

English



Fujitsu Server BS2000 SE Series

# Cluster Solutions for SE Server

Whitepaper

---

Valid for:  
M2000 V6.5  
X2000 V6.5  
HNC V6.5

Edition December 2023

## Comments... Suggestions... Corrections...

The User Documentation Department would like to know your opinion on this manual. Your feedback helps us to optimize our documentation to suit your individual needs.

Feel free to send us your comments by e-mail to: [bs2000services@fujitsu.com](mailto:bs2000services@fujitsu.com).

## Certified documentation according to DIN EN ISO 9001:2015

To ensure a consistently high quality standard and user-friendliness, this documentation was created to meet the regulations of a quality management system which complies with the requirements of the standard DIN EN ISO 9001:2015.

## Copyright and Trademarks

Copyright © 2023 Fujitsu Technology Solutions GmbH.

All rights reserved.

Delivery subject to availability; right of technical modifications reserved.

All hardware and software names used are trademarks of their respective manufacturers.

# Table of Contents

- Cluster Solutions for SE Server - Whitepaper** ..... 5
- 1 Introduction; requirements for cluster functions** ..... 6
  - 1.1 Hardware structure** ..... 7
  - 1.2 Network connections** ..... 9
  - 1.3 CRD functionality; CRD with external configuration disks** ..... 11
  - 1.4 SEM displays and functions** ..... 14
- 2 Management Cluster - central management of SE Servers** ..... 15
  - 2.1 General information about the Management Cluster** ..... 16
  - 2.2 SEM main window for Management Cluster** ..... 18
  - 2.3 Server-overarching functions (a choice)** ..... 19
    - 2.3.1 Logging in, session and user management ..... 20
    - 2.3.2 Audit logging ..... 22
    - 2.3.3 Event logging ..... 23
    - 2.3.4 Alarm management ..... 24
  - 2.4 Functions to be synchronized - a choice** ..... 25
    - 2.4.1 LDAP ..... 26
    - 2.4.2 IP-based access rights ..... 27
    - 2.4.3 NTP ..... 28
    - 2.4.4 Location ..... 29
    - 2.4.5 Add-ons ..... 30
  - 2.5 Dealing with an inhomogeneous Management Cluster - Notes** ..... 31
- 3 SU Cluster** ..... 33
  - 3.1 General information about SU Clusters** ..... 34
  - 3.2 Functionality of an SU Cluster** ..... 36
  - 3.3 SEM main window for SU Cluster** ..... 37
- 4 Live Migration (LM) for BS2000 systems** ..... 39
  - 4.1 Usage scenarios and preconditions for LM** ..... 40
  - 4.2 Operating the LM through the SE Manager** ..... 42
  - 4.3 Operating the LM through VM2000 commands** ..... 47
  - 4.4 LM procedure** ..... 50
  - 4.5 Effect of the LM within the migrated BS2000 system** ..... 52
  - 4.6 Messages during the LM** ..... 55
    - 4.6.1 Example with console messages on SU /390 ..... 56
    - 4.6.2 Example with console messages on SU x86 ..... 58
    - 4.6.3 Messages in the Event logging of the SE Manager ..... 59
  - 4.7 Rejection and abortion of an LM** ..... 60
  - 4.8 Further notes about LM** ..... 62

**5 Restoring VM definitions after an outage ..... 65**  
**6 Glossary ..... 67**  
**7 Related publications ..... 72**

# Cluster Solutions for SE Server - Whitepaper

# 1 Introduction; requirements for cluster functions

There are two types of clusters when it comes to SE Servers:

- Management Cluster for central operation and management
- SU Cluster for Live Migration support

Every cluster is fully set up by the customer service both on the hardware and software level depending on the requirements of the customer.

The customer must provide the following server-external resources if necessary:

- Physical cables for the network connection
- Disks in external storage systems

This chapter covers the central requirements for the cluster functionalities and gives an overview of the display in SE Manager (subsequently referred to as SEM):

- [Hardware structure](#)
- [Network connections](#)
- [CRD functionality; CRD with external configuration disks](#)
- [SEM displays and functions](#)

