JMS Cache Management for Web Performance

A Financial Services Perspective

An IDC White Paper

Sponsored by SpiritSoft

Analyst: Dirk Coburn

OBJECTIVES

This white paper discusses and illustrates the uses and benefits of caching technologies to enhance performance and cost metrics of Web-enabled applications.

The application architecture explored herein is the Java-based implementation of the eBusiness Platform. This white paper discusses the roles of and approaches to caching within that platform and outlines the ways in which caching is potentially a tool to extend and maximize the productivity of network resources.

Caching has the potential to be more than a facilitating tactic in the design of systems and applications. Appropriate use of caching can be a highly cost-effective strategic factor in addressing the following design objectives:

- Performance
- Data availability
- Scalability
- Session integrity

To further illustrate caching's potential, this white paper examines SpiritCache multilevel cache management software from SpiritSoft and its use with ebusiness applications in a financial services environment.

CACHING OVERVIEW

Caching is nearly as old as computing, predating not only the Internet but also many other networking standards. It originated to help companies avoid the need to repeatedly access slow data storage technologies in standalone host computers. The speed of physical data storage devices has improved exponentially since data caching originated internally within host systems. Still, caching survives and thrives in topologies updated to the concept of a multitier, Web-connected architecture as a virtual system. The continued updating of caching to current Web needs is driven not only by the speed of physical data storage access but also by other dimensions of performance and other related needs.

5 Speen Street • Framingham, MA 01701 USA • Phone 508.872.8200 • Fax 508.935.4015 www.idc.com

IDC Analyze the Future

Performance Dimensions

Performance has been a longstanding primary concern associated with multitier applications.

Any Web-based application is subject to such performance concerns as processing time and response time, as well as additional concerns about availability and session integrity.

Traditional responses to the need for performance focus on the characteristics of the following elements:

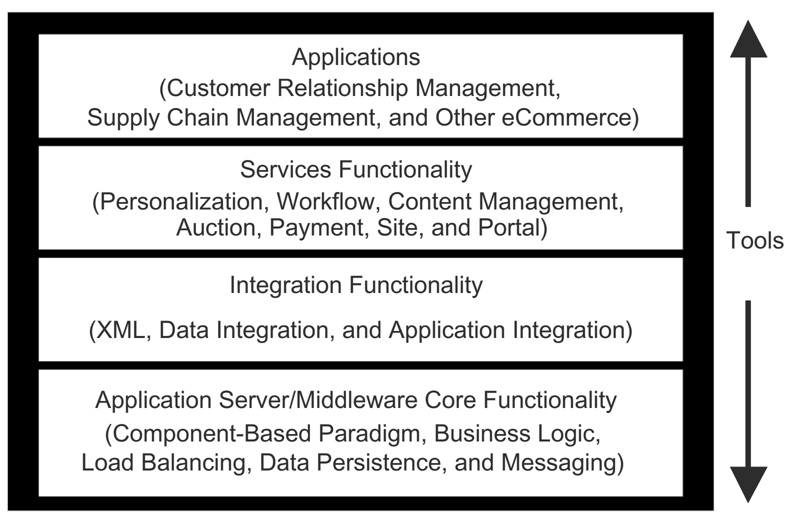
- Server hardware
- Network hardware and software
- Throughput capacity of major software components, especially database management systems and application servers, in what IDC calls the eBusiness Platform

One or more of those elements are frequently perceived as the key factors governing performance. Performance issues are often addressed with strategies that apply equipment or staff to those control factors. Such strategies typically include parallel system design, multi-threaded application design, or intensive system tuning. These strategies may be combined, for example, when technologies are acquired to monitor parallel resources and to make dynamic adjustments.

The eBusiness Platform

Figure 1 illustrates the high-level functions of the eBusiness Platform.

