

DATABASE SOLUTIONS

a white paper covering

Market, Competitive and User Trends and Issues for

**Data Warehousing • Decision Support •
Business Intelligence • Knowledge Management**

By Michael P. Burwen

Palo Alto Management Group, Inc.

July 1998



This document is copyright but may be freely distributed and quoted from as long as attribution is given to Palo Alto Management Group, Inc.

INTRODUCTION

The advent of reliable, affordable and highly scaleable servers, coupled with the growing popularity of the client/server paradigm and software that can scale with the hardware, is driving the rapid growth of an allied group of database-oriented applications. These are variously known as Data Warehousing, Data Marts, Decision Support (or DSS), OnLine Analytical Processing (or OLAP) and various other terms. We call them, collectively, *Database Solutions*.

While applications definitions often vary from one vendor to another, and there may be functional differences between them, in practice these applications frequently overlap. Nonetheless, the context of Database Solutions as used herein refers to the process of aggregating databases of information and efficiently extracting and and/or analyzing the data in order to support one or more business operations. Fourteen months ago, PAMG completed a landmark study focused on the high end of the business (i.e., >100 GB of useful data) entitled *Large-scale Database Solutions (LDBS)*. The present study covers applications based on as little as 10 GB of useful data and is called *Database Solutions II* to indicate that it is the second in the study series.

Regardless of the terminology, the industry has embraced this market as one that offers a very large and growing market opportunity. In the course of our work in the field, we identified more than 200 companies offering systems, software or services aimed at exploiting this opportunity. The majority of these companies are less than five years old.

Although Data Solutions is already a large market – roughly \$15 billion in calendar 1997 - , it is still in a relatively early stage of development. The market continues to be plagued with confusing claims and

counterclaims so that the true nature of the market's development and the relative position of the vendors is practically obscured from vendors and customers alike.

The primary objectives of the research conducted for *Data Solutions II* were to:

- Assess user perceptions of the major competitors
- Define user procurement decision processes and implementation plans
- Bring the definition of the market into clear focus
- Forecast the worldwide market in detail through the year 2002 for systems, software and services
- Furnish the market shares of the major competitors

The research for the study focused on three things:

- ***An in-depth Market Forecast*** through the year 2002, based on a formal I/O (Input/Output) econometric model
- ***Competitive Analysis*** focused on those companies having a 10% or greater share of the systems, services or packaged software market segments.
- ***An in-depth web-based user survey*** addressing subjects such as future plans and budgets, issues concerning procurement, implementation and technology, and perceptions of vendors.

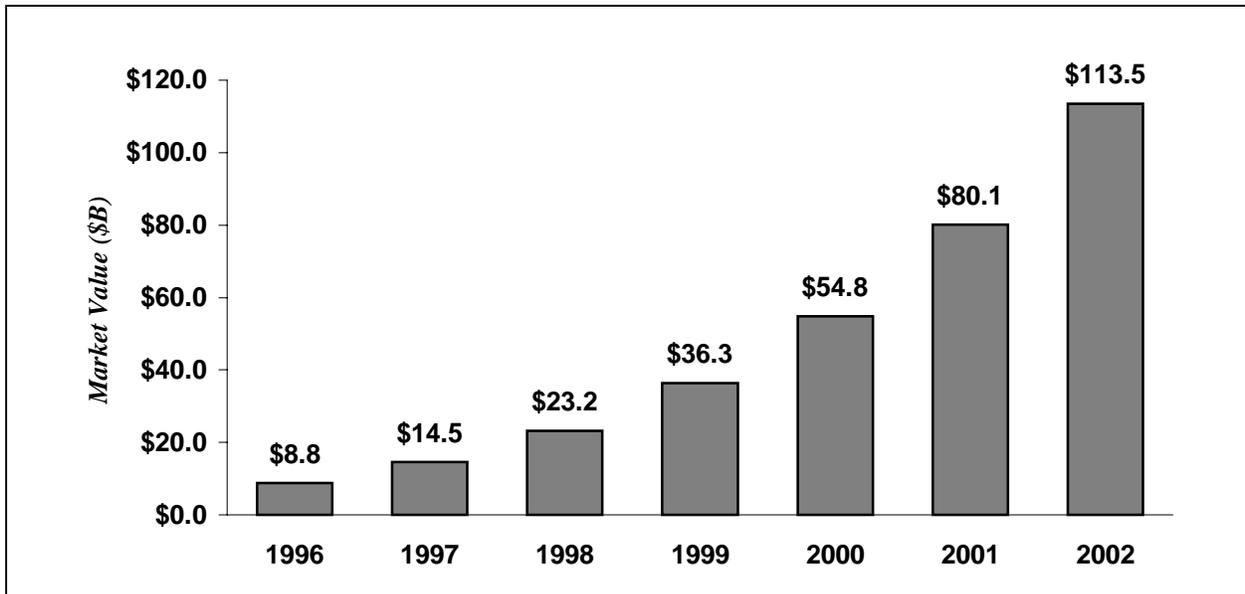
The scope of the study is worldwide, and original, primary research was carried out in North America, Europe and Japan.