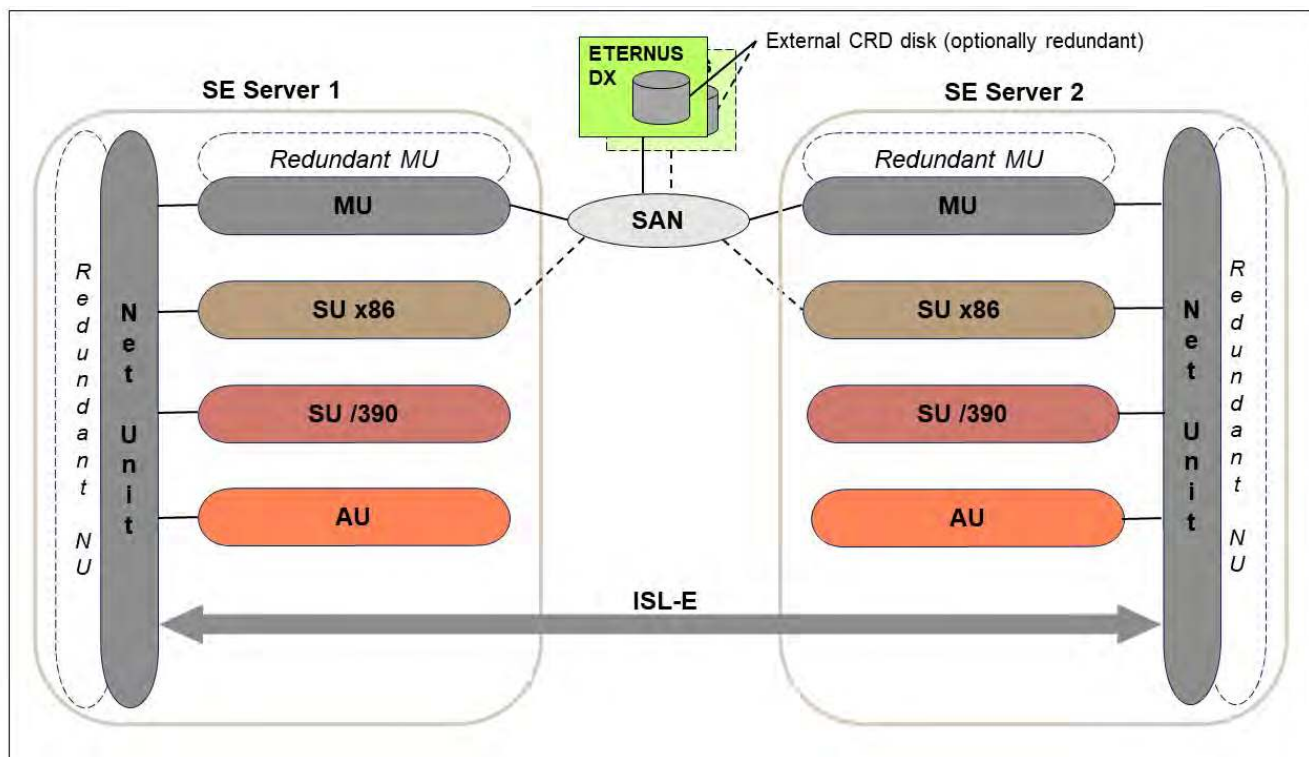
## Changes to the predecessor document:

- 10 Gbit/s ISL-E connections are supported via SFP modules in the NU.
- Updates notes for the administration of an inhomogeneous management cluster (V6.5/V6.4)  
In particular: It is ensured that in an inhomogeneous management cluster always one MU with V6.5 becomes the Master MU.
- New simplified authorization concept for BS2000 operators  
If BS2000 operators are to be granted authorization for live migration of a BS2000 system, they only need to be granted access authorization to this BS2000 system.

## 1.1 Hardware structure

The following image contains the basic HW structure as prerequisite for cluster functions:

- Two SE Servers
- ISL-E connections between the two SE Servers
- CRD disks (configuration disks) which are accessible through SAN from all units on which they are needed.



This is the basic HW structure for the following cluster functions:

- Management Cluster (SE Cluster)
- SU /390 Cluster (SU Cluster with SU /390)
- SU x86 Cluster (SU Cluster with SU x86)

**i Notice - Management Cluster with  $\geq 2$  SE servers on 2 locations**

- Also, more than two SE servers can be connected to a Management Cluster.  
The participating SE servers must be distributed over a maximum of two locations.
- The basic HW structure must be extended accordingly.  
As of three connected SE servers new questions regarding the optimal network connection arise, that must be solved individually.  
If needed, the customer service will support you doing this.
- In case of a larger Management Cluster, the menu of SEM is extended according to every SE server.  
Because there are no functional differences, the descriptions and examples in this document always assume a Management Cluster with two servers.

