Dynamic Business Intelligence

Leveraging mainframe assets to help steer the business

Tony Lock and Dale Vile, Freeform Dynamics Ltd, September 2007

One of the most versatile assets many organizations have in their IT infrastructure is the mainframe, which in recent years has evolved to deal even more effectively than it has in the past with a wide range of specific workloads. For users of mainframe systems one workload that they may wish to consider again is data warehousing. In this paper we look at how the latest IBM System z capabilities could allow the mainframe to play a more central role in a coordinated business intelligence strategy.

KEY POINTS

Process latency in the delivery of business information is no longer acceptable

Gone are the days when organizations could rely on a handful of analysts to periodically process "the numbers" on business performance to generate monthly management reports. In today's business environment, relying solely on this approach cannot maximise efficiency or competitiveness, as those driving the business need to know about important developments before it is too late to act upon them. There is a need nowadays for an on demand, holistic, real-time view of business performance. As a consequence, the modern business intelligence (BI) infrastructure must deliver consistent information, proactively, across the organization in a timely, secure and cost effective fashion.

Flatter organisations, empowered workforces and more accessible "small footprint" technology have conspired to create a range of issues

The flattening of organisation structures and the devolving of decision making responsibility means that more people today require management information to do their jobs. The problem is that most organisations are not well geared up to cope with this "democratisation of business intelligence" - their core infrastructures have not been designed to deal with the rapid, secure and broad delivery of information. Numerous "small footprint" solutions have sprung up independently at departmental, workgroup and even individual employee levels, causing a high degree of information fragmentation, in turn leading to manual overheads, incomplete views of the business, multiple versions of the truth, user frustration and ineffective processes.

A more coordinated architectural approach is required, but it's not that easy

The time has come for many organizations to take a more coordinated architectural approach to business intelligence. While deploying yet another local, independent data warehouse may appear a good idea, the risk is perpetuating the problem of fragmentation. There is much sense in considering a centrally managed approach to ensure that the enterprise operates using a secure, cost effective, single version of the truth rather than with every system creating its own story.

The answer for many might already be sitting in their data centre

Against this background, and for those that have them, a question arises: '*Do we have an existing asset that can be further exploited to provide the joined up capabilities we currently lack, and could our mainframe be that asset?*' Recent advances in the IBM System z architecture, coupled with associated advances in software and tooling plus significant changes in mainframe pricing options can now provide a linchpin to enable a more coordinated, secure, cost effective architectural approach to solving the business intelligence challenge, especially in scenarios where much of the business data resides on the mainframe thereby avoiding the costs and risks associated with data movement.

This paper was commissioned by IBM. It was authored on an independent basis by Freeform Dynamics drawing upon relevant findings from recent research as well as the extensive experience of the Freeform analyst team. Thanks also to the Macehiter Ward-Dutton (www.mwdadvisors.com) team for input in specific areas while preparing this document.

Contents

Introduction

Over the years, most large organisations have spent significant sums of money on business intelligence solutions, from huge centralised data warehouses to a wide range of more distributed solutions based on both commodity components and specialist tools. Yet despite this investment, there are still widespread complaints from business managers and other users that require business performance information to do their jobs. Even those industries best known for their advanced use of IT, such as financial services, are grappling with some serious challenges in this area, as illustrated by the results of a recent survey of senior decision makers in the City of London[1] (Figure 1).

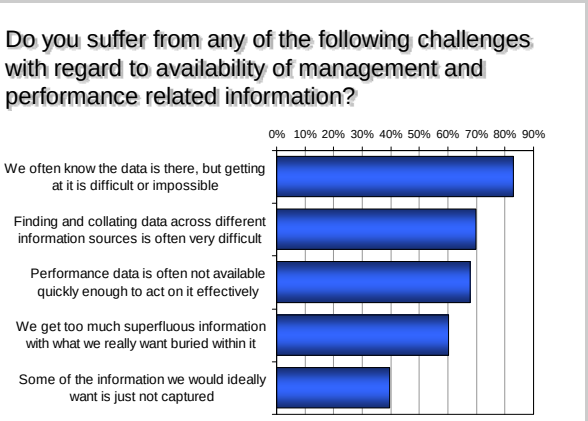

Figure 1

Business intelligence related challenges highlighted by 52 senior business managers from the financial services sector.

Source: Freeform Dynamics, 2007

This picture identifies a state of affairs that is not particularly impressive, but is indicative of a general problem that touches pretty much every large organisation.

With such challenges in mind, this paper considers the nature of the issues, how they have arisen and then goes on to provide some advice and guidance on strategies for moving forward based on a more coordinated architectural approach to addressing business information requirements in a more robust, secure and future proof manner. As part of this, we shall be paying particular attention to the potential role that an existing asset that many organizations have in their armoury, the mainframe, could play. Indeed, despite recent developments which have added many BI / DW capabilities, this might not be the most obvious area to explore. But before getting into this, let's look at some of the practical considerations in more detail.