Figure 1: The eBusiness Platform: Functional Concept



Source: IDC, 2002

Key functions of the eBusiness Platform are as follows:

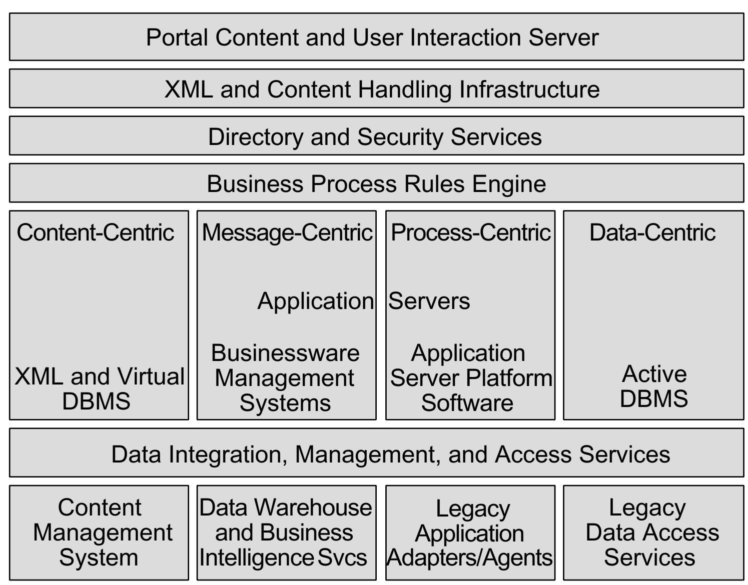
- Provides access to database and/or assets
- Incorporates the concept of application servers that enable it to deliver more than just static views of information over the Web
- Leverages investment through integration among information resources and applications
- Allows value-added services such as business process intelligence, security, accessibility, personalization, and other qualities of a robust, enterprise system that takes advantage of the power of the Web
- Provides support for a focused business process or diverse enterprisewide application needs

The eBusiness Platform offers these functions in a component-based environment.

Complementary Technologies

Figure 2 presents a more concrete view of the software technologies and products used to deliver functionality.

Figure 2: eBusiness Platform Technologies



Source: IDC, 2002

Java offers the fullest, most robust implementation of eBusiness Platform technologies. In fact, Java provides the architecture for plural interoperable approaches to the eBusiness Platform. By complying with specifications of the Java Community Process (JCP), even software vendors that provide many or, in some cases, nearly all of the platform technologies can interoperate with third-party component technologies where compelling advantage exists to do so.

The portfolio of software technologies in the eBusiness Platform operates in a distributed manner across multiple tiers from back-end enterprise data sources to edge servers providing services for content presentation and user interaction. A case can be made that Java Messaging Service (JMS) is the glue that holds this distributed, component-based environment together. JMS solves the needs of asynchronous component coordination, especially where the components are physically remote. The very usefulness of JMS makes it a prime candidate for the application of Web caching technologies.

Alternative Approaches

As noted above, there are many ways to design a Web-based system for performance. The discrete specifications that can be considered include the following:

1. Parallel storage hardware
2. Parallel processing hardware
3. Parallel database repositories
4. Additional network bandwidth
5. More efficient network delivery through compression or other technologies
6. Caching to minimize the extent and frequency of use of system resources to read data or content multiple times

Where more than one approach can yield effective results, a well-designed system applies lower-cost factors first. For this reason, the options that involve expensive parallel specification of hardware or of foundational software (options 1, 2, and 3 above) are often not optimal first considerations as performance control factors. The expense of those options often extends beyond acquisition to include fees and staff required for maintenance.

The extent to which performance can be achieved through enhancing network capacity and efficiency (options 4 and 5 above) can be limited for two reasons:

- Transmission of network messages may not be on the critical path to improving the performance of a particular system, especially if the server or client nodes have heavy processing loads.
- Network capacity and efficiency characteristics are typically planned and optimized as a matter of enterprise infrastructure investment, so sponsorship of improvements may be off-limits or prohibitively expensive for sponsors of an ebusiness project.

IDC believes that caching, which is sometimes introduced as a tactical workaround for a particular constrained resource in a networked system, should be a primary consideration in strategic performance planning for ebusiness systems.

Tight Mapping Versus Loose Coupling

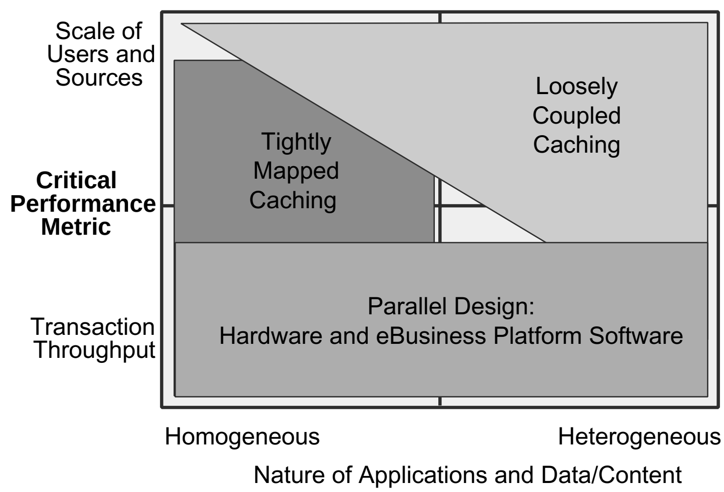
Caching can be used to enhance system performance in different ways, depending on the nature and purposes of the business processes and applications that are supported.