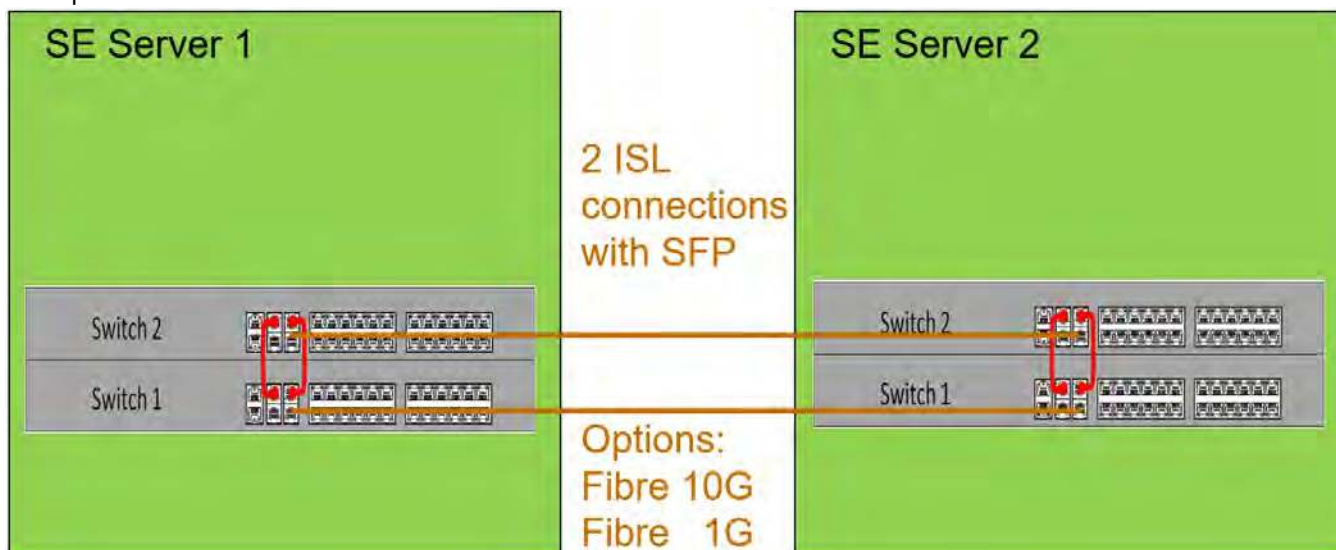
## 1.2 Network connections

The communication between MUs and SUs within the Management Cluster runs through the private (internal) management network -> MCNPR.

During Live Migration in the SU Cluster the data of the main memory is also transferred between the participating Server Units through this connection.

The Net Units (their internal LAN switches) of the connected SE Servers are linked with one another with ISL (ISL: Inter Switch Link) connections for this.

Principle structure for 2 SE servers:



The private management and data networks configured on the ISL connections are virtual LANs.

If an ISL connection is directed through the customer's infrastructure then the VLAN IDs 600..608 (for private management networks) and 800..899 (for private data networks) must be provided. Other connections are possible, but they must be verified in the course of a special release or be implemented as a customer project.

### i Recommendations and notes

- For the Management Cluster functionality a data transfer rate of 1 Gbit/s and a round trip time < 5 ms is recommended.
- Also a maximal distance of 15 km between the locations is recommended.
- Longer distances and lower data transfer rates are conceivable but the realization must be tested in practice in the customer's specific configuration.
- A configuration with a redundant NU and a redundant ISL-E connection is recommended so that a single error does not lead to a complete failure of the connection.
- Note:  
The cluster software does not check whether the above conditions are met. This has to be ensured by the customer through a suitable configuration of the infrastructure.
- As of three connected SE server new questions regarding the optimal network connection may arise, that must be solved individually.  
Please address the customer service for support.

Among others, the private management network serves the so-called -> [Master MU](#) to monitor units which are relevant for the cluster configuration. The Master MU regularly queries these units on availability and correct function by the so-called Network heartbeat.

Depending on the result of these queries a cluster status is calculated and displayed by the SE Manager in the cluster main windows (see the cluster main windows in the following chapters). The cluster status affects the availability of the cluster functions.

Further details about the networks of the SE Server can be found in the “Administration and Operation“ [1] manual.

## 1.3 CRD functionality; CRD with external configuration disks

Central data of the SE Server configuration is stored (irrespective of the existence of a cluster) on a so-called configuration disk (or CRD disk) of a unit (MU, SU x86, HNC):

- General data of the SE Server:
  - Model, name and location
  - Unit-overarching data
- Unit-specific data the content of which must be available also after an outage or after the unit (MU/SU /390, SU x86) has been turned off:
  - Model, SW version and hostname
  - IP configuration
  - FC configuration
  - VM data for BS2000
  - Storage of the last CSR backup
- Current configuration of the Net Unit switches

By default the above mentioned data are written locally on an internally mirrored disk of the MU/SU/HNC – the internal configuration disk (CRD disk).

