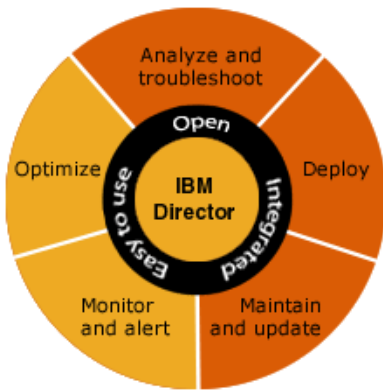




The following information is to support the requested information to answer a question why IBM System X verses white box when cheap isn't the only criterion.

The blending of Director and the hardware improves system availability beyond the sum of the parts (1+1=3).

This section combines the abilities of hardware and management to blend in such away to provide high availability and minimize the concern on geographic location of components.



The key to success is knowledge and to have that knowledge at your finger tip changes your model from reactive to proactive. The following is a very brief view of Director's ability to change Centralized/Decentralized to just managed infrastructure. This is accomplished by using standards. So using a Common Information Model (CIM), this standard was adopted and evolved by the Distributed Management Task Force (DMTF). DMTF is a published systems management standard and was developed in open forum by

DMTF member companies. Defined and promoted as an industry standard for managing systems, CIM was designed to be used for describing management information between differing management applications, running in many different operating environments, including Microsoft Windows and Linux.

CIM provides a common definition of management information for systems, networks, applications and services, and allows for vendor extensions. CIM's common definitions enable vendors (IBM) to exchange rich management information between systems throughout the network. CIM is composed of a specification and a schema. The schema provides the actual model descriptions, while the specification defines the details for integration with other management models.

CIM is used extensively throughout IBM Director. In fact, the capability in IBM Director to manage Level-1 systems is based on CIM instrumentation and providers

IBM enables the Intelligent Platform Management Interface (IPMI) a standardized, abstracted, message-based interface developed by Intel that defines records for describing platform management devices and their characteristics. This interface allows for standard communication between systems management software such as IBM Director.



To instrument this IBM is blending its hardware, Director and IPMI-compliant system management to its hardware IBM Baseboard Management Controllers (BMCs) and Remote Supervisor Adapter (RSA).

There are an abundance of technologies and products available at the operating system level to help network managers maximize uptime of their servers, but these typically come with a high purchase and/or operational cost. In response, IBM and standards organizations have been working hard to develop common management standards that can help. The Intelligent Platform Management Interface (IPMI) is one key open standard that is most likely included with your server today. It runs on a dedicated chip/controller known as a BMC (Baseboard Management Controller).

When considering server software management, typical solutions have tended to focus on loading agents on the Operating System of the server. This is sometimes referred to as 'Agent-Based'. A complementary and additional approach is to consider exploiting agentless management. One standard that is used at hardware level is IPMI. IPMI, utilizing a BMC, defines how administrators monitor system hardware and sensors, control system components and retrieve logs of important system events to conduct remote management and recovery. IPMI monitors hardware health conditions like temperature, fan, voltage, hardware errors (memory, network, etc.) and chassis intrusion.

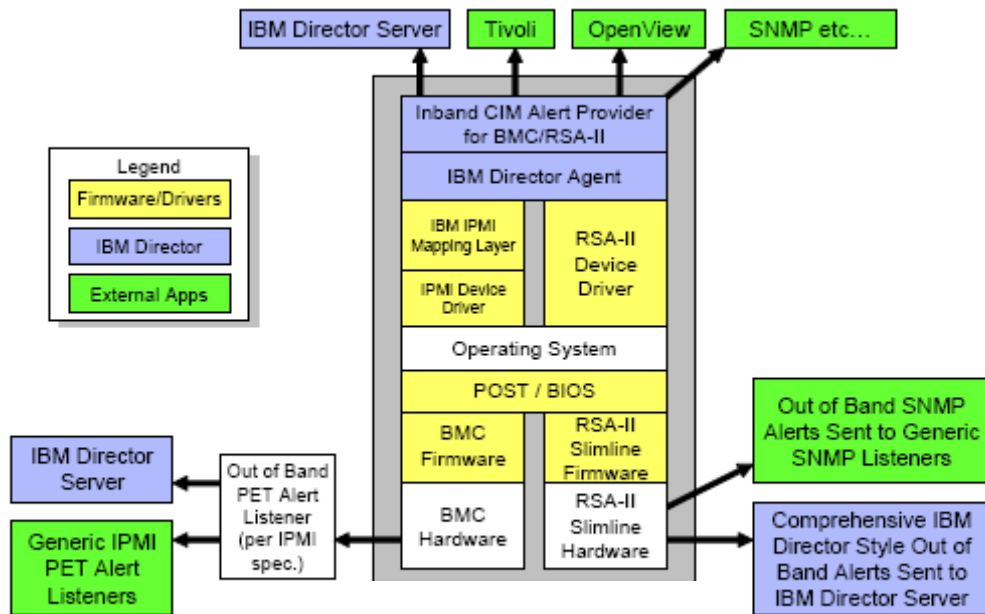
Introduced in 1998, IPMI was created by the IPMI forum - an industry-wide initiative that today has over 170 vendors – including Avocent. They work together to continually update and implement this hardware management specification for servers and other systems such as storage devices, network and telecommunications equipment. In it's third major release, IPMI 2.0, includes enhancements to, among others, Security, VLAN and Blade support.