Companies should consider the following criteria when evaluating what approach is best for a given need:

- What are the metrics that matter with regard to performance?
- What is the nature of the data and/or content that is transported through the distributed eBusiness Platform in the course of use of the system?

Figure 3 diagrams the appropriate application of caching and of parallel system designs to these criteria.

Figure 3: Cost-Effective Options for Web Performance



Source: IDC, 2002

Where data is of a homogeneous nature, typically relating to uniform transactional operations, tightly mapped caching is feasible. Such caching can be particularly effective for systems in which raw transactional throughput is the critical performance metric, since tight mapping allows a cache to be preconfigured to handle just the essential changeable data in a highly structured information environment. We note that even tightly mapped caching may solve only part of the performance need if very high transaction throughput is the primary criterion. After the benefits of such caching are achieved, a system planner needs to look at successively more expensive control points, including parallel hardware and software, to boost performance. In that case, the appropriate approach is to use the caching and parallel resources together, since caching is still likely to reduce the degree to which more expensive resources are required.

That said, the nature of much ebusiness operation is procedurally diverse and involves a heterogeneous mix of structured data and unstructured content. Loose coupling where there is not homogeneity simply cannot offer the throughput advantages of tightly mapped homogeneity. However, loosely coupled caching in a heterogeneous environment can create significant opportunity for cost containment, network productivity, and expanded automation capacity from existing resources, as we discuss in the following sections.

For an even more broadly flexible advantage, a loosely coupled caching technology could incorporate features that isolate only those elements that change or anticipate probable caching needs to make optimal use of nonpeak processing.

MAXIMIZING NETWORK RESOURCES

The advantages of loosely coupled caching can be derived not only from the design of ebusiness systems but also from improvements in cost-efficient operation of network infrastructure. The result can also include more opportunity for productive automated activities.

Cost Reduction

Two areas are prime targets for cost reduction when caching is used to relieve system loads. Back-end database management systems may be able to run on smaller systems or with fewer multiple processors. Also, reduced latency in most cases makes servers in the middle tiers more productive, resulting in a reduced need for parallel equipment.

The savings includes acquisition costs if the caching architecture is introduced before a planned service expansion. In any event, the savings also includes staff time, maintenance contracts, and renewals and support for software licenses. This savings is the most tangible benefit against which one can justify costs of cache management software and staff time to implement it.

Network Productivity

Another benefit is network productivity, which assumes that a profitable return can be determined for certain activities that occur on the network. If caching-driven performance efficiency removes a constraint on the rate at which such activities occur, the profit attributable to activity beyond the former constraint contributes to the benefit from caching.

In following this methodology, companies should refigure the overhead attributable to the activities if there is a cost reduction in system operation as discussed above.

Expanded Automation Frontier

In organizations with constrained ability to invest in expanded network infrastructure, ebusiness projects queue for consideration when the infrastructure will support them. This process can seem glacial to managers whose otherwise approved projects (often for their expected return on investment) queue in a backlog to await the required network expansion. By applying caching to existing network resources, an organization can realize a significant opportunity to reap returns on new automation. Loose coupling is most likely to be appropriate for this purpose for two reasons:

- The nature of the needs of various projects are likely to be heterogeneous.
- Network infrastructure planning tends to focus more on users and servers than on transactional volume, except where a network is dedicated to transactional systems of high volume and velocity.

This advantage does not apply to projects whose constraints involve development skills or application administration and maintenance capacity.

SPIRITCACHE

SpiritSoft, based in Massachusetts with European headquarters in London, offers the SpiritArchitecture suite of Java technologies. The suite comprises messaging-centric tools and technologies for deploying Web applications. The most recent addition is SpiritCache, a cache management software product that provides loosely coupled caching capability to multitier architectures.

