CLABBY ANALYTICS

Advisory

IBM's System x with eXFlash: Workload Optimized, Pre-Integrated Server/Storage

Introduction

In this *Advisory*, *Clabby Analytics* takes a closer look a class of servers known as "pre-integrated" servers. These types of servers are usually assembled by information systems vendors that have deep system/subsystem knowledge — giving them the ability to highly integrate a system's central processing unit (CPU) with underlying subsystems (memory, storage, input/output [I/O]). The end result of this deep integration is an extremely high performance system design upon which a variety of workloads can be mounted.

> ***To better describe what a pre-integrated server is, we focus on IBM System x eX5 models. We use this example because the System x eX5 models provide a fine example of a pre-integrated system design that tightly optimizes memory, storage and I/O (input/output) within a single server chassis.***

The benefits of using a pre-integrated system design include:

- Systems vendors perform deep system/subsystem integration, meaning you don't need to spend your valuable information systems resources trying to tune and optimize your systems/subsystems and related infrastructure;
- These vendors have been known to achieve levels of performance that your own engineers may never be able to match;
- There can be major cost advantages involved when deploying pre-integrated servers versus implementing build-it-yourself solutions;
- Pre-integrated server environments may offer significant power consumption advantages as compared with build-it-yourself distributed system designs;
- The performance advantages delivered by pre-integrated servers often make it possible to execute certain tasks ***exponentially faster*** – enabling users to perform their work more quickly (this improves worker productivity, improves customer service response time, and can potentially help create competitive advantage). Further, these performance advantages may also free up computing resources such that they can be used to work on other tasks (this improves your return-on-investment on systems purchases).

Workload Optimization and IBM's System x eX5-class Servers with eXFlash

System x is the branding of IBM's x86-based server line. It is comprised of several offerings – from blade and rack servers through tower servers and workload optimized systems.

IBM's Workload Optimized Systems feature:

- New database optimized models with support for 200 GB Solid State Drives
- New virtualization optimized models

- New Workload Optimized Solutions for SAP In-Memory Appliance, SAP HANA

When we look closely at IBM's System x servers, several differentiators quickly surface as compared with other x86-based servers. System x server designs offer:

- Greater scalability than competitors such as dual node x3850, MAX5 memory expansion to x3850 and x3690, and ability to combine two MAX5 with two x3850 for eXA Scaling with FlexNode partitioning capability
- *More highly-advanced subsystems integration* (particularly when it comes to memory management, internal storage integration, and communications fabric support)
- Extraordinary reliability features such as Double Device Data Correction (double chip kill), single device data correction, four full-width Intel QuickPath Interconnects and memory mirroring

Major System x Design Differences: Highly Integrated Subsystems: Storage, Memory, I/O
If you want to understand what a pre-integrated system is, take a close look at what IBM has done to integrate its storage, memory, and I/O subsystems with System x Intel-based x86 microprocessors.

Let's start with IBM's eXFlash storage subsystem:

> ***eXFlash is an internal storage unit that can support up to eight hot-swappable SSDs (solid state drives). These SSDs can take advantage of one or two performance-optimizing storage controllers that enable major SSD performance acceleration. eXFlash is supported on IBM System x3690 X5,x3850 X5, and x3950 X5 servers as well as other servers in the System x portfolio.***

eXFlash can be differentiated from other vendor's solid state storage in the areas of RAID (redundant array of inexpensive disks); by the controllers that IBM has built; and by the use of NAND flash memory. In the areas of RAID and controller design, other vendors use slower PCIe controllers (and those controllers don't offer RAID striping). Further, IBM offers level cell NAND flash memory to ensure high SSD reliability. The combination of these three technologies server to make IBM's SSD offerings faster than competitors' offerings — and make IBM's x86 server offering more reliable.

Now let's examine IBM's System x eX5 memory management environment:

> ***For years, IBM has offered a memory management subsystem that features a distinct memory controller (the current generation is known as eX5) that features the use of a dedicated application specific integrated circuit (ASIC) that off-loads the x86 CPU from having to handle a variety of communications tasks (for instance, the management of virtual server connections). We see the use of this hardware component coupled with the use of IBM's advanced memory management software as another distinct differentiator for IBM.***

From a communications subsystem perspective, we are most impressed with System x network interface controller (NIC) pricing, related switch pricing, its open fabric offering, and the way IBM handles resource virtualization.

