

IBM P 系列和 X 系列服务器解决方案简述

1. IBM P 系列服务器虚拟化应用解决方案

1.1 服务器虚拟化解决方案简述

在数据中心的转型过程中，客户需要增加创新投入，快速地部署新的应用来适应市场需求，但是往往受到现有的 IT 基础设施的束缚。面对不断增加的服务器数量和日益低下的设备利用率，管理维护开销的增长速度远远超过系统采购成本。源自数十年来主机虚拟化技术的积累，如今 IBM 也可以为开放平台的用户提供丰富的服务器虚拟化解决方案，从而降低管理成本，提高数据中心的灵活性、响应速度和可用性。用户可以将不同的服务器形成统一的资源池，每台服务器支持多个操作系统映像，从而实现快速部署，并支持丰富的自动化功能。

在 UNIX 平台上，用户可以选择 PowerVM 实现服务器虚拟化，并通过 IBM 提供的虚拟化服务器管理软件实现统一管理，能够集成服务器，操作系统，应用平台，最大化资源的利用率，实现资源动态调整、虚拟高可用性、虚拟机动态部署迁移、虚拟备份，全方位虚拟化及集成。从而简化 IT 基础设施，降低成本。

1.2 PowerVM 技术介绍

PowerVM 是在基于 IBM Power 处理器的硬件平台上提供的具有行业领先水平的虚拟化技术家族。它是 IBM Power Systems 虚拟化技术全新和统一的品牌（逻辑分区, 微分区, Hypervisor, 虚拟 I/O 服务器, APV, PowerVM Lx86, Live Partition Mobility)。

PowerVM 提供的虚拟化解决方案可运行在业界最为广泛的操作系统上, 包括 AIX、Linux 和 System i 客户使用的 i5/OS 系统。

1.2.1 PowerVM 主要功能介绍

1) Micro-partition(微分区): 芯片级虚拟化技术, 实现 1/10 个
cpu/>cpu/>cpu/>cpu/>cpu/>cpu/>cpu/>cpu/>cpu/>cpu/>cpu/>cpu/>
cpu/>cpu/>cpu/>cpu 为单位划分 CPU 资源, 以 1/100 单位调整资源。

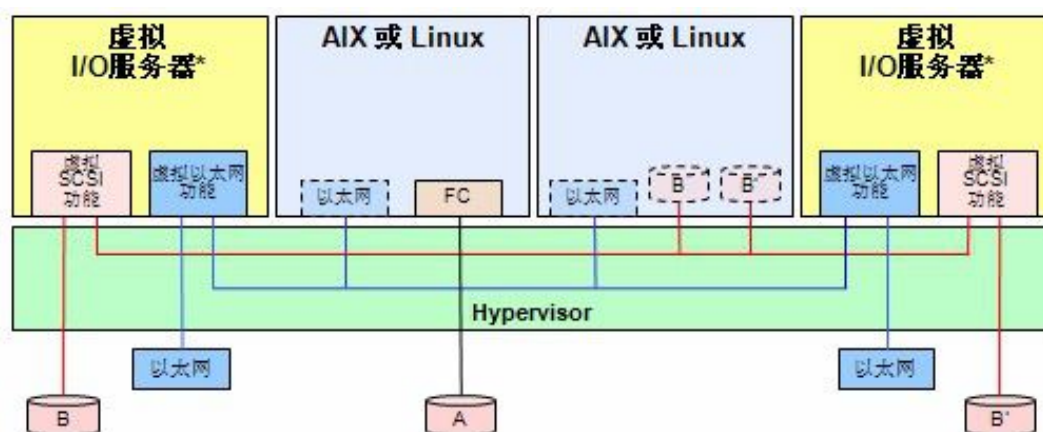
IBM POWER6、POWER7+支持完全意义上的虚拟化技术，并引入了微分区技术。由于动态逻辑分区的资源调整功能让系统管理员可以自由添加、删除或在

分区之间移动系统资源，例如 CPU、内存、I/O 适配器的分配，而不需要像原来修改之后重新启动分区。这样，微分区技术的引入，更使得动态逻辑分区的资源调整功能不但可以移动物理资源，还可移动、增减虚拟资源，具有广阔的应用场景。这样系统管理员就可以根据分区系统负荷和分区业务运行特点，随时将资源动态分配到需要的地方，从而大大提供资源的利用效率和灵活性。

