

Mainframe Role TCO

Robert Frances Group
July 2006

Table of Contents

Table of Contents.....	1
Executive Summary	1
Consolidation	2
TCO Conclusions: Mainframes versus Hewlett-Packard (HP), Sun, Wintel, and Specific Workload Hosting.....	3
Current Mainframe Misconceptions	6
Summary	10

Executive Summary

Robert Frances Group (RFG) has witnessed an emerging trend in relation to mainframes. Indeed, more and more IT organizations are reconsidering the mainframe for more diverse types of workloads. RFG views the new wave of server and storage consolidation, partitioning, and virtualization as an affirmation of the mainframe model. Moreover, it is a move away from the distributed one-application-per-server mentality that has driven the cost and complexity of computing higher during the past several years. Today's organizations are now facing dramatically increasing staff needs for network and database administrators, as well as system administrators. Increasingly large server farms create consistency and service issues, such as versioning, patch management, and associated outages. Security issues are becoming increasingly important based on hacker threats, regulatory compliance, etc. These are not easily addressed in distributed environments, but the mainframe offers a superior capability that includes centralized administration.

The multi-purpose role of mainframes, with their inherent superiority in partitioning and virtualization, provides the same options now being considered on non-mainframe platforms, but with stronger and more proven technology. IT executives should consider the value of the 20-year advantage that mainframes have over other platforms in this area. RFG believes the mainframe technology investments that IBM has made (and will continue to make) will ensure technology superiority for the foreseeable



120 Post Road West, Suite 201
Westport, CT 06880
<http://www.rfonline.com>

future. With this, in addition to the power and cooling advantages of the mainframe, IT executives and their staff should consider the mainframe as a "Tier-1" option for hosting new applications and acting as a central hub for security, server pool management, and consolidated workloads/data.

The philosophical difference in choosing mainframes versus distributed platforms stems from the "scale up, scale out" conundrum. The fact is "scale out" strategies do not reduce unit costs as the capacity grows, because the pricing for software and people in distributed platforms is almost linear. Unit costs for software and people go down as the workloads and capacity grow on the mainframe, a function of the software pricing model and the level of automation for systems and network management. Distributed platform software costs for database and other integral components are one-time and per-server, versus IBM's monthly license charge (MLC) for its software. Upgrade charges on the mainframe can be expensive when it comes to independent software vendor (ISV) software (data showed \$3,000 to \$5,000 per MIPS). However, these charges are offset by the ongoing savings where distributed workloads require more and more servers.

Consolidation

Consolidation of servers becomes the next logical decision point, which provides a way to reduce the number of servers in the distributed environment. Unfortunately, client data indicates that software costs can actually be higher in these cases, since partitioning of larger servers results in the same number of "logical servers." In addition, the software costs tied to larger servers are higher among most software providers. The other issue with consolidation of distributed platforms is the two-fold problem of Unix and Wintel having poor partitioning (e.g., VMware) capability and the inability to run non-homogeneous workloads (e.g., transactional versus batch). These two factors lead several clients to the point of having limited success

with consolidation and a complication of degraded service if not done correctly.

When considering hosting decisions such as these, RFG advocates a model that encompasses the "seven Ps" (i.e., RFG's P⁷ model and framework) – people, process, platform, product, project, planning, and portfolio. The people and process dimensions are the ones most commonly addressed, though not accurately in many cases due to shared resources. The platform "P" represents the heart of this discussion (i.e., which platform should workloads reside upon), so goes without saying. Among the other "Ps," the planning function should include "how and where" reviews that ensure optimum hosting decisions, with associated compliance around architectural standards and guidelines. This would ensure a balanced view of where the best fit exists for each workload. The product "P" is important for many reasons. Products include the services around them (e.g., managed server), and "productizing" mainframe options for workloads that make sense (e.g., Information Services, Linux S/M/L, Tier-3 database serving) offers an easy way to forecast, plan, and ensure proper use of existing technologies.

