

# Data Warehousing on a Shoestring Budget



NATHAN RAWLING, ISA CONSULTANT

*This article originally appeared as a three-part series in TDWI's BI Weekly*



Do you believe that data warehouse (DW) projects must be multi-million dollar undertakings? Don't. As this whitepaper will demonstrate, any organization, regardless of size, can benefit from data warehousing technologies, even when working with a limited budget.

Although seemingly difficult, you can make choices, which allow for the beneficial realization of data warehousing while also minimizing costs. By balancing technology and carefully positioning your business, your organization can quickly create cost-effective solutions using data warehousing technologies.

There are a few simple rules to help you develop a data warehouse on a shoestring budget:

- Use what you have
- Use what you know
- Use what is free
- Buy only what you have to
- Think small and build in phases
- Use each phase to finance or justify the remainder of the projects

In this whitepaper, we explore the traditional sources of data warehousing costs as well as how your company can best position itself to properly execute data warehousing projects, regardless of budget. By fully understanding these issues, your organization will be better equipped to achieve your profitability goals.

## **Traditional Data Warehouse Costs**

Typical data warehouse project costs derive from two sources: technology and staffing. The balance between these costs varies by a project's environ-

ment, but every project includes both. Technology costs often include system hardware, database-related software, and reporting software. Staffing costs address the individuals who gather requirements as well as those who model, develop, and maintain the data warehouse.

In reviewing the marketplace for data warehousing hardware and software, organizations can easily "break the bank" when procuring these technologies. As tempting as it may be to seek out a large SAN with clustered high-availability database servers and farms of Web servers to support a data warehouse, it is not always appropriate or necessary. It is quite possible, and sometimes far more effective, to build a data warehouse solution on relatively conservative hardware.

Further, although there are a number of outstanding software solutions that address each step in the data flow, it is critical to first find the right toolset before selecting an expensive or feature-rich product. A common organizational mistake in data warehousing is to ignore the currently owned technology and software capabilities and move ahead quickly, if imprudently, to purchase a different product. An organization often has the toolset in place to effectively meet its data warehousing demands but simply needs the right partner to fully implement its current technology.

## **Storage**

It is not uncommon for a large data warehouse installation to employ a SAN with high-performance, fibre-channel communication for primary storage. If your organization has a SAN in place, utilizing it for data warehouse storage may, in fact, be a logical choice. However, it is not necessary nor is it often most beneficial for an organization to build an SAN

solely to support a data warehouse. The benefits of SAN architecture for an organization are often more recognizable when a single storage system can be used to support a significant number of diverse systems rather than a single data warehouse or project. Networked-attached storage (NAS) creates a similar architecture to a SAN. Yet, instead of employing fibre-channel connections, which appear to the server as local disk, the server communicates to the disk via network communications protocols, such as Server Message Block Protocol (SMB) or Network File System (NFS).

This architecture can pose challenges in data warehousing applications, notably the database systems and extract, transform, and load (ETL) software may cause performance bottlenecks. Ethernet communications to the storage device can cause considerable slowdowns that adversely affect overall system performance. If an organization has NAS storage available -- a relatively inexpensive option -- it can still build data warehouses successfully through careful network management and database tuning.

Locally attached storage, with or without hardware RAID (redundant array of independent disks), is among the most inexpensive and attractive storage options because of its outstanding performance. In recent years, as disk storage prices have declined, disk storage availability in each server has increased. Although generally advantageous to separate operating system files from database and application-related files, an entire data warehouse system can be run adequately and efficiently on a single drive. Data warehouses tend to be disk intensive, therefore the addition of an extra drive can provide supplemental storage for database files, application files, and temporary storage while simultaneously improving overall performance.

### **Servers**

Currently, the database and reporting server architecture of choice is fault-tolerant clustering servers with automatic load-balancing and failover. Unfortunately, while such an environment is highly desirable for maximum performance and uptime, it is also extremely expensive and can be difficult to maintain. Although it can be rewarding to design and work within such an environment, this

type of implementation is an unlikely choice for the majority of organizations, but especially for those seeking to successfully complete a data warehousing project on a tight budget.

With fault-tolerant clustering servers an unrealistic option for most, a strong middle-ground option is a server-per-service architecture. This environment performs very well, allowing organizations to independently scale each application while simultaneously keeping costs down. A typical data warehouse and business intelligence environment might require two or three servers in this type of architecture: a database server, a reporting application server, and perhaps an additional Web server. This approach reduces the cost and still provides many of the benefits available in the higher-availability solutions. For example, in the case of a failure, it is possible to restore services on the remaining systems. Further, troubleshooting is typically easier than in a single system design and fits better into security models, which require separation of Web and database services.

