



# 富士通基于Itanium 2的新型高端PRIMEQUEST服务器为Linux和Windows应用提供关键业务托管

2006年7月

## 为Fujitsu编制

### 目录

摘要	1
引言	2
适用于哪些操作系统?	3
更多的操作系统性能	4
富士通与开源组织	
在Linux方面的合作	4
Red Hat的diskdump及其他	4
Red Hat和Novell SUSE 的增强功能	5
IDEAS概要	7
尾注	8

### 深度研究

本白皮书由Ideas国际 (IDEAS)编制,详细介绍了有关富士通新近推出的基于双核Intel Itanium 2 PRIMEQUEST的服务器的操作系统性能。三本姐妹篇白皮书对以下方面提供了类似的详细说明和深度研究:高可用性、系统架构、以及系统管理。一本概览白皮书提供了有关PRIMEQUEST服务器的功能概要。了解更多详情,请参见<http://www.fujitsu.com/global/>。

### 摘要

富士通于2006年7月发布了3款基于双核Intel Itanium 2处理器 (Montecito) 的新型关键业务企业服务器。PRIMEQUEST 520 (8个CPU, 16个内核)、540 (16个CPU, 32个内核)和580 (32个CPU, 64个内核)型号都运行Linux和Windows。富士通将其描述为基于Intel的关键业务服务器产品线,这种高端平台能满足用户的以下需求:能够提供高可靠性企业级服务器,能够从业界标准软硬件实现经济收益的系统。520、540和580型号补充并升级了富士通的早期PRIMEQUEST产品——基于Itanium Madison的420、440和480 (都设计成单核CPU,分别具有8、16和32个微处理器)型号。

富士通的服务器与其他基于IPF的平台之间的区别是,特别注重关键业务设计。富士通利用其成熟的大型机和PRIMEPOWER设计经验 (采用Solaris操作系统和SPARC微处理器),开发出了基于Itanium的系列产品,满足了客户在业界标准服务器上运行关键业务负荷的需求。

2006年9月底正式版发布之时,PRIMEQUEST 500系列服务器 (520、540和580型)将支持:

- » Red Hat Enterprise Linux Version 4 (RHEL4)。不久就会支持Updates 4和5以及RHEL5 (Red Hat发布之后)。
- » SUSE Linux Enterprise Server Version 9 (SLES9)和SLES10 (当SLES10发布之后,这取决于Novell)。
- » Itanium专用的Windows Server 2003, Enterprise Edition (最多8个CPU/16个内核,最大内存为256MB),当SP2于2007年上半年发布之后将会整合SP2 (这取决于Microsoft)。
- » Itanium专用的Windows Server 2003, Datacenter Edition (最多32个CPU/64个内核,最大内存为2TB),当SP2于2007年上半年发布之后将会整合SP2 (这取决于Microsoft)。

将来的操作系统升级和服务包版本可能包含特定的增强功能,可以优化PRIMEQUEST服务器运转,但它们最后会成为通用的Windows Server分发版本中的一部分。一般而言,PRIMEQUEST 500系列一开始会采用各个分发版本的企业版,随后再添加特定的扩展功能 (例如富士通添加的功能)。

本档由Ideas International (IDEAS) 版权所有,受美国和国际版权法和相关协定保护。未经IDEAS书面同意,不得拷贝、复制、在检索系统中存储、以任何形式传输、张贴于公共或私有的网站或公告板、或授予第三方许可。不得从本白皮书上隐匿或删除版权信息。本白皮书中所指的所有公司和产品的商标和注册商标均受保护。

本档基于认为可靠的信息和来源。本档应“保持原样、不予改变”进行使用。IDEAS对于内容的数据、主题、质量和时效性不作任何保证和陈述,并且不承担任何责任。本档中的数据可能有变动。IDEAS不负责向读者通知数据变更。另外,IDEAS对于本档所描述的产品、服务和公司的观点可能有所改变。