Projects (the sixth "P") are the way that these new workloads are implemented and should include check points, to ensure that the standard, pattern-based configurations are used appropriately based on business requirements. The portfolio "P" is often the last to consider, but can be most important. A portfolio of existing workloads and their associated infrastructure provides a basis for evaluating existing investments for making future decisions. The portfolio is also a key vehicle for interacting with business units, to convey information for collaborative decision-making.

The mainframe has undergone dramatic change in the past few years, capable of accommodating workloads such as Java, Linux, and WebSphere, but providing greater availability, security, etc. IT executives should ensure that re-use and leverage are maximized in any hosting decision, as well as direct architecture and infrastructure groups to make unbiased decisions about hosting that include the mainframe as an option.

Given the recent (and ongoing) improvements in granularity, price/performance, reduced software costs, and virtualization, the opportunities are even greater today in certain areas. To make a proper hosting decision, several key criteria must be evaluated, but these factors often are not even applied to mainframe hosting, based on the notion that workloads could be run better, faster, and cheaper on distributed platforms. The considerations that are most important, and which should be applied to all platforms, include:

- If it is a packaged application, where does it run, and with what level of support?
- How well can existing infrastructure and skills be leveraged?
- How much reserve capacity (infrastructure/skills) exists in each environment?
- How much integration/access into other applications will the application require?
 - Where are the other applications currently residing?
- What are the ROI and total cost of ownership (TCO) results for each platform being considered?
- What level of service (and security) is required?
- Where is the data and how is it accessed and shared?

TCO Conclusions: Mainframes versus Hewlett-Packard (HP), Sun, Wintel, and Specific Workload Hosting

The research done for this paper included a number of discussions with clients, where specific cost comparisons were made between specific applications and general workloads (e.g., transactional) hosted on mainframes versus something else. In a couple of cases, these were simply "paper" exercises, but in others, actual implementation data was assessed.

Comparing configurations from one vendor to another, however, is difficult. The mainframe

has unique technology and characteristics (i.e., being the strongest commercial platform for generalized computing, with multiple internal processors for handling input/output (I/O), etc.). While quantitative data was used for anecdotal evidence, the conclusions made here are necessarily qualitative based on the limited number of clients included in the study. A second paper will address the TCO issue in more detail and for specific workloads. The TCO data used for this study was shared by clients on a confidential basis, so conclusions are emphasized over specific cost data.

DB2: DB2 is a long-term winner in the database management system (DBMS) world. The idea of utilizing DB2 for most customer decisions comes down to the hosting decision for the application (i.e., whether it will be on the mainframe) and the current presence and skills around DB2. If DB2 is already installed and supported, there is tremendous opportunity to take advantage of its new capabilities that address data serving and the greater mission of information on demand (IoD). In addition, there is an opportunity to look at ways to bring data together corporately via DB2. Based on a number of discussions with clients, as well as analysis of several clients' TCO data for specific workloads, the following conclusions are drawn:

- DB2 workloads running on the mainframe generally make sense to keep on the mainframe. The strengths of the backup and recovery, functionality, and security warrant keeping the data there, not to mention better pricing in comparison to Oracle on HP or Sun configurations.
- DB2 workloads not running on the mainframe generally make sense to migrate to the mainframe, depending upon factors, such as data access, geographies of users, integration needs with other data and/or applications, and security. IBM has gradually improved its non-mainframe UDB version to be roughly 95-percent compatible with mainframe DB2, making it easier to convert/migrate.
- Migrating from Oracle, SQL Server, or any other database to DB2 is hard to evaluate on a generalized basis, but generally, the cost

drivers are licensing status, network and storage access, and size. The TCO data does not show a definitive cost advantage either way, although the greatest difference is software licensing, which can skew the advantage to DB2 significantly.

