

WHITE PAPER

VM2000 V9.0 BS2000/OSD Virtualization within the Dynamic Data Center

Issue May 2009

Pages 8

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.

In addition to supporting new BS2000/OSD-BC V7.0 and V8.0 functions, VM2000 V9.0 introduces enhanced automation capabilities and improved service level management for large configurations.



Contents

VM2000: The basis for multiple system environments	2
Classic benefits of VM2000 in a service data center	3
VM2000 added value in high-availability network	4
Functional enhancements in VM2000 V9.0	5
Version overview	8
Summary of VM2000 benefits	8

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 SX series business servers and supports the current versions of the BS2000/OSD operating system as guest systems.

VM2000-Linux extends the VM2000 virtual machine system of BS2000/OSD so that Linux systems can also run on the servers. Up to fourteen Linux systems can run simultaneously on selected BS2000/OSD business servers of the S series (S120, S140, S145, S155, S165, S170, S180, S190 and S200). Alternatively, several BS2000 systems and several Linux systems can run concurrently in parallel with and independently of one another (max. 15 systems in total).

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

Simultaneous operation of many systems

The number of guest systems depends on the type of server. The S server firmware supports as many as 15 guest systems. On the current SX servers, up to 15 guest systems are released. Technically, the maximum number of supported guest systems on SX servers is 99 as of VM2000 V8.0.

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. New S and SX models offer the Capacity-on-Demand feature: 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.

With VM2000 V8.0 the usage scenarios for service data centers were extended. These features are being enhanced once more with VM2000 V9.0:

Limitation of the CPU power for a group of VMs

A data center now can offer type 2 pricing models to customers with more than one guest system. A two-step assignment of CPU power is now 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

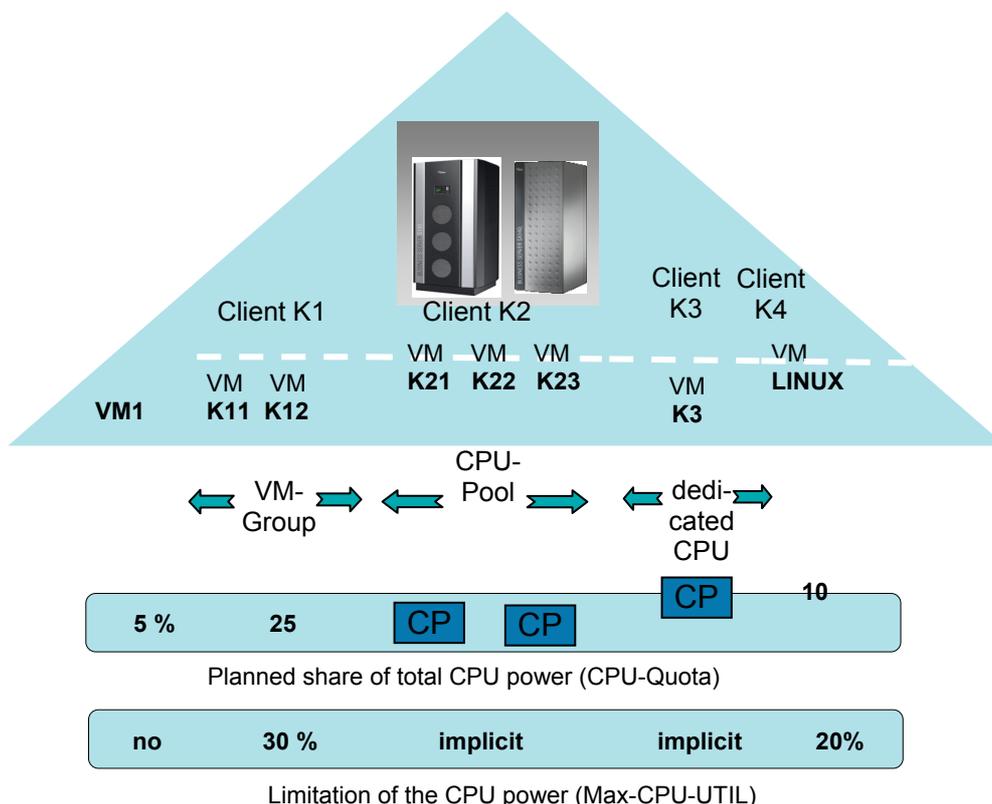
CPUs can be combined to form CPU pools. You can arrange which guest systems should run within such a pool.

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.