IDEAS对于基于本档信息以及读者试图复制性能效果和其他结果所作出的决策不承担责任。本档也不用于预测未来的价格或性能等级。本档不构成本白皮书中所讨论的产品和服务和所讨论的供货商提供的产品和服务的担保。

## 纵向领域与 ISV

PRIMEQUEST 500系列服务器将在全球销售。这种广泛分布的一种结果是，需要确定将要根据地区来对待的纵向应用领域。但一般而言，金融、电信、制造业和政府是重要的目标纵向领域。

富士通正在与这些领域的应用子集中的重要ISV协商。对于那些提供ERM、CRM和SCM（如MySAP）中的广泛的横向业务（horizontal business）应用性能的ISV，富士通也在与他们合作。此外，富士通正在致力于提供适用于PRIMEQUEST服务器环境的业界标准的数据库、中间件和工具。

一些重要的公司正在讨论、致力于或者实现了将他们的产品向PRIMEQUEST平台移植实现，这些公司包括BEA、BMC、CA、EMC、HP、IBM、SAP、SAS、VERITAS、VMware以及Oracle。

富士通也在将它的许多受亚太地区欢迎的产品移植到在亚太地区销售的PRIMEQUEST服务器中。这些产品包括Symfoware数据库、用于业务整合的Interstage中间件以及Systemwalke系统管理软件。

PRIMEQUEST服务器将可以使用多种关于磁盘、磁带、网络等的硬件组件，注意到这一点也很重要。

此外，富士通和Itanium领域的其它重要厂商都是Itanium解决方案联盟（Itanium Solutions Alliance, ISA）的成员，而且表现很活跃。ISA致力于Itanium生态系统的开发。

PRIMEQUEST服务器将会给Windows和Linux操作系统提供富士通独创的增强功能。若可能，普通用户将可以使用这些扩展功能。例如，当Microsoft公司的Vista（以前称为Longhorn）发布后，PRIMEQUEST服务器将支持这种操作系统。Vista可能包含支持PRIMEQUEST的特定的增强功能，但它们将是Windows Server的普通特征集合中的一部分。类似地，富士通也在与开放源码开发实验室（OSDL）合作，直接或通过Red Hat和Novell给Linux添加增强功能，改善所有Linux分发版本的可扩展性、RAS（可靠性、可用性和适用性）、动态重配置，虚拟化（virtualization），分区（XPAR）等等。

本白皮书更详细地描述了PRIMEQUEST服务器上可用的操作系统以及提出的改进措施。其他姐妹篇白皮书主要讨论其他的PRIMEQUEST属性，如系统架构设计、高可用性特征和系统管理性能。白皮书“PRIMEQUEST 概览”也已经发布。

## 引言

通过与Intel、Microsoft、Red Hat和Novell公司的紧密合作，富士通开发了基于双核Intel Itanium 2处理器（Montecito）的PRIMEQUEST 520（8个CPU，16个内核）、540（16个CPU，32个内核）、580（32个CPU，64个内核）型服务器，现在这些服务器都推向市场了。具有16和32路Itanium微处理器配置的高端PRIMEQUEST服务器。这些服务器不仅运行最新的Itanium 2版本，而且还将运行Itanium Montvale微处理器（2007年某个时候发布这种处理器之时）。此外，通过多个分区和域，它们还可以具有SMP或向外扩展（scale-out）性能。

PRIMEQUEST服务器目前采用RHEL4（2006年下半年或2007年出现Updates 4或5时，也能支持它们），2007年出现RHEL5时，也将支持RHEL5。它们还支持SLES9，当SLES10于2006年下半年出现时，也可以运行SLES10。新型PRIMEQUEST服务器运行Windows Enterprise Edition 2003和Windows 2003 Datacenter Edition。2007年Vista（以前称为Longhorn）发布之后，这些服务器将会采用Vista。所有的Linux和Windows系统的升级包和补丁包一发布之后就会被整合进去。