MARKET TRENDS

The worldwide market for Database Solutions expenditures is forecast to grow at an AAGR (Average Annual Growth Rate) of 51% over the period 1997-2002 to more than \$113B as shown in Exhibit 1. This forecast is consistent with the findings reached in last year's study that considered only those systems having at least

100 GB of useful data. Thus, we continue to believe that this sector will become one of the most important application markets in the computer industry, particularly in the segment that relies on high-end server-class platforms and large-scale DBMS software.

Exhibit 1
Worldwide Database Solutions Market



The view of the market taken in this study is from a user/customer (not end user) perspective. We view the market as made up of four segments which we call *Product Types*. The sum total of the markets for these four categories is equal to the amount of money spent by customers on solutions, not including the money spent on end users' equipment, interface or time. The four categories are:

1. **Systems** including platform hardware and operating system.
2. **Packaged Software** including DBMSs, Query tools such as SQL, OLAP,

ROLAP and Data Mining tools, backend tools for tasks such as data cleansing, extraction and transformation, and Vertical Applications software. The latter category frequently includes packages from the other categories wrapped in an industry- or function-specific front end.

3. **Services** including business consulting, design and implementation (software development), technical support consulting including customer training and education and systems integration services including installation.

- 4. **In-house expenditures**, which may also represent a service, market opportunity.

Exhibit 2 shows how we forecast the relative value of these segments to be distributed in 2002. Software will be the largest, followed by systems and services. The services share could be substantially larger if services vendors succeed in displacing more in-house activity than we project, or smaller if customers decide not to outsource as much as we think they will.

Exhibit 2
Worldwide Data Solutions Market by Product Type, 2002

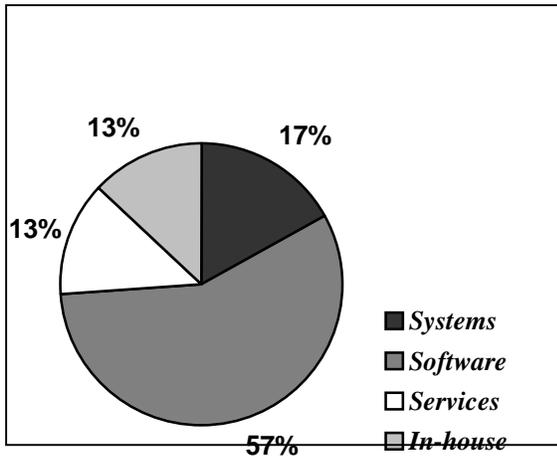


Exhibit 3 contains the forecast breakdown by geographic region projected for the year 2002.

