ITG Executive Summary

November 2008

VALUE PROPOSITION FOR IBM SYSTEMS DIRECTOR: CHALLENGES OF OPERATIONAL MANAGEMENT FOR ENTERPRISE SERVER INSTALLATIONS

Economics Benefits for IBM System x and BladeCenter Deployment

Challenges and Opportunities

At the end of 2003, the typical U.S. Fortune 1000 corporation had less than 3,000 servers. On current trends, by the end of 2008 this will have increased to more than 4,800. The vast majority of these – almost 88 percent – will be x86 platforms running Windows or Linux operating systems.

The implications of this trend have been widely documented. Management overhead has expanded. A growing number of data centers are facing capacity constraints, pressures to upgrade energy and cooling infrastructures, or both.

Growing adoption of x86 server virtualization has begun to slow growth in physical server bases. But, as organizations have moved beyond pilot projects and begun widespread deployments of VMware, Microsoft, and Xen tools, they have discovered that new challenges are posed. A completely new layer of software is added to enterprise x86 environments, and management difficulties escalate.

In most organizations, moreover, numbers of virtual images are expanding more rapidly than physical server bases. If current growth rates continue, by the end of 2013 the average Fortune 1000 corporation will contain more than 6,500 physical and between 5,000 and 10,000 virtual x86 servers.

Over the next five years, users will thus face not only an accelerated rate of server proliferation, but also the need to deal with significantly greater software complexity, and to manage increasingly mixed bases of hardware and software platforms. Clearly, new strategies must be developed and implemented now, while x86 server virtualization is still at a comparatively early stage of adoption.

This report deals with a solution set – IBM Systems Director 6.1, IBM Systems Director Active Energy Manager, and System x and BladeCenter servers – that can play a major role in meeting these challenges. The opportunity exists to put software and hardware infrastructures in place that will improve the efficiency of existing x86 server base and ensure that future growth is managed in a cost-effective manner.

Specifically, this report looks at the potential for cost savings through use of IBM Systems Director 6.1 and Active Energy Manager to manage large numbers of physical and virtual servers. This potential extends across the full range of IBM x86 servers supporting Windows and Linux operating systems in native mode, and as VMware, Microsoft Virtual Server, and Xen guests.

Savings may be realized in the following areas:

- ***Server administration costs***. In six composite profile installations of x86 server bases in large organizations, potential three-year savings in personnel costs ranged from $7.02 million to $15.22 million. The average for all six installations was $8.14 million, representing a 37.1 percent reduction in server administration costs.

Installations contained between 603 and 3,862 physical and logical Windows and Linux servers, and included rack, blade, and tower hardware and VMware, Microsoft, and Xen virtual images.

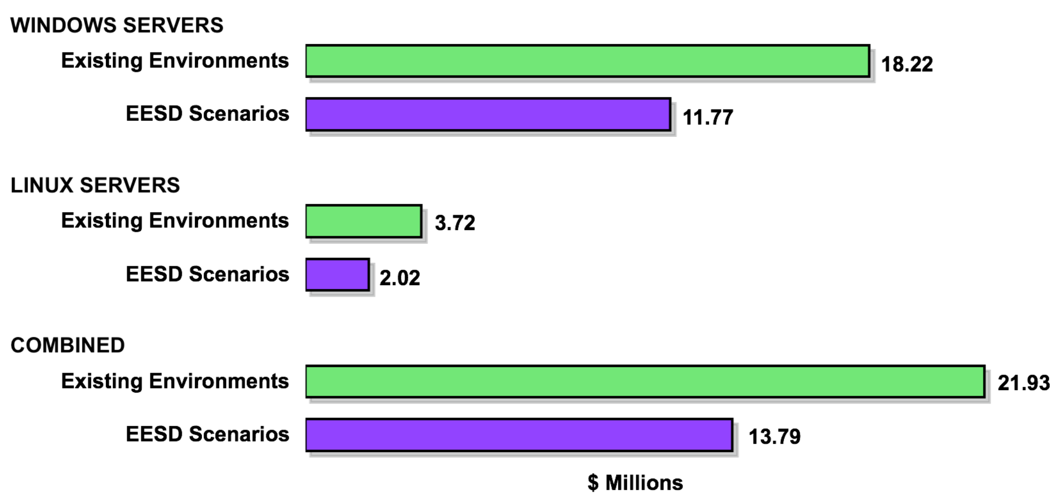
Profile installations are for large financial services, insurance, manufacturing, retail and telecommunications companies, and a state government agency. Installations are composites constructed using data supplied by 17 organizations in these industries.