Challenges and imperatives in more detail

In order to appreciate some of the later discussion on solutions and architectures, it is useful to have a clear understanding of the nature of the beast we are dealing with from a challenges and imperatives perspective. The major challenges faced by organisations in the areas of BI and data warehousing include the fact that data and hence information are now stored in multiple locations making it difficult to obtain a consistent, accurate picture of what is going on at any particular time. This is a matter of great concern as BI is deployed to enable organizations to make forward looking business decisions as a routine part of everyday operations. In addition it is clear that BI can no longer be considered only as a backwards looking review of operations; the need now is for organizations to exploit information to push the business forward. The critical nature of the business decisions being made using BI ensures that the security and accessibility of the information generating systems is now a matter of great importance.

Information, information everywhere!

At the heart of many of the problems we discussed above is the tactical approach taken over the years to addressing business intelligence requirements to serve the local needs of individual departments, workgroups and even single users. Fuelled by the challenges associated with "boil the ocean" corporate data warehousing initiatives that were popular in the 90's, and coupled with the availability of relatively cheap and accessible small-footprint technologies that provided an alternative "fast track" route to delivery, we have seen the uncoordinated proliferation of data warehouses and

marts across many organisations. However we have now entered a phase where consolidation, both physical and logical, is playing a key role as organizations look to obtain greater insight from the mass of data they hold whilst simultaneously reducing operational costs and business risk.

And the end result? Well, apart from the obvious problem of the wheel being reinvented time and time again from a development and integration perspective, the business problem can be summed up in one simple phrase – "information fragmentation" coupled with high, though often hidden, costs.

We can see this very clearly from the issues highlighted on the above chart. Fragmentation means it is difficult for those who need information to find it, and find it quickly enough. Often, they know the data is there somewhere, but locating it is a real issue. There are then the time delays as people in various parts of the business manually collate information for analysis from disparate systems. Then lastly, there is the problem of data inconsistency, which so often leads to "multiple versions of the truth" supporting business management and decision making processes, all delivered with high associated manual labour costs.

Pleasing all of the people all of the time

There is a quote attributed to Abraham Lincoln "You can fool some of the people all of the time, all of the people some of the time, but you can't fool all of the people all of the time". But try telling that to someone trying to maximise performance against aggressive targets when they can't get hold of the information they need. The fact is that today, with the flattening of organisational structures and both responsibility and empowerment being pushed deeper into and more broadly across and outside the business, the audience that needs to be served by business intelligence capabilities is much greater than it used to be.

A major consideration here is the nature of the user. Unlike in the past when business intelligence was about feeding data to a relatively small team of business analysts, today we talk very much about the "democratisation of business intelligence". What we mean is the need and expectation of the broader community of people participating in management and decision making activity to access directly the information they want when they want it. It goes without saying that they want to access the information via familiar mechanisms such as their desktop office tools or their browser, as they can't (or won't) spend time learning how to use complex specialist analytical tools. Further, such democratisation is not restricted to within the enterprise itself as today it is often the case that customers and partners now routinely demand access to information which until recently was confined solely to internal constituents.

Tell me what's driving the business now, not what drove it last quarter

The time factor has been mentioned a number of times so far, which is apt because this is one of the most prominent complaints from users of business intelligence solutions. In order to understand what is behind this, we must consider the concept of cause and effect. Most organisations are pretty good at monitoring overall financial performance retrospectively on a periodic basis, e.g. through monthly management reports. They are not as good, however, when it comes to monitoring the factors that drive that overall performance on a continuous basis, which means that events happening within the business, whether related to problems or opportunities, are often not acted upon until it is too late.

Just tell me what I need to know, when I need to know it

Of course one of the pitfalls of delivering management information at a more granular level is that users can become swamped in data. Indeed, we saw in Figure 1 previously that 60% of respondents complained of the information they really require being buried in superfluous detail.

What these complaints translate to is a requirement for intelligent and proactive delivery. Filtering and exception reporting, based on business rules embedded in data warehouses and other business intelligence components, can hone the information delivered to the subset that is relevant to the user at that time, simultaneously drawing the user's attention to that which matters the most and is most likely to require acting upon.

Taking pro-activity to the next level, selected information may be actually "pushed" to the user in the form of status updates or alerts delivered via email, a mobile device or whatever means is most appropriate and this, along with some of the other mechanisms we have mentioned really does seem to be appreciated by users (Figure 2). Indeed today, we are now beginning to see information being utilised directly into business processes in near real time to help line management make better, faster operational decisions.