Exhibit 3
Worldwide Database Solution Market Distribution by Geographic Region, 2002

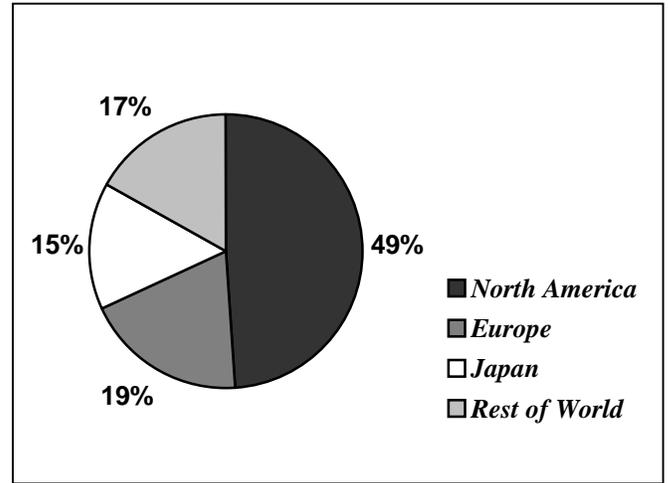
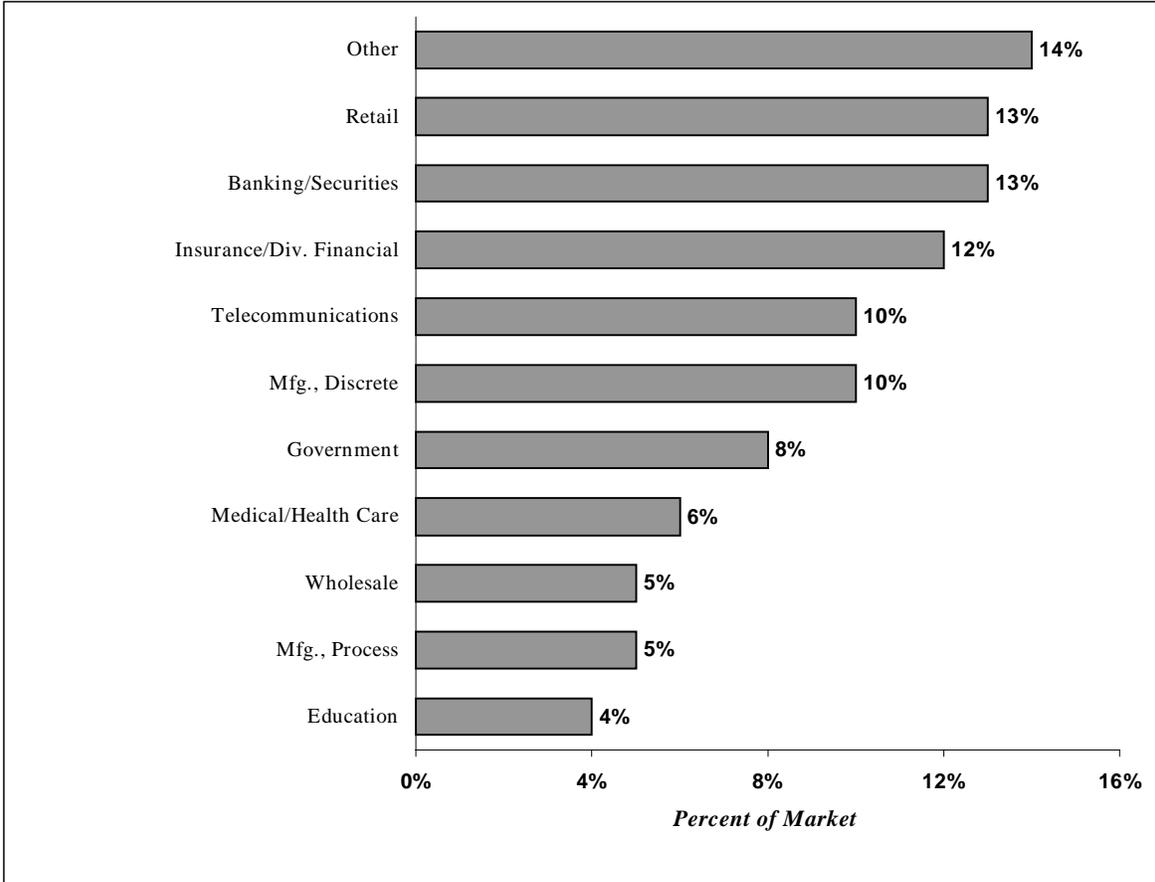


