Bruce Ullrey of Rumpke Consolidated Companies outlines his 11-point strategy for successfully deploying a data warehouse

Wasting No Time

BY JIM UTSLER

"This story isn't so much about the why," says Bruce Ullrey, director of IT with Rumpke Consolidated Companies, "but the how. That's what a lot of companies have difficulty with—and that's where they get into trouble."

Indeed, most organizations know why they're deploying data warehouses for the purpose of business intelligence (BI) (e.g., to aid in the decision-making process, to create a better understanding of the company, to increase competitiveness) but not necessarily the best way to do it. This isn't because they're unskilled or lack knowledge, but rather because no one has taught them the most expedient way. Sure, eventually, they'll deploy a data warehouse, but it may take a long time and the effort might be cursed with unexpected problems along the way.

Ullrey is no stranger to establishing databases, having been in the technology field for nearly 40 years. That's why he's happy to share his success story, not just as a depiction of transpired events, but as an instructional how-to.

"There's a lot of who, what, when, where and why information out there, but there's little how stuff," he says. "That's why I've established and documented a pretty comprehensive process; I want to teach the younger IT professionals out there. After all, we're all in it together."

The Why

Located in Colerain Township, just north of Cincinnati, Rumpke has been in operation since 1932. Originally a junkyard business based in Carthage, Ohio, it has since morphed into a waste-hauling company that has grown to include more than 1,800 employees and a fleet of almost 2,000 vehicles. One of the largest privately owned waste companies in the United States, the company now owns or operates landfills, transfer stations and recycling centers in Indiana, Kentucky and Ohio.

Operating behind the scenes of the business are two IBM* @server iSeries* systems: an 825 and a 720, both running OS/400* V5R2. The former acts as the company's production

box, running a variety of applications, the most notable of which are, according to Ullrey, "the hauling-operations software, SoftPak, and the financial back-office software, which is from Infinium."

The latter system acts as the data warehouse, running IBM's DB2* Universal Database* (UDB), access to which is gained via Windows* technology-based query and reporting tools from New Generation Software (NGS).

Before establishing a data warehouse, Rumpke was experiencing fairly typical data-usage issues, including a diffusion of data sources. In fact, the data itself wasn't necessarily a problem (financials had theirs and operations had theirs), but the integration of it was. Employees in the financial department, for example, were gathering information from Infinium while operations employees were gathering data from SoftPak. Similarly, metrics used by different operating business units was a problem, with the units often not coming to a consensus on crucial business data. As Ullrey explains: "We had multiple groups looking at different sets of numbers, and nobody could agree on what was really happening in the company, from either a cost or sales side."

Ullrey's proposed solution was to create a single data warehouse and offshoot data marts from which the entire company could derive its information. This would effectively negate the problem involving the "different sets of numbers," giving everyone a single view of the organization no matter the business unit they belonged to. He also wanted to develop agreement company-wide on a single set of measurement metrics to help Rumpke determine its operating success. This would help both mid- and executive-level decision makers make sure they were on the same page at the same time, giving them a bottom-up view of the entire company.

"The data warehouse enables this," Ullrey says. "For example, when different groups sit down to look at the company reports, they know when they're talking about truck costs or about roll-off sales and exactly how those numbers are generated."

Realizing the benefits of a data warehouse, Ullrey began the process of establishing one in the second quarter of 2004, starting with business modeling and metric creation. Within a year, he had a working warehouse that included data marts, easy-to-use reporting tools and an educated user base.

The How

Ullrey took a very deliberate approach. He already knew the benefits to developing a data warehouse (as many other companies do), and based on his years of experience, he also knew many of the best ways to deploy it, devising a general 11-point outline of the process.

The first of these points is remembering the due diligence of the IT profession: That is, identifying the system development life cycle (SDLC) to be used and developing a project plan.

"We also had to demystify our 70-plus years of tacit knowledge," Ullrey says. "That started with me asking for SOPs, but there weren't any. People just did things the way they were taught, with that knowledge being passed down like an oral history of sorts. Nothing was documented. So we set out to create those SOPs and a sense of company-wide consensus."

Ullrey's second process point involves modeling the business. Several methods of doing this are possible, but he says the first step would be to understand the management structure. By doing this, information from the lowest level can more effectively flow up the command structure, from operational supervisors to market-segment managers to executive staff to the ownership level.

"Now, cost and sales data can be rolled up to any level it needs to be, even on a daily basis, for better decision making," he says. "Think of this as an information-reporting and business-intelligence system for all levels of management and supervision in the company."

The third point of Ullrey's how-to database-deployment outline has to do with defining the required metrics. Without strict data definitions, different groups within the organization won't agree on what specific numbers mean or understand how to establish critical success factors. For example, the company can now easily establish metrics related to the cost of running a truck, average revenue per load and average disposal cost per load.

"One problem was that the level of detail in the operations software was much more sophisticated than the level of detail in the financial statements," Ullrey recalls. "So the operations guys were looking at a set of reports out of SoftPak and the finance guys were looking at Infinium reports—and nobody could agree on the numbers. By establishing metrics of costs, whether for truck maintenance or in-house supply costs, the numbers all of a sudden began agreeing with each other."

