FUJITSU

White paper
Fujitsu BS2000/OSD VM2000 V10.0

Virtualization of BS2000/OSD within the Dynamic Data Center.

Introduction

A virtual infrastructure like VM2000 reduces IT costs by increasing efficiency, flexibility and response capability. It provides IT resource allocation on-the-fly in response to new business requirements and service requests. Extremely high levels of server utilization are a byproduct.

VM2000 supports the simultaneous operation of different, totally segregated system environments on one server. The CPU power and main memory of one real server can be distributed across up to 15 virtual servers. This distribution and the allocation of peripheral devices can be modified dynamically. The configuration of peripherals, including their connections (channels), and other devices can be modified or extended during live operation.

The advantage of using VM2000 as compared with the use of multiple servers is the possibility of consolidation with the aim of providing more efficient use of hardware resources, human resources and infrastructure.

VM2000: The basis for multiple system environments

Customers are increasingly faced with the need to handle different system environments simultaneously on one server in order to cope most effectively with the wide variety of IT tasks they have to deal with.

The reasons:
- Optimization of costs
- Simple and uniform handling and administration
- Parallel operation of production, development, test and version updates
- Automation and operational reliability
- Complex systems, for example for service data centers
- Availability for backup systems
- Separation of sensitive applications

VM2000 runs on the current BS2000/OSD S series and SQ series business servers and supports the current versions of the BS2000/OSD operating system as guest systems.

VM2000 can fulfill your requirements flexibly thanks to the following features:

Simultaneous operation of many systems
The maximum number of supported guest systems is 15 .

Full separation of guest systems
Access to memory areas on the other guest systems is protected. Faults in operation on one guest system do not affect the other guest systems, even if these errors cause the system to crash.

Flexible assignment of resources to the VMs
Memory, devices, CPU power and global store can be assigned to VMs "on the fly". The granularity of assignment is very small.
The Capacity-on-Demand feature is offered: The administrator can switch on extra CPUs for a certain time, so CPU power can be increased to cover a peak load period.

Increased reliability and availability
When the guest system used for administration fails, it is automatically restarted. A manual restart of this system can also be initiated. This does not affect the remaining guest systems.
When one CPU fails, VM2000 automatically activates the available spare CPU (S140, S145, S155, S170, S180 and S190), and system performance remains unimpaired. The same recovery is done for an involved guest system: a virtual spare CPU will be switched on – the guest system performance remains the same. With this technique, the availability of mono guest systems is equal to the availability level of MP guest systems.

BS2000/OSD guest systems have the same functionality as systems in native mode
The instruction set, network communication options as well as test and diagnostic utilities of all guest systems running under VM2000 correspond to operation without VM2000.

Performance of guest systems is comparable with native mode
The guest systems access the CPUs directly, with only minor emulation required. The memory is assigned permanently to the guest system and necessary address conversion is done by hardware. The devices are assigned directly to the guest system; virtual I/O operation is only necessary for shared pubsets and disks.

Simple system operation and configuration

Performance-hungry guest systems can be prioritized, thus enabling a flexible response to customer requirements. Administration of I/O peripherals is possible for VM2000 as a whole.

Full integration of BS2000/OSD guest systems into the automation and high availability concepts of BS2000/OSD

- HIPLEX-AF for application monitoring
- HIPLEX-MSCF for clustering

Classic benefits of VM2000 in a service data center

VM2000 allows data center service providers to install one or a small number of high-performance business servers that can run several operating systems for a variety of external customers.

This enables detailed capacity planning throughout the organization. Obvious knock-on effects of this include cost savings in relation to operating staff and space requirements for computers. The virtualization of resources such as CPU, main memory and global storage guarantees a high level of efficiency and optimum use of resources.

The billing of the consumed CPU power can be done in two different ways:
- Usage based
  VM2000 writes VM-specific accounting records. They show the consumed CPU and the time periods of resource assignment.
- Service level agreements
  A fixed amount of CPU power is agreed with the customer. The amount of CPU power used can be limited using the VM2000 function MAX-CPU-UTILIZATION.