PRIMEQUEST服务器与PRIMEPOWER和PRIMERGY产品一同完善了富士通的服务器产品线。PRIMEPOWER适合UNIX（基于SPARC微处理器的Solaris），而PRIMERGY适合业界标准32位/64位（EM64T）Intel处理器和Windows/Linux。

PRIMEQUEST服务器是富士通利用长期的大型机、向量处理器和UNIX经验开发出来的，将在全球市场与HP、SGI、IBM、NEC和Unisys的机器进行直接竞争。它针对的应用包括OLTP、决策支持/数据仓库、高性能计算、转移旧平台（legacy migration），以及服务器/数据库/应用整合。在适当时可以使用SMP和集群配置。

## 适用于哪些操作系统？

2006年9月底按计划推出之时，PRIMEQUEST 500系列服务器（520、540和580型号）将支持：

- » Red Hat Enterprise Linux Version 4 (RHEL4). 不久将支持Updates 4 和5以及 RHEL5（在Red Hat发布之后）。
- » SUSE Linux Enterprise Server Version 9 (SLES9)和SLES10（在SLES10发布之后，这取决于Novell）。
- » Windows Server 2003, Enterprise Edition for Itanium（最多8个CPU/16个内核，最大内存为256MB），当SP2于2007年上半年发布之后将会整合SP2（这取决于Microsoft）。
- » Windows Server 2003, Datacenter Edition for Itanium（最多32个CPU/64个内核，最大内存为2TB），当SP2于2007年上半年发布之后将会整合SP2（这取决于Microsoft）。

不支持Windows Server 2003 Standard Edition和Web Edition，因为Microsoft提供的这些版本不支持Itanium。

将来的操作系统升级和Service Pack版本可能包含特定的增强功能，可以优化PRIMEQUEST服务器运转，但它们最后会成为通用的Windows Server分发版本中的一部分。一般而言，PRIMEQUEST 500系列一开始会采用各个分发版本的企业版，随后再添加特定的扩展功能。

PRIMEQUEST服务器预计会给Windows和Linux操作系统提供富士通独创的增强功能。如果可能的话，普通用户将可以使用这些扩展功能。例如，当Microsoft公司的Vista（以前称为Longhorn）发布后，PRIMEQUEST服务器将支持这种操作系统。Vista可能包含支持PRIMEQUEST的特定的增强功能，但它们将是Windows Server的普通特征集中的一部分。类似地，富士通也在与开放源代码开发实验室（OSDL）合作，直接或通过Red Hat和Novell给Linux添加增强功能，这将改善所有Linux分发版本。

具有相应的Service Pack的Windows 2003以及Red Hat Enterprise Linux Version 4 (RHEL4)、Novell SUSE Enterprise Linux Version 9 (SLES9)都支持Itanium中的IA-32 Emulation Layer。因此，IA-32应用软件无需修改就可以在PRIMEQUEST服务器上运行。但是，Emulation Layer通常会降低性能，因此不要广泛使用这项功能。

PRIMEQUEST 500系列系统将经过Microsoft认证，适用于Windows Server 2003 DataCenter、Enterprise Editions的单机和集群配置。实际上，Microsoft一直在变更它的认证系统，对Datacenter 和Enterprise配置而言，PRIMEQUEST将获得统一的Microsoft认证。富士通的PRIMECLUSTER集群软件也可以运行在采用Linux的PRIMEQUEST系统上，将来还计划支持采用Windows的PRIMEQUEST系统。

当然，为了使PRIMEQUEST平台能够吸引客户，操作系统性能只是必要条件，而不是充分条件。ISV和IHV支持也很关键。补充说明“纵向领域与ISV”概述了富士通在这方面正在进行的努力。