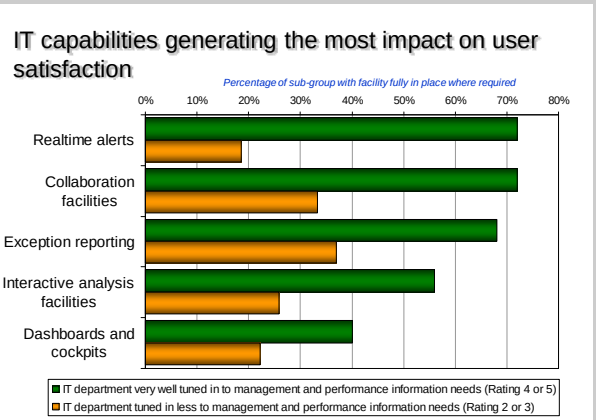

Figure 2

IT departments who are considered well tuned in to information needs are delivering real-time alerts, collaboration capability, exception reporting and interactive analysis facilities.

Source: Freeform Dynamics, 2007

So what are the implications of the challenges and imperatives we have been discussing from a business intelligence infrastructure evolution perspective?

Infrastructure evolution

The evidence from our research suggests that few reading this paper will fail to recognise at least some of the above challenges and imperatives. Looking around at your infrastructure, however, probably utilizing a mixture of IBM, Oracle and Microsoft database technologies running data warehouses and integration software on a variety of hardware, you may be wondering where to begin the journey to a more coherent and integrated business intelligence approach.

Possible approaches

At the highest level, you have three options:

1. Continue as you are, solving business intelligence problems at a local tactical level, in which case many of the problems identified previously are likely to persist;
2. Take measures to proactively rationalise and/or integrate existing data warehouses and other information stores into a logically and physically coherent architecture;
3. Implement a more gradual evolution programme, defining criteria and guidelines for new projects, effectively halting the fragmentation and improving coherency over time.

The last option is really a “slow burn” version of the second one, applicable when time or resources to spend on general modernisation and improvement are limited.

For those who regard options 2 or 3 as desirable, and those that already have an improvement initiative in place, it is worth considering your scenario and the options open to you. We will look at some common scenarios in a little while, but first let us reflect on some of the key considerations and requirements for a future proof business intelligence infrastructure.

Key platform infrastructure considerations and requirements

When looking at the requirements we have been discussing, a clear set of considerations emerge with regard to the business intelligence infrastructure, particularly in terms of data warehousing platforms and associated integration and management tools, so let’s now introduce this systems dimension into the equation.

Upping the level of currency and responsiveness

As we have seen, there is increasing evidence, both empirical and circumstantial, showing that the modern business must significantly change the way that it uses business intelligence / data warehousing solutions.

In the past it was commonly the case that the data warehouse was regarded as simply a big bucket into which data was periodically placed, with some degree of data transformation and aggregation taking place along the way, usually via ETL technology. This methodology has been in use for the last 10-15 years utilizing reporting tools from organizations such as Business Objects, Hyperion, Cognos, SAS, etc. Such approaches have been very much the norm when it comes to driving management reporting. Indeed, many enterprises have employed such static warehouses, typically combined with batch oriented refresh and report generation, perhaps based on daily, weekly or potentially monthly cycles, as the foundation for supporting business intelligence requirements. Whilst such systems are widely deployed the value they deliver is variable and the growing demand for access to dynamic information means that for many organizations the static data warehouse approach may no longer be entirely suitable.

This change in the need to access business information on a more on-demand basis requires that data warehousing systems contain current, or at least near current, information at all times. Equally it is necessary that the reporting, query and analysis tools themselves can be used by decision makers in real time and that business processes be able to access such information in real time to enable greater business responsiveness. This move has placed new demands on the business intelligence infrastructure in terms of core warehousing capability to handle “in flight” updates and refreshes of data, and an increased emphasis on integration and BI process management automation that are essential to tapping into source systems in an effective and efficient manner. Even real-time “drill down” from data in the warehouse into operational information stores has become a requirement.

In practical terms, such requirements, coupled with the need to keep the data pool on which decisions are taken up to date and as free as possible from holding contradictory data, creates an obvious case for keeping warehouse systems logically in close proximity to as much of the source data as possible. In such situations it becomes possible to avoid the risks associated with the proliferation and movement of data, especially information integrity and security. In an ideal world one would want to access the information in place, but this may not always be practical. However it is important that information doesn’t undergo more transformation, aggregation etc. than is absolutely necessary. Multiple transformations of data can have unanticipated effects on data integrity (similar to the “Chinese whispers” effect). It is also clear that in scenarios where the root data is held on disparate platforms that the BI / DW hub must be well integrated with the original data sources.