Exhibit 4 is a market breakout by industry sector. This breakout varies, of course, by region and even within countries depending on the economic mix.

Exhibit 4
Data Solutions Market by Industry Sector, 2002



The total market is broken out by *system architecture* into the following four categories.

- **Central Data Warehouse.** Sometimes called an "Enterprise Data Warehouse," this configuration contains a large data warehouse that typically contains data from several operational database sources and serves the needs of a broad spectrum of users.
- **Distributed System.** The data warehouse is distributed across multiple platforms often remotely located from one another.
- **Data Warehouse + Data Marts.** Data from a Central Data Warehouse is

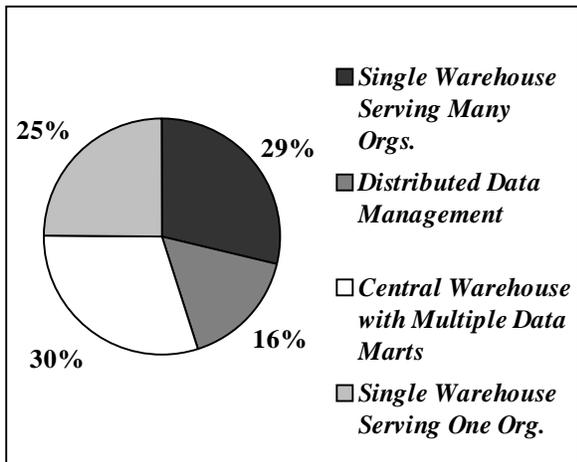
extracted, transformed and fed into smaller systems called Data Marts, each designed to serve a single user or small subset of users, typically a single department or function.

- **Single Warehouse or Data Mart.** A system usually fed by a single source of operational data and designed to serve the needs of a single department or function within an enterprise. Such systems are typically relatively small, but are sometimes quite large.

Exhibit 5 shows how the market will be divided into these four architectural categories. The data shows that systems based on central data warehouses will dominate the market. At first

glance, one may think that this flies in the face of what seems to be an increasingly popular notion; namely, that much of the market is moving to small data marts because of their relatively low cost and easy implementation. However, while data marts might cost, say \$500,000, big data warehouse installations cost many millions, sometimes as much as \$100,000,000. It takes a lot of little data marts to equal the market value of one big data warehouse. Furthermore, with experience, it will get easier and faster to implement big data warehouses, so that some of the motivation for installing data marts will diminish.

Exhibit 5
Worldwide Data Solutions Market by Architecture, 2002

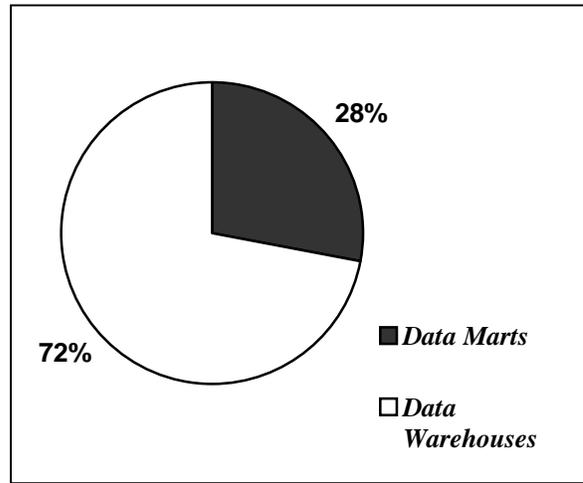


The trade press in general and some of the vendors in particular has made much of the “phenomenon” of the data mart. An impression is often created that marts are easy to implement, while data warehouses take a long time to implement, are expensive, and, are at best risky propositions. These arguments are most often self-serving and have little basis in hard facts. In fact, we view the expression “Data Mart” as a marketing term without any technological basis to distinguish it from a data warehouse.