SpiritCache is positioned to support broad caching solution functionality, including:

- Page-level caching at the front end for Web services
- "Last image" caching of dynamic content, such as market data or business traffic metrics, supplied to a network
- The ability to isolate, at the message level, information that has changed
- Multitier caching services for mobile client devices
- Event notification, providing alerts to applications when underlying data changes

SpiritCache is a pure Java product; therefore, it is independent of the operating environment in which it runs. Also, SpiritCache was designed for heterogeneous object management. Caching technologies designed to support specific application products tend to be more tightly mapped. Support for XML, Java, relational, and other types of cached objects supports application flexibility. The application programming interface (API), based on the JCP specification for JCache, is conducive to skills sourcing and interoperability with J2EE and J2ME environments.

The use of advanced hierarchical clustering, developed for the SpiritWave JMS-compliant messaging product, supports fault tolerance, recovery, and scalability. Therefore, SpiritCache can be implemented tactically or in a broad, multilayer deployment for tunable, distributed caching. In larger architectures, highly skilled thought and work go into such implementations; on the brighter side, the tuning of such implementations is abstracted from the application, reducing the costs and risks of application maintenance.

Loose coupling, a standard Java API, and the ability to manage a broad range of objects position SpiritCache well for a key role in facilitating distribution of heterogeneous business content over the network. Such content may be unstructured, or a mix of structured and unstructured, and may be generated by more than one application. SpiritCache does not integrate such multiple applications. Rather, it supports enhanced performance, especially with regard to the number of users and the response latency they experience. SpiritCache can also support higher-level technologies in integrating multiple applications. To that end, the event notification feature is of particular interest.

Deployment Case Study: E*TRADE Financial

A particularly strategic deployment of SpiritCache supports what E*TRADE Financial calls its Application Storage Array Network (ASAN).

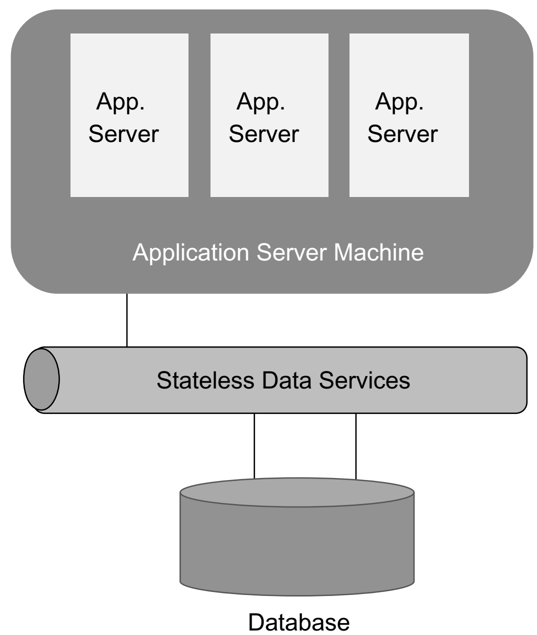
Founded nearly 20 years ago as a provider of online investing services, E*TRADE Financial has evolved into a fully integrated financial services provider that offers a diverse range of products and services personalized to meet the multiple lifetime investment needs of today's financial services consumer.

As E*TRADE Financial's offerings grew more comprehensive, systems grew correspondingly more complex. Decoupled, federated applications needed to share data, increasing the demand for database management services on the network. Business requirements drove systems to become increasingly multilingual and graphical. The network grew larger, with more "moving parts." The need for high availability spurred the acquisition of large amounts of database transaction software, machine hardware, and storage hardware.

Josh Levine, E*TRADE Financial's chief technology officer, envisioned the concept of a continuous data fabric to help E*TRADE Financial respond to the challenge of brute-force scaling. Key to the continuity of the fabric is distributed data that is consistent across machine boundaries and application federations.

Before the continuous data fabric was conceived, E*TRADE Financial's architecture was stateless at the front end (see Figure 4).

Figure 4: E*TRADE Financial's Stateless Architecture Before SpiritCache



Source: E*TRADE Financial, 2002

In the words of Chris Berry, a senior principal engineer in E*TRADE Financial's Architecture, Consortium, and Exchange (ACE) group, that fact was "critical to serving E*TRADE Financial's exponential growth" during the rapid adoption of online brokerage services by consumers in the mid- to late 1990s; statelessness also meant scalability.