### 更多的操作系统性能

另一项能吸引客户的PRIMEQUEST操作系统性能是服务器能够混合使用不同分区中的操作系统，并使得它们能协调运转。例如，Linux能运行在跨越两个系统板的一个逻辑分区上，而Windows能运行在另一个分区的一块单独的系统板上。

此外，可以在不同的分区上运行不同的Linux版本，进行实际运营和测试。但这些Linux版本必须能够提供PRIMEQUEST服务器需要的特征和功能。实际上，PRIMEQUEST服务器需要Linux 2.6内核（或更高）才能正常运转。这种内核级别具有大型多处理器服务器所需的可扩展性增强性能。

Linux 2.4内核不足以满足像PRIMEQUEST这样强大的服务器。例如，2.6内核与2.4内核相比，能处理的CPU总数为32 CPU vs 8CPU；最高的物理内存容量为2TB vs 64GB；虚拟地址空间为16TB vs 4GB；能容纳的设备数量为数万 vs 256。此外，2.6内核还具有一个逻辑卷管理器（logical volume manager），具有更大的文件系统性能以及更好的文件I/O。

### 富士通与开发源码组织在Linux方面的合作

富士通正在与Red Hat、Novell和OSDL合作，以改善Linux的可靠性、恢复能力、动态重配置性能、虚拟化性能和可扩展性。这种合作意味着，将来的Linux分发版本和相关的Linux软件最终会包括富士通的贡献，这些贡献会改善PRIMEQUEST服务器环境及其在以上这些方面的性能。

### Red Hat的diskdump及其他

例如，考虑系统转储（system dump）功能的问题。如果出现故障，系统挂起，或者内核失败（kernel panic），系统转储可以从PRIMEQUEST内存获得有用信

息, 确定故障原因并确保该故障不会再次出现。对Red Hat Linux而言, `diskdump`是实现这种转储并将转储信息转移到硬盘驱动器的标准工具。

Red Hat也提供标准的`netdump`工具, 它通过网络连接方式将系统运转期间的错误转储到一个转储服务器。Linux高性能计算应用中通常会采用`netdump`。富士通开发了`sadump`工具以进一步增强`diskdump`性能。`sadump`收集`diskdump`未涉及到的错误。这些错误包括`diskdump`收集信息期间的故障和启动错误。当操作系统开始引导的时候, `sadump`收集所有的启动错误。富士通现在希望在一张单独的CD-ROM (不在Linux分发版本中) 上为PRIMEQUEST系统提供`sadump`。

### Red Hat和Novell SUSE 的增强功能

富士通和其他厂商给开发源码组织提供建议和咨询。此外, 他们还提供开发人员和工程师与该组织合作, 改善Linux和相关的开发源码软件。富士通在这个领域地做了大量有组织的工作, 投入数百个软件开发人员改善Linux可靠性、性能和鲁棒性。

由于这些组织的努力, RHEL4和SLES9 Linux操作系统 (两者均基于Linux 2.6内核) 都包含了许多新特征。预计RHEL5将在2007年第一季度发布, 而SLES10将在2006年下半年发布。

在这些新特征方面, RHEL4和SLES9之间将存在某些区别, PRIMEQUEST服务器是否能够具有这些特征取决于使用了何种Linux分发版本。在Linux 2.6内核中, PRIMEQUEST 500系列服务器使用的一些重要的新特征包括:

**用于故障处理的Dump性能。** 在RHEL4中可以使用`diskdump` (上面讨论过), 在SLES9中则可以使用类似的工具`LKCD` (Linux Kernel Core Dump)。如上所述, 富士通还使得PRIMEQUEST服务器可以使用类似于大型机的`sadump`工具, 因为它具有更广泛的转储能力。富士通也将`sadump`提供给Linux组织。此外, SLES9还可以使用开放源码的`Flight Recorder`工具。这种工具可实现连续地将日志记录到一个循环缓冲区中, 以便可以更方便地分析转储信息。

