keynote

Industry System Design:

Influencing Factors for consideration of Future Generation System Design

Benjamin DeLuca
International Business Machines
Austin, Tx

(1999 IRW keynote summary)

Design of successful future systems requires a good understanding of customers wants and needs. Many factors influence the customer expectations and requirements. We need to understand the current business environment and the system characteristics that are required to support this environment. System characteristics like availability and reliability continue to be very desirable for most customers. As we design future systems we need to carefully consider the system characteristics desired at design concept time in order to drive an implementation that will have the most success in the marketplace.

Currently one of the key influences for business is the growth of the digital economy. You can not pick up a newspaper or a magazine today without seeing examples of how the internet is changing the way we do business. One example of a customer that is exploiting electronic commerce is Charles Schwab on-line brokers. Currently they handle 30% of the worlds on-line trading. They have needed to add significant capacity to their systems over the past year to keep up with customer demand. The customer needed a system that provided flexible growth and had solid availability characteristics. An example of growth worldwide is that on-line trading increased from 253,000 trades per day to 340,000 trades per day from 3Q98 to 4Q98, up by over a third in only one quarter. Another example is consumer on-line holiday shopping. The estimates from 1997 through 1999 show greater than a 200% annual growth rate. The number of internet users will conservatively grow from about 142 million in 1998 to more than 500 million in 2003, with 200 million people on-line by year end 1999. As we look at the trends in electronic commerce Forrester predicts an increase in this area from $36 Billion in 1998 to over $1.3 Trillion in 2003. One additional fact here is that this is just the beginning, only 5% to 10% of the business transactions are on-line. The business to business transactions are 20 times more of the electronic commerce than the business to consumer transactions. Currently only 200,000 businesses out of the 40,000,000 businesses worldwide are on-line with electronic commerce. We are also seeing more and more devices, like cell phones, personal digital assistants connected to the internet. This trend has already begun to change how people search for information . More and more people are using the internet for information on news, weather, and travel information. As we look at internet transactions it is important to consider the series of background activities that a single purchase will drive. The anatomy of a transaction may consist of authentication, authorization, order processing, inventory checking, build instructions, credit checking, payment and shipping. This discussion provides an introduction to the environment that is driving system characteristics for computer systems.

Transactions are complex, the volumes are unpredictable and peak loads can strain business systems. Customers need to have systems that provide flexible growth. They need systems that are easy to use and have good system management characteristics. One of the most important factors is availability, because if your system is not available you can not execute business transactions. To calibrate, look at the difference a nine can make. If your system is available 99% of the time that translates to 3.6 days of outages per year, 99.9% translates to 8.8 hours, 99.99% it's 58 minutes and at 99.999% it is unavailable only 5 minutes a year. Downtime can be caused by many reasons and it can also be very expensive. An example of a complex system will help illustrate the level of robustness that may be required to provide high levels of availability. The Nagono Olympics provides a good case study to help illustrate what demands a computer system might need to support. These games occurred over 16 days with 154 major

[Visuals of the Keynote Address can be found at irps.org/irw/99 and on the CDROM in the HANDOUT.]

events, during this period 2 blizzards, 2 thunderstorms and one 5.0 earthquake occurred. Users of systems included 3,000 athletes, 36,000 volunteers, 70,000 participants, 1.5 million spectators, and 2 billion viewers. The computer systems need to be designed with redundancy, and configured for high availability characteristics to insure solid support of these games. This design included monitoring devices to indicate when preventative maintenance needed to take place. The results were solid, 200,000 hits per minute peak, 650 million hits over the 16 days, 321,000 e-mails. These results were achieved due to robust design of the base systems and by establishing the proper configurations which could tolerate individual elements failing without taking down the complex. Let's reference a 1998 Standish Group Research Note to help categorize the causes of unavailability. The largest cause is system hardware at 21%, next planned maintenance at 19%, application bugs at 17%, operator error at 15%, system software at 13%, environmental conditions at 11% and other at 4%. It is generally not a surprise that system hardware is the largest cause, however, it is very important to recognize that is only 21% of the pie, such that we must address the other wedges if we need to get higher levels of availability. The same Research Note provides estimates of various applications unit cost/ minute for unavailability, which range from $27,000 to $1,000 and average $10,000 per minute. These are only examples, the important point being each customer has their unit rate for their application and will likely vary, but for many customers it is clear that unavailability can be very costly. Considering various industry platforms Gartner Group in 1998 reported on levels of unavailability. The highest rank was the S/390 Parallel Sysplex at 99.998%. A few other examples include S/300 non-Sysplex at 99.9, UNIX systems at 99.73% and NT systems at 97.44%. As a customer is making a buying decision it is often important to illustrate that they need to consider not only direct purchase costs but also the "costs" of lost business due to differences in availability in the various platforms. Using the average cost application data one finds a Parallel Sysplex unavailability cost annually would be $100,000 while a UNIX server would have an unavailability cost of over $14 million. This cost can easily shift the balance significantly over the differences in the initial purchase cost.