In a pinch, however, it is possible to build an entire data warehouse and business intelligence solution on a single server, with or without virtualization. At the outset, this option may appear to save your organization valuable funds because only one server must be purchased. Yet, be mindful -- this singular purchase demands that the server selected must be requisitely powerful. It must handle the load of all processes: database, ETL, and reporting. This architecture can yield acceptable performance, but it can also be quite difficult to troubleshoot. Further, the environment is also difficult to integrate into organizations with well-developed security policies, which often require strict separation of database and Web servers.

Data appliances are receiving tremendous attention and are gaining market share because they can provide solutions for midsize-to-large volume data warehouse markets by offering low-cost performance -- most commonly on data volumes in the terabyte to petabyte range. ISA Consulting has recognized a significant increase in the number of clients -- traditionally, a conservative crowd -- utilizing DW appliances.

One of the benefits of DW appliances is the lower total cost of ownership (TCO) compared to traditional methods. The total cost of ownership of a data warehouse consists of initial entry costs, maintenance costs, and the cost of increasing capacity as the data warehouse grows. DW appliances offer low entry and low maintenance costs with initial costs ranging from \$10,000 to \$150,000 per terabyte, depending upon the size of the DW appliance installed. The resource cost for monitoring and tuning the data warehouse is a significant aspect of the TCO, often more than 50 percent.

DW appliances reduce administration for operations, setup and integration, and many also offer lower costs for expanding processing power and capacity. With the increased focus on controlling costs, combined with tightening IT budgets, data warehouse managers must reduce and manage expenses while leveraging their technology for full maximization. These demands make DW appliances, notably Oracle, Netezza, and Greenplum, a natural solution.

An interesting trade-off with hardware is that often one of the least burdensome components of a data warehouse is upgrades or replacements. Upgrading disk storage, adding memory or processors to a server, or simply installing additional servers to an environment rarely cause as much work as fully changing a database or reporting platform. This is not an excuse to intentionally under-power the hardware environment -- a situation likely to cause complications. Rather, this should temper the tendency to overbuild system environments for data warehouses. If cost is a driving force, it is important to recognize that real and significant savings can be achieved simply by purchasing the hardware needed today rather than the hardware that might be needed next year.

### **Software**

Although selecting an operating system (OS) and selecting hardware should be considered concurrently, the OS is nonetheless true that it is a software component. Many systems integration professionals have a personal preference regarding the ideal operating system for data warehousing and business intelligence. The reality, however, is that almost any common platform can be successful for all but the most extensive data warehouses.

The best way to maximize a limited budget for operating systems is to use the system your organization is most familiar with and select the minimal feature-set necessary to support the additional applications within your environment. Branching out into an entirely new platform for your data warehouse implementation will increase expenses for migration tasks such as staff training.

The database is the heart of the data warehouse. Similar to the choice of an operating system, professionals tend to advocate a particular vendor when an organization must choose a database. Not every data warehouse can run on any/every database platform, yet all major database platforms are sufficient for the majority of data warehouses. If your organization has a well-developed relationship with a particular vendor, maximize the opportunities afforded by that relationship can help to minimize technology, training, and implementation costs. In fact, many of the major database vendors offer different levels of software with additional features at various price points.

Carefully examine and review exactly which features you need prior to ordering the most expensive version; in some cases, it is quite possible to simply purchase the additional features you need without buying a product bundle.

All major database vendors offer ETL and reporting products and your organization can and should take advantage of product bundles whenever possible. If a database product you own comes with integrated ETL software, it makes little sense to buy an additional ETL product until your organization is confident that its current software will not successfully meet demands. The same rule can be applied to purchasing reporting tools, especially since the market for these products continues to be very competitive.

Another typical DW expense is for database modeling software; many include sophisticated features, including cross-platform support, version control, and reverse/forward engineering of database objects. In the modeling of a large transactional system or a complex data warehouse, the advantages of such tools are evident, because they can easily manage large amounts of metadata. Unfortu-

nately, such features come at a dear price. It is critical that the software products selected do meet your organizational needs: unintelligible database model diagrams can slow development. There are free and inexpensive commercial modeling software products which save valuable resources, such as Computer Associates' ERWin and Embarcadero's ER/Studio. Both are outstanding tools. However, Microsoft Visio Professional is quite capable of producing effective star and snowflake schemas at a quarter of the cost.