Limitation of the CPU power for a group of VMs

A data center can offer type 2 pricing models to customers with more than one guest system on S servers. A two-step assignment of CPU power is possible. In the first step you decide how much power is given to the group and how the consumption of the group is limited. In a second step you decide how to distribute this power within the group. Power not consumed by group members will preferably be offered within the group.
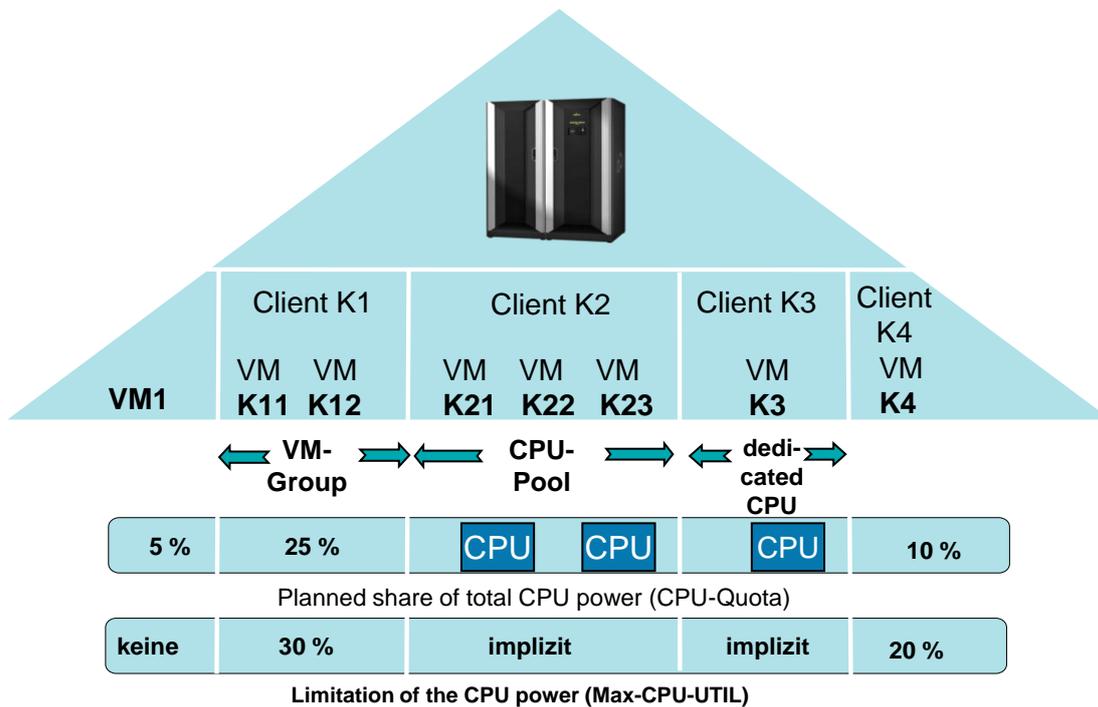
Assign a subset of the server to customer

CPUs can be combined to form CPU pools. You can arrange which guest systems should run within such a pool.

Dedicated CPUs

The data center can offer quasi-native running guest systems to customers with very high performance requirements. The virtual CPUs of such guest systems will be bound in a fixed manner to real CPUs. So the Hypervisor overhead and the indirect overhead will become extremely small for the VM.

Very fine CPU-QUOTA and UTILIZATION

These scheduling parameters can now be specified by decimal numbers. So it is possible to specify CPU assignments for big servers which correspond to one-digit RPF values.

VM1	Client K1		Client K2			Client K3	Client K4
	VM K11	VM K12	VM K21	VM K22	VM K23	VM K3	VM K4

VM-Group ← → ← CPU-Pool → ← dedi-cated CPU →

5 %	25 %	CPU	CPU	CPU	10 %

Planned share of total CPU power (CPU-Quota)

keine	30 %	implizit	implizit	20 %

Limitation of the CPU power (Max-CPU-UTIL)

VM2000 added value in high-availability network

The general aim of a high-availability BS2000/OSD cluster is to be able to switch to redundant hardware and operating systems in the event of a server or application failure. In the event of a fault, business-critical applications on the failing system are exported together with their resources to the surviving system (less important services on the standby system may have to be delayed).
HIPLEX (Highly Integrated System Complex) is Fujitsu's clustering concept, designed to support an availability cluster comprising several BS2000/OSD business servers. A HIPLEX with VM2000 servers provides the following benefits:

System availability
If a guest system on which an application is running fails, the entire system environment with the application can be made available without delay on the second business server.

Application availability
If the application fails, it can quickly continue running on the same server in a backup guest system or it can be switched to a second server. The second system is in this case actively redundant. An example of an application that requires maximum availability would be automatic teller machines (ATM) in different bank branches: The service is available around the clock and the transactions are monitored centrally. Two business servers running VM2000 are provided for this purpose at the bank's head office as well as the BS2000/OSD operating system and the relevant "ATM" application software. The required availability is provided by a HIPLEX with VM2000 running on every server.