Let me try to illustrate some of the system design considerations required to develop a robust enterprise server. There are a significant amount of reliability, availability and servicability (RAS) features that must be designed into the system to provide the desired system characteristics. In general we have approached this effort in a multi-tiered fashion. Our RAS initiatives include unexpected outage reduction, system availability extensions, system problem determination enhancements, system complexity reduction and ease of use enhancements, and lastly installation, upgrade and maintenance. To improve problem determination effectiveness our diagnostic strategy has moved from failure re-create to first failure data capture. This technique is more effective because many complex problems are very difficult to duplicate. If you are unable to duplicate a problem your customer will likely experience additional outages prior to a solution being found. In order to capture first failure information unique sense information is added to the system. This sense instrumentation is included in the design of the basic silicon circuit design where fault isolation registers are imbedded to capture machine state of critical internal busses. Each system has an added service processor function which among other things enables the customer engineer to interrogate the isolation registers to diagnose the cause of failure. There are also several user functions that the customer can set as desired. A few examples are security passwords, reboot policies, dial out capability for remote support in addition to viewing the error log. In addition to the service processor which is largely inward looking we have a Service Director application which provides some level of predictive problem analysis and notifies the customer when preventative maintenance is suggested. This application monitors the information provided by the service processor, analyzes severity, determines disposition and notifies the appropriate site personnel. We also use a layered approach to provide enhanced system management. Beginning with the system error log, this information is then provided to the next layer of utilities that monitor health and history of the machine or system. The next layer provides access to information for additional diagnostics or assessment. The top layer can be designed as a graphical user interface (GUI) across clusters of systems that can be administered by the customer and also routed to the supplier for ongoing real time assessment.

Physical packaging layout of an enterprise server can be designed to support enhanced RAS. Design should accommodate N+1 blowers, and N+1 power and also support hot plug of these elements to provide uninterrupted service. In high end systems "book" packaging can be used to provide physical protection, stabilize electronics, minimize pin damage on insertion, and provide secure retention. Another approach to minimize customer outages besides providing redundant power and cooling to provide a redundant system that is managed through High Availability Cluster Multi-Processing (HACMP) . This technique is used to minimize single points of failure for example having two paths to the communication network. If one path has a failure the system can still function properly by using the alternate path. A broader step extends HACMP across a large geographic area through High Availability Geographic Cluster (HAGEO) which can be used to protect systems by mirroring transactions across a large geographic area with the view that if an earthquake were to destroy a data center in California the mirrored geographic cluster in Colorado would be able to carry on business transactions uninterrupted.

There are two aspects of availability impact that need to be addressed, first, the impact of failures. The goal is to actively reduce the failures the cause the highest impact to customers and drive failures as much as possible to be transparent or recoverable. The second aspect of availability that needs to be addressed is the customer impact of repairs. The goal here is also to drive impacts away from unscheduled interrupt repair actions (UIRA) to more concurrent repair or deferred repair. Our target for UIRA reduction is to reduce them from 80% of repair actions (RA) in 1997 to 5% of RA's in 2001, at the same time reducing RA's 15-20% year to year. To accomplish these reductions we have planned several phases of RAS actions for the next few years. In 1997 and 1998 we introduced first fail data capture (FFDC) and broad usage of ECC in CPU caches and array. The second phase in 1999 and 2000 extends FFDC to I/O devices and allows us to support concurrent PCI hot swap. The third phase for 2001 an 2002 addresses PCI bus error recovery, logical partitioning and dynamic reconfiguration. Current measurements have shown positive results in both RA and UIRA reductions.

Systems generally consist of a number of discrete boards which are interconnected to form a system. The boards are placed in a chassis with power and cooling and surrounded by a set of covers to complete the package. The goal of our designs is to keep cables to a minimum to eliminated that source of failures. The example here is a system structure which consists of a discrete processor/cache board , a discrete mother board for memory that has connectors for a maximum of 16 memory DIMMS, a single board for I/O with 9 slots for PCI devices, and a single system backplane which contains memory controllers and 2 connectors for memory boards and 2 connectors for the CPU boards. As the failures are categorized the largest percent contributors are I/O adapters for 41% and disks for 17% of UIRA's. Approximately 24% of the failures are attributed to the cards and boards with the remaining percentages split between power, cooling and software. The learning point here is that our strategy to address PCI bus error recovery will have a significant impact on reducing UIRA's. To address disk failures most customers use mirrored disks or redundant arrays of disks. The electromechanical nature of disks makes them more susceptible to failure versus solid state devices. In the first design release of the processor board we projected the highest failure contributor to be a voltage regulator. As we completed the second design release we added N+1 regulators which eliminated this element as a contributor to failures. The I/O backplane contains the most number of discrete logic modules and, as such, we find that approximately 60% of the failures are attributed to logic modules. To reduce these failures we attempt to find modules that consolidate more function to reduce the logic module count on the board. The system backplane contains a large amount of connectors with a large pin count. This drives connectors to be the most significant contributor at %54 of the failure rate. The memory mother board also has a large amount of connectors for DIMMS, and we find that 84% of the failures here are due to connectors. As you add memory modules to the mother board the failure rate contribution of the memory is about 34% and the connector rate reduces some to 55%. Basically the approach here is to continue to drive suppliers for higher reliability connectors and memory. In the example of discrete I/O drawers, we have added hot plug PCI, N+1 fans, N+1 power and hot plug disks to improve the RAS characteristics of this element. The performance projections for this subsystem show the highest percentage contributor to be PCI Adapters at 71%. This is another example where PCI bus error recovery will result in a significant reduction to UIRA's.

A few other factors for consideration are that since the computer system business is under continuous price margin pressure it is vital that we find cost effective solutions which improve RAS characteristics. In this regard we also need to keep component power levels down as much as possible, since more power drives additional expense for cooling and power supplies. In summary, the growth of internet use is a key driver for customer requirements. Availability is the key system characteristic to support business success. As future systems continue to be improved, key elements that must be integrated into our designs are robust hardware, software and system management.