If we use the definition of a mart as a single-purpose warehouse serving a single allied group of users, however, we can make some estimates as to the relative sizes of these two markets. The

data is presented in Exhibit 6. It shows that, even at the end of the forecast period, marts will account for only a little over a quarter of the market.

Exhibit 6
Marts vs. Warehouses, 2002

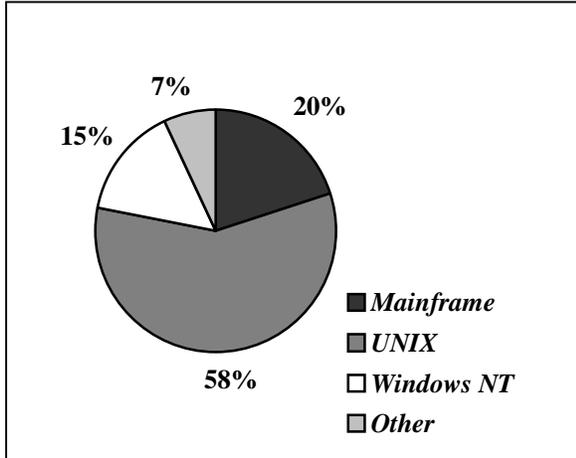


The market is cut by operating system platform into the following four categories:

1. **Mainframe** (principally MVS and plug-compatible variants)
2. **UNIX**
3. **Windows NT**
4. **Other**

Articles in the trade press and speeches at trade conferences would have one believe that the Database Solutions market is today basically a UNIX-based business with Windows NT to come on strong in the next couple of years. The forecast presented in Exhibit 7 strongly refutes those notions. The data shows that mainframes currently have nearly one-quarter of the market, and, even by 2002, will still account for about 20% of the market!

Exhibit 7
Worldwide Database Solutions Market by Operating System, 2002



Consider the following:

- The market is mostly big-system based, and mainframes still dominate the big system market.
- It is a path of least resistance for many existing mainframe sites to add a processor complex or two rather than to bring in a new technology.
- Mainframes address the availability issue better than any other platforms except fault tolerant systems, and availability is a key purchase criterion.
- The Price/performance and scalability of mainframes is approaching that of other high-end architectures
- The cost of the hardware platform is less than 25% of the total solution cost. Thus, even if the hardware/OS cost of a mainframe solution is twice that of an alternative solution, the impact on total cost will only be in the 10-15% range, and the cost of retraining personnel might well offset that difference.

Nonetheless, UNIX is the single largest OS platform, accounting for 59% of the value of

1997 expenditures. That share is projected to decrease only slightly to 58% by 2002.

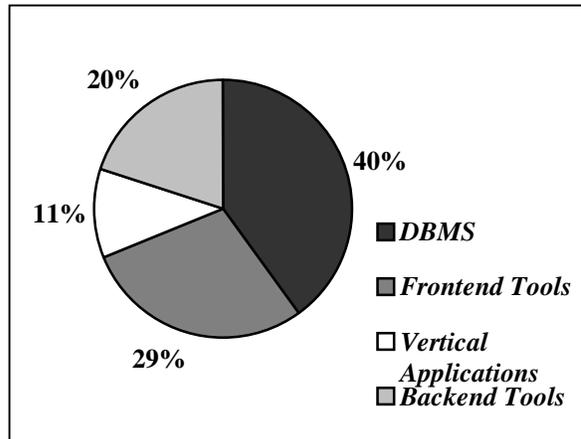
"Other" systems, e.g., Tandem, Teradata, White Cross, AS/400, etc., currently account for 13% of the Database Solutions market. That share will gradually decline to about 6% by 2002. By that time, however, much of the proprietary system business will consist of upgrades of existing systems as opposed to new installations.