The new requirements for information to be available to very large numbers of people and processes in near real time has in turn generated a number of technical demands that modern BI systems must be able to meet to ensure that accurate, timely decisions can be made. Amongst these there is a clear need for BI / data warehousing systems to be able to process data in parallel from numerous sources coupled with the requirement for sophisticated, centralised meta data management and scheduling. These in turn require that adequate auditing, monitoring and staging capabilities, amongst others, be available and manageable from a central perspective.

Dealing with access needs of a broader audience

As has already been mentioned, there is a clear trend for formerly specialist “Business Intelligence” activity to move much more closely into line of business job functions.

One of the implications of this in infrastructure terms is a need for warehousing systems to be accessible from a range of familiar front end tools, as well as the more specialist analytical software that would typically be run by the business analyst. Perhaps of yet more importance is the need for business applications and operational processes to be able to utilize BI generated information in real time. These necessitate creating systems that interoperate through a combination of open standards support, e.g. published APIs, service-oriented interfaces to information assets, etc. to allow easy and flexible integration with user’s tools of choice.

Data warehouse performance, scalability, resilience and security

As the volume of information routinely acquired by organizations continues to grow at rapidly accelerating rates, and the need to support many more users monitoring and managing business performance at a more granular level continues to increase, there is a clear need to be able to cope with very large scale information access.

These factors have combined to create the requirement that data warehouse / BI systems not only be capable of scaling to handle large data volumes, but that such systems be highly performant, available “on an ad hoc basis” and able to meet increasing demand in a predictable fashion. The

watch-words here are performance, security, scalability and resilience. Gone are the days when the data warehouse being down was just an inconvenience. With more organisations looking to embed business intelligence enabled decision making into their day to day operational processes, the BI system(s) itself becomes a business critical component of the overall IT infrastructure.

It almost, though not quite, entirely goes without saying that given the nature of the information being processed and the use to which such systems are now put, the operation of these systems is a matter of concern, especially as the requirements of users of these BI systems, inside and outside the organization, clearly pose a wide range of security / data protection challenges. When coupled with the need to ensure data integrity and that access to sensitive information is restricted to those with "a need to know", the operational security of BI / data warehousing systems becomes highly visible and sensitive.

It is widely accepted that the mainframe platform is still considered by many to be the "gold standard" server platform in terms of ease of management as workload increases and, even more importantly, the most secure computing platform available today, a fact enhanced by the robust management processes employed in its daily operations.

Dynamic Data Considerations

Risks associated with data proliferation and movement to other platforms from the Mainframe:

- *Integrity*
- *Security*
- *Access to Skills / Knowledge*
- *Cost of Management / Ownership*

What everyone really wants is to build a BI / data warehousing architecture that delivers the information business users require, when they need it and one in which it is straightforward to manage complex, and sensitive, security issues in as simple a fashion as possible. With security management long baked into its soul, the mainframe clearly holds great potential in these scenarios. It is possible to secure distributed systems to an appropriate degree, but the effort involved far exceeds that necessary to secure a mainframe. Here again the low incremental TCO of the mainframe is likely to prove valuable.

In summary, the increasing importance of factors dictating that information be well secured and highly available make the mainframe a very attractive proposition deploying dynamic data warehouse / BI systems, especially those that are required to keep business processes operating effectively in real-time environments. In addition the mainframe's low cost of ownership should ensure that BI and the mainframe be viewed as naturally complementary in many scenarios.

BI / Data Warehouse Evolution

It is clear that over the last few years database and business intelligence systems have slowly changed in their fundamental usage patterns and this is now driving architectural changes. Originally database systems created to handle OLTP workloads were optimized to cater for real time access demands coupled with the need for them to be highly available, scalable and secure. On the other hand traditional data warehouse systems were optimized to handle read performance related issues and were typically addressed using complex data partitioning techniques coupled with sophisticated workload management tools.

These changing demands are reflected in the way that database / data warehouse / BI solutions have developed. In the first phase of their use, Query and Reporting were the main functionality delivered to provide, for example, periodic financial or sales reports. In the second generation of such systems the business demands were met via data mining type operations to perform multi-dimensional analysis to help optimize inventory management and support merchandizing. The latest business requirements focus around active business management and real-time decision making. This in turn promotes a need for a third generation of BI / Data warehousing systems that can generate a wide variety of information, potentially from multiple data sources when the user needs it, i.e. now.