- Though no actual data was available, customers are excited about the advent of using the System z Integrated Information Processor (zIIP). zIIP is expected to lower costs by reducing the central processing unit (CPU) overhead of DB2, which relegates it to a specialized processor engine that is not subject to software costs, and improve performance.

CICS: No comparison data was available for this study between CICS and alternative online transaction processing (OLTP) systems on Unix or Wintel platforms. Several clients provided cost data for CICS on a total and per-transaction basis; however, with limited cost data received on Unix transactional processing on a total basis. Discussions with clients indicated a general view that the pricing of CICS (size versus per-server licensing, which is higher), robustness, and superior service and security characteristics outweigh cost decisions, especially for larger applications. Java-based OLTP on Unix and Wintel is viable (and usually cheaper than CICS) for small, one-off applications that have lesser management, security, and service requirements. RFG comparisons to CICS, however, indicate a cost advantage for larger systems versus other platform alternatives. In addition, the use of CICS is a staple for many, which provides a strong foundation for support and the ability to remain up-to-date. Overall, CICS is a superior transaction monitor, and has cost advantages as well. Where messaging-based (i.e., store and forward) approaches are workable for the application in question, there are often CICS adjunct transactions that are used to support the processing. In addition, the stability of the environment, along with the existing skills set that is in place among companies running CICS on a mainframe, are compelling reasons to leverage it.

WebSphere and Java: Customers generally do not have cost data to compare WebSphere and

Java workloads on the mainframe versus distributed platforms, simply because nobody has done it both ways. The general view and conclusion of the clients interviewed for this paper felt that the mainframe costs for these workloads were generally higher than that of distributed platforms. But as usual, RFG analysis indicates that the higher the volume of transactions, as well as transaction intensity (arrival rate), the more the mainframe looks favorable. However, there are clearly many factors such as the size of the platform and its usage, especially for the development environment. The TCO for hosting the execution environment favors Wintel and small Unix servers, both with Sun and HP boxes compared. What is interesting in RFG's research, albeit not quantified, is the fact that unit costs tend to go down when parts of the WebSphere and/or Java 2 Platform, Enterprise Edition (J2EE) environments are already installed. For example, the execution environment is simplified if the development environment is also there, lowering people costs and slightly lowering software costs when leveraging software already installed. The larger servers vary among cost comparisons reviewed for this paper, with software costs a major wild card (per-server charges multiply quickly in the Unix and Wintel environments). People costs favor the mainframe.

Enterprise Resource Planning (ERP) Systems:

The economies of running ERP systems on the mainframe generally favor distributed platforms. However, several clients indicated that the "Tier-3" data-serving role for the mainframe was utilized and very successful in the contexts of service and cost. Once again, the location of key data and the access characteristics drive such decisions. Moreover, the new data-serving capabilities recently announced by IBM make it even more attractive. Finance and HR systems often appear on the mainframe separately, but the economies of these versus hosting on other platforms are not clear given the data available.

Data Serving: Data serving on the mainframe is a tremendous opportunity for those that have critical data there. The new announcements around DB2, its specialty engine (zIIP), and its inherently attractive pricing, make use of DB2 a

unique option to provide data-serving capabilities. The recent announcements in this regard show the value of leveraging the mainframe and providing additional "information services" with minimal cost increments. Obviously, other DBMS and file structures can be leveraged in this regard, but the key is to have a clear data architecture and associated infrastructure design established.