Windows NT will capture 15% of the market by 2002. This forecast presumes that NT will soon be much improved in its ability to scale, in availability characteristics such as fault resilience and in the area of system management/administration.

The prevailing wisdom suggests these enhancements will not become generally deployable before late 1999. Thus, the forecast for NT may well be generous. It is based, in part, on Microsoft's well-known marketing skills. Remember that this discussion is focused on the server side. NT will, of course, be the major player on the client side.

Exhibit 8 contains the packaged software expenditure distribution forecast for the year 2002. The DBMS category leads with a 40%

Exhibit 8
Worldwide Database Solutions Packaged Software Expenditures Distribution, 2002



share followed by Frontend Tools (29%), Backend Tools (20%) and Vertical Applications (11%). This split is based on data from the user survey.

This said, the forecast for Vertical Applications is probably quite conservative. The reason is that business solutions are nearly always become more function-specific and "verticalized" over time. One can see this happening from the responses of some of the leading system and DBMS vendors which have formed solutions groups to support specific industries and or cross-industry functional areas, and by the leading services vendors, many of which have formed industry-specific "practices."

The software product segment will become the largest product segment in market terms by the end of 1998. It is the fastest growing segment and is expected to exceed \$60B by the end of the forecast period.

Exhibit 9 is the forecast breakout for four categories of services. System Integration is the leading category, accounting for 40% of revenues - \$7.4B in 2002. This position is comparable to the one that the category occupies in other areas of large system and network deployment, and reflects the current trend toward outsourcing.

The second largest category is Implementation Consulting. This area is driven by several factors, including:

- System complexity
- The need to accommodate an existing operational data infrastructure
- Complications in the data cleansing, extracting and transformation processes
- Moving to new platforms; e.g., a site with its operational data in a mainframe environment and its data warehouse in a UNIX environment.

The Technical Support and Business Consulting categories each hold a 15% share of the forecast services revenues.

Exhibit 9

Worldwide Database Solutions Services market,

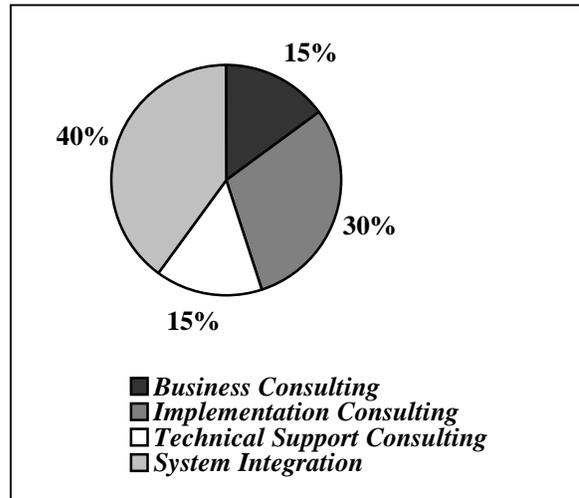


Exhibit 10 shows how the market breaks out by application. It shows that the largest applications fall into the financial and marketing areas.

Exhibit 10
Worldwide Database Solution Markets by Application, 2002

<i>APPLICATION</i>	<i>PERCENT OF MARKET (%)</i>
Financial Analysis	11
Customer Marketing	10
Product Performance	8
Sales Force Analysis	8
Quality	8
Promotion/Marketing Campaigns	7
Risk Management	6
Vendor Performance	6
Churn Analysis	6
Category Management	6
Market Basket Analysis	6
Supply Chain Analysis	6
Fraud Detection	5
Other	3
Yield Analysis	3
Network/Traffic Pattern Analysis	2

COMPETITION

There are literally hundreds of vendors vying for a piece of the Database Solutions market. In 1997, the leading vendor worldwide was IBM with an estimated \$2.4 billion in revenues, up from \$1.3 billion in 1996. Oracle, HP and NCR, each with revenues in the \$750-800M range, follow IBM. These companies participate in the systems, packaged software and services segments.