### **CRD with external configuration disks**

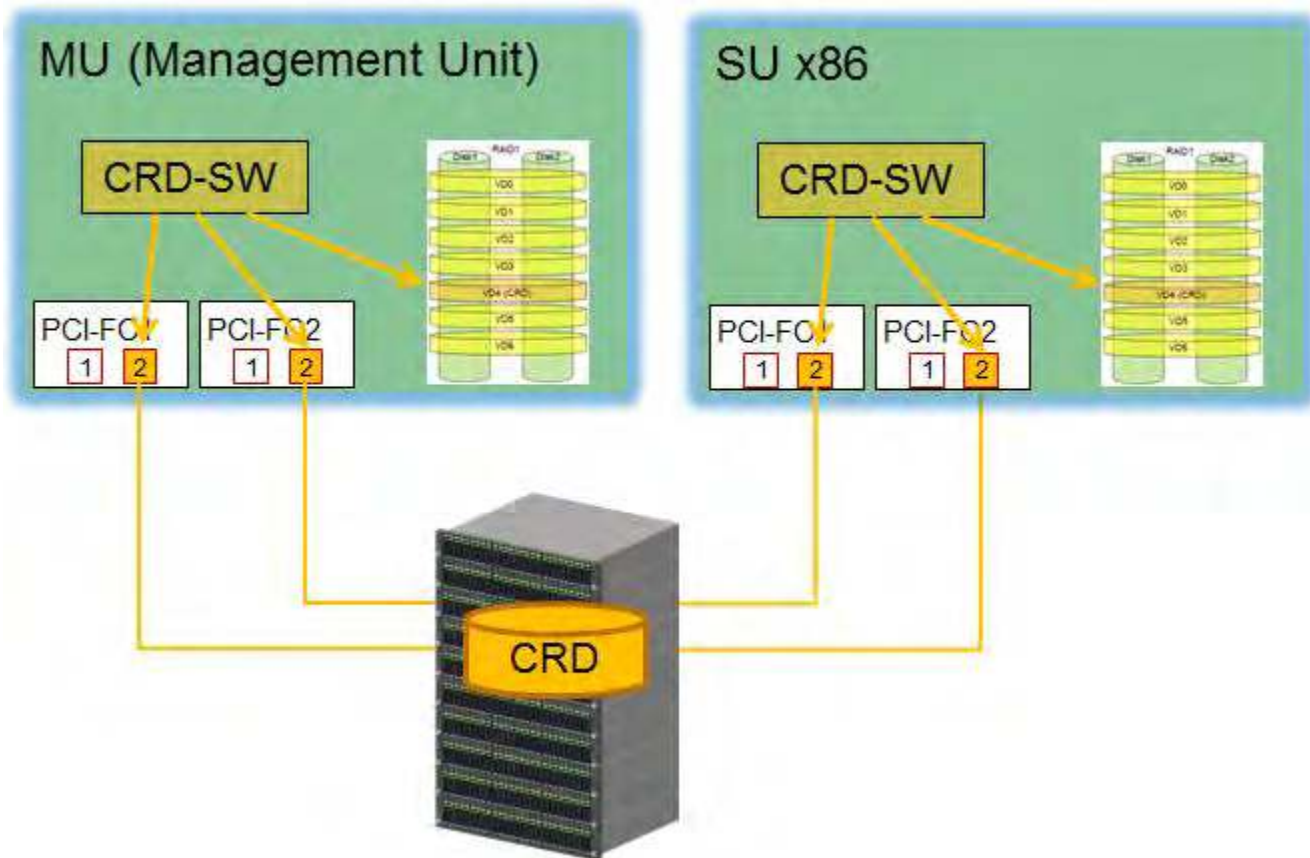
Additionally to the internal CRD disk, up to two external CRD disks can be configured situated on external FC RAID systems to which all MU and SU x86 have access through a redundant connection.

The internal and the external CRD disks are combined to a logical unit („logical disk“) via mirroring in SW-technical way and built up to the so-called "CRD" ("Configuration Raw Device") which provides the "CRD functionality" to all MU and SU x86.

The above mentioned central data of the SE Server configuration are „written on the CRD“ and read from there.

With the CRD functionality every MU and SU x86 has the same view on the data of the SE Server and the actions on these units can be done in coordinated way. (unit-overarching consistence)

Example with CRD principle:



External configuration disks (also called CRD disks) are required in the following cases:

MU redundancy for SE /390	-> mandatory <ul style="list-style-type: none"> <li>The external CRD disks must be configured on both MUs of the SE server.</li> </ul>
MU redundancy for SE x86	-> recommended <ul style="list-style-type: none"> <li>The absence leads to some restrictions (see "Administration and Operation" [1])</li> <li>The external CRD disks must be configured on both MUs of the SE server.</li> </ul>
Cluster	-> mandatory for every type of cluster <ul style="list-style-type: none"> <li>Management Cluster: The external CRD disks must be configured on all MUs of all SE servers.</li> <li>SU /390 Cluster: The external CRD disks must be configured on all MUs of all SE servers.</li> <li>SU x86 Cluster: The external CRD disks must be configured on the participating SU x86 and on all MUs of all SE servers.</li> </ul>

Technically, it is external disks with FC connection and a capacity of at least 8 GB that have to be used for external CRD disks.

The MU and the SU x86 access the external CRD disks directly through FC. SU /390 access the external configuration through the BS2000 subsystem REWAS and one MU (in normal operation through a server-local MU).

### **i Recommendations and notes**

- To avoid a „single point of failure" a configuration with two external CRD disks in different storage systems with redundant FC connections is recommended.
- Every external CRD disk must be accessible from every MU (X-cabling or SAN-X-configuration of the RAID periphery).
- In case of a SU x86 cluster the external CRD disks must be also accessible from every SU x86.
- A maximal distance of 15 km between the locations is recommended.
- Note:  
The cluster software does not check whether the above conditions are met. This has to be ensured by the customer through a suitable configuration of the infrastructure.

The CRD disks are provided by the customer and are set up by the customer's service.