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

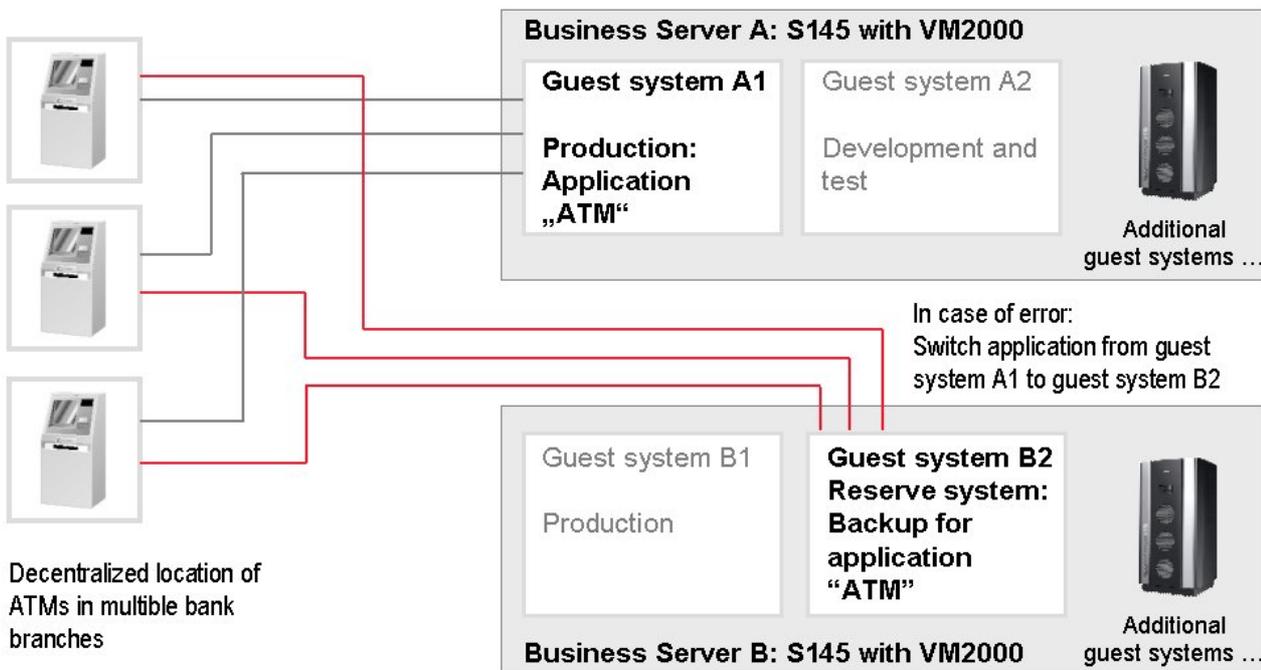
If a guest system on which an application is running fails, the entire system environment with the application can be made available without delay on the second business server.

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 V9.0

VM2000 V9.0 is the VM2000 version featuring support for the new BS2000/OSD-BC V7.0 and V8.0 functions.

VM2000 V9.0 supports

- On S servers: OSD V6 , V7 and OSD V8 as monitor system, OSD V5 – OSD V8 as guest systems,
- On SX servers: OSD V6- , V7- or OSD V8-based OSD/XC packages as monitor system, OSD V5- to OSD V8-based OSD/XC packages as guest systems.

VM2000 V9.0 requires X2000 as of V3.0 on SX servers.

The main new functions in VM2000 V9.0 are:

Easier VM assignment of pubsets

The VM2000 device pool includes all the devices identified at the startup of the monitor system. After a VM has been initialized, peripheral devices can be assigned to this VM either explicitly or implicitly:

- Explicitly by the VM2000 or VM administrator using /ADD-VM-DEVICES or /SWITCH-VM-DEVICES,
- Implicitly upon ATTACH-DEVICE by the guest system, assuming the VM has been issued with ASSIGN-BY-GUEST privilege and the ASSIGN-BY-GUEST device attribute has been assigned to the corresponding devices by means of the MODIFY-VM-DEVICE-ATTRIBUTES command.

Assigning the devices of a pubset to a VM is greatly simplified. Pubset reconfigurations (e.g. expansion) are automatically taken into account. Command procedures for setting up VMs are shorter and more clearly structured, and no longer need to be adapted, in particular when the SPACEPRO provisioning tool (new in OSD V7) is used.