In the event of errors, operation of the same production application (in A1) is guaranteed on the second business server (B) under VM2000. A guest system (B2) with a basic memory configuration is set up for backup purposes in this instance. If errors occur, the main memory of the backup system is extended dynamically using the resources of the lower-priority guest systems, the production application is moved out to the backup system and can continue working immediately. The data belonging to the production application must be transferred, provided it is not stored on shared data media (shared pubset).
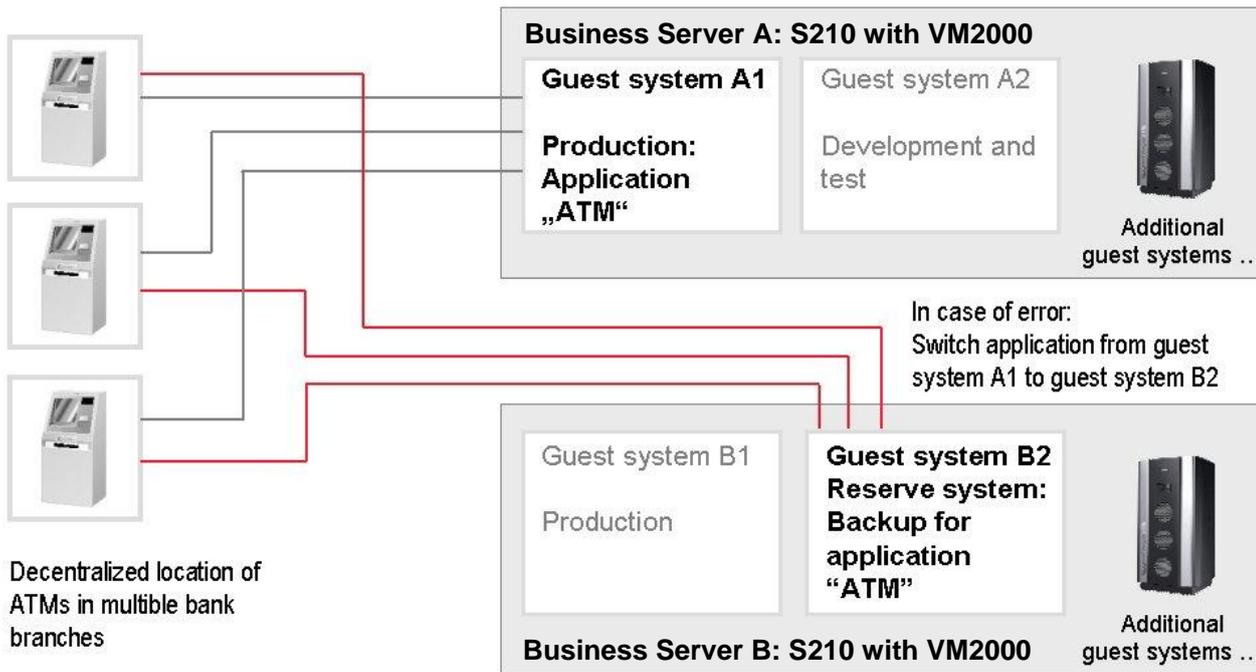
Functional enhancements in VM2000 V10.0

VM2000 V10.0 featuring support for the new BS2000/OSD-BC V9.0 functions.

VM2000 V10.0 supports

- on S servers: OSD V8 and OSD V9 as monitor system, OSD V6 to OSD V9 as guest systems,
- on SQ servers: OSD V8- or OSD V9-based OSD/XC packages as monitor system, OSD V8- to OSD V9-based OSD/XC packages as guest systems.

VM2000 V10.0 requires X2000 as of V5.3 on SQ servers.
SX servers are supported no longer.

The main new functions in VM2000 V10.0 are:

Consolidation of HSI lines

After version V9.0 (S/SX) and V9.5 (SQ only) VM2000 V10.0 is a joint version, which will again support all the current HSI lines (S/SQ). SX servers are officially no longer supported.

Support of ETERNUS DX disk systems

Eternus DX disk systems are supported to the same extent as EMC Symmetrix disk systems by VM2000 (PAV, Snap & Clones).