**快速的高可用性集群故障切换。** RHEL4和SLES9中都包含用于快速集群故障切换的四种新性能。这些性能包括: 当有一个`netdump` (尽管如上所述, SLES9没有实现`netdump`) 警告集群软件出现了一个节点故障时, 就会触发快速故障切换。这种性能可以更快速地将功能转移到另一个集群节点。

另外两种性能 (两种分发版本中都有) 是在进程失败之前通知节点故障, 以及通知应用程序故障。这些性能是通过一个钩子 (`hook`) 实现的, 就如同富士通PRIMEQUEST服务器中为Solaris实现的方式一样。

两种分发版本中都能够通过从一个活动节点获得的通知终止一个故障节点。这项性能是从PRIMEQUEST架构开发出来的, 用以改善集群完整性, 现在已经捐献给Linux组织。

**Scalability.** 两种Linux分发版本都具有一些新性能，从而增强系统可靠性。例如，给Linux 2.6内核添加库，允许容纳更多的系统设备。PRIMEQUEST服务器和其他大型的Linux系统能使用数百台（或更多）系统设备。

此外，Linux glibc工具现在允许线程安全（thread-safe）操作。Glibc是通过标准C库为Linux实现的业界标准开放软件。它有助于系统的可扩展性，因为现在可以将程序安全地分发给多个线程，而无需考虑线程之间的资源交互/争夺。而且，系统的稳定性也得到了改善，即使线程操作繁多也不会造成系统故障。

此外，两种Linux分发版本现在都具有设备名称一致性。支持添加、拆除和热切换设备的环境（如PRIMEQUEST）需要这种一致性。现在，通过用于一致性参考的一个一致性名称就可以识别设备，而不是通过模糊的主/从设备映射（可以相互转换）。

**其他增强功能。** 其他增强包括易管理性方面的改善，如管理用于长期运行系统进程的具体资源。为了获得进一步的易管理性，RHEL4和SL3S9已经具有了sar和xxstat（用于报告内核统计）。

未来的增强功能会进一步改善基于内核的资源管理，以便安排商业经营环境中的长期运行的程序和系统资源。

**未来的重要增强功能。** 当前，Linux缺少动态重配置性能。这是开放源码组织中的一个有争议的问题，关于如何最佳地处理这个问题存在很多不同的观点。

富士通和其他厂商正在共同努力处理动态重配置问题，这个问题对于不中断客户应用服务的前提下进行热插拔、并发维护和分区调整来说极其重要。本系列中的关于高可用性以及系统架构的白皮书也讨论了动态重配置。<sup>1</sup>

Windows Longhorn edition可望提供动态重配置功能。富士通正在给Microsoft提供关于这个主题的信息。

**最佳案例之动态重配置开发。** 目前我们预期（这种期望极可能随着操作系统厂商的交付计划的变化而变化，这里只是作为一种“最佳案例”考虑提出来。）PRIMEQUEST 500系列服务器将具有下列动态重配置性能：

- » PCI-X设备热插拔
- » PCI-EX设备热插拔
- » 系统板热添加
- » 系统板热卸取（Hot Remove）
- » I/O单元热插拔

**Windows 和 Linux在 PRIMEQUEST系列 500 型服务器上运行良好, PRIMEQUEST 500 利用其类似于大型机的特征, 将支持有更多操作系统性能。**

**最佳案例之虚拟化开发。**目前我们预期(这种期望极可能随着操作系统厂商的交付计划的变化而变化, 这里只是作为一种“最佳案例”考虑提出来。) PRIMEQUEST 500系列服务器将具有下列虚拟化性能:

- » Linux环境的PRIMEQUEST虚拟化将基于SLES10和RHEL5中包含的开放源码Xen软件。富士通将会开发PRIMEQUEST需要的某些附加特征。正如为提升性能开发的那些特征一样, 那些为关键业务操作开发的特征将会移交给Xen/开放源码组织。富士通开发的管理特征和PRIMEQUEST专用的代码将不会移交给Xen/开放源码组织。
- » 根据系统管理程序(Hypervisor) PRIMEQUEST 500系列服务器虚拟化功能将需要使用Itanium专用的Intel Virtual Machine Assist硬件。主操作系统可以是RHEL5或SLES10。客户操作系统可以是RHEL4或RHEL5, SLES9或SLES10, Windows 2003 Enterprise或Datacenter Editions。所有的这些操作系统都将在即将来临的2007年实现。