Specifically, device assignment for pubset devices is extended as follows:

- Explicit assignment of devices to VMs will be possible at pubset level (via specification of the pubres device or the catid)
- Assigning the "ASSIGN-BY-GUEST" device attribute will be possible at pubset level.
- In a pubset reconfiguration, the VM assignment and the attributes of the implicit assignability for the affected devices are updated automatically. A similar dynamic update takes place during an ATTACH pubset for new devices added to the pubset in the interim. This automatic update is performed only on pubset reconfiguration or pubset attach by a guest system as of OSD V7.

This new functionality offers the following customer benefits:

- The command input for explicit assignment of pubset devices or for preparing the implicit assignment will be shorter and therefore easier. It is sufficient to specify the pubres rather than the list of all devices in the pubset.
- The follow-up action to a pubset expansion or pubset reduction, i.e. to update the VM assignment or the device attributes of the affected device in the VM2000 command files, is no longer required.
- For the current VM2000 session, assignment and device attributes of the reconfigured devices are updated automatically.
- In the next VM2000 session, all then current pubset devices are automatically registered by the corresponding command.

Finer privilege levels for implicit device assignment

The implicit assignment of devices to a VM (during ATTACH for the device in the guest system) requires the VM to be issued with the appropriate privilege and the device to be released for the function.

VM2000 V9.0 enables

- devices to be subdivided into ASSIGNMENT sets for implicit assignment
- corresponding VM privileges to be conferred.

This function opens up a variety of usage scenarios:

- An exclusive set of devices/pubsets for implicit assignment can now be specified for a VM or a VM group.
- The finer ASSIGN-BY-GUEST privilege assignment permits this function to be used when there are a number of strictly segregated VMs or VM groups present.
- VM assignment for pubsets which are used by the SPACEPRO product is implicit. A pool pubset of this kind can therefore only be released for a selected set of VMs.

Support for Snapsets

EMC provides the Timefinder/Snap local replication mechanism for Symmetrix DMX storage systems.

Timefinder/Snap produces a virtual copy of a logical volume at any snap(shot) time as follows: Timefinder/Snap saves the content of the volume by writing a copy of the original block into a special Symmetrix space, the so-called Save Pool, when a block of the original volume is modified for the first time (copy-on-first-write technique).

For snap-based recovery, the blocks are copied back from the Save Pool to the original volume.

The principal deployment scenario for Timefinder/Snap is fast and efficient backup and restore to/from disk.

BS2000/OSD-BC as of V7.0 supports snapshot-oriented backup/restore scenarios in DMX configurations. The pubset copy that can be used for restore consists of the simultaneously generated snap units for all volumes of the pubset. The snap unit is a virtual device that contains the pointers to the original data. For unchanged data, the pointers address the original volume. For changed data, the pointers address the Save Pool.

A pubset copy of this kind from snap units is referred to below as a "snapset". The BS2000/OSD-BC V7.0 functions for snapset support are based on the functions for TimeFinder/Snap control in SHC-OSD as of V6.0.

VM2000 V9.0 supports the use of snapsets on guest systems running BS2000/OSD-BC V7.0 or higher.

- VM2000 displays snap units in its SHOW commands.
- Under VM2000 V9.0, the AUTO-SNAP-ASSIGNMENT privilege permits the guest system on a VM to assign itself snap units of a snapset implicitly without VM and device being prepared for the implicit device assignment (i.e. no ASSIGN-BY-GUEST privilege or attribute for VM and device).

A VM is granted the AUTO-SNAP-ASSIGNMENT privilege by default at /CREATE-VM time.

BS2000/OSD-BC V8.0 supports snapshot-oriented backup/restore scenarios in FibreCAT CX configurations on SX servers in the context of snapshots (based on the functions for SnapView Snap control in SHC-OSD V7.0). VM2000 V9.0 supports the use of SnapView-based snapshots on guest systems running BS2000/OSD-BC V8.0 or higher.

Support for the I/O Resource Manager (IORM)

IORM comprises the following functions to control I/O resources in an autonomous dynamic manner (devices, controllers, channels, paths): I/O Priority Handling for Tasks (IOPT), dynamic I/O load balancing for disks on FC channel (DPAV), optimized device selection in CentricStor operation under VM2000 (DDAL), compression switching for LTO tapes (TCOM,) and I/O consumption limit for particular VM2000 guest systems (IOLVM).

IORM was developed for BS2000/OSD-BC V7.0. Functional subsets of IORM were released for operation with BS2000/OSD-BC V6.0B and BS2000/OSD-BC V5.0C.