For each installation, three-year costs were first calculated for full time equivalent (FTE) personnel for existing environments. These typically involved the use, with varying degrees of efficiency, of a mix of server management tools from multiple vendors. In many cases, manual techniques were employed. Such environments are common in large organizations.

Numbers of FTEs were then reduced based on International Technology Group (ITG) estimates of efficiency gains that could realistically be achieved in each installation through effective enterprise-wide use of Systems Director 6.1 (EESD scenarios). Three-year personnel costs were then calculated for the new staffing levels.

For all six installations, potential savings in server administration costs averaged 35.4 percent for Windows servers, and 45.7 percent for Linux servers; i.e., a great deal of time was saved by administrators dealing with both environments. Figure 1 summarizes these results.

Figure 1
Server Administration Costs for Existing Environments and EESD Scenarios: Averages for All Installations



Comparatively high efficiency gains for Linux platforms reflected generally weaker use of management tools and practices for these in existing environments. In Windows environments, there was more widespread use of management tools from Microsoft and others. Windows server management practices, however, were often fragmented and incomplete.

The ability of Systems Director 6.1 to integrate management of physical and virtual servers proved to be significant. FTE reductions, and corresponding savings in personnel costs, were particularly significant where Systems Director 6.1 was employed to manage physical server bases supporting VMware Virtual Machines (VMs), as well as Microsoft and Xen equivalents.

Management economies for blade servers were also realized through use of the IBM BladeCenter Open Fabric Manager (BOFM) solution with Systems Director 6.1. This contributed to lower administration costs for EESD scenarios for installations employing blade systems.

Additional information on profile installations, along with FTE staffing and cost breakdowns and details of methodology may be found in the Detailed Data section of this report.

- ***Energy costs***. For the same installations, energy consumption was also compared for existing environments and EESD scenarios in which Active Energy Manager was employed. Three-year energy costs were then calculated for both.

  Potential three-year savings in energy costs ranged from $65,455 to $323,944. The average for all six installations was $208,623, representing a 13.3 percent reduction.

  Figure 2 summarizes these results.

Figure 2

Energy Costs for Existing Environments and EESD Scenarios: Averages for All Installations

  Most of the organizations upon which profiles were based did not employ server energy management tools, or employed them only in pilot projects. The remainder employed such tools for small segments of their overall server bases. Even where organizations had put energy conservation programs in place, these were typically at an early stage of implementation.

- ***Overall impact***. The overall impact of savings in server administration as well as energy costs through use of Systems Director 6.1 and Active Energy Manager in profile installations is summarized in figure 3.

Figure 3

Overall Impact on x86 Server Operating Costs

CATEGORIES	SAVINGS	FACTORS
Server administration	Range: 33.8% – 42.1% Average: 37.1%	Reduced full time equivalent (FTE) staffing & personnel costs for server administration tasks
Energy consumption	Range: 11.3% – 18.4% Average: 13.3%	Reduced energy costs through improved monitoring & control of server energy usage
Combined cost impact	Range: 32.9% – 40.2% Average: 35.5%	Combined server administration & energy consumption cost savings

Cost savings may be realized in other areas. For example, a well-managed server environment will typically experience fewer outages.

Systems Director 6.1 monitoring capabilities will tend to reinforce this effect by allowing administrators to identify and resolve problems at an earlier stage than would be the case in a conventional server environment.

Costs of downtime measured in terms of lost productivity, business disruption, and damage to customer relationships would thus be lower. Fewer outages, as well as reduced end user exposure to hardware and software problems, would also result in FTE savings for help desk and other support personnel.

While the potential for cost savings is significant, it is only part of the value proposition for Systems Director 6.1. Effective use of Systems Director 6.1 and related IBM offerings across organizational server bases may play a further role in enabling the realization of broader enterprise management strategies.

Server View

Concentration and Virtualization

Key x86 server technology shifts, such as growing adoption of blades and virtualization, offer the potential to address physical server proliferation. But they create new challenges.

The potential of virtualization to enable physical server consolidation, for example, is accompanied by new problems. As multiple applications and workloads are concentrated on fewer physical platforms, system and workload management challenges increase. Larger, more powerful N-way platforms are required to achieve high levels of virtual server concentration.

Higher levels of concentration, moreover, require greater attention to such issues as availability (single outages may disable multiple applications), security (single security or malware exploits may affect multiple virtual servers), and backup and recovery (processes become more complex than for conventional servers).