每个 POWER5/5+/6 的 CPU 可以划分为 10 个微分区（Micro- Partition），并且按照 1% 的颗粒度调整 CPU 资源，使管理员可以将资源分派给处理器的一小部分而非整个处理器，大大提高了系统的使用率，并且资源分配的更改对于用户是透明的。IBM 高级虚拟化技术支持虚拟以太网卡和虚拟 SCSI 卡，在微分区中，不需要配置物理网卡和 SCSI 卡就可以支持网络连接和存储系统。

2) Virtual I/O Server（虚拟 I/O 服务器）：VIOS 实现以太网、SCSI 和光纤通道磁盘的共享。

虚拟 I/O 服务器用于为客户机分区提供虚拟 I/O 资源。分配给一个分区的物理适配器可以由一个或多个其他分区共享，使管理员可将用于各个客户机的物理适配器的数量降到最低。因此，虚拟 I/O 服务器通过消除对专用网络适配器、磁盘适配器和磁盘驱动器的需求，以帮助客户降低成本。跟其他虚拟化技术不同，高级 Power 虚拟化不需要虚拟化所有设备。这些设备可以是实现最高性能而分配给分区的专用设备组合；也可以是用于虚拟 I/O 主管分区的设备，它们由多个分区共享以提高资源和适配器的使用效率。



3) Integrated Virtualization Manager (IVM，集成虚拟化管理器)：IVM 是中低端 RISC 服务器虚拟化的利器。

IVM 是 IBM 为了将虚拟技术带给更多中低端 UNIX 服务器用户,帮助他们充分利用虚拟技术简化系统管理和提高系统利用效率,而推出的一款经济易用的集成虚拟化管理器。

IVM 的管理功能是 HMC 功能的一个子集。如果服务器的物理资源有限,又没有 HMC,但用户仍然希望将服务器划分成多个分区、安装不同操作系统以供多个应用使用时,就可以采用 IVM。

IVM 进行虚拟化设置时,首先需要安装一个虚拟 I/O 服务器分区 (VIOS),这个 VIOS 拥有整个服务器的物理 I/O 资源,并占用很小一部分处理器和内存资源进行自身的运行管理。在 VIOS 上有一个虚拟管理通道 (Virtual Management Channel) 的设备,它可以在 IVM 和 Power 固件层 (Hypervisor) 之间进行通讯,这样 IVM 就可以管理整个服务器的物理资源。用户在 VIOS 上通过 IVM 界面进行系统资源的划分、用户分区的创建和管理。处理器和内存通过固件层直接分配给用户分区,而物理磁盘和网络设备则在 VIOS 上经过虚拟化之后,以虚拟设备的方式供各个用户分区使用。

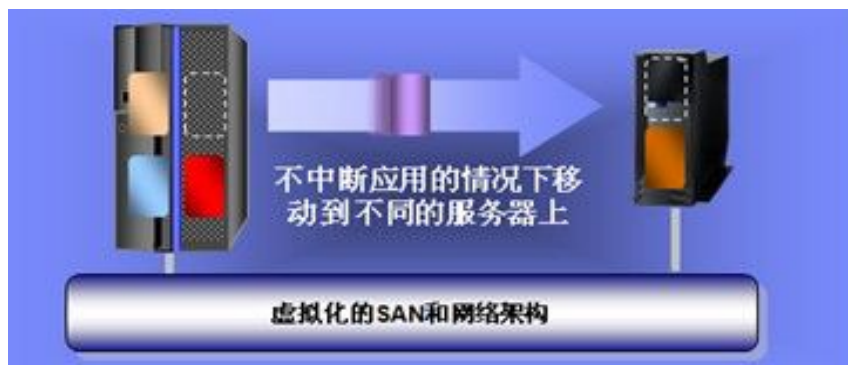
IVM 可以支持的型号包括, p5/5+ 560Q 以下的小型机, 以及 JS21/JS12/22, p505, p510, p520, p550。

注意: IVM 仅支持一个 VIOS。

4) Lx86: Lx86 使 x86 Linux 应用运行在 Unix 服务器上。

PowerVM Lx86 是在基于 IBM Power 处理器的服务器上产生一个虚拟的 x86 Linux 应用运行环境。大部分的 32 位 x86 Linux 上的应用可以不重新编译就可以运行在 IBM Power Systems 服务器上。

