key areas that data warehousing technology can play a key role.

One such example that is already in use is at the US Air Force Materiel Command. Those who have trouble keeping track of things on their desk should try tracking 6,138 United States Air Force planes and their millions of sophisticated parts. The task is made even more complicated by the sheer logistics involved in monitoring traffic at each of 184 air bases worldwide, logging every air mission, scheduling and recording maintenance, and inventorying parts. When lives depend on having planes where you need them, when you need them—and in good working order—you can't afford to miss a beat.

Before implementing the now twenty-five terabyte Air Force Enterprise Data Warehouse (AF EDW), there wasn't visibility on a timely basis. It could take weeks or months to extract information manually out of all the stovepipe OLTP transaction systems and put it together, and even then, reliability was suspect.

But after visiting Wal-Mart's Arkansas headquarters to view its enterprise data warehouse, senior Air Force officials quickly realized what state-of-the-art data integration could mean for "war fighters." They wanted to provide ready access to any member of the USAF who needs combat support information to perform analysis that supports the Air Force mission. The USAF needed a single place for data comparison to determine which planes are mission capable and what percentage of the fleet is available.

The vastness of its data warehouse is indicated by the 70 million updates it gets each month, in near real time, on matters such as aircraft status, maintenance actions, mean time between failure and the number of sorties a plane flies. The AF EDW also receives 50 million batch updates every month from procurement, finance, supply and maintenance systems. The process of consolidating such vast amounts of information from multiple business systems into a one-stop enterprise information shop has already paid off.

To efficiently get parts where they need to be in Iraq, the Air Force needs complete visibility of the supply chain, movement of materiel, information and money. As an example, the process of tracking a single part through the myriad repair processes, maintenance systems, procurement systems and shipping routes.

Multiply that exponentially, and you can see how complicated things can become. Every day there are thousands of parts needing repair coming out of Central Asia, and thousands of repaired parts going back. This, in and of itself, brings about a multitude of questions:

- Where are the parts?
- What is their status?
- What transportation systems are moving them along?
- Is there money to fix these parts at the depot?
- Are they in the maintenance production line?
- What are the competing priorities of other parts in that maintenance line?
- How will this affect mission capable readiness?

Instead of dozens of systems to track down, it's much better to get to the point of typing in a part number like UPS or Federal Express does to see exactly where a part is and what's going on with it. Hence, the data warehouse solution.

The AF EDW Program Office recognizes the future potential of the Air Force's newfound enterprise view, not the least of which will be greater aircraft availability because of a growing ability to predict equipment failures. This will allow parts to be purchased and shipped ahead of time. As the system grows, it will also alert key personnel when maintenance or performance thresholds are reached. On the weapon system acquisition end of things, the AF EDW will manage a vast portfolio of information related to weapon design, production, upgrade and delivery. Data will include personnel records, financial records and cost-and-schedule briefings (i.e. earned value management and assessment of progress over quarters).

Simple queries can reveal whether money is being spent at the mandated standards rate, who the contractors are and who within the Air Force is using a specific weapons system. There are plans to make the system helpful for formal inquiries into contract disputes. You could look at the histories of all the competing companies— of their compliance and timely performance on prior projects—and in a day you could turn around and give an answer to the inquiring congressional staffer. Before, it might take weeks.

Cost-effective comparisons become more realistic when deciding things like whether to upgrade an older plane or go with a new design. Hence, arguments in the appropriations process over spending priorities should carry additional weight, and ... improved efficiencies and lower costs will have a better focus.

The enterprise data warehouse's potential should become even more apparent as the system matures. The Air Force plans to integrate data from 22 combat support functions, such as health services, legal, munitions, transportation, services, personnel and logistics plans.

The Air Force predicts a 144% return on investment over six years and financial self-sufficiency after that, thanks to cost reductions in other areas as a direct result of the AF EDW. There will also be considerable savings in man-hours, not to mention the advantages that come with instant access to accurate data for decision-making.

Future Direction

The compilation and analysis of un-structured data is just beginning to take hold. This capability will allow the analysis of structured (numeric integer data) along side semi-structured (MS Word, PDF files, etc) and unstructured (JPEG, MPEG, GIF, TIF, etc) files within the data warehouse. Advances within the RFID

industry will also require data warehouses to hold vast amounts of data related to asset inventory and logistics. While the data warehouse has evolved over the past twenty years, its evolution is not anywhere near complete.