To process data from source systems into the data warehouse, organizations must use ETL. There are specialized products to simplify development, ranging from simple automation tools to full-fledged suites of applications, with extraordinary capabilities. These specialized tools, however, have expensive price tags that vary from tens of thousands to millions of dollars, depending on the number and type of servers. There are four ways in which an organization can potentially avoid paying costly licensing fees:

- Do not specifically use a tool. Instead, use the procedural code built in to your database for ETL development. Keep in mind, though, that working to manually code ETL logic in a procedural language may take far longer than developing ETL by simply using one tool within the bundle.
- If you are able, use a tool already bundled with your database. Microsoft SQL Server 2005 comes with SSIS and SSAS; Oracle Database versions 10g and later come with Oracle Warehouse Builder. There are other bundles available today; it is simply a matter of fully investigating all of the options.
- Check to see if a department in your organization has previously licensed an ETL product, and if your project could use those licenses or buy additional licenses at a reduced cost.
- Some organizations are building internal "centers of excellence" that offer free or low-cost infrastructure, licenses, and data warehousing expertise to the organization. Many times these groups are not well-known throughout the organization, but can be extremely helpful. Before making a substantial new investment in DW technologies and staff, make

sure you check inside your organization for such a group.

If circumstances are such that you need to purchase an ETL tool, make sure to look at ETL tools such as Oracle Warehouse Builder (OWB), Cognos 8 Data Manager, and Microsoft Integration Services, all of which seek to challenge the market leaders, including Ab Initio, IBM DataStage, and Informatica PowerCenter. The latter tools are excellent "work-horses" for enterprise data integration, but your organization may not need to begin with one of these tools for a small project.

### Reporting

Presumably, once data is incorporated into the data warehouse, users in your organization will want to access and utilize it. Such a goal implies that a reporting software package might be required. Similar to the specialized ETL tools, it is possible to spend a significant portion of your organization's budget on the most highly functioning reporting solutions. Still, the options for reporting software on a limited budget remain much the same:

- Buy a commercial product
- Use a reporting tool bundled with the database
- Build the reporting tool from scratch

If complex, formatted reports are needed, however, it may be necessary to purchase an advanced commercial package. If the reporting needs are more modest, a less expensive solution may suffice.

When considering software and vital tools such as ETL or a reporting system, organizations must consider the timing and duration of the data warehousing project. The complexity of the project should dictate and justify the purchase price of a given tool. The most important points for a data warehousing project on a limited budget are assessing an organization's technological needs, understanding the intricacy and density of the project, and making informed, cost-effective decisions regarding servers and software.

Organizations must be prepared to balance the cost savings of purchasing less-expensive tools by accepting the increased development effort and staffing requirements associated with a data

warehousing project. In the next section, the human costs derived from a data warehousing project will be examined. We will offer advice for how employees and a budget can be maximized to successfully complete this initiative.

As explained in the first part, organizations can maximize a limited budget when purchasing hardware and software simply by staying focused on immediate organizational needs and wisely selecting tools that meet these exact requirements. However, there is another important aspect of a data warehousing project where organizations must be accountable and, frankly, careful, in their budgeting -- the human costs.

What are human costs? Simply, they are costs inherent in needing and having individuals select data warehousing tools, implement these tools, and train staff in the proper use of this technology. To support your data warehouse project you will need to find individuals to fit these roles in your organization, hire additional staff, or outsource these needs.

### **Individual Roles**

Traditionally, a data warehouse implementation requires many roles: a sponsor, business analysts, database administrators (DBAs), data modelers and/or data architects, ETL developers, reporting/business intelligence developers, and testers. In a complex and sizeable project, many individuals may be required in each role. For a smaller project, however, a large staff clearly would not match the project's needs. It is simply enough to ensure that every role is covered with sufficient personnel to complete the effort in a timely manner. In planning a data warehousing project, these roles and subsequent responsibilities must be considered when balancing your human costs.

**Sponsor:** A DW sponsor serves as the business representative from the office or department of an organization that seeks to benefit from the data warehouse. The sponsor is active throughout the effort but is focused on facilitating the relationship between the warehouse builders and the end users of the project. The time required for this important task cannot be underestimated. From organizing interviews for requirements gathering to reviewing

proposed designs to approving testing paradigms and organizing training, the sponsor is responsible for some of the most critical activities within a data warehousing implementation.