Equally, blade systems require less rack capacity and reduce energy consumption. However, I/O and network interconnect management tasks tend to be more complex than for most conventional servers, and require use of specialized fabric management software. For both reasons, it may be difficult to integrate blades into broader server management systems and processes.

Blade configurations also concentrate heat within smaller physical spaces. Higher temperatures increase failure rates. Many organizations have found that they cannot fully populate blade enclosures without risking heat-related problems, or have been obliged to install specialized cooling devices, or both.

Systems Director 6.1 and Active Energy Manager can play major roles in meeting these challenges. Their effectiveness, however, is materially reinforced by the capabilities of underlying IBM System x and BladeCenter platforms.

System x and BladeCenter

Like other x86 server vendors, IBM offers a broad range of rack-mounted, tower, and blade models incorporating Intel as well as Advanced Micro Devices (AMD) Opteron platforms.

However, the company has consistently differentiated itself against "commodity" hardware competitors through design features that leverage its expertise and experience in such areas as system architecture; semiconductor design (IBM is a major developer and manufacturer of microprocessors and memory chips); and reliability, availability, and serviceability (RAS) and energy and cooling engineering.

These strengths are reflected in the following:

- ***X-Architecture***. This is an Intel Xeon-based server design, currently in its fourth generation, that is optimized for high levels of system-wide performance and scalability. The design, which is based in part upon IBM mainframe architecture, enables x86 servers to scale in a manner that maximizes not only processor, but also memory and I/O performance.

  The design incorporates a number of distinctive technologies. These include a customized chipset built around Intel Xeon processors that incorporates embedded performance accelerators. These include a 324 MB "snoop" filter that increases multiprocessing efficiency (standard Intel designs support only up to 64 MB); a large Level 4 cache (standard Intel designs support only levels 1 to 3); and Advanced Buffer eXecution (ABX) technology that reduces buffer latency.

  The latest generation of IBM eX4 systems offers scalability of up to 16 Xeon processors, for a total of 96 cores using latest-generation Intel six-core models, along with one terabyte (TB) of main memory, one gigabyte (GB) of L4 cache, and 28 PCIe slots, including eight hot-pluggable PCIe slots. Upgrades from 2- to 16-way configurations are simple and rapid.

IBM is the only first-tier x86 server vendor to offer a Xeon-based platform that scales beyond four-way configurations. Eight-way, 16-way and, in some cases, larger configurations have been routinely deployed by customers for a wide range of applications.

These have proved particularly popular for business-critical applications and as server consolidation platforms. One organization that contributed to the profiles presented in this report, for example, had deployed more than 50 VMware virtual machines on a single 8-way X-Architecture server, and another had deployed 80 on a comparable 16-way machine.

The latest-generation IBM X-Architecture high-end model, the x3950 M2, is offered with an embedded hypervisor capability based on VMware ESXi. As a result, virtualized applications can be deployed "out of the box"; i.e., without server modifications.

- ***BladeCenter systems***. These include a full range of Intel- and Opteron-based blades, as well as the flagship 14-bay (BladeCenter H) and compact 6-bay (BladeCenter S) chassis and specialized models for telecommunications carrier applications.

  BladeCenter servers incorporate X-Architecture features. Like System x models that implement this architecture, BladeCenter systems are designed and engineered for high levels of performance and scalability.

  The overall design, however, conforms to industry standards. For example, a variety of IBM and non-IBM switch modules for 4 gigabits per second (Gbps) Fibre Channel, Ethernet (10 Gbps, 1/10 Gbps Uplink, and Layer 2/3) and N Port ID Virtualization (NPIV) networks are supported.

  Key capabilities are provided by Advanced Management Module (AMM) firmware and BladeCenter Open Fabric Manager, which is generally regarded as one of the industry's most functional and scalable blade fabric management solutions. BOFM handles I/O and network interconnects for up to 100 BladeCenter chassis with up 14 servers each; i.e., for a total of 1,400 servers. Administrators may employ a single log-in to all of these.

  Combined use of AMM firmware and BOFM also enables detection of failures, failover to alternate blade servers that may be anywhere within the same 100-chassis domain, and reassignment of network addresses from the failing server to the new server. These processes are automatic, and may occur with minimal interruptions of service.

- ***RAS capabilities***. IBM strengths in design and engineering for reliability, availability, and serviceability are reflected in numerous System x and BladeCenter features.