Exhibit 11 contains a list of eight vendors that had revenues of \$300 million or more derived from the Database Solutions market in 1997.

Put together, these vendors accounted for 46% of user expenditures for Database Solutions, and 64% of all user expenditures excluding the in-house category. It is likely that these vendors will continue to be major players for several years to come.

Nonetheless, there are many other companies that might become significant players in the market, even if they don't make the top player list. These vendors are listed in Exhibit II-12. Note that the figures for Compaq include Tandem and Digital.

Exhibit 11
Leading Database Solution Vendors, 1997

VENDOR	ESTIMATED 1997 DATABASE SOLUTION REVENUE (\$Millions)			
	TOTAL	SYSTEMS	SOFTWARE	SERVICES
IBM	2400	1800	360	240
Oracle	800		600	200
NCR	750	300	150	300
HP	750	600		150
Compaq (inc. Tandem, DEC)	700	450	50	200
Sun	450	350		100
SASI	400		400	
Andersen Consulting	400			400
Total Leaders	6650	3500	1560	1590
1997 Worldwide Market Value	14460	4125	3485	2790
% held by leaders	46	85	45	57

Exhibit II-12

Vendors That Could Become Top Tier Data Solutions Vendors

<i>VENDOR</i>	<i>SYSTEMS</i>	<i>SOFTWARE</i>	<i>SERVICES</i>
Arbor Software		✓	
Brio Technology		✓	
Business Objects		✓	
Cognos		✓	
Computer Associates		✓	
Coopers & Lybrand			✓
CSC			✓
Data General	✓		
Deloitte & Touche			✓
EDS			✓
Ernst & Young			✓
Fujitsu/ICL/Amdahl	✓		✓
Hitachi/HDS	✓		
Informix Software		✓	
KPMG Peat Marwick			✓
Microsoft		✓	
Microstrategy		✓	
NEC	✓		
Perot Systems			✓
Pilot Software		✓	
Platinum Technology		✓	
Price Waterhouse		✓	✓
Red Brick Systems		✓	
Sequent Computer	✓		✓
SHL Systemhouse			✓
Silicon Graphics	✓	✓	✓
SNI	✓		✓
Sybase		✓	
Unisys	✓		✓

USER ISSUES

COST ISSUES AND DRIVING FORCES

The average annual cost of a DW/DSS solution is \$1.8 million. The figure is higher in North America and Japan and lower in Europe.

Expenditures for all product types are expected to increase during the forecast period, but a significant number of users stated that they expected that expenditures for outside services would stay relatively flat. With roughly 30% of user expenditures currently allocated to in-house expenditures, this finding has to be taken in context; i.e., many users are concerned about having their in-house personnel displaced by outside service providers and so tend to minimize the importance of outside services.

Driving much of the anticipated growth in expenditures is the anticipated growth in the size of the systems that will be required to meet future needs. For example, although the average size of the database reported by respondents was 272 GBytes, respondents reported that they expected the size of that database to increase by an average factor of 24 in three years! If they are right, the average data warehouse size at the end of the year 2000 will be 6.5 TBytes! (A July '98 news item wherein NCR announced that over forty of its customers have TByte+ size databases.) gives this statement some credibility.

This same group of respondents reported expecting that the average number of users accessing their data warehouse would increase by a factor of 42 over three years from 2176 to 91,392, and the average number of users concurrently accessing the warehouse would increase by a factor of 17 from 779 to 13,243. The reason is that many users see the warehouse as a focal point of information distribution for employees, customers and suppliers, in addition to the “knowledge workers” who are today’s most frequent users of Database Solutions. These numbers are up considerably from those obtained in our first survey.