5) POWER6 动态分区迁移 (Live Partition Mobility): LPM 实现把一个正在运行的逻辑分区实时迁移到另一台物理服务器上。



动态分区迁移 (Live Partition Mobility) 允许您将正在运行的 AIX 和 Linux 分区, 以及它们所承载的应用程序从一台物理服务器迁移到另一台物理服务器, 而不会对基础服务产生任何影响。这项迁移操作将传输整个系统环境, 包括处理器的状态、内存、附加的虚拟设备, 以及连接的用户。该操作只需要花费几秒钟的时间即可完成, 可以维护整个系统事务的完整性。POWER6 动态分区迁移功能有助于减少计划内宕机影响; 在不影响应用的情况下, 动态负载均衡, 合理利用资源; 在不中断服务的情况下, 采用新技术; 移除低利用率服务器上的负载, 并关闭空闲服务器, 节省能源。

6) Workload Partition (工作负载分区)

AIX6 提供的工作负载分区 (WPAR) 技术是一种全新的软件虚拟技术, 用于隔离用户和应用程序。它不依赖于硬件特性, 是多种 AIX 核心技术的综合体现。

WPAR 是由软件创建的、AIX 6 映像中的虚拟化的操作系统环境。对于所承载的应用程序来说, 每个工作负载分区都是一个安全的、隔离的环境。WPAR 中的应用程序认为, 它正执行于自己的、专门的 AIX 实例中。对于大多数应用程序来说, WPAR 就好像是一个 AIX 的启动实例。通常, 在 WPAR 中, 应用程序无需修改就可以运行。

WPAR 与 LPAR 的区别

	WPAR	LPAR
操作系统版本	所有 WPAR 都运行在同一版本的 AIX 中。	每个 LPAR 都有自己的 AIX 版本。
可用性	所有的 WPAR 都依赖于 AIX 的单一实例。	每个 LPAR 都有自己的操作系统。
隔离	所有 WPAR 都共享同一内核。	每个 LPAR 都有自己的内核。
部署速度	WPAR 的部署速度非常快。	LPAR 需要更多的资源调配步骤, 这是因为需要对整个操作系统和所有相关的资源进行资源调配。

	WPAR	LPAR
资源使用	WPAR 使用的资源较少，如内存和网络。	LPAR 使用的资源较多，这是因为每个 LPAR 都有一个隔离的操作系统。
维护	对于共享同一 AIX 版本的所有 WPAR 来说，只需要维护一个 AIX 实例。	对于每个 LPAR，则需要维护一个操作系统实例。

7) Live Application Mobility (动态应用程序迁移)



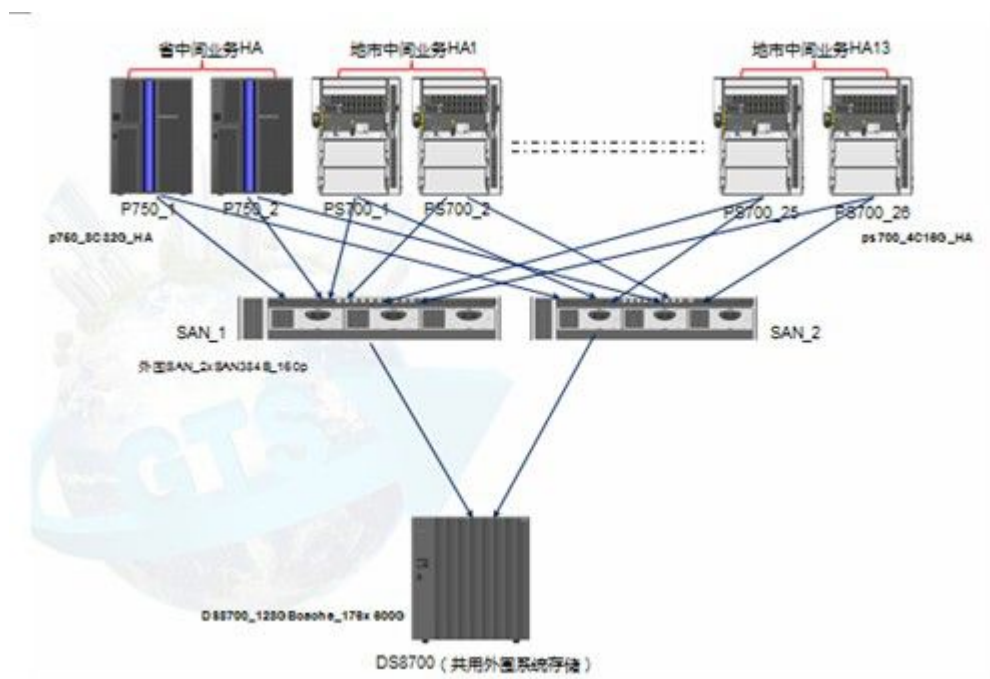
动态应用程序迁移是 AIX 6.1 的一项独特创新，是其 WPAR 策略的组成部分，它允许您移动工作负载，而不是整个分区。实际的目标 WPAR 可以是不同的服务器，但并不一定必须是不同的服务器。它比较灵活，因为您可以在一些具备混合物理体系结构（Power5 和 Power6）的环境中使用它。它可以将应用程序从需要计划停机以便进行维护的系统中移动到别处。通过将工作负载从不堪重负的服务服务器移动到负载较轻的服务器，它还可以用于提高性能。而且，通过采用这种方式移动工作负载以允许物理服务器在非高峰期间进入睡眠状态，它还可以帮助节约电能。