One of the biggest challenges for distributed systems managers and administrators — and especially blade administrators — is the constant provisioning/reconfiguring of systems to adjust for changing workloads. To provision and configure blades, administrators need to

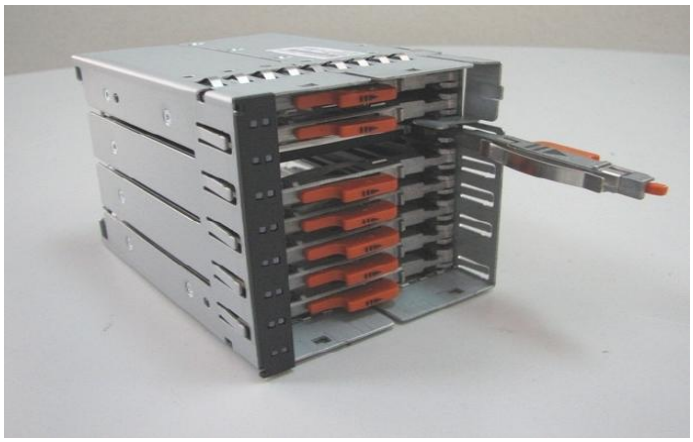
constantly assign and manage a myriad of MAC (Media Access Control — or Ethernet Hardware Access) addresses and World Wide Name (WWN) identifiers — a monumental task in large–scale environments. After basic hardware set-up, systems administrators are called upon to assign LAN connection addresses and then SAN connection addresses. And then these administrators must repeat this procedure for every server that is being installed.

IBM, with a product called IBM BladeCenter Open Fabric Manager, has found a way to mix data and storage into a common virtual pool of I/O resources that can be shared across a set of blade servers. Instead of physically mapping each blade to external networks every time a new or replacement server is needed, IBM's BladeCenter Open Fabric Manager automatically maps new or reconfigured servers with a process referred to as I/O Address Virtualization). Systems administrators also have the ability to pre-assign all LAN and SAN connections and failover scenarios across the entire enterprise portfolio of switches and pass-thru modules. By using IBM's Open Fabric Manager, blade managers can cut configuration time down from days to minutes, saving their enterprise substantial budget by eliminating significant and expensive administrative manual labor.

Let's Take a Closer Look at eXFlash

As described in the previous section, eXFlash is a combination of SSDs and specialized storage controllers designed to deliver enhanced performance and higher reliability over other internal PCIe-based solutions. eXFlash drives are offered in 50 GB and 200 GB capacities – and up to eight 1.8 inch SSD units can be configured in hot swappable drive bays (see Figure 1).

Figure 1 – An 8 Bay SSD eXFlash Cage



Source: IBM Corporation – December, 2011

As for access, two hot swappable eXFlash disk cages can fit in the front bay of IBM x3850 or x3950 servers; while three eXFlash disk cages can fit into the front bay of IBM x3690 servers. The total number of drives an IBM eX5 server can support is 24 – delivering access to up to 4.8 TB of local, internal super-fast storage.

As for access, two hot swappable eXFlash disks can fit in the front bay of IBM x3850 X5 or x3950 X5 servers; while three eXFlash disks can fit into the front bay of IBM x3690 X5 servers. The total number of drives an IBM eX5 server can support is 24 – delivering access to up to 4.8 TB of local, internal storage. And, IBM X5 and other System x platforms are expected to soon to support over 12TB of internal SSD storage.