  Basic capabilities include high levels of component reliability and redundancy; hot-add and hot-swap functions (i.e., the ability to add or replace components without taking servers offline); and extensive hardware- and software-based diagnostic, and fault isolation and resolution facilities.

  X-Architecture systems also implement a multi-level memory protection system. This includes Chipkill Memory, an IBM-developed error checking technology that is significantly more reliable than conventional error correction code (ECC) techniques, along with memory scrubbing, memory mirroring, and Memory ProteXion. These technologies provide multiple forms of redundancy and recoverability for on-chip data.

  Other key capabilities include Predictive Failure Analysis, which reduces the potential for unplanned outages; Light-Path Diagnostics, which reduces the amount of downtime required for on-site maintenance actions; and 24x7 electronic service coverage.

  These and other RAS capabilities are summarized in figure 4.

Figure 4

System x and BladeCenter Reliability, Availability and Serviceability Capabilities

MEMORY SUBSYSTEM (IBM ACTIVE MEMORY)	
Chipkill Memory	Creates duplicate checksum copies of data within memory subsystem, disables inoperative memory chips in the event of a failure (server remains online), & reinstates data copy. Aids in correcting single- & multiple-bit errors using standard memory chips.
Memory scrubbing	Performs automatic daily test of all the system memory. Detects & reports developing memory errors.
Memory ProteXion	Enables data stored on potentially damaged memory chips to be rewritten to an alternative location. Activated if memory scrubbing determines an error is recoverable (a.k.a. redundant bit steering).
Memory mirroring	Provides RAID 1-equivalent function that writes data to multiple memory cards & enables it to be rebuilt in the event of a failure. Allows data to be rebuilt on hot-swapped cards.
Advanced Buffer eXecution	Provides buffer redundancy across up to four memory cards.
OTHER CAPABILITIES	
Hot-add/hot-swap components	Memory cards, disk drives, I/O adapters, power supplies, & fans may be added or replaced without taking server offline.
Disk data redundancy	ServeRAID MR-10k controller supports RAID levels 5, 6, 10, 50, & 60.
Solid state storage (BladeCenter only)	BladeCenter option offering approximately three times the reliability of disk drives (solid state devices have no moving parts).
Automated failover (BladeCenter only)	Advanced BOFM capability enabling automated failover to alternate blade servers within the same chassis or elsewhere within BladeCenter domains. Transfer of Ethernet Media Access Control (MAC) &/or Fibre Channel worldwide name (WWN) addresses & activation of alternates are also handled automatically.
Predictive Failure Analysis	Provides up to 48 hours advance warning of component failures. Illuminates the appropriate light path diagnostics indicator &/or messages system administrators. Enabled components include processors, memory cards, disk drives, I/O slots, power supplies, fans, & voltage regulator modules.
Light-Path Diagnostics	Identifies component failures & illuminates LEDs on server panel indicating those that require service. Monitored components are the same as for Predictive Failure Analysis.
IBM Electronic Services	24x7 online service that collects & analyzes error & inventory data, analyzes it, & notifies system administrators &/or places service calls directly to IBM if an actual or potential problem is identified. Interfaces to onboard system monitoring & Predictive Failure Analysis. Enabled by Electronic Service Agent on server.
Dynamic System Analysis	Software tool that collects & analyzes system information to aid in diagnosing complex problems. Interfaces to Electronic Service Agent.

IBM also works closely with suppliers of conventional high availability cluster solutions, including Microsoft (Microsoft Cluster Service), Oracle (Real Application Clusters), Steeleye Technology (Lifekeeper) Symantec (Veritas Cluster Server), VMware (VMware HA) and others to ensure that these are fully compatible with and supported for System x servers.

- ***Energy and cooling capabilities***. System x and BladeCenter servers incorporate a number of industry-leading features that reduce energy consumption, increase cooling efficiency (which lowers failure risks and further contributes to energy savings), or both.

  In addition to Advanced Buffer execution, which can reduce memory energy consumption by up to 37 percent, these include high-efficiency IBM-designed power supplies and IBM Calibrated Vector Cooling, which employs advanced airflow design and fan control technology to deliver exceptionally effective air-cooling.

The IBM Rear Door Heat eXchanger (RDHX), an inexpensive chilled water system that can remove up to 50,000 British Thermal Units (BTU) of heat, may also be employed to cool System x and BladeCenter racks. A key benefit is that this system attaches to existing racks and does not require dedicated data center space.

These capabilities, and their implications for data center energy and facilities costs, are discussed in more detail later in this report.