The fourth step involves sourcing the data, which allows Rumpke to understand from which applications the data is being derived and at what level. Of course, end users don't care about the sourcing, but if they're going to create a bigger picture of company operations, they need access to the data, no matter the derivation. As Ullrey explains, "Because we're de-normalizing the data, we need to know where it's kept, where the details are, so we can put a person's name or truck description into the database, and people who need access to it can access everything they need in one place—it's right there for them."

Step five of the how-to guide involves purchasing the toolset. This may sound like a no-brainer, but this isn't necessarily the case. As Ullrey points out, the proper tools should fulfill many requirements, including comprehensive filtering, metadata capabilities, graphics ability and a simple online analytical processing (OLAP) engine. He further notes that many BI tools are "over engineered," referring to the oft-quoted 80/20 rule, which states that 80 percent of users will only use 20 percent of a tool's capabilities and 20 percent will use 80 percent.

"It's not enough to just give them a spreadsheet; you actually have to go through the spreadsheet with them and make sure they understand how to use the information." —Bruce Ullrey, director of IT, Rumpke Consolidated Companies

With that in mind, Rumpke decided to go with a suite of reporting tools from NGS; he was already familiar with and had been using NGS software. As he notes, "The capability and methodology of NGS's toolset is absolutely enough to do what we needed. Others often involve more costs, more training and more mental gymnastics—things you don't need to get to the substance of running the business."

Steps six and seven are building the warehouse and data marts, which Ullrey cautions should come after the data sourcing and the determination of measurement metrics. The warehouse itself is the central data repository, linking the sourced data with other bits of sourced data. Of course, wading through the warehouse can be a chore, even with the simple-to-use query tool of NGS. Hence the creation of data marts, which Rumpke has aligned with specific lines of business.

"If somebody wants to look at roll-off sales, for example, they don't have to wade through all of the other 13 lines of business to get to those records," Ullrey notes. "They can go specifically to roll off or rear load or front load or recycling data that are specific to a line of business."

The OLAP-template creation is step eight. Rather than relying on users to muck about with the query tools, leaving them frustrated with the process and unlikely to continue use, Ullrey suggests building the initial templates for individual users; they can then run the reports they need based on the templates and, as he puts it, "immediately begin to realize the benefits of the data warehouse." He notes, however, that he learned a lesson during this process, realizing that while creating the templates did indeed simplify tool use, it needs to be backed up with user training regarding how to use the information. "It's not enough to just give them a spreadsheet; you actually have to go through the spreadsheet with them and make sure they understand how to use the information."

Which leads to step nine: training the users. Before the toolset can be deployed (step 10), users must understand what the warehouse is, how to access it and how they can benefit from it. As part of the training process, Ullrey hosts sessions with individual business units. ("But only for an hour or two," he notes, "before their eyes glaze over.") He wrote a 150-page instruction manual to supplement his training sessions. And if that fails, users have his number. "They can reach me anywhere, and I can walk them through any problems or issues," he says.

Step 10, as indicated, involves deploying the tools, which Ullrey simply sums up by saying, "That's just going through the implementation." He adds, however, that the choice of tool depends on the level of complexity the user company wishes to endure. Some may have more robust needs, but others not. For his part, Ullrey went to other companies in the Cincinnati area to evaluate the tools they were using. Based on this experience, he found that "the results of the bigger toolsets have often been disastrous because of the complexity of the implementation and the training. Typically, you just need something simple."

Even with the simplest tools, however, technical support, as point 11 of Ullrey's warehouse deployment outline, must remain available. "That's why everyone knows my telephone extension," he says. He also goes out of his way to visit user sites, for instance, driving five hours round trip to train one user, "because that's how much I believe in the value of the data warehouse."

The Results

Ullrey admits that the deployment of a data warehouse isn't in itself "very innovative," but he does tout its benefits, especially as they pertain to understanding the business, creating additional efficiencies and increasing competitiveness. To him, however, the most important part of the deployment process is, in fact, understanding the process itself; the whys are self-evident, but the hows, he says, are not.

The 11 Steps to Building a Successful Data Warehouse

Step 1: Develop a project plan.
Step 2: Model the business.
Step 3: Define the metrics.
Step 4: Source the data.
Step 5: Purchase the toolset.
Step 6: Build the warehouse.
Step 7: Build the data marts.
Step 8: Create the OLAP template.
Step 9: Train the users.
Step 10: Deploy the toolset.
Step 11: Be available.

"I just want to show people how to do this without them having to go through some of the learning pains I've had to go through in my many-year career." —Bruce Ullrey

"Us old guys," he says, "have to teach the young guys how to do this and keep them from making some of the same mistakes we've made. I just want to show people how to do this without them having to go through some of the learning pains I've had to go through in my many-year career. Who knows? There may be a book in here somewhere."

First-Hand Knowledge

YOU CAN LEARN MORE ABOUT THE experience of Rumpke Consolidated Companies by catching Bruce Ullrey's session at the fall COMMON conference in Orlando, Fla. "Data Warehousing 301: Implementing a Solution—A Customer's Perspective" is Tuesday, Sept. 20, at 11 a.m.

Jim Utsler, @server Magazine *senior writer, has been covering the technology beat for nearly a decade. His responsibilities include writing case studies and other articles of general computing interest. Jim can be reached at jutsler@msptechmedia.com.*