A Closer Look at eXFlash Performance/Cost Characteristics

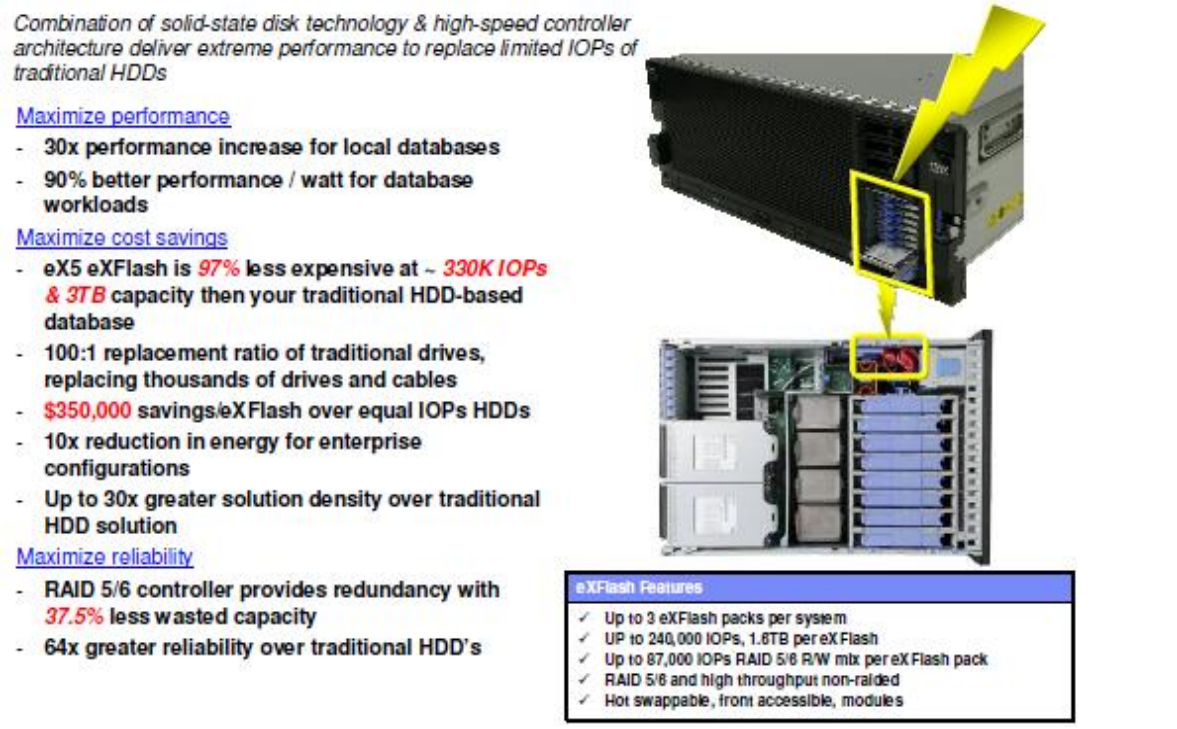
In short, the closer you locate your storage to your CPU, the faster the data can reach that CPU to be processed. In the past we've written about how more and more enterprises are placing data in very large memory (VLM) such that it can be processed more quickly than doing read/writes to mechanical disk. But, with the advent of affordable SSD technology, SSDs are now helping to offset the cost of buying loads of expensive memory – and accordingly, the integration of SSD storage subsystems is becoming a new means to build high-performance systems for data-intensive and parallel processing applications.

With eXFlash, IBM has made it possible to achieve significantly higher IOPS (I/O per second) and deliver more megabytes per second (MB/s) than can be achieved using farms of SATA/SAS mechanical hard drives. Using integrated solid state disk, *IBM can deliver the performance of hundreds of mechanical disks in a single server*. Using eXFlash, IBM has been able to achieve IOPS performance that is six times faster than HDD mechanical disks – offering up to 240,000 IOPS per eXFlash unit. Further, IBM has been able to increase storage capacity within a chassis by a factor of 4X (currently offering 1.6 TB [terabytes] of storage per eXFlash cage – scaling to 3.2 TB per cage).

> ***IT buyers can expect further expansion of IBM eXFlash capabilities for both IOPS and throughput performance and capacity over time as SSD technology improvements are developed and adopted by the industry.***

For a more concrete example of the kind of performance advantages and cost savings – as well as some of the reliability benefits that can be expected using a pre-integrated eX5/eXFlash system design – consider the following (Figure 2).