The requirement today for systems that are capable of meeting the business needs highlighted in this document is forcing BI / DW architectural evolution. In essence, systems capable of operating in a "dynamic data warehouse" environment must combine the characteristics from both the OLTP and traditional DW systems. However, it is difficult for systems built with read optimization in mind to adjust to cater for the more interactive transaction processing type demands. Thus it is, at least architecturally, more straightforward for OLTP systems to handle sophisticated BI demands when built on a platform designed to handle these workloads, such as the mainframe. To the ability to handle

these workloads can now be added the fact that the operational licensing costs traditionally associated with running such interactive BI demands can be significantly reduced by use of the new IFL, zIIP and zAAP specialty engines that have been introduced in recent years.

Mainframe Architectural options

The Evolution of the Mainframe as a Cost Effective Enterprise Platform

Before we examine the three potential architectural solutions utilized in BI and data warehouse environments it is worthwhile examining the architectural development of the Mainframe over the course of the last few years, many of which have potential use in modern BI / DW environments. Many of the benefits enabled by the use of the specialty engines are described in more detail in the following section, but it should be noted that in terms of operational licensing costs engines such as the IFL, zIIP and zAAP effectively enable an organization to build and operate an "open" environment coupled to the Mainframe in a single, secure and cost-effective System z platform.

In addition it is worthwhile recognizing that as workload demands increase the requirement for additional physical space, power supply and associated cooling scale much more slowly using a System z Mainframe than alternative Unix, Linux and Windows servers. The requirement for additional, expensive operational administrative resources to manage the increased workload are also minimised as the Mainframe was designed from its birth to be manageable as it scales. [2]

The Mainframe, Data warehousing and BI

Architecturally, there can be considered to be three major elements to a BI and data warehousing environment:

(1) the reporting and query interfaces

(2) the analytical engines and / or data integration engines

(3) the data store(s)

Experience has indicated that user interaction is ideally suited to graphical user interfaces. Thus, we can assume that the reporting and query interfaces will operationally focus on the desktop / workstation.

The analytical engines provide a bridge between the reporting and query interfaces and the data stores. The choice of platform is dependent on two factors: (1) high speed access to the source data and (2) a cost-effective compute intensive environment. Today, all of the widely used solutions in this area easily connect with the underlying data stores, regardless of platform. Thus, the decision will typically be based solely on which platform provides the most cost effective operational profile for the organization or on operational platform preferences / experience.

The selection of the optimal platform or platforms for the data stores centres on factors that have already been introduced: speed, scale, security and reliability, as well as cost. This decision can be more complex as several factors impact the choice of platform.

Offload Engines

Customers making use of zIIP processors free up capacity in their System z9's general computing resources by effectively pushing eligible database oriented workloads (such as certain business intelligence and data warehousing queries) onto the dedicated zIIP engines instead of running on the main System z9 processors. Whilst each zIIP engine needs to be purchased in addition to the System z9 processors, IBM has chosen not to impose any software charges on zIIP capacity. This licensing decision saves customers money compared with standard licensing models which impose software charges based on MIPs.

In addition, zIIP engines were designed to enhance the performance of business applications. IBM DB2 for z/OS version 8 was the first IBM software fully capable of exploiting zIIP and workloads using select query processing (e.g. BI, ERP or CRM network-connected applications) and business intelligence application query processing utilizing the DB2 star-schema parallel query capabilities. These queries along with DB2 utilities performing index maintenance structures see significant performance and price-performance benefits.

A second workload engine exists, IFL, to run Linux applications on the Mainframe. This option may prove to be attractive to organisations that utilize Linux applications as part of their BI systems but who wish to exploit the low cost of management and ownership associated with the Mainframe along with the associated benefits of security and availability.

Taking the IBM System z as our reference point for mainframe computing, the architectural options considered as ways to meet future BI and warehousing needs may be broadly classified for the purposes of discussion in this paper as:

- System z-centric BI / DW: locate the operational data store alongside the data warehouse on the mainframe
- Hybrid BI / DW: mix data stores on the mainframe with those on distributed platforms
- Non-System z BI / DW: store all data off the mainframe

We will now review the nature of each of these and the relevant merits of each model as the data store foundation for a forward looking business intelligence infrastructure.

System z-centric BI / DW

In a System z-centric environment, all of the fundamental information components of the solution are hosted on the System z server, including the operational data store (ODS), the data warehouse (DW) itself and any data marts utilized in the system. This model is well suited to address strategic business needs described earlier in the paper, centred around the demand for real-time or near real-time access to information to support key decision making or business process requirements.

Typically the systems will be based around the IBM DB2 offering, functioning under the z/OS operating system. Any operational information sources not hosted on the System z platform are integrated using specialist connectors to provide a single hub of information utilised in the BI system. In this scenario the tools utilized by end users are either hosted on their workstations or run on distributed application servers.

The qualities inherent in Data Warehouse and BI systems utilizing this “pure” System z architecture are very much those intimately associated with the System z mainframe platform itself. Put simply the mainframe is usually regarded as the litmus test of servers when it comes to discussing factors such as security, scalability, reliability and availability. In addition, this model allows information to be accessed in rapidly varying business scenarios where key processes or decision making actions are based on having accurate and up to date information available.