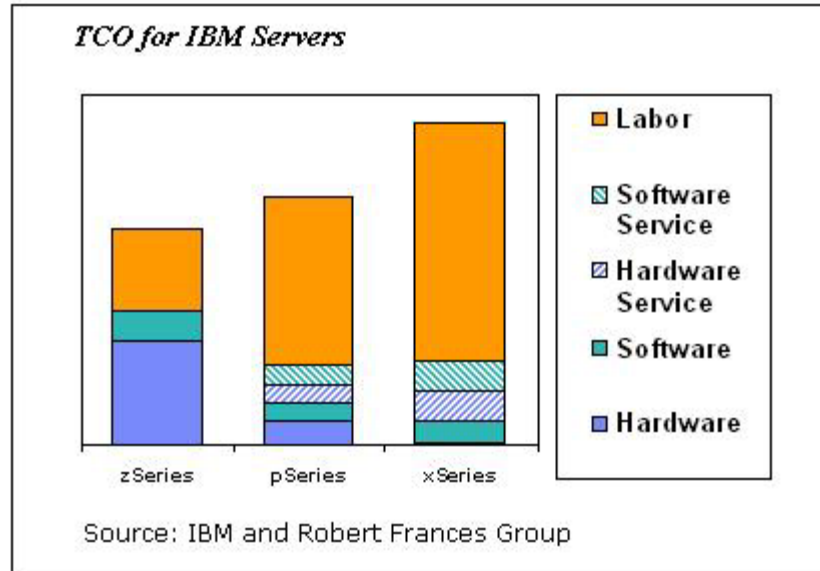
Data copying and syncing across environments is complex and prone to failure. Several small ISVs make a living from software revenues tied to managing backups and copy/sync operations. The most troublesome factor is the cost, since RFG finds that these additional operations, from a TCO perspective, represent up to 30 percent of operational costs, which can be avoided with centralized data. The idea of centralized data is difficult to sell for a number of reasons, such as customer ownership issues, performance issues, security and audit, etc. However, the trend toward bringing customer and other data together corporately (e.g., master data management or MDM) is encouraging in the context of recentralization of data and mining. The software costs and people costs that are normally high on other platforms are minimal for the mainframe in this expanded role. The specialty engine further reduces costs by virtue of eliminating DB2 functions from billable CPU totals.

Linux: Customers that have compared running Linux on mainframes versus other platforms have mixed results. The use of mainframe Integrated Facilities for Linux (IFL) processors fares better the larger the environment. Cost data does not show much difference in overall cost among the categories of people and software, but Linux management software is often lacking and represents a significant skewing of cost data. Where there are a large number of Linux environments to be consolidated, use of virtualization on the mainframe should prove more cost-effective than virtualization on blade servers. In one case, a customer migrated a number of AIX/RISC Linux instances to the mainframe, and yielded a 30-percent ongoing savings (it took six months to breakeven, based

on one-time costs). Again, the economies are based largely on people cost savings, which are much lower on the mainframe.

HP and Sun Servers versus Mainframes: TCOs for mid-range to high-end Unix servers in comparison to the mainframe are generally inaccurate, since they do not consider all factors. Doing such a comparison is also very difficult, considering the difference in processor capacity units (e.g., MIPS versus MHz), discrepancies in software and people tallies, and lack of functionality in distributed platforms that is on the mainframe. The industry benchmarks (e.g., TPC-C) are also misleading, since they assume things like Unix servers running at 100-percent utilization, which is unrealistic. These benchmark results are generally driven by the database throughput performance, which is also complicated because of the drastic differences that partitioning and caching make. Thus, the majority of clients are left with vague notions that the cost is higher on mainframes (perpetuating the myth), and that the pain of migrating workloads to the mainframe is too great. Nonetheless, these workloads can be isolated to their own partition and managed accordingly with great availability, performance, reliability, and security.

The other striking difference between Unix versus mainframe servers is the I/O processing. In the mainframe environment, separate processors are provided to support an I/O director, which avoids the need for the primary server capacity to be used for I/O. This both provides more capacity for application processing and enables use of the advanced I/O technology inherent in the mainframe environment. FICON (fibre channel) on the mainframe is a superior technology as compared to traditional I/O processing in the Unix server environment, and the configuration options are more flexible and redundant in the mainframe environment as well. The figure below illustrates that the TCO for IBM's pSeries and xSeries platforms is higher than for the mainframe, especially in the context of labor costs.