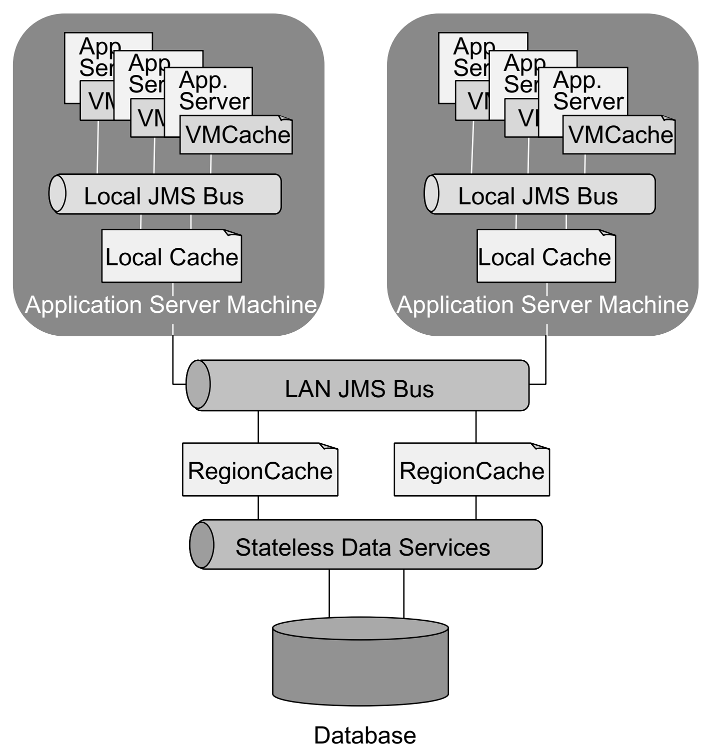
The ACE group evaluated technologies to support the continuous data fabric concept. It decided early on to look at JMS as the pillar of the architecture. The relevant systems are still not completely written in Java, but they are all Java on the front end and tend increasingly toward Java throughout. Also, the team noted that messaging-oriented middleware (MOM) vendors have widely embraced the JMS standard. JMS would help E*TRADE Financial avoid the need to acquire specialized adapter tools to interoperate with different MOM implementers and, therefore, avoid exposure to vendor lock-in.

The ACE team evaluated some database-centric technologies but ruled them out as serving only specialized purposes, such as requiring use of Entity Enterprise Java Beans (EJBs). The lingua franca in the continuous data fabric is, in Berry's words, “pellets of data” to be shaped and delivered through a service-oriented architecture. Currently, the pellets are based on a proprietary marshalling language, but the architecture is moving toward marshalling XML, eventually interpreted at the browser level by Extensible Stylesheet Language Transformations (XSLT).

In addition to Web application servers, the company needed an architecture that could interface with voice application servers and smart mobile devices. Meeting this design requirement would allow E*TRADE Financial to abstract form factors and to cache at different tiers for mobile devices, depending on the status of connections and synchronizations.

SpiritCache is the enabling technology used to fulfill E*TRADE Financial's vision of a continuous data fabric. Figure 5 diagrams the software's architecture.

Figure 5: SpiritCache's Distributed Cache Architecture

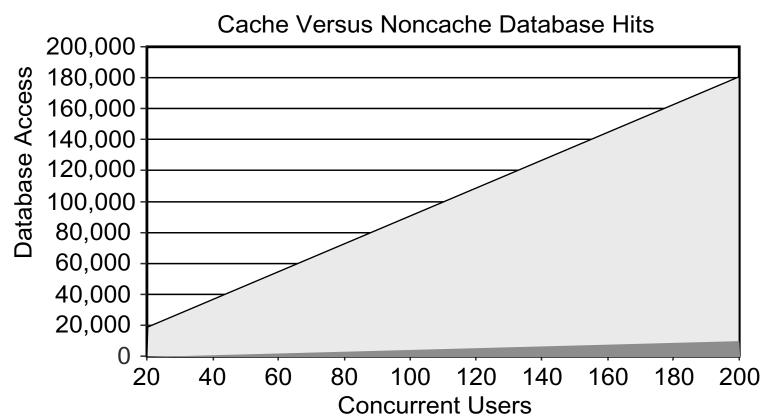


Source: E*TRADE Financial, 2002

The preliminary results of ASAN's distributed caching architecture are promising.