下面的表格介绍了动态应用程序迁移和分区迁移之间的一些基本区别。

动态应用程序迁移和分区迁移之间的区别		
类型	动态应用程序迁移	动态分区迁移
OS	AIX 6.1	Linux、AIX 5.3、AIX 6.1
硬件	PowerPC® 970、POWER4™、5™、6™	Power6

粒度	WPAR	LPAR
----	------	------

1.3 某农信中间业务测试环境 PowerVM 解决方案



解决方案描述：

- 1.采用 PowerVM 虚拟化技术，利用有限的硬件资源（两台 P750)创建了 15 个分区满足测试需求；
- 2.采用 PowerVM 微分区技术，使用较少的 CPU 满足大量测试分区需求；
- 3.采用 PowerVM 高可用技术，保证整个系统无单点故障。

解决方案优势：

- 1.采用 PowerVM 虚拟化技术，在节约成本的同时能够快速、轻松地满足后期测试需求；
- 2.采用 PowerVM 微分区技术，能够提高系统整机利用率；
- 3.整合服务器和虚拟化资源，减少硬件、软件、管理和能源成本。

同时，P750 内置硬盘将被 VIO 服务器管理后，划分个每个 LPAR，作为 AIX 的 ROOTVG,外置 DS8700 被 VIO 服务器管理后,划分给每个 LPAR 作为数据盘。

2. IBM X 系列服务器应用解决方案

2.1 IBM X 系列服务器的领先技术

IBM 企业 X 型架构展示了一个巧妙构思的演进方法如何创造创新功能。企业 X 型架构采用三种行业标准服务器技术组件—处理器、内存和 I/O—并使用先进的特性来进一步增强，旨在把标准系统提高到一个新的水平。

企业 X 型架构为业界标准服务器带来了以前仅向大型机和其它高端系统用户提供的功能。这些新功能，结合现有的 X 型架构技术创造了革命性可扩展性经济、无与伦比的灵活性和新级别的可用性和性能。通过简化管理、降低成本和提高可用性，使客户感到满意的关键特性包括：

XpandOnDemand 可扩展性、系统分段、PCI - X I/O 子系统、Active PCI - X、I/O、Memory ProteXion、Chipkill 内存、内存镜像、热添加/热插拔内存 (即将推出)、XceL4 服务器加速器缓存

在下面的内容中，会按照服务器的扩展性、四级缓存、内存技术、I/O 这四个方面的内容做详细的介绍。

企业 X 型架构: XpandOnDemand

供助于其灵活的模块化设计，企业 X 型架构为服务器创造了一个革命性新经济模式：客户不再需要预先购买他们可以支付的尽量多的服务器，以确保未来容量增长。您可以边增长边付费。我们称这为创新式 XpandOnDemand 可扩展性。

企业 X 型架构技术使用称为 SMP 扩展模块的增强型高性能 4 - 路 SMP 标准构件。使用这些 4 - 路模块作为可扩展的企业节点，IBM SMP 扩展模块允许从 4 - 路有效扩展到 8 - 路再到 12 - 路—甚至 32 - 路系统，通过一个高速 SMP 扩展端口将它们连接在一起。因此，如果客户最终需要更多的处理功能，可以增加一个备用 4 - 路模块，结合简单布线创建一个 8 - 路服务器。如果这一 8 - 路服务器不能提供足够的插槽和托架，它们可以通过插件式外置远程 I/O 扩展单元（稍后说明）和远程存储单元，如 IBM EXP500，以进一步增加 I/O 插槽容量。