Because IPMI operates independent of the operating system, when sending commands to the BMC over IP, it provides administrators with the ability to monitor, manage, diagnose and recover systems, even if the operating system has hung or the server is powered down. IPMI also includes alert notification and recovery capabilities that enables an administrator to monitor and react before hardware problems occur. IPMI's hardware monitoring features also provides additional levels of security. Chassis intrusions can be detected by configuring IPMI to detect if the server has been opened. And, the use of multi-layer privileges and passwords together with authentication and on-the-wire encryption lets IT managers allow or deny access to specific IPMI features securely.

Agentless technologies and standards like IPMI are easy to exploit as they generally come pre-integrated within the server or device. And best of all, they are free. They also fill the gap left by Agent-based management by being available even if the OS has hung. In this way, they are very complementary to



your existing management toolkit. For effective server management you need both agent-based on agentless approaches.



The capabilities of IBM Director can be divided into the following categories:

Inventory

A critical first step in any systems management strategy is to understand exactly what hardware exists in the environment and how it is configured. IBM Director performs a thorough inventory scan of each managed system it discovers. Hundreds of hardware and software data points are collected and stored in the IBM Director database. Inventory collection can be repeated both manually and through multiple automated processes.

Hardware status

The moment you install IBM Director, it starts working to let you know about hardware problems that occur on managed IBM System x servers. If there is a problem with a power supply, fan, voltage regulator module (VRM), network interface card (NIC), or other hardware, IBM Director will let you know what the problem is and which system is affected.

Event management

At the heart of any systems management solution is the ability to alert IT staff in the event of a system problem. IBM Director provides a unique and very powerful



method of alerting called Event Action Plans, which enables you to define event triggers independently from actions that might be taken. Then, simply combine these two types of items into customized plans for action and assign them to individual or groups of managed systems.

Process management

Using the Process Management task in IBM Director, you can keep track of all important processes running in your environment. IBM Director can alert you if any monitored process starts, stops, or fails to start.

Resource management

Resource management is an important aspect of keeping an IT environment running at peak efficiency. It is important to know whether any given system is overloaded and not able to keep up with the workload demand. IBM Director provides the ability to monitor hundreds of system resources, set individual or group thresholds for these resources, and alert you in the event that a resource threshold has been exceeded.

Remote management

IBM Director was built to perform remote management. Any management task that can be performed on a local system can also be performed on a system thousands of miles away, provided network connectivity is available. In addition, the Remote Control task in IBM Director allows you to take control of any managed system in your managed environment.

Update management

The Update Manager, in IBM Director, provides update management through a native IBM Director task. Update functions include tasks for creating profiles, downloading updates, comparing updates defined in profiles against systems, and generating reports.

Mass configuration

One of the advantages of managing systems using IBM Director is in its ability to make certain configuration changes on multiple managed systems at once. Even in a dynamic host control protocol (DHCP)-enabled environment, many critical servers tend to use static addresses. Using Mass configuration profiles, you can, for example, change the IP address these managed systems use to locate their primary DNS server, all without having to physically visit each system.

SNMP management



In addition to the sophisticated management capabilities IBM Director enables for systems running the IBM Director Agent, any SNMP device can be discovered and managed as well. IBM Director can send and receive SNMP traps and convert these traps into native IBM Director alerts, delivering more helpful information than a raw SNMP trap normally can provide.