Performance metrics from a pilot project at E*TRADE Financial are shown in Figure 6. These metrics demonstrate a great improvement in database load and, therefore, in leveraged database scalability. Database access in ideal cases dropped by more than 95% in a test involving 200 concurrent users per application server machine. The anticipated benefits to E*TRADE Financial are reduced data services and database infrastructure costs and improved response time.

Figure 6: Reduction of Database Load



Source: E*TRADE Financial, 2002

CONCLUSION

The E*TRADE Financial case study shows SpiritCache to be a promising supporting technology for Web applications. It can be used not only as a flexible tactical tool for loosely coupled caching but also as a central technology for an architecture that expands the frontier of cost-effective Web application performance.

IDC Worldwide Offices

CORPORATE HEADQUARTERS

IDC
5 Speen Street
Framingham, MA 01701
United States
508.872.8200

NORTH AMERICA

IDC Canada
36 Toronto Street, Suite 950
Toronto, Ontario M5C 2C5 Canada
416.369.0033

IDC California (Irvine)
18831 Von Karmen Avenue
Suite 200
Irvine, CA 92612
949.250.1960

IDC California (Mountain View)
2131 Landings Drive
Mountain View, CA 94043
650.691.0500

IDC New Jersey
75 Broad Street, 2nd Floor
Red Bank, NJ 07701
732.842.0791

IDC New York
2 Park Avenue
Suite 1505
New York, NY 10016
212.726.0900

IDC Texas
100 Congress Avenue
Suite 2000
Austin, TX 78701
512.469.6333

IDC Virginia
8304 Professional Hill Drive
Fairfax, VA 22031
703.280.5161

EUROPE

IDC Austria
c/o Loisel, Spiel, Zach Consulting
Mayerhofgasse 6
Vienna A-1040, Austria
43.1.50.50.900

IDC Benelux (Belgium)
Boulevard Saint Michel 47
1040 Brussels, Belgium
32.2.737.76.02

IDC Denmark
Omøgade 8
Postbox 2609
2100 Copenhagen, Denmark
45.39.16.2222

IDC Finland
Jarrumiehenkatu2
FIN- 00520 Helsinki
Finland
358.9.8770.466

IDC France
Immeuble La Fayette 2
Place des Vosges Cedex 65
92051 Paris la Defense 5, France
33.1.49.04.8000

IDC Germany
Nibelungenplatz 3, 11th Floor
60318 Frankfurt, Germany
49.69.90.50.20

IDC Italy
Viale Monza, 14
20127 Milan, Italy
39.02.28457.1

IDC Netherlands
A. Fokkerweg 1
Amsterdam1059 CM, Netherlands
31.20.6692.721

IDC Portugal
c/o Ponto de Convergancia SA
Av. Antonio Serpa 36 - 9th Floor
1050-027 Lisbon, Portugal
351.21.796.5487

IDC Spain
Fortuny 18, Planta 5
28010 — Madrid
Spain
34.91.787.2150

IDC Sweden
Box 1096
Kistagangen 21
S-164 25 Kista, Sweden
46.8.751.0415

IDC U.K.
British Standards House
389 Chiswick High Road
London W4 4AE United Kingdom
44.208.987.7100

LATIN AMERICA

IDC Latin America
Regional Headquarters
8200 NW 41 Street, Suite 300
Miami, FL 33166
305.267.2616

IDC Argentina
Trends Consulting
Rivadavia 413, Piso 4, Oficina 6
C1002AAC, Buenos Aires, Argentina
54.11.4343.8899

IDC Brazil
Alameda Ribeirao Preto, 130
Conjunto 41
Sao Paulo, SP CEP: 01331-000 Brazil
55.11. 3371.0000

International Data Corp. Chile
Luis Thayer Ojeda 166 Piso 13
Providencia
Santiago, 9, Chile
56.2.334.1826

IDC Colombia
Carerra 40 105A-12
Bogota, Colombia
571.533.2326

IDC Mexico
Select-IDC
Av. Nuevo Leon No. 54 Desp. 501
Col. Hipodromo Condesa
C.P. 06100, Mexico
525.256.1426

IDC Venezuela
Calle Guaicaipuro
Torre Alianza, 6 Piso, 6D
El Rosal
Caracas, Venezuela
58.2.951.1109