企业 X 型架构 SMP 扩展模块包括处理器、内存、I/O 支持、缓存、存储和其它设备，可以像其它服务器一样单独运行。每个模块可以运行一个不同于其它模块的操作系统，或者需要时通过系统分段将多个模块分配给一个操作系统版本。使用系统分段，一个系统可以配置作为一个共享 16 - 处理器的内存系统，

或分成多个分段。最终，当支持所有 EXA 功能时，一个分段就像一个处理器一样小。

模块通过称为 SMP 扩展端口的专用高速互连设备相互连接，共享资源以实现近乎线性的可扩展性，从而用户能够适应用于运行多个节点作为一个大型耦合单元，或作为两个或多个小型单元—甚至根据需要稍后重新安排配置。

EXA 技术还提供所有处理器和所有内存之间的接入，与它们各自的节点无关，从而减少连接。借助于每个增加的节点，您还可以增加芯片集、前端总线、PCI 总线和其它资源，以共同分担数据流量。更多的节点等于更多的系统带宽。想像一下您在传统的 16 - 或 32 - 路 SMP 系统中遇到的冲突和资源问题。

同样，通过故障切换支持连接服务器的一个群集就像连接两个、三个或四个 4 - 路节点一样简单。您可以在节点之间使用相同的系统扩展端口布线以进行群集互连。对于可扩展的群集来说，无需复杂的以太网设置就可创建一个高速互连，因为通过 SMP 扩展端口它已经存在。此外，以太网 PCI - X 插槽对其它 I/O 开放。

SMP 扩展模块技术：XcelL4 服务器加速器缓存

企业 X 型架构 (EXA) 支持的一种先进特性是一种巨大的第 4 级(XcelL4 服务器加速器缓存) 系统缓存，这种特性确保 SMP 扩展模块内存性能技术的正常运行，在基于 Itanium 的服务器中每 SMP 扩展模块为 64 MB 400 MHz DDR (双倍数据传输速率)高速 ECC 内存，而在 Xeon 系统中为 32 MB。

通过在处理器和主内存之间使用高速 DDR 内存，XcelL4 缓存可以大大提升处理器和 I/O 设备的性能。性能提升多少呢？在厂商自夸比竞争对手超过 2% 的性能优势的行业中，XcelL4 缓存可以提高所有服务器吞吐量高达 15% 到 20%。

Intel 32 - 位和 64 - 位处理器包含相对小规模(128 K 到 4 MB, 取决于处理器)的第 1 级、第 2 级和 (使用 Itanium) 第 3 级内置缓存内存。内置缓存的数量受到处理器模块内部可用空间的限制。高速缓存内存越大，处理器更频繁地查找它需要的数据，同时更少地必需接入速度较慢的主内存。(处理器速度正在以远快于主内存速度的速率增加；每年缓存遗漏时必需接入主内存的次数增加。) 大内存容量

Active Memory 是企业 X 型架构中大量内存技术突破，设计用于增加容量、提升性能和可靠性。其中一项技术是支持大内存容量的能力。

虽然某些服务器仍旧受到它们可以安装的内存插槽数量的限制，其它服务器受到服务器正在使用的芯片集可以支持的最大内存的限制。由于这些原因，大多数服务器的内存限制为 16 GB RAM 或更少。企业 X 型架构打破了这一障碍，允许在一个基于 64 - 位 Itanium 的服务器中使用多达 256 GB 的 RAM (在基于 32 - 位 Intel Xeon MP 处理器的服务器中为 64 GB)。

Memory ProteXion

Memory ProteXion 帮助保护不出现由于硬内存错误导致的突发故障。它的工作方式与 Windows NTFS 文件系统中热备用磁盘扇区有些类似，如果操作系统检测到磁盘上有坏扇区，它将把数据写到备用扇区以实现这一目的。