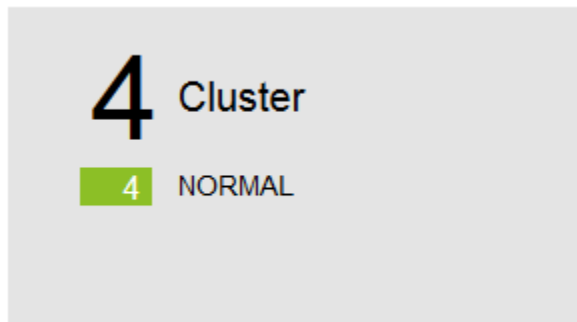
The unit specific CRD disks and their current status are displayed under

*Hardware -> Units -> <server> (SE<model>) -> <unit> (MU) -> Information tab CRD disks* in the SE Manager.

Their current status is displayed in the main window of the cluster for which it is relevant.

## 1.4 SEM displays and functions

The cluster states are summarized by the SEM in the dashboard tile „Cluster“:



In the list view of the dashboard the Management Cluster and the SU Cluster are differentiated:

▼	4	Cluster	4	NORMAL
	1	Management Cluster	1	NORMAL
	3	SU Cluster	3	NORMAL

The cluster overview main window is linked in the dashboard and contains the overview of all clusters and their states:

The screenshot shows the SEM interface. On the left is a sidebar menu with the following items: Dashboard, Systems, Applications, Performance, Devices, Hardware, Cluster, and Overview (selected). The main content area is titled 'Overview' and contains a 'Cluster overview' window. This window displays a table with the following data:

Name	Type	Function	Status
Management Cluster	SE	Management	✓ NORMAL
orngrwn	SU	Live migration	✓ NORMAL
redpuma	SU	Live migration	✓ NORMAL
SE1SE2SU390	SU	Live migration	✓ NORMAL

Below the table, it indicates 'Total: 4'.

The main windows for the Management Cluster contains only displays and is described in the chapter -> [SEM main window for Management Cluster](#). The central management (as a cluster function) is implemented in the whole SEM.

Aside from displays, the main window for the SU Cluster contains also the action to activate/deactivate the SU Cluster and is described in chapter -> [SEM main window for SU Cluster](#).

The SU Cluster function LM (Live Migration) can be found in the SEM main window "Operation" for the individual BS2000 systems and is described in chapter -> [Operating the LM through the SE Manager](#).

## 2 Management Cluster - central management of SE Servers

In the following chapter the most important subjects regarding the central management of SE Servers in the Management Cluster will be described:

- [General information about the Management Cluster](#)
- [SEM main window for Management Cluster](#)
- [Server-overarching functions \(a choice\)](#)
  - [Logging in, session and user management](#)
  - [Audit logging](#)
  - [Event logging](#)
  - [Alarm management](#)
- [Functions to be synchronized - a choice](#)
  - [LDAP](#)
  - [IP-based access rights](#)
  - [NTP](#)
  - [Location](#)
  - [Add-ons](#)
- [Dealing with an inhomogeneous Management Cluster - Notes](#)

## 2.1 General information about the Management Cluster

If two or more SE Servers are connected to one another, then this becomes a “Management Cluster” (also called “SE Cluster”).

A Management Cluster is configured by the customer service on customer's request and serves for both operating and managing of the participating servers.

Configuring a Management Cluster does not require any own maintenance windows or downtime.

There are several crucial requirements for the structure of a Management Cluster: a Net Unit connection between the participating SE Servers (ISL-E) must exist for communication purposes, as well as a so-called CRD with an external configuration disk (or two) must be accessible by all MUs for managing special global data.

In relation to administration and operation, all MUs of the Management Cluster are equal. This means that one can centrally administrate and operate all objects of the whole SE Server configuration on every MU.

The SE Servers remain operational as long as at least one MU is functional.

Exceptions:

- For SVP operation of an SU /390 and its correct HW display, a functional MU of the own SE Server is required.
- The operation of a BS2000 VM of an SU /390 (BS2000 console and BS2000 dialog) is possible only with a functional MU of the own SE Server.
- As the SKP functionality is always delivered by a server-local MU, there are still two MUs in an SE /390 required for the redundant SKP functionality.
- The complete administration of the BS2000 VMs is possible only with a functional MU of the own SE Server.

Note:

An inhomogeneous management cluster is a management cluster consisting of an SE server with all units (MU, HNC, SU x86) with one version (e.g. version V6.5) and another SE server with all units with another version (e.g. version V6.4).

In such a case certain restrictions have to be considered.

See for this chapter [Dealing with an inhomogeneous Management Cluster](#).

### Preparing measures in regard to global data

If the SE Servers that are already operational and have customer-specific configurations are to be connected to a Management Cluster, it needs to be taken into account that there are server-overarching “global data” on all participating SE Servers. These data can come into conflict during a merger.

As it is not easily determinable which objects of the global data would “win”, the following measures should be undertaken before connecting to the Management Cluster:

- All global objects must be made unambiguous before a Management Cluster is set up.
- Unambiguousness can mean:
  - Sameness:
    - The objects must be configured in the same way on all servers (including old attributes).
    - Example: The password for the account “admin” is set on all servers to the same value.