IBM strengths in energy efficiency and cooling systems are leveraged in a customized IBM offering, the ***System x iDataPlex***. This employs an innovative airflow design and RDHX technology to deliver twice the compute density of conventional 42U industry racks.

The 100U iDataPlex rack has 84 horizontal 1U server slots and 16 vertical slots for switches, power distribution units (PDUs), and other devices. It has proved popular among corporate users and service providers for high-volume Internet applications.

Similar capabilities are offered by some competitive systems. The overall portfolio of IBM capabilities, and the extent to which these are integrated and optimized in System x and BladeCenter servers are, however, unmatched by any other vendor.

Enterprise View

For more than a decade, there has been a growing recognition that all components of organizational IT environments have grown more complex and interdependent. It has become clear that full range of IT resources must be managed more effectively at the enterprise level.

The challenges, however, are daunting. Enterprise management capabilities must extend across servers and storage as well as networks, applications, and middleware. They must integrate traditional disciplines such as systems management, operations, asset management, change management, and service-level management. Security, compliance, and energy conservation processes must also be incorporated.

Solutions designed to meet these challenges have been available for more than a decade. But, although effective management systems have often been put in place for specific disciplines and IT resources, in most organizations, enterprise-level integration remains a remote goal.

One reason for this is that solutions are themselves often technically complex. Implementation challenges are magnified by the diversity of applications, platforms, and technologies in enterprise IT environments, and by the rate of change in these. Deployment initiatives must also deal with shifting business demands and with the need to create management processes that are highly interdependent and complex.

The most effective approaches have proved to be those that focus first on a subset of IT environments and then build toward broader structures. There are a number of potential candidates for this initial focus. But server bases are the obvious starting point.

In most organizations, servers account for a significantly higher percentage of IT spending than other infrastructure components. Even when they have been physically concentrated in data centers, server infrastructures remain dominated by small machines and diverse operating systems and software versions. Overall capacity utilization is typically the lowest, and downtime the highest for any major IT resource.

Early successes in server management thus offer the potential for larger savings and greater benefits in service quality than is the case for other segments of IT infrastructures. They may also act as a catalyst for, and may materially facilitate the realization of enterprise management goals.

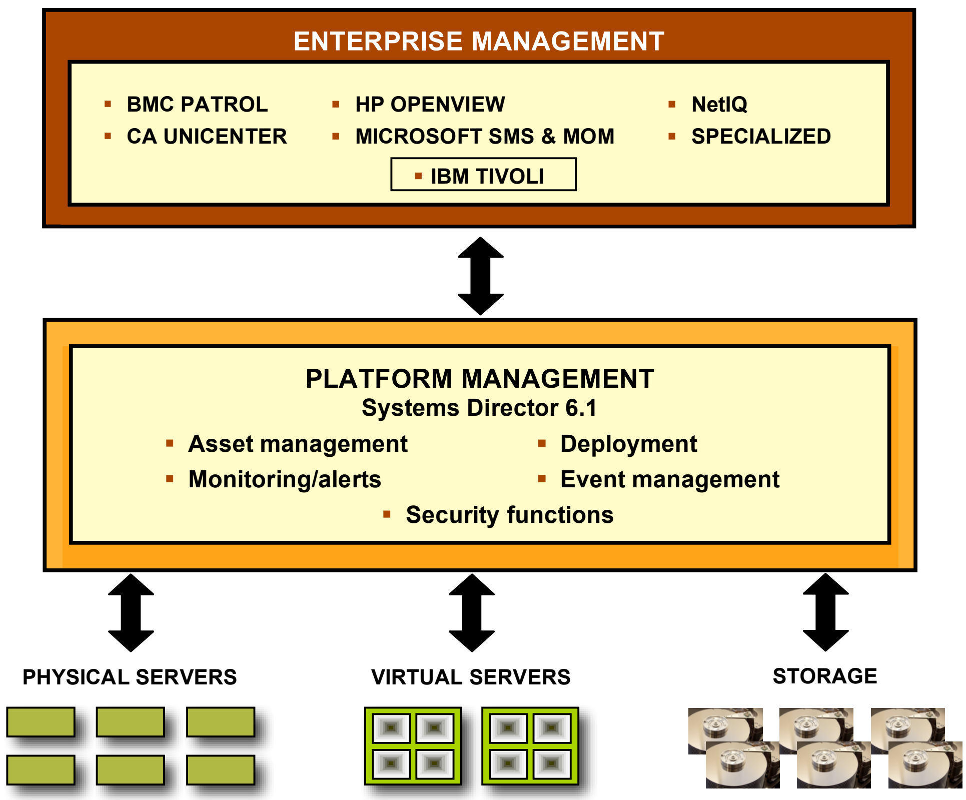
The odds that this will occur are significantly increased if server management solutions are closely integrated and highly synergistic with those that address management challenges at the enterprise level. This is the case for Systems Director 6.1.