**Business Analyst:** The tasks of gathering requirements from the business, assimilating them, and producing coherent documentation for the team members is traditionally performed by a business analyst (BA). In some environments, the BAs can take on even more responsibility, such as data profiling, data mapping, and testing. In many ways, however, the role of the business analyst is most critical in successful data warehousing projects because they translate the needs of the business to the technical team.

**Database Administrator (DBA):** The database administrator manages the database and, in some cases, the server which houses the database. In a DW project, the DBA is likely to perform capacity planning, create databases and schemas, and recommend/create indices. DBAs are also generally responsible for the security of the database, so the DBAs need to be involved in developing and implementing access requirements and policies.

**Data Modeler/Data Architect:** In a DW project, data from many source systems is taken and delivered into the warehouse or data mart. Understanding the source systems and developing the target schema requires a strong background in data structures and knowledge of data warehousing strategies for data storage (star and snowflake schemas). Depending upon the complexity of the source systems, it's likely that an organization will need a small effort from this role relative to other tasks. Most DW projects will not require a full-time data architect, but instead will require data architecture activities be completed at certain points throughout the project.

**ETL Developer:** The ETL developer creates the data transformations using the requirements from the business analyst. Again, depending upon the complexity and number of the necessary transformations, this is frequently one of the longest-running efforts in a DW project. Use of ETL tools such as Informatica or Data Stage can certainly speed development but can drive up software

costs as well. Conversely, manually coding ETL using procedural languages can be very slow. Carefully balance these two expenses to minimize the overall total cost for the organization.

**BI Developer:** The BI or reporting developer, the counterpart for the ETL developer, is responsible for the front-end reporting solution that delivers data to the end users. Using the requirements provided by the business analyst, the BI developer designs and develops reports, analytical tools and dashboards so that end users can easily access the data warehouse. Frequently, this activity is second only to the ETL development in effort.

**Tester:** The ETL and BI developers generally perform unit testing -- testing the individual components in development. However, they rarely perform end-to-end validation, which is testing of the whole system from source to final output reports. Such stringent testing requires test plans and careful execution to ensure that the DW solution effectively meets the business requirements, which is where the position of tester comes into play on the DW team. In many cases, testing does not require substantial effort, but unfortunately the amount of time allocated for testing is nearly always underestimated. Sufficient testing to ensure that the requirements have been met exactly presents a significant undertaking.

**End User:** While the role of the end user seems obvious, it is not uncommon for them to participate in requirements gathering sessions, late acceptance testing and systems training. Although important, time spent to aid in DW implementation can certainly add up and reflects time taken away from other job responsibilities.

### Combining Roles

Small teams can produce stunning results when faced with appropriately sized projects. The key to success, though, is that the implementation team must have the right composition, with the right mix of skills, and time to complete all necessary tasks. When attempting to consolidate roles into individual team personnel, the most important considerations are the following:

- Total effort required by role in current project
- Concurrency of efforts

- Available skill combinations
- Traditional responsibilities/skill overlaps

Three activities within the typical data warehousing project constitute the bulk of both effort and time: requirements-gathering and analysis, ETL development, and BI development. ETL and BI developments cannot truly begin until the requirements are gathered completely. Therefore, it is possible to have the same resources perform the requirements-gathering and other roles in the implementation without necessarily impacting the project timeline. However, it is not uncommon for the ETL development and BI development to execute concurrently to speed project completion. You eliminate this type of project plan if you combine the ETL developer and BI developer roles into a single team member.

When building a small team to implement a data warehouse, it is possible to effectively complete this type of project with as few as two people in addition to the sponsor and the end users. Yet, while it is possible to run a single-person project, it is rarely advantageous to do so, especially since timelines tend to be very long.

**The DBA/Architect/ETL Developer:** Aside from the previously mentioned option of ETL developer and/or BI developer, a common scenario is to combine the various intensive database roles. Many ETL developers have some prior experience with database administration and data warehousing methodologies. However, not every ETL developer has all of these skills. Yet, finding one that does possess these skills can potentially save adding two full-time employees to your payroll. The typical data warehouse schema is not a complex data model to design and the task of creating tables within the database is typically within the capabilities of most developers.

There are risks to this approach, however. If a complex database issue arises, it will take valuable time away from ETL development. It may also simply be beyond the database administration ability of an ETL developer. Furthermore, even as the target schema in a DW implementation tends toward star-schema simplicity, the source databases can have varying degrees of complexity, and therefore an experienced data architect might be needed.

Still, in a typical data warehousing project, merging these roles makes sense if you do not require a full-time data architect and DBA on your project.

