FUJITSU

White paper
FUJITSU Software BS2000 VM2000 V11.0

Virtualization of BS2000 within the Dynamic Data Center.
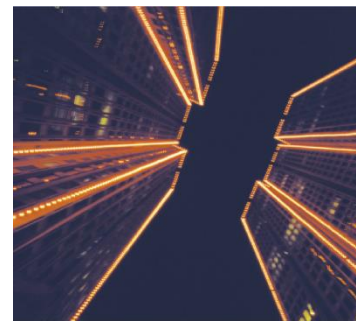
Introduction

A virtual infrastructure like FUJITSU Software BS2000 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 operating resources (CPU power and main memory) of one real server can be distributed across up to 32 BS2000 guest systems. This distribution 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 as base for different 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 are:
- Optimization of costs
- Simple and uniform handling and administration
- Parallel operation of production, development, test and version updates
- Automation and operational reliability
- Differentiated systems, for example for service data centers
- Availability of backup systems
- Separation of sensitive applications

VM2000 V11.0 is exclusively released for the FUJITSU Server BS2000 SE series and supports the current versions of the BS2000 operating system BS2000 OSD/BC as guest systems.

The provision of different system environments fulfills VM2000 in a flexible manner by the following features:

Simultaneous operation of many guest systems

On server unit x86 up to 32 BS2000 guest systems (incl. monitor system) can run simultaneously (on server unit /390 the number of guest systems is limited to 15).

Full separation of guest systems

Access to memory areas and devices of other guest systems is not possible. 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 guest systems

Memory, devices, CPU power and global store can be assigned to guest systems "on the fly". The granularity of assignment is optimally small. The Capacity-on-Demand feature is offered: through connection of extra CPUs on the fly, CPU power can be increased for a certain time.

Increased reliability and availability

When the guest system (or the monitor system) fails, it can be automatically restarted. A manual restart of the monitor system can also be initiated. This does not affect the remaining guest systems.

When one CPU fails, VM2000 automatically activates the available spare CPU (on servers unit /390), system performance remains unchanged. The same applies to possible affected guest systems: 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 multiprocessor guest systems.

BS2000 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 guest systems execute the IO's normally directly.

Simple system operation and configuration

Important guest systems can be prioritized, thus enabling a flexible response to customer requirements.
The management of I/O peripherals can be done VM2000 spanning: the reconfiguration and the dynamic expansion of peripheral objects is done in common for all guest systems from the monitor system.

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 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
  the customer is guaranteed a certain CPU power, for which a RPF-based constant pricing is determined. 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 guest systems

A data center can offer type 2 pricing models to customers with multiple guest systems 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 dynamically CPU pools. For such a pool VMs can be determined, which only can use this pool. The CPUs and VMs of the pool form a part of the server unit, which is provided to a customer.

Dedicated CPUs

If the number of connected real CPUs in a CPU pool is greater or equal to the sum of connected virtual CPUs, then VM2000 assigns each virtual CPU of a VM one real CPU.
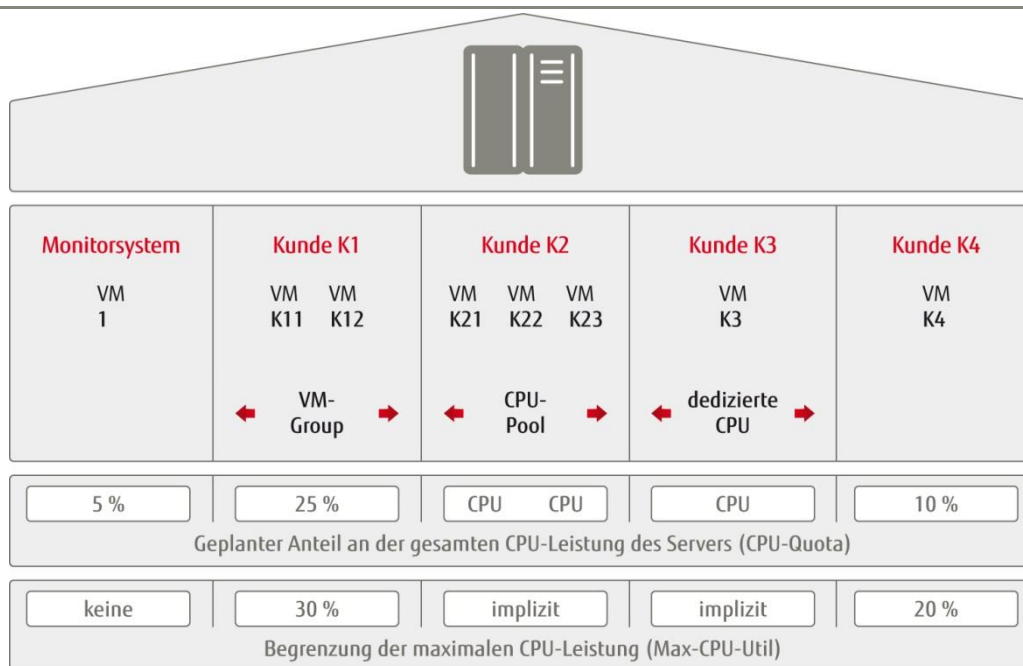The solid CPU assignment is optimal for a sufficient number of real CPUs in terms of performance, because each virtual CPU is always running on one and the same real CPU.