In addition to high levels of integration with IBM Tivoli solutions, which are discussed later in this report, Systems Director 6.1 interfaces to the industry's major non-IBM enterprise management offerings.

These include BMC Software's Patrol, Computer Associates' Unicenter TNG, Hewlett-Packard's OpenView, Microsoft's System Systems Management Server (SMS) and Microsoft Operations Manager (MOM), and NetIQ tools. Interfaces to specialized third-party tools, such as Altiris Deployment Solution, are also supported.

This approach, which is illustrated in figure 5, ensures functional consistency and technical compatibility between the enterprise and server levels of the management environment.

Figure 5
Systems Director Relationship to Enterprise Management Solutions



In relation to enterprise management solutions, Systems Director 6.1 provides a subset of key operational management functions for servers running AIX, Windows, and Linux operating systems, including Linux versions supported by IBM for x86, Power, and mainframe platforms. Also supported are the System i platform, server-based storage, and low-end and midrange IBM disk systems.

Management processes are standardized across all of these platforms, and a common Web-based interface is provided for administrators dealing with them.

Conclusions: Saving the Lobster

There is an old saying that a lobster can be boiled without realizing it. As the temperature increases only one degree at a time, the lobster does not realize what is occurring until it is too late. Arguably, something similar has occurred in the IT world.

The growth of infrastructure inefficiency has been a gradual process. As organizations moved away from traditional centralized computing models in the 1980s and 1990s, IT infrastructures became dominated by small servers. The impact of the Internet and intranets and the growing complexity of software architectures and applications in the late 1990s and 2000s reinforced this trend.

More widespread use of management tools and consolidation initiatives has led to some improvements. In most organizations, however, physical server bases remain fragmented, and their use is characterized by levels of inefficiency that would not be tolerated in other areas of the business.

Use of x86 server virtualization allows users to achieve higher levels of consolidation than have been realized in the past, and offers improved provisioning flexibility and other benefits. But virtualization enablers pose new management and service quality challenges that are becoming increasingly apparent as organizations move toward large-scale deployment.

Organizations are also encountering another effect: growth in numbers of virtual servers rapidly outpaces reductions in physical servers. There is a risk that virtual server proliferation will create problems even more serious, and more difficult to resolve than was the case for growth of physical server bases.

Server virtualization offers part of the solution. But to preempt the effects of virtual server proliferation, virtualization initiatives must be coupled with aggressive deployment of advanced server management tools an early stage. If this is done, major gains in infrastructure efficiency may be realized across physical as well as virtual resources.

Potential savings in server administration and energy costs documented in this report, in practice, could translate into reductions of between 15 percent and 25 percent in overall server costs in large organizations.

What might the effects be in any organization if these savings became available for new, high-value application deployment projects? Major gains in business agility and competitiveness might be realized.

The tools and technologies to realize such gains are becoming available. What is needed now is the ability to recognize and act upon the opportunities that they represent.

Additional Information

This ITG Executive Summary is based upon results and methodology contained in a Management Brief released by the International Technology Group. For copies of this Management Brief, please email requests to info-itg@pacbell.net.

International Technology Group

4546 El Camino Real, Suite 230
Los Altos, California 94022-1069
Telephone: (650) 949-8410
Facsimile: (650) 949-8415
Email: info-itg@pacbell.net

Copyright © 2008 by the International Technology Group. All rights reserved. Material, in whole or part, contained in this document may not be reproduced or distributed by any means or in any form, including original, without the prior written permission of the International Technology Group (ITG). Information has been obtained from sources assumed to be reliable and reflects conclusions at the time. This document was developed with International Business Machines Corporation (IBM) funding. Although the document may utilize publicly available material from various sources, including IBM, it does not necessarily reflect the positions of such sources on the issues addressed in this document. Material contained and conclusions presented in this document are subject to change without notice. All warranties as to the accuracy, completeness or adequacy of such material are disclaimed. There shall be no liability for errors, omissions or inadequacies in the material contained in this document or for interpretations thereof. Trademarks included in this document are the property of their respective owners.