Memory ProteXion (在其它系统上也称为冗余位调整)最初开发用于 IBM 大型机并且多年来用于 zSeries 和 iSeries 服务器。

经验显示 **Memory ProteXion** 保护的服务器比一台使用标准 ECC 内存的服务器出现故障的几率要少近 200 倍。ECC (错误检测和纠正) DIMM 包含 144 位，但只有 140 位用于数据，剩余的四位未使用。**Memory ProteXion** 仅是将数据重新写到其中一些备用位，而不是迅速禁用 DIMM。这种方法允许 **Memory ProteXion** 每 DIMM 纠正四个连续的位错误—每内存控制器八个连续的位错误 (一台服务器可能有多个控制器)。这种先进的技术可以帮助减少服务器宕机，从而形成一个更强大的客户机—服务器计算平台。这在大型数据库环境中尤其重要，在崩溃的数据库恢复正常运行之前，服务器之间的事务处理/回滚、重新索引和数据同步会导致数小时的丢失。如果一个内存控制器在备用位外运行，它继续作为第二条防线 **Chipkill memory**。

仅在一台服务器在短时间内遭遇如此多的错误且 **Memory ProteXion** 不能解决的情况下，**Chipkill ECC** 内存 (现在业界标准计算机中为第三代) 才发挥作用。

内存镜像

防止由于内存故障导致的服务器宕机的第三条防线是内存镜像。在该项技术中，内存的管理方法与 RAID 配置中磁盘镜像非常类似。在这种情况下，主内存条上数据的精确映射被镜像到备用或热备份内存条。结果是，如果一根内存条出现

故障，镜像的内存条将变为主内存条。在更换了故障内存条之后，主内存条内存中的数据将镜像复制到新内存条。

PCI - X I/O 系统和 Active PCI - X

当前最新的 PC I/O 总线允许多个 64 - 位 66 MHz PCI 总线分段，支持每个分段提供 400 到 500 MBps。这种带宽不够支持新兴的 10 Gbps (GB 每秒) 或更高 I/O 环境。

如果没有其它的性能提升，PCI 将很快成为阻碍这些高速网络以网络最大速度连接服务器的瓶颈。I/O 瓶颈已经阻碍了行业标准服务器成为均衡的系统架构，这是基于高速 Intel 的服务器和大型机系统的一种特性。因此，为了解决这些性能问题，行业开发了一种称为 PCI - X 的增强型总线，旨在延长 PCI 的使用寿命，直到如 InfiniBand 等下一代串行 I/O 架构准备就绪为止。

PCI - X 允许所有目前的 32 - 位和 64 - 位 66 MHz PCI 适配器在 PCI - X 总线中正常运行。PCI - X 适配器充分利用新 100 MHz 和 133 MHz 总线速率，这允许单个 64 - 位适配器每秒传送多达 1 GB 的数据。此外，PCI - X 在一条总线中支持两倍于 PCI 的 66 MHz 64 - 位适配器。

Active PCI - X 允许您无需关闭服务器就可添加或更换 Active PCI 和 Active PCI - X 支持的卡。旨在提升服务器整体可用性的 Active PCI - X 特性分类如下：

热插拔 允许您更换故障或即将发生故障的适配器，无需重启 热添加 提供简便升级，允许您在服务器运行时添加新适配器(IBM 是业界第一家提供这一特性的厂商) 故障切换 在主适配器故障时允许备份适配器负责运行所有正在处理的业务。

2.2 IBM X 系列服务器解决方案

2.2.1 IBM X 系列服务器应用

在网络的关键应用（EPR、SAP、财务系统、CRM 等）上，利用 IBM X 系列服务器建立了高性能的双机热备份解决方案。

方案 a:



主要产品：System X + Windows server 2008 Enterprise + DS3500

方案说明：为客户提供高可用性保护平台，最重要应用系统提供实时在线保护，确保系统停机时间最少

方案描述：

以 X 系列机架(X3850X5) 服务器作为主要服务器，两台服务器作为双机高可用性群集，外挂 DS3500 作为在线备份存储。

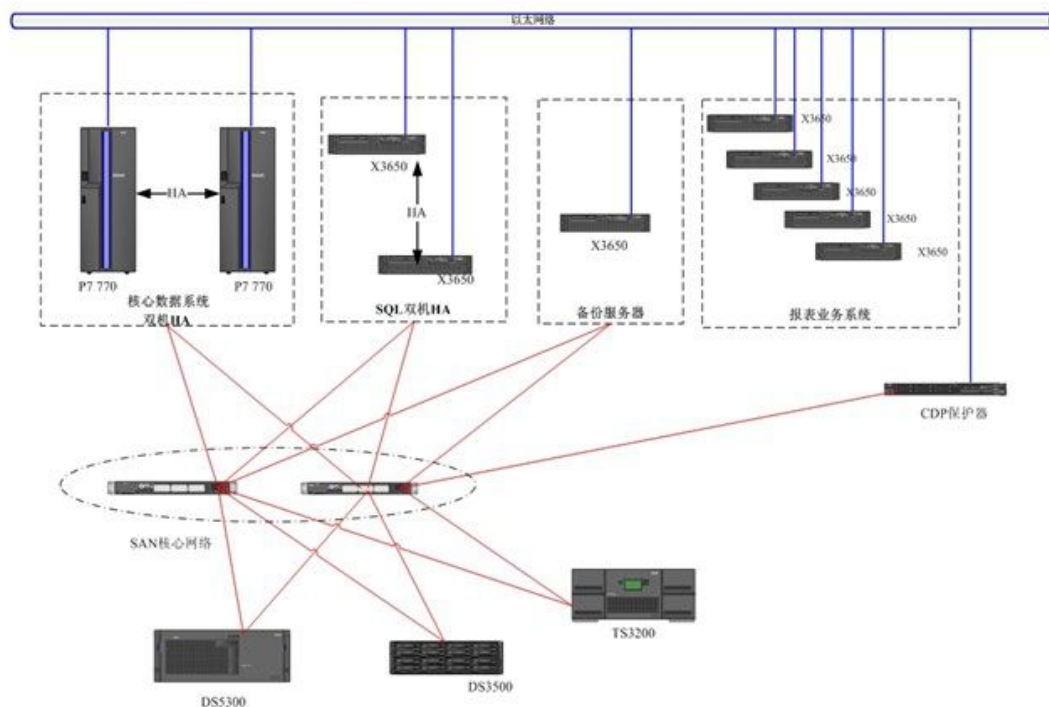
方案 b:



主要配置：2 台 IBM X3850 X5 服务器；2 台 IBM SAS B24-4 光纤通道交换机；1 台 DS3500 双控制器存储（FC 接口,最低配置 4 块硬盘）及相关配件。

方案说明: 两台高性能服务器希望连接具有 RAID 特性的光纤盘柜实现双机热备的数据库解决方案（需要建立自己的 SAN 存储网络）。

2.2.2 某保险公司 IBM 服务器应用



方案说明：

1. 两台 Power7 770 服务器作为核心数据库系统，为保证业务连续性，两台机器之间做双机，接入 SAN 网络。
2. 两台 x3650 作为 SQL Sever 的双机数据库系统，两台之间 HA 保证业务高可用，接入 SAN 网络。
3. 一台 x3650 作为备份服务器。
4. 报表业务系统采用 5 台 x3650。
5. 存储采用 DS5300 作为核心存储，本地采用 CDP 保护器，将数据连续保护到 DS3500 中；同时构建以磁带库 TS3200 为备份介质的备份系统，利用 Lanfree 的备份技术，将生产数据备份到磁带库中。
6. 影像平台数据放在存储 DS3500 上，共用这个连续数据保护存储，既满足了要求，又节省了 IT 的投入。
7. 核心 SAN 网络采用两台 B24 构建。

2.3 IBM X 服务器选型规则

服务器是系统中至关重要的核心设备，其作用是为各类应用提供硬件运行平台。对服务器的选择不仅仅是满足当前已开展的各项业务需要，更要着眼于未来。

而在企业的发展历程中，其不可避免地要经历创业期、发展期与成功期三个阶段。相应地，在市场上，我们可以根据规模将企业划分为三个类别，分别是中小型企业、大中型企业和大型企业。不同类别的企业在进行信息化建设时所采取的策略也不尽相同，只有根据自身的特色和规模，本着“需求至上”的原则，使平台的搭建运转与企业发展水平相贴合，才能使其为企业的成长发挥最大的作用。

主要考虑到因素有：

- 1、服务器的使用环境（物理环境：确定选择塔式或者是机架式）；
- 2、服务器的应用环境（确定选择服务器的档次，主要为服务器的功能）
- 3、服务器所要承受的并发连接数量：

$X \leq 100$ 一路服务器或者具有扩展到两路的低主频 2 路服务器

$100 \leq X \leq 300$ 两路服务器，根据情况选择合适的主频及内存配置等

$300 \leq X \leq 500$ 高主频的两路服务器，或者低配置的 4 路服务器

$500 \leq X \leq 1000$ 四路服务器（6 核和 8 核可选）

$X \geq 1000$ 选择据有扩展性的四路以上的服务器