Figure 2: eXFlash Performance, Cost, and Reliability Benefits in a Nut Shell



Source: IBM Corporation – December, 2011

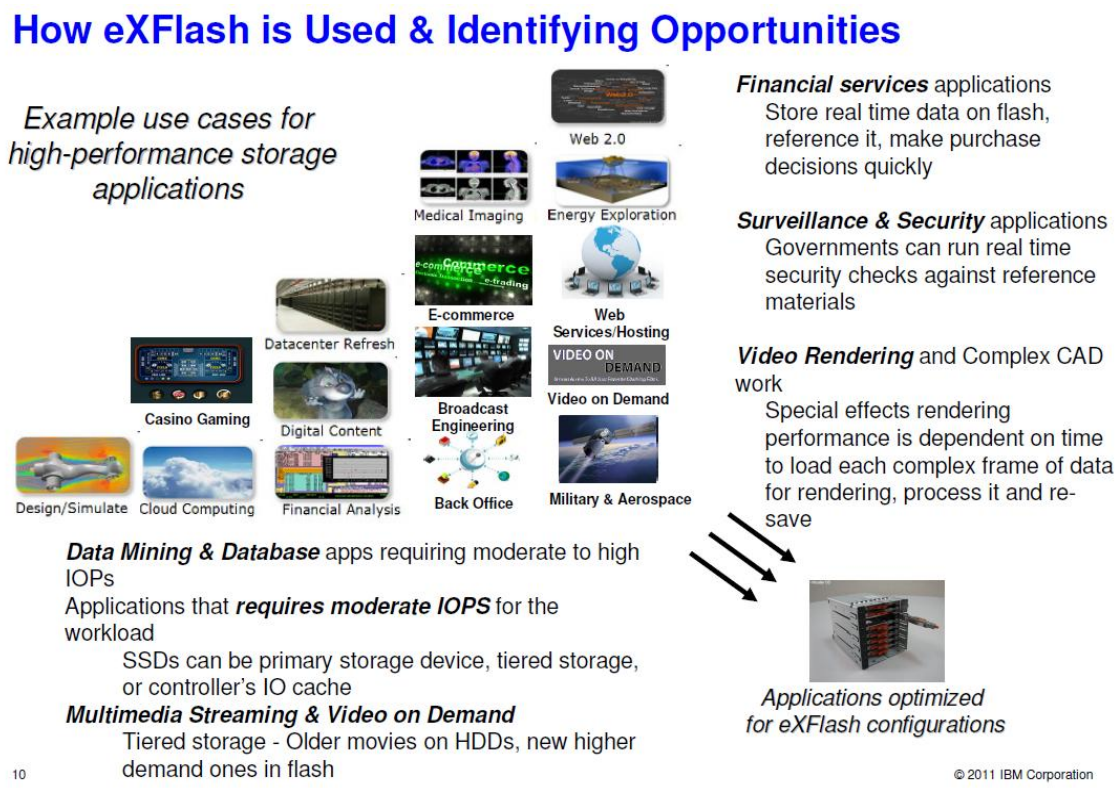
Exploiting Pre-integrated Systems: A Wealth of Applications

The previous sections defined workload optimization, provided system decision making advice based upon workload considerations, and explored four different system designs. We then focused on pre-integrated server configurations, showing how to drive down costs, drive up performance and capacity – while improving reliability.

But great technology does not a solution make. Solutions are delivered at the application and database layers – and IBM's eX5/eXFlash solutions offer a wealth of pre-qualified application/database solutions as illustrated in Figure 3 (next page).

What is important to note in Figure 3 is how applications are able to exploit eXFlash technology to achieve new business results using high performance storage. For instance, data mining and some database applications can exploit the increase number of IOPS that eXFlash offers to achieve faster results. And eXFlash SSDs can be used to accelerate multimedia streaming, or to better service video-on-demand applications (older movies with less demand could be delivered on existing HDDs, while newer, higher demand movies could be delivered to more people more quickly using SDDs). The financial services industry could benefit by storing real time data on flash – and then by referencing that data to make purchase decisions more quickly. And surveillance and security applications can also exploit real time data in new ways to deliver improved security results. Video rendering is also an ideal application for eXFlash – exploiting instantaneous rendering rather than having to constantly read from and write on mechanical disks. And this same concept applies to computer aided design (CAD).

Figure 3 – eXFlash: Achieving New Business Results



High performance, lower cost, pre-integrated storage (along with advanced memory management and fast I/O) can all be combined to deliver exponentially better performance to the business – enabling the business to deliver existing solutions in new ways; enabling the business to make decisions more quickly (and thus service customers more quickly); and enabling the business to establish significant competitive advantage over competitors who stick to the old build-it-yourself, distributed, external HDD approach.

Selecting the Best Type of Server for the Job

When choosing an information system to process your work, your goal should be to pick one that can optimally process your workload. The way we see it, there are four basic systems designs from which to choose:

1. *General purpose* (a design that can run serial, data-intensive and parallel applications all at once);
2. *Optimized* – systems in this category are usually tuned to run a single workload optimally (check out the myriad industry standard benchmarks for examples of this type of system) ;
3. *Pre-integrated* – this type of system is usually hyper-tunes underlying subsystems for ultimate performance (this type of system is the topic of this report); and,
4. *Hybrids* – this is an architecture that unifies different system types using common governance.