- Unique existence:
  - Unwanted objects are deleted.
  - Example: The mail configuration for the Alarm Management is deleted on all servers except one.
- The following global objects must be taken into account for this:
  - Account (local and LDAP) as well as its properties: passwords and password properties (only for local accounts), multi-factor authentication (MFA), individual rights, descriptions
  - Alarm Management with mail configuration, mail recipient and trap recipient
  - Configuration der private IP networks DANPRxx and MONPRxx
  - Configuration of the FC networks – the FC switches and the properties for monitoring
  - SE operation state
  - Settings for the remote service access
  - If given, server-internal SU x86 cluster

**Role concept:**

- The scope of every account (regardless of its role) is always the whole Management Cluster or all objects of the Management Cluster relevant for the role.  
Limitation to one SE Server is not possible.
- The main windows Overview and Management Cluster in the category “Cluster” can be accessed in SE Manager by all users with the roles Administrator, BS2000 administrator and BS2000 operator.

## 2.2 SEM main window for Management Cluster

The overview in the first section shows central status values of the Management Cluster:

- Status summary (see chapter -> [Network connections](#))
- IP Networks ISL-E (see -> [ISL-E](#))
- Master MU: name of the Management Unit which acts as -> [Master MU](#) currently in the Management Cluster, and the status of the internal resources of the Master MU.

The following table informs about all Management Units of the Management Cluster displaying these data:

- Name of MU
- Name of SE Server to which the MU belongs
- Power Status of the MU (see -> [Power Status](#))
- Status of the CRD disks of the MU (see chapter -> [CRD Functionality; CRD with external configuration disks](#))
- Status of the network heartbeat of the MU (see chapter -> [Network connections](#))

Example:

The screenshot shows the 'Management cluster' overview in a web interface. On the left is a navigation menu with 'Management cluster' selected. The main content area has two sections:

**Management cluster summary**

Status summary		✓ NORMAL
IP networks ISL-E		✓ NORMAL
Master MU	abgse4mu2-1	✓ NORMAL

**Management cluster overview**

Management Unit	Server	Power status	CRD disks	Network heartbeat
<i>Filter</i>	<i>Filter</i>	<i>All</i>	<i>All</i>	<i>All</i>
abgse2mu1	SE-Servre-2	▶ ON	✓ NORMAL	✓ NORMAL
abgse2mu2	SE-Server-2	▶ ON	✓ NORMAL	✓ NORMAL
abgse4mu1-1	SE-Server-4	▶ ON	✓ NORMAL	✓ NORMAL
abgse4mu2-1	SE-Server-4	▶ ON	✓ NORMAL	✓ NORMAL

Total: 4

## 2.3 Server-overarching functions (a choice)

In this chapter a choice of most important server-overarching functions of the SE Manager is described shortly in regard to the Management Cluster:

- [Logging in, session and user management](#)
- [Audit logging](#)
- [Event logging](#)
- [Alarm management](#)

Note:

The messages in the Dashboard as well as in the global overviews through systems and units are both server-overarching.

Example of a global overview through the systems:

Overview Pending BS2000 messages C2H data

Systems

Per page 32

Name	Type	Operating system	Server	Unit	Description	Status
MONI	All	Filter	Filter	Filter	Filter	All
MONISE2	VM2000	BS2000 V21.0A	SE-Server-2	EM2	ABGSE211; Home Pubset E22N	▶ RUNNING
MONITOR	VM2000	BS2000 V21.0A	SE-Server-2	su300se2	ABGSE201; Home Pubset E21N	▶ RUNNING
MONITOR	VM2000	BS2000 V21.0A	SE-Server-4	EM1	ABGSE1BS; Home Pubset VP10	▶ RUNNING
MONITOR	VM2000	BS2000 V21.0A	SE-Server-4	su310se4	ABGSE401; Home Pubset E24N	▶ RUNNING
MONITOR	VM2000	BS2000 V21.0A	SE-Server-8	su310se8	ABGSE401; Home Pubset E24N	▶ RUNNING

Total: 5 of 217

Example of a global overview through the units:

Units Update overview

Units

Name	HW model	Chassis	Server	Power status	System status	HW status
Filter	*SUJMUJHNC	Filter	All	All	RUNNING	NORMAL
su390-se-2	SU700	-	SE-Server-2	▶ ON	▶ RUNNING	✔ NORMAL
abgse2mu2	MU M1	-	SE-Server-2	▶ ON	▶ RUNNING	✔ NORMAL
hnc2-se2	HNC M2	-	SE-Server-2	▶ ON	▶ RUNNING	✔ NORMAL
su390-se-4	SU710	-	SE-Server-4	▶ ON	▶ RUNNING	✔ NORMAL
abgse4mu1-1	MU M3	-	SE-Server-4	▶ ON	▶ RUNNING	✔ NORMAL
abgse4mu2-1	MU M2	-	SE-Server-4	▶ ON	▶ RUNNING	✔ NORMAL
hnc2-se4	HNC M3	-	SE-Server-4	▶ ON	▶ RUNNING	✔ NORMAL
abgse8mu1	MU M3	-	SE-Server-8	▶ ON	▶ RUNNING	✔ NORMAL

## 2.3.1 Logging in, session and user management

### Single sign-on

Users have to provide their username and password only once on any MU and receive access according to their authorization to all functions of all SE Servers, including -> [add-ons](#).