It is worthwhile noting that the mainframe has long been a cornerstone of business activity. Over its forty plus years of development the mainframe has become synonymous with availability and reliability. The fact that System z is designed for continuous operations is a characteristic that has, perhaps, fallen out of the minds of many. Unplanned downtime on System z is almost unknown and planned system interruption is in most organizations a very rare occurrence. Equally the very high security classification that the mainframe routinely delivers is another capability that is now taken for granted. However, these qualities will be valuable to many organizations, especially as data / information security is a matter of paramount importance to every business executive. In this context it is worth remembering that the IBM System z9 has long incorporated numerous capabilities to secure transactions and the information held on the platform. Capabilities including encryption, sophisticated identity management and authentication along with a host operating system that has been refined to handle the most complex and sensitive security requirements that extremely demanding customers can place upon it.

Another factor to consider lies in the ability of BI / DW solutions based on System z to scale in a straightforward fashion. Several studies have shown[2] that as System z solutions grow the management effort required to keep them operational does not scale in anything approaching a linear fashion. The manageability of System z solutions helps ensure that the cost of ownership of the Mainframe compares very favourably with those of other platforms.

Studies have shown that in terms of cost per user per year, System z can be lower to operate than other platforms[3]. This factor has been greatly assisted over the last couple of years with the advent of mainframe specialty engines. The latest of these, the System z9 Integrated Information Processor (known as zIIP), has been designed to run key database workloads. zIIP is the second such specialty engine on the mainframe and follows on from the zAAP processor that handles Java workloads.

In many scenarios where the System z platform is important in the BI / Data Warehousing solution it is important to ensure that the complex matters surrounding data integration are addressed, especially in dynamic business situations. The mainframe has established great interconnectivity with a very wide range of alternative platforms and now boasts sophisticated software tools making it well able to

address data integration questions. A final factor that is becoming increasingly important is the ability of the System z to manage highly variable workloads in a secure fashion without imposing untenable demands on systems administrators. This matter is of no small consequence as data warehousing and BI systems are now forced to handle not only continuous, fairly predictable workloads generated by routine report generation, but are now faced with handling on-demand analysis and management information generation on an ad-hoc basis by growing numbers of business users. The highly variable and rapidly changing demand for service created by these mixed workloads plays well with the capabilities of System z whose Workload Manager has been designed and refined over many years to ensure that access, control and service levels are maintained in line with defined business needs automatically as workload demands vary.

BI / DW on Hybrid System z Architectures

In so-called Hybrid systems the operational data store (ODS) and the data warehouse (DW) is hosted on the IBM System z Server running z/OS and IBM DB2. Statistical data marts and end user tools are operational on Unix, Linux or Windows platforms. The end user systems may frequently only contain Web browser tools to provide access to the systems, possibly with local data stores and local manipulation / presentation products.

Hybrid systems share many of the benefits associated with the homogeneous architecture highlighted previously as much data storage and processing remains on the mainframe. The architecture comes into its own when the analysis tools that are utilized have not been written for the z/OS operating platform. In these scenarios the analysis applications continue to run on the platforms for which they were developed, but can "transparently" make use of data resources held on the mainframe. It should be noted here that the mainframe platform also supports the Integrated Facility for Linux (IFL) and ICFs specialty engines in addition to zIIP and zAAP. The fact that IFL specialty engines allow Linux workloads to run on the System z platform without incurring any IBM software licence charges potentially adds great attraction for organizations operating BI / Data Warehousing in these hybrid architectures. For example, the use of an IFL could permit an organization to run any Linux dependent components of the BI solution on the mainframe with all of the attendant security, reliability, manageability and TCO advantages without losing the benefits bestowed with the use of Open Source tools.

Why Choose a Hybrid Solution?

Hybrids involve users with data on distributed and mainframe platforms. Making the implementation mainframe-centric or distributed is a decision that organizations will base on a number of factors including:

- *Volume of data on the various platforms*
- *Skills associated with managing the data on the various platforms*
- *Security issues*
- *Capacity issues*
- *Scalability*
- *Availability of resources, power consumption, heat generation, cooling etc.*

BI / DW not utilizing System z

In this architecture the System z platform plays no role at all, other than as a potential source of data to be accessed as required. These solutions typically utilize Unix, Linux and Windows platforms and may also make use of dedicated data warehousing appliances. Such organisations may not employ Mainframe systems at all in their environments.