Wintel versus Mainframes: Comparisons of Wintel versus mainframe servers are especially difficult, since the applications are often the major determinant of where the platform should run. In addition, ISVs writing software for Wintel often do not have the capability to provide a mainframe version. In the context of TCO, the main advantages for the mainframe are the people costs, as well as the superior management tools and capabilities to ensure good service. The Wintel platform, however, is the widest known for its low average utilization. Indeed, many customers are paying for boxes that are averaging 15-percent utilization or less. Beyond that, the data and other resources are not being leveraged, and copying/syncing operations are necessary to ensure good backup and data protection. The availability of Wintel platforms is also much less on average than the mainframe, so the lower levels of service and the associated revenue and productivity impacts must be subjectively included in any comparison.

From a utilization standpoint, Unix servers average 35 percent to 60 percent, while mainframes run at 85 percent or more. This is more emphatic on the Wintel side, where utilization averages about 15 percent. One client took four Intel server applications and

consolidated them on a blade server with VMware, with an additive utilization expected at 60 percent. The actual result was 40 percent, indicating that the system overhead on Intel boxes is significant. To extrapolate, 1,000 Intel servers, averaging 15 percent and adding significant overhead unto themselves, are at best averaging in the area of 10-percent or less application utilization. That means that of the 1,000 servers purchased, as many as 800 of them would be idle if they could be utilized in the same manner that a mainframe is utilized. The associated cost of maintaining these additional servers is significant as well.

Current Mainframe Misconceptions

Numerous misconceptions exist around the mainframe, despite the years of experience among most companies with distributed (Unix and Wintel) environments. These misconceptions have led some companies to focus only on supporting workloads that already run there. Most financial services and retail companies, however, do not share that view, and have done innovative things to leverage the mainframe for new

workloads. These efforts have helped to dispel the myth about mainframe disadvantages. Some key misconceptions are discussed below.

Misconception that Charge-Back Costs are Higher: For years, the cost of distributed environments was subsidized by mainframes based on the rates used for charge-back and cost recovery. These rates were necessary due to the inability to measure and charge back for distributed environments. Another common item that mainframe rates are burdened by is the overhead staff. Management and administrative staff are often lumped into the mainframe rates, either because there are no such rates for the other platforms, or because it was always that way and never changed. Finally, facilities costs are often attributed fully to the mainframe, which can be a major factor in TCO analyses. The end result for many has been a number of lost opportunities to capitalize on mainframe strengths where it is truly cost-effective to do so.

Misconception that Distributed Systems are Simpler: For a long time, and even today, many IT executives felt that distributed platforms were simpler, and thus, easier to plan, build, and run. This has proven to be false as the distributed environments become bigger, both in configuration size and sheer number of servers. Indeed, many clients are dealing with thousands of servers, with complexity increasing exponentially. The management processes are weak in these environments, which has led to poor availability due to high error rates. Mainframe support organizations have mature processes and tools for supporting and managing workloads, which is why that lack of "overhead" made the costs seem lower than they were.

Misconception that Mainframes Require More Staff: The level of sophistication

in processes and tools described above is the primary reason that staffing for mainframe support is typically lower than that of most distributed environments. Again, the rigor of production control, change management, etc., caused a higher staffing level than alternative platforms, but that is no longer true. The need for these processes, especially in the context of recent regulatory compliance legislation, is now clear among IT executives, forcing the same rigor for all platforms, and thus, increasing the cost of ownership for these platforms. The other factor to consider, and one that has also affected the charge-back rates, is the dual role of mainframe support staff, overlapping to provide support for other platforms, but charged as mainframe support. This sort of thing is not malicious in most cases, but has contributed to missing opportunities to leverage mainframes where they make sense.

Misconception that Mainframe Migration Costs are too High: Both criticisms have been applied to migrating work either to or from the mainframe. Migrating away is too expensive unless there is a turnkey solution that is easily supported and managed, and even then, the expense does not consider the appropriateness of migrating away from a TCO standpoint. Migrating workloads to the mainframe are often viewed as too high due to misconceptions about the platform. For example, use of Linux IFLs and leveraging virtualization technologies (e.g., logical partitions or LPARs) can provide an easy way to recreate environments within the mainframe.