VM2000 V9.0 supports IORM. In VM2000 operation, IORM should operate on the monitor system and on all guest systems. The DPAV function of IORM is controlled by the monitor system. The DDAL and IOLVM functions of IORM serve to optimize VM2000 operation. In detail:

- **Dynamic I/O load balancing for disks on FC channel on S servers with DPAV in VM2000 operation**
Dynamic PAV (DPAV) assigns autonomously alias devices to those volumes that will benefit most. In VM2000 operation the switching of alias devices is coordinated and executed by the DPAV instance of the monitor system.
Under VM2000 V9.0 dynamic PAV with IORM is offered for BS2000/OSD-BC V5.0C guest systems with monitor system as of BS2000/OSD-BC V6.0B.
- **Optimized load balancing in CentricStor operation under VM2000 with DDAL (Dynamic Device Allocation)**
The use of virtual tape devices in CentricStor can be set in BS2000/OSD-BC V7.0: As an alternative to using devices as defined at generation (as previously), the system can ensure uniform utilization of all available ICPs (Integrated Channel Processors).
Under VM2000 the DDAL function extends the optimized local device selection to all BS2000/OSD-BC V7.0 guest systems of a server.
Even though the optimized local device selection is available only with OSD V7.0, it makes sense to activate DDAL also in the OSD V6.0 guest system. The reservation counters can be determined in the OSD V6.0 guest systems and can be made available to the OSD V7.0 guest system. In contrast, DDAL is not available for OSD V5 guest systems.
- **Limited I/O consumption of particular VM2000 guest systems with IOLVM (I/O Limit for Virtual Machines)**
The IOLVM function of IORM serves to limit the I/O consumption of particular VM2000 guest systems. So the operator can ensure that low-priority, yet I/O-intensive guest systems cannot affect higher-priority guest systems.
IOLVM is available for guest systems as of BS2000/OSD-BC V5.0C.
The VM I/O limit controls for IOLVM can be set in VM2000 V9.0 via the VM2000 user interface. The /CREATE-VM and /MODIFY-VM-ATTRIBUTES commands were extended appropriately. The control value is displayed by the VM2000 information commands /SHOW-VM-ATTRIBUTES and /SHOW-VM-RESOURCES. It is also shown if the IOLVM function is activated in the guest system.

Enhanced support for large configurations

Support for large configurations is optimized by the following measures:

- HIPLEX-MSCF-coordinated Move of a VM
- Very small VMs
- Option to combine VM groups and CPU pools
- Support for large VM2000 main memory

HIPLEX-MSCF-coordinated Move of a VM

/MOVE-VM is used to move a VM in VM2000 main memory. The VM is stopped by VM2000 while it is moved in main memory. Since the time taken for this operation for a large domain (greater than 2 GB) is no longer negligible (more than 10 seconds), there is the risk that a guest system in the HIPLEX-MSCF cluster will be deemed to have crashed while HIPLEX-MSCF partners are being moved. In VM2000 V9.0, the MOVE-VM operation is safeguarded via an interface to HIPLEX-MSCF for OSD V7 guests.

Very small VMs

On very large consolidation servers it is not possible at the present time to operate VMs with a CPU component of 5-10 RPF. With a 2000 RPF server, for example, the value range 1.00 – 100.00 for CPU-QUOTA, MEMBER-CPU-QUOTA and MAX-CPU-UTILIZATION allows only specifications of 20 RPF or more.

Two new functions are being introduced so that very small VMs can be defined.

The value range of the CPU-QUOTA, MEMBER-CPU-QUOTA and MAX-CPU-UTILIZATION operands is being formally extended to 0.01 – 99.99 (or 100.00 for MAX-CPU-UTIL).

The VM timeslice is being shortened for small VMs so that a VM schedule can be guaranteed ideally every 100 ms. Shorter timeslices guarantee a sufficiently frequent, albeit more strictly limited, CPU allocation for these VMs.

Option to combine VM groups and CPU pools

VM groups and CPU pools were introduced in VM2000 version V8.0A.

- **VM groups:** Multiple guest systems can be combined into a VM group for which CPU scheduling parameters can be specified across the board within a service level management framework.
- **CPU pools:** The VM2000 administrator can form disjoint sets of real CPUs and in each case specify a set of VMs that will make exclusive use of this pool. With a real MP level of 4-6- CPUs or more, CPU pools produce a lasting improvement in VM2000 performance.