In the case of server-specific LDAP configurations, the login with an LDAP account has to be executed on a MU of the suitable server.

### Global session

The session is global or over-arching. The MU can be changed without further authentication in a server-overarching way, for example when starting add-ons.

Requirement: DNS server and LDAP (if necessary) must be available for all MUs.

Example:

Displaying a global session (see the first column) for the (own, highlighted in blue) account *johndoe*



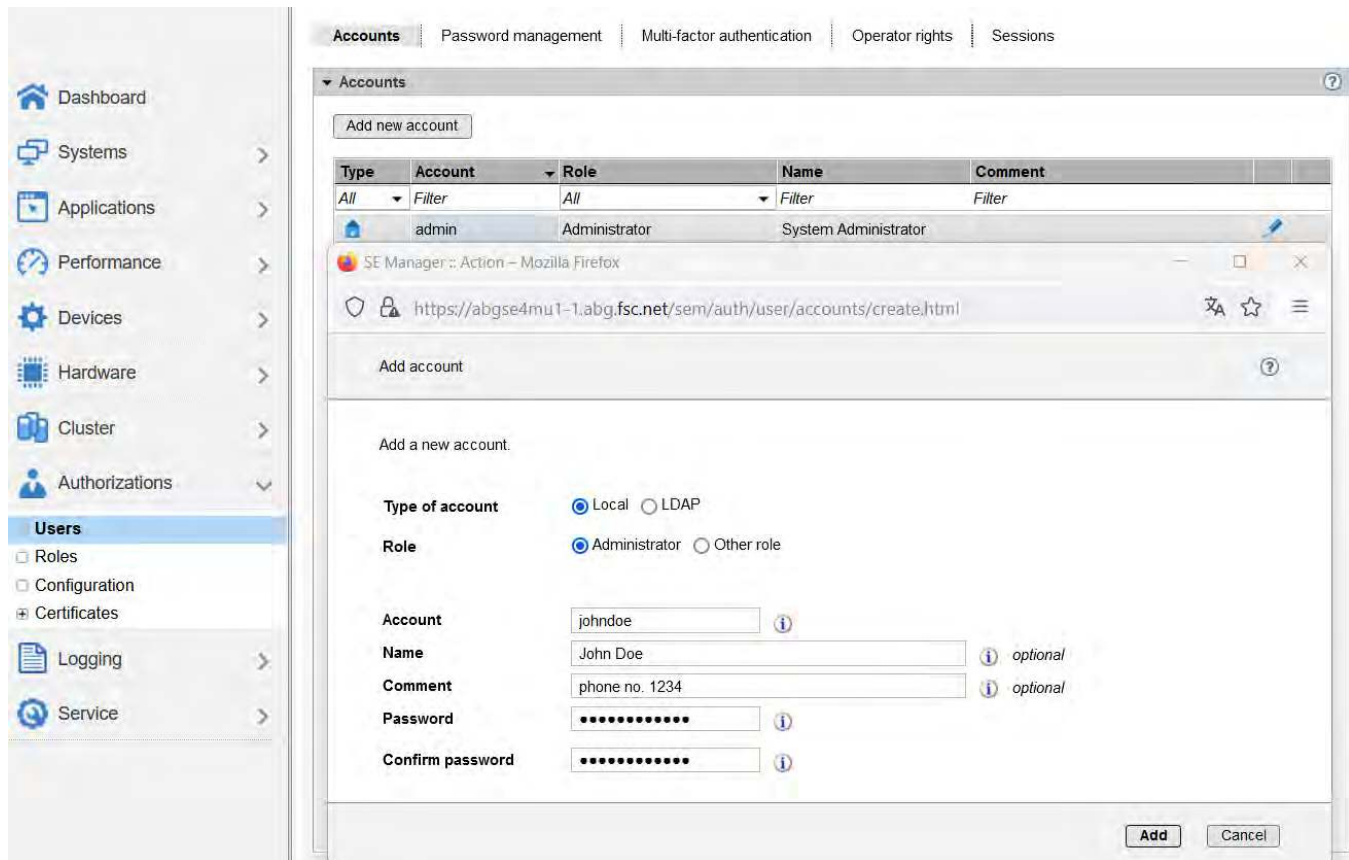
Management Unit	Account	Name	Role	IP address	Language	Autom. upd.	Timeout	Starting time
Filter	john	Filter	All	Filter	All	Filter	Filter	Filter
-(global)	johndoe	John Doe	Administrator		English	-	20 min	2023-11-03 11:44:26

### Global management of accounts

All accounts – local and LDAP – are valid for the whole Management Cluster. They are ready on every MU of every SE Server in the same measure.

Example:

The account *johndoe* is (centrally) created and managed.



The passwords (for local accounts) and other properties, the multi-factor authentication (MFA) as well as individual rights (for BS2000 operators) are also managed centrally.