Such systems can be straight forward to develop without the requirement to develop specialist mainframe skills. Until recently these "Open System" solutions often benefited from speed of development as specialist business analysis tools were developed on Unix and Windows platforms. In addition, the capital acquisition costs associated with setting up Open systems were often less costly than those associated with the mainframe and the skills necessary to operate open systems were readily available whilst mainframe administration skills may have been either unavailable or reserved for operating so called business critical applications.

It can also be argued that until recently, in particular with the development of the mainframe specialty engines, the mainframe platform did not offer similar levels of price performance to those achievable on Open systems.

Before the advent of specialty engines it was difficult to cost justify a mainframe-centric approach to BI. Previously the visibly lower hardware and software costs of distributed environments, even when allowing for redundancy and higher management manpower costs associated with distributed server farms, attracted many organizations. It is possible that over time the advent of the IFL, zAAP and zIIP specialty engines may fundamentally revise the perceived price performance attractiveness of the mainframe when coupled with its security features and low TCO, especially as the mainframe continues to enjoy much lower operational costs, especially in terms of manpower, than other platforms.

Relevance of options

The important question to consider now is where and when the three architectural approaches are most appropriate given the attributes, benefits and limitations we have discussed. We will now look at each of these in turn then pull the various threads together into some recommended next steps.

Where is a mainframe centric approach to BI / Data Warehousing appropriate?

There are several factors that indicate when a mainframe centric approach to BI / DW should be considered. Situations that point towards a System z approach can be found below:

- Where significant sources of data (e.g. data warehouses, transactional, operational data stores etc.) are held in System z data sources including DB2, VSAM, IMS amongst others;
- There are existing System z / DB2 skills available and the organisation is prepared to continue to invest in them/expand them;
- Mission critical warehousing where “Management”, “Security” and “Risk” drive corporate policies;
- Organisations where System z is operationally connected to major data repositories;
- Scenarios with highly variable demand for standard reports, ad-hoc intelligence analysis, OLTP etc;
- Where continuous access to “analytical data” BI / DW resources and reports is essential for people, other systems and business processes to operate effectively;
- Situations where the existence of multiple DW / Operational Data Sources / Data Marts causes conflict.

Where is a System z / hybrid approach to BI / Data Warehousing appropriate?

- Keeping Data Warehouses updated in scenarios where the majority of source information is held on platforms other than System z.
- When geographic distribution significantly improves performance for users who are remote from the centralized mainframe resources.
- Scenarios that incorporate analytical tools that are not available on native z/OS.
- When a cost/benefit analysis determines that the complexity of a multi-platform environment is offset by the mixed price/performance profiles of the systems involved. In these situations it is now possible that the use of mainframe IFL engines could provide an alternative to traditional hybrid approaches.

> ***The IFL Advantage***
>
> *With the addition of the IFL the mainframe effectively becomes a logical hybrid architecture combining the best of Linux and z/OS in a single centrally managed system that may be a lower cost solution to a traditional hybrid architecture.*

Where the mainframe approach to BI / Data Warehousing does not fit?

- Customers committed to utilize database platforms other than IBM DB2.
- Organizations where there is little / no System z experience / expertise.
- Organizations seeking to reduce mainframe workload.

Recommended next steps

There are essentially two main parameters to consider when looking at next steps – the current landscape in terms of existing data warehouse installations and the mix of sources being tapped into. These together give rise to a range of scenarios, and recommended next steps (at least at a high level) for each are laid out in the table below. Where the scenario is against a green background, this indicates a good starting point for moving forward in strategic terms. An amber background means there may be some strategic adjustment required and red indicates a possible strategic misalignment that represents a potential future exposure of the business.

APPRAISING YOUR POSITION AND OPTIONS

The following is clearly no substitute for a full structured appraisal of your position and objectives, but identifying where your organisations sits in the table below may help in providing some high level pointers and ideas to ensure the optimum approach is being taken from the cost, value and risk perspectives.