Electronic Service Agent

- Electronic Service Agent (ESA) monitors hardware events and transmits system inventory to IBM. Electronic Service Agent has two key functions:
- Automatic hardware problem reporting - Hardware errors that meet certain criteria for criticality are automatically reported to IBM and a service request is generated.
- Inventory collection - Performs hardware and software inventory collections, and reports inventory changes to IBM. All information sent to IBM is stored in a secure IBM database and used for improved problem determination such as down-level firmware or software drivers.

PowerExecutive

PowerExecutive enables you to manage power and thermal needs of IBM enhanced hardware through IBM Director and PowerExecutive.

IBM PowerExecutive enables you to monitor *actual* power draw and thermal loading information which can help with:

- More efficient planning of server location such as datacenter, CO or closet construction and or modification
- Proper power input sizing based on physical systems
- Capping of power when resource demands are exceeded or in the event of a power crisis
- Justification of incremental hardware purchases based on available input power capacity
- Better utilization of existing resources

PowerExecutive works at the Physical Platform Management Object (PPMO) level and queries power and thermal information through the following interfaces:

- Baseboard Management Controller
- Remote Supervisor Adapter II
- BladeCenter Management Module



Remote Deployment Manager

Remote Deployment Manager (RDM) enables you to remotely configure, deploy and retire systems, including updating system firmware, changing configuration settings, installing operating systems and applications, backing up and restoring partitions and securely erasing data from disks

RDM utilizes the Preboot Execution Environment (PXE) and therefore requires Dynamic Host Configuration Protocol (DHCP) to be enabled for Remote Deployment Manager target systems.

Predictive Failure Analysis

Predictive Failure Analysis® (PFA) gives key components in IBM System x servers the ability to monitor their own health and generate an alert up to 48 hours before failure occurs. This allows the system administrator to either hot-swap the component (if applicable) or schedule downtime at low-impact times for the component to be changed or refreshed.

PFA code monitors certain subsystems within the component and if tolerances exceed a pre-determined range an alert is automatically generated.

For example, in hard disks, PFA code monitors by Director includes:

- Read/write errors
- Fly height changes (The height of the disk head above the platter)
- Torque amplification control (The amount of power used to keep the drive spinning at a constant speed)
- IBM implements PFA on more server components than any other vendor.
- System x components currently protected by PFA are:
 - Hard disk drives
 - Fans
 - Power supply units
 - Memory
 - CPUs
 - Voltage regulator modules



Mean Time Between Failure (MTBF) and Mean Time Between System Failure (MTBSF)

Where is MTBF important for the machine system designers needing to evaluate which components will best fit a specific machine design or application, it's standard practice to compare each unit's mean time between failures (MTBF) rating. Most believe that MTBF is the number of operating hours that will elapse before a unit fails. In reality, MTBF is the total functional life of a system component divided by the number of failures a measurement of reliability.

So, I will focus on key design points that will enhance the availability and reliability of the system (MTBSF). The key components that are most likely to fail are power supplies, disk drives, fans and memory. So what was done at the hardware layer to minimize these failures?

Fans are placed in the system and are redundant if one fails the remaining have the ability to keep the box from over heating. When second power supply is installed additional fans are available to remove heat added by the power supply converting electrical state.

Disk Drives have MTBFs listed publicly at 1 to 1.5 million hours. These are large numbers and could be misleading. To reduce the potential of outage due to numerous factors can be reduced even more by mirroring or depending on spindle count using a RAID level 5 or 6 to minimize exposure to spindle failure.

Power Supplies have public numbers of 100,000 hours (MTBF); These are large numbers and could be misleading. The power supply can be impacted by power quality coming in, environmental including heat. To help minimize the impact of an outage add a redundant power supply. This allows system availability even if a supply fails.

Memory has been, is and will be in its current form a major component of system failure. What are some of these failures and why. As memory density goes up, so does the impact of alpha particles, failed spots in memory structure, and soft miss reads or writes to good address memory. IBM supports chipkill memory, memory scrubbing and SEC ECC.

Example of this design improvement model

We will compare a 32MB parity only memory, 1GB SEC ECC and 1GB chipkill.

- 32MB system will experience 700 failures per 10K systems
- 1GB SEC ECC will experience 900 failures per 10K systems
- 1GB Chipkill will experience 20 failures per 10K systems

See appendix A for more detail on Active memory.



These memory features are collectively known as *Active Memory*:

- Memory ProteXion

Memory ProteXion, also known as “redundant bit steering”, is the technology behind using redundant bits in a data packet to provide backup in the event of a DIMM failure.

