Rock Solid Enterprise Solutions with POWER Systems and WebSphere

Andre Metelo
IBM SWG CPO, IBM Power™ Systems
404-935-6804
metelo@us.ibm.com

Last update:01/26/2010

1. Introduction

The world is going through a big change. Individuals are getting more connected, more demanding and more dependent on technology. Users expect to be able to access online services and information around the clock, from a variety of locations, and through a myriad of different devices. To further complicate things, devices are communicating to each other – becoming smarter. These are generating more data that has to be stored and processed by the systems executing the online services. To put more pressure on the IT infrastructure, tools that can mine the stored data are asked to be smarter and capable to find more intelligence out of data then ever before.

This change is forcing businesses to modify both their services and their business models to comply with new services requirements. Businesses depend on their infrastructure to provide 24x7 services. It becomes obvious that highly available servers and applications are most in order to enable data-centers to provide a service level that is compatible the requirement of “always on” operations.

To further complicate this problem, in a modern data-center environment there are no windows that allow a service to be stopped for a planned maintenance event. The days of “Let's stop the server during the night shift” are over. It becomes even worse when taking into consideration the problems of an unplanned server crash that brings down multiple workloads.

In fact, the Meta group has sponsored a study on the costs of server downtime on different industries - IT Performance Engineering & Measurement Strategies: Quantifying Performance Loss, October 2000. The study indicates that the average cost per hour of outage is over one million US Dollars and, in some industries, it can be over two and half million US Dollars. This data alone is a good enough reason to convince corporations to aim at servers and software that provide the features required for high availability solutions, as a server and/or software failure will be very costly.

It pushes enterprises, and their staff, to understand what makes a platform a good choice from the availability point of view. Although there are a very diverse set of features and design choices that creates high availability, one thing is easy to understand – the less downtime a server and its software stack has during its lifetime, the more available it is.

This paper discusses the reasons why the IBM Power Servers, and the upcoming POWER7 based servers, when combined with IBM WebSphere Middleware can easily deploy solutions that provide better availability than solutions offered by Oracle/Sun and Microsoft/Intel. We will discuss the difference in the RAS (Reliability, Availability and Serviceability) philosophies and practices from IBM when compared to the vendors mentioned above, and highlight the key features from POWER6 and POWER7 systems, as well as what WebSphere offers that enables data-centers to experience fewer outages and less downtime.

2. What Is An Availability Strategy?

Understanding the concept of service availability can be very intuitive. One is looking to reduce the downtime of a service to as close to zero as it makes sense financially to do. It does not matter if this downtime is planned – where the system becomes unavailable in a controlled manner – or unplanned – where the system abruptly goes off-line. The objective is to have the service available for its users for as long as possible without interruptions.

However, what is NOT intuitive or trivial for many to understand is exactly how to create and deploy a highly available service within a data-center. Such a plan is called an "availability strategy". An availability strategy is usually the result of a multi-faced approach to the availability components combined with a solid disaster recovery plan – see Figure 2.

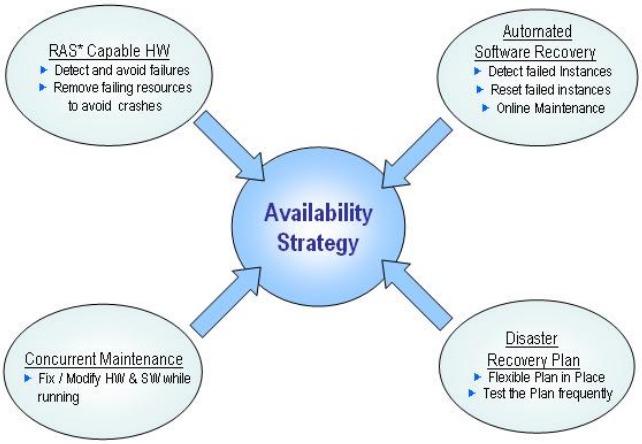

Figure 1 – Key Facets of an Availability Strategy

Of the four components depicted in Figure 1, three are linked to the architectural choices for the service platform – Server Hardware, Operating System and Middleware Stack. In the following sections we will discuss how combining Power Systems and the WebSphere offerings make these choices simpler and easier to deploy, while at the same time making the service more reliable.

However, it is fundamental to understand that the disaster recovery plan is as important as the other availability elements. Unlike the other facets, it is typically a function of proper planning and documenting the process that handles the uncommon situations that typically go far beyond the problems caused by a Single Point of Failure (SPOF) in the service daily operation.