**The Business Analyst/BI Developer/Tester:** Although ETL developers tend to have a somewhat stronger technical background, the BI developers tend to come from a business background. For this specific reason, it is sometimes simpler for BI developers to take over the responsibilities of interacting with the business. Moreover, as end-to-end testing cannot take place until the reporting solution is completed, it makes the most sense to allocate that specific role to this person as well. When troubleshooting, the combination of these roles will be beneficial to the organization as these resources, typically, must work together to correct any deficiencies.

When a DW project involves multiple, complex source systems or a sophisticated business transformation, it is not always appropriate to ask a business focused individual to complete all of the requirements' analysis outside their area of expertise. In those circumstances, assigning requirements-gathering duties to the database person or sharing those responsibilities will likely yield superior results.

### **Outsourcing**

Going outside of the organization is a different yet excellent way to access DW professionals. Although per-hour rates can seem steep at times, the savings from hiring a dedicated team can be substantial. Outsourcing can be a terrific way for an organization to begin a data warehousing initiative as it later takes a strong first step toward building an in-house team to manage the solution.

Another scenario where outsourcing is beneficial is when an organizational project requires extra help to fill a particular role or requires overflow capacity. If you have a business analyst, DBA, ETL developer, and BI developer available, but your organization still needs a data modeler who is capable of mapping complex source systems, outsourcing makes sense. Similarly, if the ETL development will be more substantial than your existing resources can accomplish, bringing in additional experts to supplement your team members can allow your organization to meet demands within the project

plan without being concerned about having excess/ idle team members at other times in the project.

Checkpoint consulting (consulting provided by an expert on a part-time basis over the life of the project) has proven to be an excellent way of benefiting from the vast experience of an enterprise DW team while paying for only a fraction of a consultant. An experienced consultant with broad and deep knowledge and experience can evaluate your existing data warehouse and business intelligence solutions, make recommendations, and, ultimately, help you leverage your existing assets.

Each of these roles is critical to the proper implementation of a data warehouse, but each is also a key aspect of the human costs associated with each DW project. Clearly, these roles are important and as much as an organization attempts to maximize its budget through cost-effective decision-making, one area where an organization must be willing to make an investment is in its DW implementation team. This is an investment, although seemingly too significant at first, which makes all of the difference when considering a project's success. When operating on a limited budget, it is the organization's responsibility to find a way to maximize its people for the benefit of a project as opposed to the detriment of it.

Making informed decisions about hardware, software, and staffing, as detailed in the prior sections parts of this paper, are important facets within every data warehousing project. However, a final (and critical) aspect of completing a data warehouse on a shoestring budget is choosing the right project for your organization.

There are data warehouse requirements that are beyond simple limits -- a project with several source systems on different mainframe platforms; data with multiple languages and character sets; extreme availability requirements; and vast data volumes. These are just a few examples of such overwhelming requirements, particularly in combination, which make shoestring budgets impractical or unsatisfactory. For most organizations, though, there are several opportunities for relatively straightforward data warehousing projects that do not require a substantial investment to produce results.

## Success Factors

An organization building a data warehouse on a limited budget must select projects that meet four requirements:

- Provide a clear benefit to the business
- Utilize a small number of source systems
- Work with source systems that are well understood
- Allow creation of straightforward data maps

The goal must be to build a data warehouse that will allow your organization to begin realizing the benefits of such an implementation with a minimum investment. By building slowly and experiencing a few small successes, organizations will be able to better justify making larger investments in data warehousing and pursuing future projects of larger scope.

When success factors are not considered and/or pursued, a common “horror story” of data warehousing implementations often emerges. The process seems to be endless, utilizing precious budgets and resources, without bringing any real benefits to your organization. Perhaps it is the result of scope creep, technical breakdowns, shifts in organizational direction, or simply communication failures. Regardless, addressing success factors and breaking down projects into smaller phases can help you avoid the “warehouse without end” scenario and allow for ready execution of projects in a timely, cost-effective manner. A careful look at the business benefits of the data warehousing project will help you to choose an appropriate project.

## Business Benefits

A data warehouse is rarely an end in itself. It is a tool to help your organization handle data more effectively, answer questions, analyze data, and achieve success. A thriving data warehousing project must bring data together or reorganize data so it can be used in new and innovative ways.

Consider the real life importance of the data warehouse. If customers are stored in one system, but sales are stored in another system, it can be difficult or impossible to report on sales for each customer or per channel. Similarly, if sales are stored separately from production costs, calculating the unit margins can pose a real challenge. Simply put,

the goal of the data warehouse is to make an organization operate better.