CENTRAL AND EASTERN EUROPE

IDC CEMA
Central and Eastern
European Headquarters
Male Namesti 13
110 00 Praha 1
Czech Republic
420.2.2142.3140

IDC Croatia
Srednjaci 8
1000 Zagreb
Croatia
385.1.3040050

IDC Hungary
Nador utca 23
5th Floor
H-1051 Budapest. Hungary
36.1.473.2370

IDC Poland
Czapli 31A
02-781 Warszawa, Poland
48.22.7540518

IDC Russia
Suites 341-342
Orlikov Pereulok 5
Moscow, Russia 107996
7.095.975.0042

MIDDLE EAST AND AFRICA

IDC Middle East
1001 Al Ettihad Building
Port Saeed
P.O. Box 41856
Dubai, United Arab Emirates
971.4.295.2668

IDC Israel
4 Gershon Street
Tel Aviv 67017, Israel
972.3.561.1660

IDC South Africa
c/o BMI TechKnowledge
3rd Floor
356 Rivonia Boulevard
P.O. Box 4603
Rivonia 2128, South Africa
27.11.803.6412

IDC Turkey
Tevfik Erdonmez Sok. 2/1 Gul
Apt. Kat 9D
46 Esentepe 80280
Istanbul, Turkey
90.212.275.0995

ASIA/PACIFIC

IDC Singapore
Asia/Pacific Headquarters
80 Anson Road
#38-00 IBM Towers
Singapore 079907
65.6226.0330

IDC Australia
Level 3, 157 Walker Street
North Sydney, NSW 2060
Australia
61.2.9922.5300

IDC China
Room 611, Beijing Times Square
88 West Chang'an Avenue
Beijing 100031
People's Republic of China
86.10.8391.3610

IDC Hong Kong
12/F, St. John's Building
33 Garden Road
Central, Hong Kong
852.2530.3831

IDC India Limited
Cyber House
B-35, Sector 32, Institutional
Gurgaon 122002
Haryana India
91.124.6381673

IDC Indonesia
17th Floor, Tower 2
Jakarta Stock Exchange
Jl. Jend. Sudirman Kav. 52-53
Jakarta 12190
62.21.515.7759

IDC Market Research (M) Sdn Bhd
Jakarta Stock Exchange Tower II
17th Floor
Jl. Jend. Sudirman Kav. 52-53
Jakarta 12190
62.21.515.7676

IDC Japan
The Itoyama Tower 10F
3-7-18 Mita, Minato-ku
Tokyo 108-0073, Japan
81.3.5440.3400

IDC Korea Ltd.
Suite 704, Korea Trade Center
159-1, Samsung-Dong
Kangnam-Ku, Seoul, Korea, 135-729
822.551.4380

IDC Market Research (M) Sdn Bhd
Suite 13-03, Level 13
Menara HLA
3, Jalan Kia Peng
50450 Kuala Lumpur, Malaysia
60.3.2163.3715

IDC New Zealand
Level 7, 246 Queen Street
Auckland, New Zealand
64.9.309.8252

IDC Philippines
703-705 SEDCCO I Bldg.
120 Rada cor. Legaspi Streets
Legaspi Village, Makati City
Philippines 1200
632. 867.2288

IDC Taiwan Ltd.
10F, 31 Jen-Ai Road, Sec. 4
Taipei 106
Taiwan, R.O.C.
886.2.2731.7288

IDC Thailand
27 AR building
Soi Charoen Nakorn 14,
Charoen Nakorn Rd., Klongtonsai
Klongsan, Bangkok 10600
Thailand
66.02.439.4591.2

IDC Vietnam
Saigon Trade Centre
37 Ton Duc Thang Street
Unit 1606, District-1
Hochiminh City, Vietnam
84.8.910.1233; 5

IDC is the foremost global market intelligence and advisory firm helping clients gain insight into technology and ebusiness trends to develop sound business strategies. Using a combination of rigorous primary research, in-depth analysis, and client interaction, IDC forecasts worldwide markets and trends to deliver dependable service and client advice. More than 700 analysts in 43 countries provide global research with local content. IDC's customers comprise the world's leading IT suppliers, IT organizations, ebusiness companies and the financial community. Additional information can be found at www.idc.com.

IDC is a division of IDG, the world's leading IT media, research and exposition company.

02-063SOFTWA3282
April 2002

IDC
Analyze the Future

www.idc.com