A Disaster Recovery Plan does not have to be complex, but it has to define:

- What is to be protected – data, applications, user space; recovery objectives
- How much time will it take to be back online
- How much data loss is acceptable
- A Scope of the recovery plan – level of disaster that the plan will cover.

Although the recovery plans are very diverse, a pattern can be seen: the more you want to protect and the faster you want to recover, the more resource intensive the plan becomes.

3. Power Systems and Availability Strategy

The ITIC Group recently published the results of a survey taken by over 400 different participants companies spread over 20 countries that shows the measured average downtime of different server offers that exist in the surveyed companies' data-centers. Figure 2 shows the results from the study.[1]

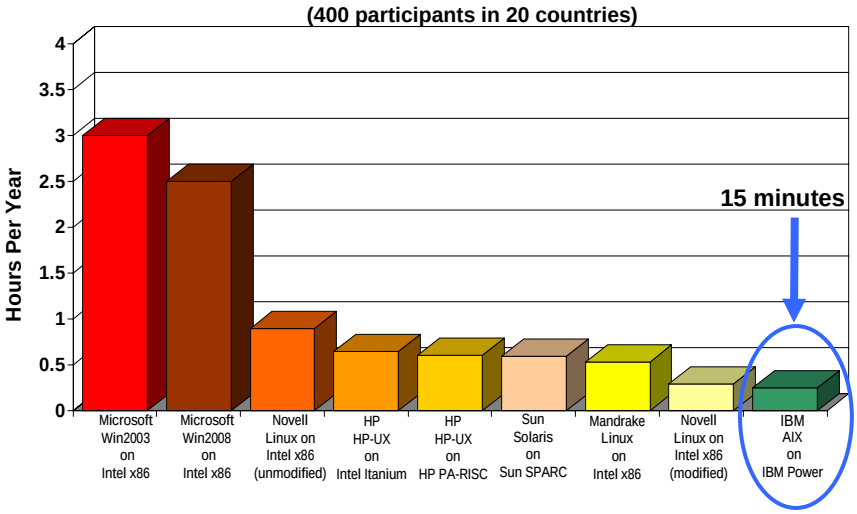

Figure 2 – ITIC that AIX is leader in System availability – ITIC, July 2009

These results are not accidental. One of the key design requirements for the Power Systems product line is industry leading RAS (Reliability, Availability, Serviceability). Because of these requirements, the Power Systems offers include exclusive features that are intending to improve server availability that is built-into them.

[1]Source: ITIC: ITIC 2009 Global Server Hardware & Server OS Reliability Survey; http://itic-corp.com/blog/2009/07/itic-2009-global-server-hardware-server-os-reliability-survey-results ; Results are measured in minutes per year. Not All operating systems included in the survey are included in this chart.

The list of availability features that are available in a POWER6 and POWER7 servers is very extensive: Detect and remove memory chips that are about to fail; Passive System Backplane (High-end servers); Detect, Isolate and Correct PCI errors; Hot Pluggable Disks, Fans, Power, PCI slots; ECC Memory on All Buses; DLPAR Operations and mobility; Online Firmware Updates; Workload isolation – WPAR; Dynamic Kernel Tracing – Probevue; OS level error logging and notification; Online Kernel Patching.

Covering all these features in details is outside the scope of this paper, but several Redbooks and RedPapers are available at the IBM's Redbook website[2] that provide in-depth discussions on all the features mentioned above, and many more.

However, the following three features are unique to the Power Systems and are worth further discussion: *First Failure Capture; CPU recovery and deallocation; Memory Protection Keys*. These features are key elements of the Power System servers design. Such features improve server uptime in the context of an availability strategy as discussed in section 2.

3.1. First Failure Data Capture

First Failure Data Capture (FFDC) is one of the pillars of Power Systems availability capabilities. FFDC enables systems to store information about its operation and system component's condition during run time. This data allows the service processor in a Power System server to isolate and gracefully remove from operation a failing component before a catastrophic error and/or defect happens. And at the same time, the service processor calls back IBM and request specific parts that require replacement. With this feature, there is no need to go through a problem recreation process to identify hardware problems in the Power Servers.

In the end, this capability is able to remove a large amount of unplanned outages from the service, and considerably reduce the amount of time that a planned outage for maintenance takes.