How do you sort out which system is best for deployment? The way we see it is:

- In some cases it makes sense to run a mix of serial/data/parallel jobs on a single system design. For instance, if all (or most) of the workloads an enterprise needs to run must be highly reliable/available/secure – and if these workloads are input/output intensive (a communications subsystem requirement), we usually recommend that an enterprise consider using a mainframe computer. But remember, all microprocessors can be deployed as general purpose computers. So if the workload mix is highly parallel, if there are less stringent QoS (quality of service) requirements, and there are a few serial or data intensive tasks that need to be run, we might recommend an x86-based server to handle this workload mix at this QoS level.
- In other cases it makes sense to run a single workload on a server that has been highly tuned and optimized to serve it (we call this an optimized design). For instance, an enterprise may choose to tune and optimize a server to handle one workload extremely well (such as an SAP or Microsoft Exchange workload). Look at computing industry standard benchmarks and you will find dozens upon dozens of examples of general purpose processors that have been optimized to perform specific tasks. In general, this class of server leaves a lot of the tuning up to the buyer (in other words, the buyers, with some guidance from their vendors, take on most of the customization and performance tuning tasks).
- Another category of systems design is pre-integrated servers. *These are servers that offer extremely highly-tuned components and subsystems that deliver exponential performance increases as compared to build-it-yourself designs.* The reason that these performance increases are so impressive is that the vendors who

make these systems tightly integrate their processors, communications subsystems, memory, and storage subsystems – and, in some cases, their operating environments, systems software (infrastructure), and applications/databases deliver extreme performance improvements. Note that with most pre-integrated systems, a vendor has taken on much of the customization task – delivering a pre-integrated solution that may require minor tweaking for specific industry applications in order to delivery extraordinary performance.

Note that the better integrated the CPU is with underlying subsystems – the better the overall system performance will be. This is the hallmark of pre-integrated systems – better integration means potentially exponentially better performance.

- The final category of server designs is called hybrid computing. This type of environment unites various servers with various processors/operating environments onto a single, tightly coupled communications architecture that can be governed in a uniform manner. There are strong performance advantages related to choosing this architecture – and even stronger management advantages (because consistent management across heterogeneous servers can significantly drive management costs related to human labor downward).

The bottom line when making a server selection is this: match your applications to your CPU characteristics; match your QoS requirements to your server design; and closely examine the four classes of server design to determine which meets your performance, customization, and management requirements best.

Summary Observations

At *Clabby Analytics* we don't advocate one system's design over another. Instead, we look at the characteristics of the workload involved – and then at the CPU, QoS characteristics, and system design characteristics (such as the level of integration of subsystems) before determining which server architecture to recommend. In this *Advisory*, we have encouraged you to do the same when examining the potential benefits of SSD technology.

As we look more closely at pre-integrated server solutions, we like what we see. Vendors are building integrated hardware and software solutions that reduce the need for their customers to perform integration work – while at the same time offering exponentially better performance. This can be a big win for customers looking to reduce deployment and integration costs.

To summarize this *Advisory*, it is our belief that applications practically tell you which type of server that they should be run upon. All you need to do is match the application design (it will be either serial, data intensive, or parallel) to the processor best suited to execute that application. Then you need to match the application's QoS requirements to the server design that houses your microprocessor choice (so, if your application requires high reliability, availability, and security (RAS), you need to make sure the server/operating environment that you choose can match your application's service level requirements. Once that you've chosen the best processor to execute your task – and you've examined the RAS characteristics of your workload, your final task is to choose which system design can best execute your workload (general purpose, optimized, pre-integrated, or hybrid).

IBM's System x with eXFlash:
Workload Optimized Pre-Integrated Server/Storage

Based upon our research into pre-integrated servers, we conclude that CxOs, data center executives, virtualization managers and database managers need to become very familiar with SSD technology – and particularly with eXFlash (due to performance and reliability advantages over competitors). The reason you need to become familiar with this technology is that by continuing to buy older HDD technologies and continuing to deploy external storage networks to handle certain workloads, your enterprise stands to lose tens-of-thousands to millions of dollars of IT spend by not embracing new SDD-based technologies such as eXFlash.

Finally, we strongly recommend that IT buyers who want a more technical understanding of eXFlash view IBM's recent "Red Paper" entitled: "Choosing eXFlash Storage on IBM eX5 Servers" available at http://www.redbooks.ibm.com/abstracts/redp4807.html?Open. This document provides a more technical description of eXFlash as well as a deeper look at some of the workloads that are best suited for deployment on IBM eX5/eXFlash servers.

Clabby Analytics
http://www.clabbyanalytics.com
Telephone: 001 (207) 846-6662