VM groups

Multiple VMs on a server unit /390 can be combined into a VM group for which CPU scheduling specifications (CPU-Quota and MAX-CPU-UTILIZATION) can be effected comprehensive. The first step is to determine which CPU performance and CPU power limitation receives the VM group. In the second step, the power distribution within the VM group is determined. VM2000 provides for priority allocation of the group planned CPU power within the group. Service data centers can thus organize guaranteed computing performance for customers with multiple VMs.

Granularity setting of CPU-QUOTA and MAX-CPU-UTILIZATION

The two attributes for controlling the VM's performance can be specified with two decimal places. This means that for very big server units definitions are possible in the one-digit RPF area up to one percent of the CPU capacity.

Monitorsystem	Kunde K1	Kunde K2	Kunde K3	Kunde K4
VM 1	VM K11 VM K12	VM K21 VM K22 VM K23	VM K3	VM K4
	← VM-Group →	← CPU-Pool →	← dedizierte CPU →	

5 %	25 %	CPU CPU	CPU	10 %

Geplanter Anteil an der gesamten CPU-Leistung des Servers (CPU-Quota)

keine	30 %	implizit	implizit	20 %

Begrenzung der maximalen CPU-Leistung (Max-CPU-Util)

Functional enhancements in VM2000 V11.0

The virtual machine system FUJITSU Software BS2000 VM2000 V11.0 exclusively supports the new SE servers.
On S and SQ servers VM2000 V11.0 has not been released.
The main new features of VM2000 are:

Support of SE servers

VM2000 is available for the server unit /390 as well as the server unit x86 of the SE servers with a unified management.
On the server unit /390 up to 15 VM guest systems are released, on server unit x86 up to 32.
Basic VM2000 features to manage the BS2000 VMs are available to the administrator of the SE server via the SE manager.

VM definitions

From VM2000 V11.0 each BS2000 VM on a server unit of a SE server owns a configuration description, the VM definition. It contains among other attributes and resources that are associated to the VM or provided for the VM.
VM definitions with the attribute PERSISTENT are called persistent VM definitions; they are still available after a restart of the server unit. With their help, a persistent VM is re-established and can be started at once again corresponding to the definition.

The management of the VM definitions is task of VM2000 administrator; therefor functions for configure, update, activate, read and delete of VM definitions are available.

Timeslot extension for the nucleus lock
In order to avoid lock conflicts and the associated performance losses for a large multi-processor degree of a VM, a mechanism is implemented on the SU /390 between BS2000 and the VM2000 hypervisor via which a forced removal of a virtual CPU from a real CPU (after a time slot end) can only occur when the virtual CPU does not hold a nucleus lock.

Version overview
Server models supported by VM2000:

BS2000 Server	VM2000				number of possible VMs
	V9.0	V9.5	V10.0	V11.0	
S165	x	-	x	-	15
S175	x	-	x	-	15
S200	x	-	x	-	15
S210	x	-	x	-	15
SQ100	-	x	x	-	15
SQ200	-	x	x	-	15
SQ210	-	-	x	-	15
SE Server, SU x86	-	-	x*	x	32
SE Server, SU /390	-	-	x*	x	15

* in VM2000 V10.0 the new functions for the SE servers are not available

Version of operating systems supported by VM2000 V11.0

VM2000 version	server series	version of the monitor system		version of the guest systems	
		version of operating system	version of OSD/XC package	version of operating system	version of OSD/XC package
V9.0	S server	V7.0, V8.0	-	V7.0, V8.0	-
V9.5	SQ server	V8.0	V4.x	V8.0	V4.x
V10.0	S server	ab V8.0	-	ab V7.0	-
	SQ server	ab V8.0	ab V4.0	ab V8.0	ab V4.0
	SE server	V9.0, V10.0	V9.5, V10.0	V8.0, V9.0, V10.0	V8.5, V9.5, V10.0
V11.0	SE server	V10.0	V10.0	V8.0, V9.0, V10.0	V8.5, V9.5, V10.0

VM2000 V11.0 requires X2000 as of V6.0 on SE servers, server unit x86.

Performance information
VM2000 overhead
The VM2000 overhead, 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 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 systems, respectively on server unit x86 also Linux/Windows systems on one server
- Support for version upgrades of the operating system, system-specific software and application systems
- More flexible resource distribution than is possible on multi-server configurations
- Provision of backup systems
- Price advantages compared to several servers (consolidation)