Misconception that Mainframe Disaster Recovery/Business Continuity (DR/BC) Costs are Higher: The unfortunate misconception about mainframe DR costs being higher is largely due to the hot-site costs that were associated only

to mainframes, despite the addition of distributed systems into these agreements. The costs were high, and the true level of DR provided was questionable because of insufficient testing. This fact exacerbated the already noticeable price tag for mainframe DR, but again, the evolution of distributed systems and the increasing importance of the workloads have demanded the same rigor of DR as that of mainframes.

When looking at the DR equation for distributed versus mainframe workloads, it is critical to normalize the costs of the hot-site configuration to the size of the total workload, as opposed to distributed environments where servers are dedicated to each application. The total cost of DR may or may not be lower for the mainframe, depending upon workload size and complexity, but the larger issue is the complexity itself. The cost of managing a DR plan that includes thousands of servers can be viewed as too tenuous in comparison to a small number of mainframes. On the other hand, "by application" recovery tends to be easier in distributed environments where none of the infrastructure is shared, although the obvious loss of leverage opportunity is a downside in that sense.

Misconception that Mainframes Have Less Flexibility: One of the major reasons that distributed platforms became popular is because they were owned by the lines of business (LOB), and offered greater flexibility by virtue of support going around the central IT group. That did not last long, but it did perpetuate a model that suggests "smaller is better and simpler." The "red tape" associated with the mainframe was a huge issue for many companies, but the issue was largely due to the comparison made to distributed platform "red tape," which was minimal. The reason it was minimal, however, is that

it was inadequate in the context of compliance, governance, and processes. As these disciplines evolved in distributed environments, the "red tape" increased commensurately.

On the other hand, the mainframe environment support processes have the advantage of better tool kits, higher levels of refinement, more maturity, as well as inherent recognition that providing flexibility to customers is imperative. Thus, it is clear that mainframe support provides as much flexibility as distributed environments, excluding the rogue platforms that are implemented in a departmental context.

Misconception that Dedicated Systems are Better: LOB executives felt that having their own systems was better than having them shared on a mainframe. This gave them better control. The outcome of this has been the uncontrolled proliferation of distributed processors. This has become a cost and management challenge. Today, many companies are now looking to virtualization (a staple on the mainframe) to come to the rescue. Many executives are finding that the gains from consolidation through virtualization on blade servers are not meeting expectations, due primarily to the power and cooling issues they create. Some have spent large sums of money, or spread out servers to create more open space for cooling, to address these issues.

Misconception that Mainframes Have No ISV Mindshare: Customers have increasingly looked for commercial, off-the-shelf, packaged applications to address business needs, as well as to find a means of re-hosting applications. When considering the hosting decision for packaged applications, it is not only important that a version exists for each environment under consideration, but also that the support is "primary," and

thus reliable. Several ISVs have treated the mainframe as the "last platform to support," but have done so for a variety of reasons. The real answer about ISV mindshare for the mainframe is that mainframe support (which includes Linux-based applications) is greater than ever. The answer as to why the ISVs choose to migrate to the mainframe last is two fold. First, mainframes have the lowest number of users among the popular ISV packages (e.g., SAP). Second, the software migration effort is simplified by doing so (e.g., the add-ons and changes are simpler when the other code is built and tested on other platforms). Either way, there are numerous ISVs that have realized how standards can minimize the effort to write applications that can run transparently on multiple platforms (e.g., J2EE).

Misconception that Mainframe Software Prices are too High: This has both been a misconception, as well as a point of negotiation that can make huge differences in the TCO equation. The software costs among ISVs have been the issue, with IBM offering various ways to reduce costs through a number of pricing algorithms. The new specialty engines are a great example of IBM providing DB2, Java, and other (e.g., IFLs) processing capabilities that are not being charged for by ISVs. Another major factor is the procurement process, which generally provides better leverage and deeper discounts based on the total investment.