In VM2000 V8.0A, VM groups can only be set up in the standard CPU pool. The standard CPU pool is the CPU pool to which all normal and extra CPUs are assigned at the start of a VM2000 session.

In VM2000 V9.0, VM groups can be set up in all CPU pools, subject to the following rules:

- Each VM group is always bound to precisely one CPU pool.
- Multiple VM groups may be bound to one CPU pool.
- A virtual machine can only be assigned directly to another CPU pool if it does not belong to any VM group.
- Virtual machines belonging to a VM group change their CPU pool if the VM group is assigned to another CPU pool.

Thus, VM groups and CPU pools can be used concurrently and the benefits of the two functions – better performance and optimal service level management – can be achieved at the same time.

Support for large VM2000 main memory

In VM2000 V9.0, the VM main memory was expanded above the former limit of 2GB for VM1, and the VM main memory for the remaining VMs above the former limit of 32GB. This takes account of the large main memory configurations of today's servers (main memory for S200 max. 256 GB), and VM1 can also be optimally used as a production system. Now the main memory of VM1 can be expanded and reduced dynamically.

Shutdown for individual VMs and coordinated shutdown for the entire VM2000

The VM2000 administrator is given the option to initiate a shutdown for individual VMs. Besides the SHUTDOWN command execution, a "Shutdown" RUN file can be activated. A coordinated VM2000 shutdown between guest systems and monitor system is also supported as a further function. (Currently, the coordination is handled by the operator or by cooperating procedures in the VMs; a time-controlled function via X2000 is also possible on SX servers.)

In VM2000 V9.0, a shutdown is initiated automatically in all guest systems as a first step. Following termination of all the guest systems, a shutdown is initiated in VM1. This then also terminates VM2000. The function is available for guest systems as of OSD V7.

Improved performance monitoring with SHOW-VM-STATUS

The /SHOW-VM-STATUS command periodically outputs current plan values and current measured values of active objects during ongoing VM2000 operation. Active objects are attached normal and extra CPUs, running VMs, CPU pools with attached real CPUs, and VM groups with running VMs.

The usability of the SHOW-VM-STATUS command is being enhanced as follows:

- As an alternative to the previous merely periodic output, a one-time, synchronous output of the measured data of the immediately preceding past is also possible.
- The command for one-time, synchronous output can be used by several administrators simultaneously.

Version overview

Server line	S server						SX server		
VM2000 version	V7.0		V8.0		V9.0		V7.5	V8.0	V9.0
VM2000-Linux		V7.0		V8.0		V9.0	-		
General release	May 02	Mar 03	Mar 05	Mar 05	May 07	July 07	Jan 03	Mar 05	May 07
BS2000/OSD version of the monitor system	V4.0, V5.0		V5.0, V6.0, V7.0		V6.0, V7.0, V8.0		OSD/XC V1.x	OSD/XC V1.x, V2.x, V3.0	OSD/XC V2.x, V3.0, V4.0
Guest systems									
BS2000/OSD V4.0	x		x						
BS2000/OSD V5.0	x		x		x				
BS2000/OSD V6.0	x		x		x				
BS2000/OSD V7.0			x		x				
BS2000/OSD V8.0					x				
OSD/XC V1.0, V1.1, V1.2							x	x	x
OSD/XC V2.0, V2.1, V2.2							x	x	x
OSD/XC V3.0								x	x
OSD/XC V4.0									x
Linux-S		x		x		x			
Supported business servers									
S110,S115,S130,S135, S150,S160	x		x		x				
S120,S140,S145,S155, S170,S180,S190	x	x	x	x	x	x			
S165, S200*			x	x	x	x			
SX100-A-V-B,SX130, SX140							x	x	x
SX100-C, SX150**								x	x
SX100-D, SX160***									x

*: guest systems as of OSD V5.0 (also with VM2000 V8.0)

**: monitor system OSD/XC as of V2.x only

***: monitor and guest systems as of OSD/XC V3.0 only

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)

All rights reserved, including intellectual property rights. Technical data subject to modifications and delivery subject to availability. Any liability that the data and illustrations are complete, actual or correct is excluded.

Designations may be trademarks and/or copyrights of the respective manufacturer, the use of which by third parties for their own purposes may infringe the rights of such owner.
For further information see ts.fujitsu.com/terms_of_use.html

Published by department:

Margret Germann
Phone: ++49 (0)89 3222 2623
Fax: ++49 (0)89 3222 329 2623
Margret.Germann@ts.fujitsu.com
ts.fujitsu.com

Partner login

partners.ts.fujitsu.com