## Selecting Initial Projects

Finding the right data warehousing project to begin with is never easy. However, many common requests for a data warehouse can serve as potential starting points, such as financial reports, management dashboards, purchases of third-party data (such as market-share information), and cross-functional reporting. A major part of managing any data warehousing project is controlling the scope. These example projects all lend themselves to situations where the needs of the users can be separated into “must-have,” “nice-to-have,” and “future-capabilities” categories to isolate the immediate business needs from features with possibly limited returns. Focusing development on immediate needs make it possible to deliver a data warehouse solution with minimal scope creep on a tight budget.

## Financial Reporting

Quarterly and biannual financial reports often bring together many sources into a spreadsheet. In addition to the sheer volume of manual effort and time required to complete such reporting requirements, the approach can introduce typographical errors and frequently, the visual appeal of the final product is unsatisfactory. Often, though, the most difficult challenge in transitioning this process to a data warehouse is collecting the business rules contained within the spreadsheets. If the required raw data is available in existing source systems, the potential savings can make such a project truly shine.

Certain financial reporting projects, can realize substantial benefits to the organization through simple consolidation functionality and available analytical tools, especially when the number of sources is limited (i.e., one budget system and one general ledger). Also, the potential savings from eliminating manual processes from the organization’s account practice can justify data warehouse expenditures.

## Dashboards

Dashboards pose a unique challenge. Typically, they summarize data at a high level from source data that

is finely detailed. Sliding windows, fall-back criteria, and other rather difficult requirements can quickly complicate a dashboard view. Dashboard performance is typically a challenge as well. Running a dashboard directly from a live system may provide a popular real-time view, but to generate monthly, quarterly, or yearly statistics, the dashboard will have to scan enormous volumes of source data and aggregate it.

The visual appeal of dashboards is hard to deny, though. The key to maximum DW performance is for dashboard developers to gather detailed data into the warehouse, where it should be aggressively aggregated, rather than forcing the reporting suite to perform the aggregations in real-time. In some cases, the dashboard's source data might come from only one or two systems. However, those systems are transactional and not capable of producing answers rapidly enough to satisfy end users. In such situations, data warehousing techniques can quickly and efficiently show benefits. A report that in the past might have taken 45 minutes can be optimized to run in seconds just by creating high-level aggregate tables.

### **Third-Party Information/Unstructured Data ETL**

Many organizations receive data from third parties. Whether from electronic orders, through electronic-data-interchange (EDI), purchased data, customer profiles, or other standardized data where the format is "take what we give you," the data from third parties is almost always produced in a structure convenient to the third party. This format is not likely to be ideal for every organization, which makes third-party data tricky to work with.

Unstructured data from documents such as reports, spreadsheets, Web sites, and other non-database sources can similarly hold useful information that is difficult to access in a consolidated way. Unstructured data ETL is a relatively new feature for traditional ETL tools, but it can provide a view into this repository of business information which was previously unavailable.

However, DW technologies can combine such data with internal systems so end users can make better, more-informed decisions. The combination of internal relational data with unstructured data and/or

third-party data typically means building conforming dimensions, which is an ideal opportunity for a data warehouse.

### **Cross-Functional Reporting**

Many organizations have data stored in silos by function. Your organization may, in fact, have a database of customer contacts used by your sales team, a scheduling system used by field technicians, and a general ledger application used by the financial team.

Unfortunately, such data silos do not reflect how a business truly operates. At some point, a call to a customer generates an order; a field technician is scheduled to install a system; and invoices are sent. How can an organization figure out the average revenue per sales call or the average cost per installation? Can an organization detect fraudulent transactions in any of the three silos without looking in the others?

A data warehouse can become a powerful tool in cross-functional reporting as it can help an organization answer these types of questions and provide a more coherent view for management.

### **Review**

It is certainly possible to build a data warehouse on a shoestring budget. Discipline when choosing technology, accounting for the human costs of data warehousing, and selecting the right project can help your organization reap the benefits of data warehousing without becoming financially overextended.

Each step in the data warehousing lifecycle allows you to balance requirements and costs. Moreover, taking a fresh and realistic approach to reviewing current organizational capabilities can help you avoid spending on unnecessary technology features or selecting projects that simply are too complicated to justify until more resources are available.

Data warehousing can be an outstanding solution for organizations seeking to maximize its information. It is, however, a process that should be undertaken with a full understanding of what is needed and how organizational demands can be reconciled with the bottom line.