Currently, other industry-standard servers use 8 bits of the 72-bit data packets for ECC functions and the remaining 64 bits for data. However, the server needs only 6 bits to perform the same ECC functions, thus leaving 2 bits free. In the event that a chip failure on the DIMM is detected by memory scrubbing, the memory controller can re-route data around that failed chip through the spare bits (similar to the hot-spare drive of a RAID array). It can do this automatically without issuing a Predictive Failure Analysis (PFA) or light path diagnostics alert to the administrator. After the second DIMM failure, PFA and light path diagnostics alerts would occur on that DIMM as normal.

- Memory scrubbing

Memory scrubbing is an automatic daily test of all the system memory that detects and reports memory errors that might be developing before they cause a server outage.

Memory scrubbing and Memory ProteXion work in conjunction with each other and do not require memory mirroring to be enabled to work properly. When a bit error is detected, memory scrubbing determines if the error is recoverable or not. If it is recoverable, Memory ProteXion is enabled and the data that was stored in the damaged locations is rewritten to a new location. The error is then reported so that preventative maintenance can be performed. As long as there are enough good locations to allow the proper operation of the server, no further action is taken other than recording the error in the error logs.

If the error is not recoverable, then memory scrubbing sends an error message to the light path diagnostics, which then turns on the proper lights and LEDs to guide you to the damaged DIMM. If memory mirroring is enabled, then the mirrored copy of the data in the damaged DIMM is used until the system is powered down and the DIMM replaced. If hot-add is enabled in the BIOS, then no rebooting would be required and the new DIMM would be enabled immediately.

- Memory mirroring



Memory mirroring is roughly equivalent to RAID-1 in disk arrays, in that memory is divided in two ports and one port is mirrored to the other half. If 8 GB is installed, then the operating system sees 4 GB once memory mirroring is enabled (it is disabled in the BIOS by default). Since all mirroring activities are handled by the hardware, memory mirroring is operating system independent. Certain restrictions exist with respect to placement and size of memory DIMMs when memory mirroring is enabled, and these are system dependant.

- Hot-swap and hot-add memory

Currently, only the x445 supports hot-swap and hot-add memory. There are two configurations where you can add or replace memory while the server is still running:

- Hot-swap, where you can replace failed DIMMs of the same type, size, and clock speed without turning off the server. Hot-swap memory is operating-system independent. Memory mirroring must be enabled to use hot-swap.
 - Hot-add, where you can add new DIMMs without turning off the server, thereby increasing the amount of RAM available to the operating system. This feature is currently only supported by Windows Server 2003, Enterprise Edition and Datacenter Edition. Memory mirroring must be disabled when using hot-add and due to the way memory is implemented in the x445, the port you are adding memory to must be empty before you add memory, and DIMMs must be added in multiples of two.
- Chipkill memory

Chipkill is integrated into the XA-32 second-generation chipset and does not require special Chipkill DIMMs. Chipkill corrects multiple single-bit errors to keep a DIMM from failing. When combining Chipkill with Memory ProteXion and Active Memory, the server provides very high reliability in the memory subsystem. Chipkill memory is approximately 100 times more effective than ECC technology, providing correction for up to four bits per DIMM (eight bits per memory controller), whether on a single chip or multiple chips.

If a memory chip error does occur, Chipkill is designed to automatically take the inoperative memory chip offline while the server keeps running. The memory controller provides memory protection similar in concept to disk array striping with parity, writing the memory bits across multiple memory chips on the DIMM. The controller is able to reconstruct the “missing” bit from the failed chip and continue working as usual. Chipkill support is provided in the memory controller and implemented using standard ECC DIMMs, so it is transparent to the OS.



In addition, to maintain the highest levels of system availability, if a memory error is detected during POST or memory configuration, the server can automatically disable the failing memory bank and continue operating with reduced memory capacity. You can manually re-enable the memory bank after the problem is corrected via the Setup menu in the BIOS.



© IBM Corporation 2006
IBM Systems and Technology Group
Produced in the USA.
05-06

All rights reserved.

IBM, the IBM logo, System x, eServer and xSeries are trademarks or registered trademarks of IBM Corporation in the United States and/or other countries.

AMD and Opteron are trademarks or registered trademarks of Advanced Micro Devices, Inc.

Intel and Xeon are trademarks or registered trademarks of Intel Corporation

Other company, product, and service names may be trademarks or service marks of others.

IBM reserves the right to change specifications or other product information without notice. References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates. IBM PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions; therefore, this statement may not apply to you.