On the flip side of the equation, there are indeed areas where IBM software pricing can be dramatically lower than counterpart solutions. The two obvious examples are CICS and DB2 on the mainframe, where the equivalent pricing in distributed environments can be by user or server, whereas DB2 can support large volumes of users and transactions based on a standard monthly fee. This is

also the case with CICS, where the software supports thousands of users and millions of transactions at a lower cost than other distributed alternatives. The use of selective software stacks in separate LPARs on the mainframe helps lower the costs of CICS and numerous other program products (including ISV products) by associating a smaller capacity with the software's use.

Misconception that Data Sharing is a Bad Concept Architecturally: The mainframe remains the primary environment where data and database sharing is encouraged, mostly because of issues tied to data security, dual-phase commit, and referential integrity across platforms and multiple databases. The approach in all other platforms is to have non-shared or "shared nothing" environments, which results in added costs associated with data copying and synchronization. In addition, it has made job scheduling more complex and prone to failures. Moreover, it has resulted in enterprises having multiple "version of the truth" and no ability to create an MDM environment easily. Architecturally, the shift to distributed systems has hindered business' ability to function effectively and to cross-sell.

Misconception that Power and Cooling Economies are Better on Smaller Processors: In the 1990s, the mainframe was the hottest processor in the complex. However, times have shifted significantly to where the mainframe is now the coolest processor in the data center. The distributed blade servers have become so hot that there are solutions that bring water or other piped cooling systems back into the data center.

Competitors have attempted to take advantage of these misconceptions by denigrating the mainframe as obsolete, when in fact, if investigated, the truth underneath many of the misconceptions actually favors

the mainframe. Given this, companies who have not done so already should indeed rethink the role of the mainframe for modern workloads, and in general, how the platform can be optimally leveraged.

Summary

In today's world (21st century), IT's new prime directive is to maximize the value of existing and future investments in technology, and support staff/processes on behalf of the businesses it serves. IBM has continued to make significant investments in mainframe technology to achieve better price/performance for traditional workloads, support new-age workloads, and provide an industrial-strength solution over alternatives. The need to consolidate, reduce costs, and simplify is driving IT back toward a more centralized view of data processing. The recent emphasis on MDM has created a focus on information services, which leverages enterprise data to derive new knowledge about investment opportunities. The mainframe is architecturally and technologically the best solution to provide shared data via information services.

The advent of open service-oriented architectures (SOAs) that embrace the mainframe can help relegate critical

application logic now on mainframes to SOA-based services that can be reused. This can be a major advantage for customers that want to leverage critical applications residing on the mainframe. These are the things that will enable IT to be more business-aligned and business-focused. The architectural role cannot be underestimated here.

Most companies today emphasize information services (such as data independence), IT Infrastructure Library (ITIL)-based services, and SOAs. As such, RFG believes enterprises can exploit the mainframe as a data-serving hub and for Linux consolidation and transaction-intensive processing. Use of the mainframe as a security hub is also important, but more from a functionality and reliability perspective versus cost. An increasing number of clients are rethinking the role of the mainframe based on its ability to provide centralized administration, consolidation, information services, and superior reliability, availability, and serviceability (RAS). These steps will help maximize leverage around mainframe investments.

Copyright © 2006 Robert Frances Group, Inc. All rights reserved. 120 Post Road West, Suite 201, Westport, CT 06880. Telephone 203/429-8950. Facsimile 203/429-8930 www.rfgonline.com. This publication and all publications may not be reproduced in any form or by any electronic or mechanical means without prior written permission. The information and materials presented herein represent to the best of our knowledge true and accurate information as of date of publication. It nevertheless is being provided on an "as is" basis. Reprints are available for a nominal fee.