## IDEAS概要

PRIMEQUEST是基于Intel 64位Itanium 2处理器系列(开始是Madison, 现在是Montecito, 将来是Montvale)的高端的关键业务服务器产品线, 这种处理器系列专为Windows(包括将来的版本, 如Vista)和两种主要的Linux分发版本(包括将来的版本, 如RHEL5及其后续版本, SLES10及其后续版本)设计。PRIMEQUEST服务器解决方案允许不同数量的纵向产业(vertical industries)和应用享受业界标准硬件(Intel)和运行在关键业务硬件上的操作系统(Windows与Linux)软件的服务。这些产业也将受益于富士通正在开发的完整的PRIMEQUEST生态系统(ecosystem)。

Windows和Linux可以不经改变即应用在PRIMEQUEST服务器中。但富士通正在与Microsoft公司和开放源码组织密切合作, 将附加特征添加到Windows和Linux中, 这些特征使其具有更多的适用于PRIMEQUEST服务器和其他计算机系统的功能。

例如, 在Windows方面, 富士通正在与Microsoft公司合作, 帮助将动态重配置性能整合到Longhorn中, 这种操作系统将会运行在PRIMEQUEST服务器上。

在Linux方面, 富士通正在与OSDL、Red Hat和Novell SUSE合作, 为他们各自的Linux分发版本和全体Linux团体添加特征。这些特征将处理转储性能、高可用性集群、可扩展性、RAS、资源管理和动态重配置。RHEL4 和SLES9(Linux内核为2.6)现在已经具有了这些领域中的一些特征。其他特征将出现在RHEL5和SLES10之中。

PRIMEQUEST服务器将在高端的关键业务服务器环境中脱颖而出。它类似于大型机的特征和功能（如具有高可靠性和系统管理性能），极高的性能，对于在Itanium上运行Windows和Linux的独特贡献（未涉及UNIX），独特的架构，强大的机械设计和小底座，以及分区性能等，使得PRIMEQUEST服务器对全球用户而言具有与众不同的优势和有价值的方案。

富士通期望PRIMEQUEST服务器能提供比其他厂商的Itanium 2产品更出色的总体拥有成本（COO）和投资回报（ROI）率。这方面与本白皮书系列中讨论的PRIMEQUEST服务器特征和功能相结合，一定引来全世界客户对PRIMEQUEST服务器的关注和兴趣。

## 尾注

1 请参阅下列姐妹篇白皮书：

- » “富士通新型高端双核Itanium PRIMEQUEST 服务器提供最大的高可用性,” Ideas International, 2006年7月
- » PRIMEQUEST系统架构,” Ideas International, 2006年7月

### Americas

Ideas International, Inc.  
800 Westchester Avenue  
Suite S620  
Rye Brook, NY 10573-1330  
USA  
Tel + 1 914 937 4302  
Fax +1 914 937 2485

### Asia/Pacific and Worldwide Headquarters

Ideas International Limited  
Level 3  
20 George Street  
Hornsby, NSW, 2077  
Australia  
Tel +61 2 9472 7777  
Fax +61 2 9472 7788

### Europe, Middle East, Africa

Ideas International Europe  
1 Deanes Close  
Steventon  
Oxon OX13 6SZ United Kingdom  
Tel +44 (0) 1235 437 850  
Fax +44 (0) 1235 437 851

[www.ideasinternational.com](http://www.ideasinternational.com)