LANDSCAPE / SOURCES	*Operational sources mostly mainframe based*	*Mix of mainframe and distributed systems operational sources*	*Operational source mostly distributed systems based*
Mainframe currently the dominant architecture for data warehousing	The architectural approach is almost certainly a good match for dealing with business requirements. Explore options for providing access via portals and desktop tools as necessary, and ensure you are up to speed with zIIP and the latest integration and management tools from IBM.	Make sure you are fully up to speed on the latest integration and management tools to streamline and effectively manage the integration of distributed as well as mainframe sources. There may be opportunities in both areas to boost service levels delivered to the business.	Make sure you are fully up to speed on the latest integration and management tool options to streamline and effectively manage the integration of distributed as well as mainframe sources. There may be opportunities in both areas to boost service levels delivered to the business.
Equal dominance of mainframe and single distributed systems option	There may be advantages in developing a strategy for a structured hybrid solution approach, with the mainframe acting as a coordinating hub. This would streamline integration with data sources and avoid disruption to existing local solutions,	The main imperative here is to tackle problems of fragmentation that are likely to exist in this kind of scenario. Again, there may be advantages to considering a structured hybrid approach with the mainframe acting as a coordinating hub to keep everything in sync.	The key to moving forward from this scenario is to evaluate how the mainframe and distributed warehouses are being used. It is common in this scenario for the mainframe to have assumed the role of a large legacy "bucket", in which case a strategic review is required.
Single dominant distributed systems option for warehousing	This may be an acceptable situation, but the mainframe asset might not be fully exploited and addition of specialist offload engines could be useful. There may be advantages in exploring a strategy for a structured hybrid solution approach, with the mainframe as a hub.	The architectural approach is likely to be a good fit with business requirements. There may be advantages to considering leveraging the latest developments in mainframe warehousing and integration technology, but that depends on current capability and coherency.	The architectural approach is likely to be a good fit with business requirements. There may be advantages to considering leveraging the latest developments in mainframe warehousing and integration technology, but that depends on current capability and coherency.
No real consistency in data warehousing usage	This scenario represents a potential business risk for the future and a strategic review is highly recommended		

Discussion

A significant number of organizations today employ IBM System z mainframe computers in order to run many facets of their IT operations. The platform has merits that are widely, if often now sub-consciously, accepted, which include the ability to manage highly variable workloads comfortably, and the platform's outstanding security features and ease of security administration.

Beyond this, the ability to scale without needing significant, additional skilled management time and effort is perhaps less widely recognised, but with demands on IT escalating relentlessly while resources remain constrained, this is rapidly becoming a very attractive feature. Indeed, against the background of a frequently encountered market perception that the mainframe is "expensive" to manage, the reality is almost entirely the opposite, especially once a critical workload level has been reached, after which additional workloads can be supported with minimum additional resources. From a commercial perspective, therefore, it is fair to say that while mainframe systems are not cheap in capital terms to acquire, the operational cost per unit of workload drops consistently as workload increases.

These characteristics are worthwhile reviewing for any organisation with an existing IBM System z investment that is faced with having to deal with the changing information delivery demands highlighted in this paper. Depending on the environment and existing systems landscape, there may well be merit in integrating the mainframe into strategies and plans for business intelligence and data warehousing. In fact, the case for doing this has been strengthened by the emergence and/or enhancement of mainframe capability that directly addresses the needs of these areas. Some of these developments, as articulated by IBM, are summarised in the box-out to the right.

Continued Platform Evolution

The ability of the mainframe to operate in information delivery intensive environments has been greatly enhanced over recent years as the IBM Information Server has matured. The company has sought to make it possible to deliver, maintain and manage a range of BI and data warehousing environments employing a variety of server architectures, and when coupled with its Dynamic Data warehouse initiative it is interesting to note the potential that has been made to cater to requests for operational flexibility. The Dynamic Data warehouse initiative is designed to ensure that the mainframe has all of the functional capabilities to handle parallel data processing, meta data management, scheduling, auditing etc. that are essential to build a responsive, real-time BI system that can be administered without the need for additional management manpower resources. The addition of the IFL allows organizations to now build a logical hybrid BI solution entirely within the IBM System z9 platform combining the best characteristics of Linux for z and mainframe operations in a single, well managed, cost effective and secure platform.

In conclusion, the democratisation of business intelligence and the need for much more dynamic information delivery capability in the future means IT departments must consider all options for meeting rapidly evolving requirements. While the mainframe is not the answer to all business intelligence related problems, if you have such an asset, the chances are it will have a role to play somewhere in the equation, and could even form the linchpin for a forward looking strategy.

We hope the discussion in this paper has provided some insights and guidance to help you establish whether and how this important asset might be relevant.

References

[1] Report to be published, Freeform Dynamics (2007)

[2] IBM Scorpion Studies for Server Consolidation (2000)

[3] Arcati Research Limited, Dinosaur Myth 2004 Update (2004)

About Freeform Dynamics

Freeform Dynamics is a research and analysis firm. We track and report on the business impact of developments in the IT and communications sectors.

As part of this, we use an innovative research methodology to gather feedback directly from those involved in IT strategy, planning, procurement and implementation. Our output is therefore grounded in real-world practicality for use by mainstream business and IT professionals.

For further information or to subscribe to the Freeform Dynamics free research service, please visit www.freeformdynamics.com or contact us via info@freeformdynamics.com.

About IBM

At IBM, we strive to lead in the invention, development and manufacture of the industry's most advanced information technologies, including computer systems, software, storage systems and microelectronics.

We translate these advanced technologies into value for our customers through our professional solutions, services and consulting businesses worldwide.

For more information on IBM, please visit www.ibm.com.