Save & Restore of the VM2000 configuration

The backup of an existing VM2000 configuration is intended to enable the VM2000 administrator to recover the VM2000 configuration after the restart of the overall VM2000 system, i.e. ending all guest systems and restarting the monitor system, without having to write a suitable procedure file with VM2000 commands or adapt an existing procedure file. It should also be possible to use the backup to restore the VM2000 configuration on a backup server, on which - apart from the monitor system - no further BS2000 guest systems run.

The VM2000 administrator can generate a backup with the command SAVE-VM-CONFIGURATION. The VM2000 administrator can trigger the recovery of a VM2000 configuration through the installation of a backup with the command RESTORE-VM-CONFIGURATION.

CPU pools also for SQ servers (incl. adaptation of the multiprocessor capacity of the VMs)

The restriction that in addition to the static CPU pools no further dynamic CPU pools can be configured on SQ servers is lifted. In order to administer the CPU pools the commands CREATE-, DELETE- and SHOW-VM-CPU-POOL as well as ASSIGN-VM-TO-CPU-POOL and SWITCH-VM-CPU are available to the VM2000 administrator. In the last two commands the current multiprocessor capacity of the affected VM(s) is - as on S servers - checked and automatically modified.

32 virtual CPUs per VM for SQ servers

The maximum multiprocessor capacity of a VM on an SQ server is increased from 16 to 32.

VM2000-specific deliverables for SQ servers in a HA cluster

The high-availability & LiveMigration functionality for SQ server units is mainly implemented through MARS, X2000 and HA-Addon software. Nevertheless, deliverables are also required in VM2000, which make it possible to use VM2000 in such a HA cluster (e.g. a global view of all VMs in a cluster due to the required global uniqueness of VM names and necessary VM2000-specific tests).

Version overview

Server line	S server				SQ server	
VM2000 version	V7.0	V8.0	V9.0	V10.0	V9.5	V10.0
General release	May 02	Mar 05	May 07	Sep 12	Sep 10	Sep 12
BS2000/OSD version of the monitor system	V4.0, V5.0	V5.0, V6.0, V7.0	V6.0, V7.0, V8.0	V8.0, V9.0	OSD/XC V4.x	OSD/XC V4.x, V9.0
Guest systems						
BS2000/OSD V4.0	x	x				
BS2000/OSD V5.0	x	x	x			
BS2000/OSD V6.0	x	x	x	x		
BS2000/OSD V7.0		x	x	x		
BS2000/OSD V8.0			x	x		
BS2000/OSD V9.0			x	x		
OSD/XC V4.0, V4.1					x	x
OSD/XC V9.0						x
Supported business servers						
S110, S115, S130, S135, S150, S160	x	x	x			
S120, S140, S145, S155, S170, S180, S190	x	x	x	x		
S165, S200 [1]		x	x	x		
S175, S210			x [3]	x		
SQ100					x	
SQ200 [2]					x	x
SQ210					x [4]	x

[1] : guest systems as of OSD V5.0 (also with VM2000 V8.0)
[2] : monitor and guest systems as of OSD/XC V4.1 only
[3] : subsequent also with VM2000 V9.0 possible
[4] : subsequent also with VM2000 V9.5 possible

Performance information
VM2000 overhead

The VM2000overhead, which arises for loads with exclusive allocation of the disk peripherals in comparison with native operation, is caused by the CPU requirement of the hypervisors (HPV-ACTIVE), the paths for switching from one BS2000 guest system to the hypervisor, and as a result of the loss in hardware performance because of the sinking effectiveness of the CPU caches due to competing guest systems. The size of the overheads is essentially determined by the number of virtual processors and their relationship to the physical processors as well as by the IO load intensity.

The VM2000 overhead was measured using the TPS benchmark on S servers. It contributes more than 10% for a configuration with a quadro VM. If only mono and bi-VMs are operated, the overhead is between 6 and 6.5%, thus making it higher than the overhead measured with VM2000 V9.0 (4.1%). More frequent changes to the hypervisor due to the higher system performance were determined as the cause for the increased overhead.

The recommendations previously made for the operation of BS2000 systems under VM2000 should be observed, i.e. the usual guidelines for the load (65-70% for TP operation) should be maintained and the chosen degree of VM multiprocessors should be as low as possible.

Summary of VM2000 benefits

- Parallel operation of several BS2000/OSD systems on one server
- Support for version upgrades of the operating system, system-specific software and application systems
- Parallel operation of several different BS2000/OSD operating system versions and Linux systems on one server
- More flexible resource distribution than is possible on multi-server configurations
- Provision of backup systems
- Price advantages compared to several servers (consolidation)