3.2. CPU Recovery and Deallocation

This feature allows the system to detect CPU errors during runtime, and it reverts the processor to the state it was before the error was detected[3]. Then, it can re-execute the operation. If the result is fine processing is continued normally with any effects on the Operating System (OS). However, if the error persists, the system is able to move the state to another processor, and continue execution from there without any effects on the OS.

[2] http://www.redbooks.ibm.com.

[3] The POWER7 design adds new detection mechanisms, thus increasing the number of errors that the CPU Recovery feature can identify and act upon. This will allow POWER7 based system to further improve the availability record of the Power System product family

But, the System goes one step further, and after a predetermined error threshold is reached, it removes the CPU from the pool of available CPUs to avoid any further faults that could cause an unplanned outage, as these errors have a strong correlation with CPUs that are about to permanently fail. And, if there is any spare CPUs present in the system – through the Capacity Upgrade on Demand feature - a replacement is automatically activated, so the environment does not suffer any performance degradation. All this happens without any OS intervention, and does not cause service outages.

3.3. Storage Keys

Storage Keys is a feature made possible by combining the capabilities of the POWER6 and POWER7 CPUs with AIX. The "keys" provide additional protection for code running in the system.

With the Storage Key functionality enabled a memory space allocated to the kernel, device drivers and/or application can not be overwritten by any other process that is not authorized to do write into that memory space. This protects applications and the OS itself from buffer under-run errors and other intermittent errors that can be associated with software coding bugs. Figure 3 shows how storage keys can be used to protect different code space in a Power server running AIX.

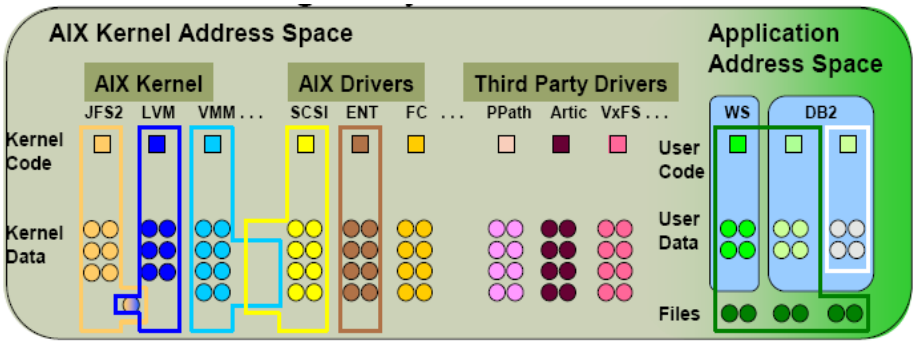

Figure 3 – AIX Storage Protection Keys

4. WebSphere and Availability Strategy

Just like the Hardware and Operating Systems (OS) offered in the Power Systems product line, the WebSphere middleware family is designed with autonomic features intended to simplify management and reduce service down time.

Due to WebSphere design, any service that is based on the WebSphere offers can take advantage of its Active/Active clustering, Online application roll out, Online Cluster performance and health monitoring and management. With such features a WebSphere based solution is able to make sure specific services are always available and complying

with service levels agreement without much manual intervention by the system administrators.

However, it is not only about the availability features above, it is about having the services available around the clock generating reliable and correct results. In a study done by the Software Group Competitive Project Office (CPO) in 2008[4], the results were surprising. In over sixty five simulated failure test cases that involved network disconnects, server power failure, database and application software crashes, only WebSphere was able to process all the transactions without loosing orders, causing data duplication issues and/or security breaches. Competitive offers from Oracle and Microsoft were plagued with one or more types of problems throughout the tests.

With this in mind, let us take a look at a few WebSphere features that are relevant for an availability strategy that focus on making their WebSphere services and applications available around the clock.

4.1. Active/Active Clustering

In WebSphere the application servers themselves can be clustered. This provides both workload management, and high availability. Multiple copies of an application server can be defined in a horizontally cloned environment, while a load balancer cluster handles the traffic dispatching. Web containers and/or EJB containers can be clustered - if an application has both EJB and Web containers, you should strive to create two separate clusters, as this will increase the overall availability of the system, and potentially increase the performance. Figure 4 gives you a typical WebSphere Clustering solution.