The highest rated driving forces for investing in Database Solutions are:

- Improve decision or management processes
- Improve customer service
- Need to keep ahead of competition
- Corporate strategy initiative
- Reduce cost of operations
- Retain customers
- Identify new customers

PURCHASING ISSUES

In contrast to many other large-scale system procurements, Database Solutions system procurements are most heavily influenced by the departments that will use the system or by line-of-business management rather than by an IT organization. Typically, the IT organization was attributed to have only about a 30% influence rating on purchase decisions.

Respondents were presented with a list of twenty-three purchase criteria and asked to rate them on a 1-5 scale. The six highest rated criteria in order of descending importance are:

- Data Integrity
- System reliability

- Query Capability
- Scalability
- Availability
- Price/performance

With one exception, this list would be very close to that for any application considered as mission-critical. That Database Solutions have or will become mission-critical is something that all vendors must take into account.

In a similar manner, respondents were asked to rate nine vendor characteristics as purchase criteria. The criteria singled out as the most important is the amount of data warehouse experience the vendor could demonstrate.

IMPLEMENTATION ISSUES

Much to-do has been made in the press and by vendor pronouncements about the provision of complete solutions. Many would have us believe that it is the Holy Grail of the Database Solutions market. It turns out that only about half the market is interested in a packaged solution, and of the half that are interested, many don't believe it is possible to buy a packaged solution that will meet their needs. The main reason for interest in packaged solutions is *faster deployment* followed by *lower risk*.

The average elapsed time reported from when the business case for a Database Solution is completed to full production status is 25 months. This is the same number obtained in last year's survey.

More than a third of the respondents have implemented or plan to implement at least part of their Database Solution on platforms used for other applications. Most of those replying in this vein were existing large mainframe users, reinforcing the claim that mainframes constitute a major portion of the Database Solution market. The majority of users, however, thought they were better off dedicating systems to Database Solutions. Management control, cost, and compatibility are cited as the main reasons for not sharing a Database Solution platform with other applications.

SOME FUTURE DIRECTIONS

Respondents were asked which technologies they deployed today and which they anticipated would be deployed in three years. Most technology categories were expected to increase anywhere from 100% to 500% over the three-year time span. These included:

- Intranet enablement
- Extranet enablement
- The ability to do data mining directly on operational data
- The ability to handle multimedia data
- Direct linkage of decision support application to operational application

Both the Internet and the Intranet are expected to have a significant impact on Database Solutions. Both are seen as the vehicles for increasing data distribution. On the input side, the Internet and Intranet are seen as vehicles for improving data collection, and, for the Internet, providing access to external information that would otherwise be difficult to collect. On the output side, both are seen as vehicles for dramatically expanding user, customer and supplier access to data. In fact, many companies see Internet/Intranet browsers as the standard front end for distributing information. Some also see the Internet as the repository for software tools, permitting users to select from a variety of simple and complex query facilities just like users today can select from a variety of Web-based search tools.

An important issue in many organizations is the question of permitting a decision issued by a DSS system to be implemented without human intervention. For example, if the system says "double the inventory," should it be allowed to automatically order the merchandise without that decision being reviewed by management? A surprising 58% of the respondents stated that this would become a future requirement.

Another issue confronting some organizations is whether or not to permit data mining directly against operational databases. Some 67% of the respondents said they planned to do this in the future. Unfortunately, this question was not pursued further, so we don't know if this is just wishful thinking or, if plans to that effect are in place, how they will be accomplished.

A final futures question dealt with the application of multimedia data. In last year's survey, most users said they would need to support multimedia data in the future, but few had any ideas about what they would do with it! This year's survey respondents were much better prepared to address this question. In fact, two-thirds of the respondents said they planned to use it for document management. Other applications mentioned frequently are marketing material, training, product displays, mapping, and sales presentations.

Palo Alto Management Group, Inc.
2672 Bayshore Parkway #701 ■ Mountain View, CA 94043
Tel: (650) 968-4374 ■ Fax: (650) 968-4245
email: datasol@pamg.com
Web: www.pamg.com