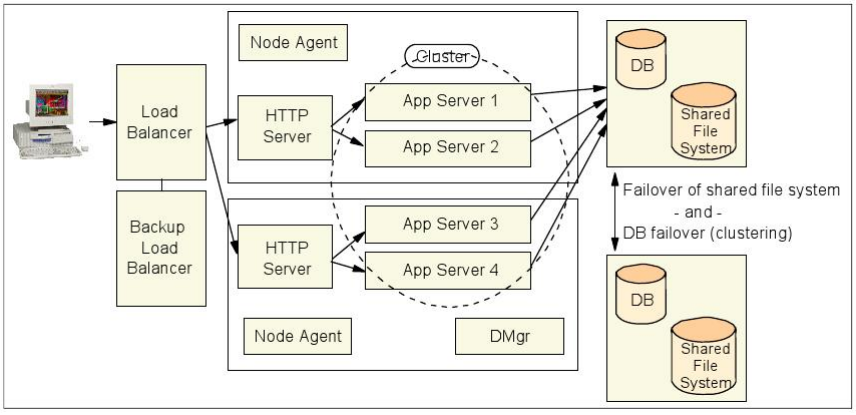

Figure 4 – WebSphere Clustering Diagram

The main reason to use WebSphere's clustering solution over the usual Hardware clustering options is the fact that the plug-in module knows the full topology of the cluster. Therefore, in the case of a failure, the plug-in marks the cluster member as

[4] IBM WAS Delivers Better Transaction Integrity than Oracle; Kovan, Gerry, Jul 2008; https://w3-03.ibm.com/sales/competition/compdlib.nsf/SearchView/7AD79133E376D1500025743B00782220?Opendocument

unavailable and stops the routing of requests to that member for a fixed interval. The plug-in periodically retries the path to that cluster member and marks the member available again, after the path has been asserted. This approach reduces the number of failed connections, and makes cluster reintegration considerably faster than the classical clustering solution offered by the Server OS itself.

4.2. Non-Stop Application Upgrade

Another benefit of using the WebSphere clustering solution is the ability to deploy upgrades to the applications on real time without any outage in the cluster.

Because the plug-in module knows the whole cluster topology, the same plug-in module that controls the cluster is able to migrate the sessions data from one node to another, then stop all traffic to the node, install the new version of the application. Once the node is upgraded it can rejoin the cluster and it starts serving the new version of the application while the other nodes go through the same process.

This whole process is done through the WebSphere console and it can be scaled to make sure that the application being update does not miss it service level agreements (SLAs).

4.3. WebSphere Virtual Enterprise (VE) and Application Health and Performance Monitoring

Once more, WebSphere provides unique capabilities[5] that take advantage of the fact that the clustering plug-in is aware of the cluster topology. WebSphere VE adds performance and health monitoring capabilities to the plug-in – search for memory leaks; page faults; application response time. With these additional capabilities added to the plug-in, WebSphere VE is able to modify the amount of resources allocated to the applications it manages in order to make sure that SLAs are achieved throughout the cluster. Further more, if the servers in the clusters have resources available, WebSphere VE is able to create new instances of the applications that are getting too close to miss an SLA by using the extra resources that are unused in different physical servers. This is done automatically, without the need of manual intervention by a system administrator.

Moreover, WebSphere VE integrates seamlessly with PowerVM in order to execute DLPAR (Dynamic Logical Partitions) operations so it can acquire resources from Power Servers that are available to handle peak periods, deploy new instances of applications experiencing high workloads. Once the peak period is over, WebSphere VE can terminate the extra instances and return them back to the original users of the resources or back to the Capacity On Demand resource pool.

Meanwhile, applications managed by the WebSphere cluster(s) are always available for customer to use them and able to comply to predetermined SLAs.

[5]Oracle and Microsoft do not have these capabilities integrated with their offers. However, part of the functionalities can be achieved by Oracle and Microsoft through the acquisition of 3rd party software to monitor and shape traffic and SLAs.

5. Conclusions

Because IBM is able to create a fully integrated stack composed by the server hardware, the operating system and the middleware the solution elements are able to better explore the capabilities provided by the whole stack. Although this paper covered a brief list of examples of such integration between Power Systems and WebSphere solutions, the same can be said for the other Middleware stacks. Just look at the recently announced DB2 pureScale clustering solution. DB2 pureScale provides unparalleled scalability and availability in the open systems world through a solution that takes advantages of the best features in the server, the storage elements, operating system, and middleware capabilities.

With such level of integration and a design process that focus on solution availability and leading performance, enterprises can combine the IBM Power Servers offers and the IBM Middleware offers to create an unbeatable availability strategy for their services to thrive in a 24X7 world. Moreover, no other vendor has the integration capabilities that IBM has to provide top notch performance combined with world class availability for solutions, and not only components of the solution as some vendors tend to offer.