## 2.3.2 Audit logging

Audit logging protocols every action that is performed on a unit (MU, SU, HNC) of the Management Cluster through the GUI (SE Manager and add-ons) or through a CLI command. This way, an administrator can always check who has performed which action on which unit with what success.

In the following example we can see the following for the ID *john DOE* (top-down):

- Logging in to SE Manager
- Entering in Alarm management (see also chapter -> [Alarm management](#))
- Change of the session timeout

The screenshot shows the 'Audit logging' window in the SE Manager. The left sidebar contains navigation options: Systems, Applications, Performance, Devices, Hardware, Cluster, Authorizations, Logging, Audit logging (selected), Event logging, Alarm management, and Service. The main window displays 'Audit logging entries' with a search period set to '2023-03-16 00:00:00 -'. The 'Oldest entry from' is '2022-04-14 13:02:48' and 'Per page' is '32'. A table lists the following entries:

Date	Unit	Account	Component	Type	Message
2023-03-16 19:36:57	bern	john DOE	SEM	OK	Action=Add mail receiver; Mail receiver=johndoe@fujitsu.com; Component=ANY; Weight=ANY -> Mail receiver johndoe@fujitsu.com for component ANY successfully added.
2023-03-16 19:36:57	bern	john DOE	SEM	Start	Action=Add mail receiver; Mail receiver=johndoe@fujitsu.com; Component=ANY; Weight=ANY
2023-03-16 19:36:21	bern	john DOE	SEM	OK	Action=Change individual settings; Type=local; Account=johndoe (John Doe); Update cycle=0; Session timeout=0; Menu=Standard; -> The individual settings for local user johndoe (John Doe) have been changed successfully.
2023-03-16 19:36:20	bern	john DOE	SEM	Start	Action=Change individual settings; Type=local; Account=johndoe (John Doe); Update cycle=0; Session timeout=0; Menu=Standard;
2023-03-16 19:36:01	bern	john DOE	SEM	Login successful	Login to SE Manager successful; Account=johndoe; IP address=172.26.205.133

Total: 7 from 45408

Further details can be found in the manual “Administration and Operation“ [1] and in the online help of the SE Manager.

## 2.3.3 Event logging

In event logging, events that occur in the whole Management Cluster are protocolled and shown to the administrator.

Example: The status changes of systems (for example the start of a BS2000 VM), starting a unit (SU, MU, HNC), the status changes of clusters (Management and SU Clusters).

All events in event logging are prioritized according to their meaning.

After assessing and processing (if needed), the current events should be acknowledged. This will cause them to disappear from the display of current events.

Hint: the current events cannot be acknowledged individually. They must be acknowledged in total.

The number of unacknowledged events as well as their priority is displayed in the Dashboard in their own tile.

In the following example events can be seen in context (bottom-up):

- The MU abgwhite is powered off and reports itself as deactivated (component M2000)
- As the MU abgwhite does not answer to heartbeat queries, it is reported with the priority WARNING (component Cluster)
- As the functionality of the Management Cluster is affected by powering down the MU abgwhite (loss of redundancy), it is reported with the priority WARNING (component Cluster).  
The status of the Management Cluster changes to WARNING as well.
- As the functionality of the SU /390 Cluster is affected by powering off the MU, it is reported with priority WARNING (component Cluster).  
The status of the SU /390 Cluster also changes itself to WARNING.

The screenshot shows the SE Manager interface with the following details:

- Header:** SE Manager, John Doe, Log out, FUJITSU
- Navigation:** Dashboard, Systems, Applications, Performance, Devices, Hardware, Cluster, Authorizations, Logging, Audit logging, **Event logging**, Alarm management
- Current events:** All events
- Event Log Table:**

Timestamp	Weight	Unit	Component	Message
2017-04-26 07:51:55	WARNING	abgblue	Cluster	Cluster state of 'SU390' changed from 'NORMAL' to 'WARNING'
2017-04-26 07:51:55	WARNING	abgblue	Cluster	Cluster state of 'Management Cluster' changed from 'NORMAL' to 'WARNING'
2017-04-26 07:51:55	WARNING	abgblue	Cluster	Network heartbeat of unit 'abgwhite' changed from 'NORMAL' to 'NOT_ACCESSIBLE'
2017-04-26 07:45:27	NOTICE	abgwhite	M2000	M2000 deactivated
- Additional Info:** From: 2017-04-26 07:45:27 To: 2017-04-26 07:51:55, Oldest entry from: 2016-06-17 06:54:37, Total: 4

Further details can be found in the manual “Administration and Operation“ [1] and in the online help of the SE Manager.

The list of possible events can be found in an attachment (see “General information”) of the online help of the SE Manager.

### 2.3.4 Alarm management

The administrator can configure automatic messages per mail or SNMP trap for all events of a certain priority.

Example:

- SNMP traps are sent for all events (component ANY and weight ANY) to a management station.
- The selected user receives mails for every event of the Sys-Mgmt component (system management of the SUs) with the weight >=WARNING.

**Alarm management**

SNMP trap receivers

Add new trap receiver

Trap receiver	Trap community	SNMP version	Component	Weight			
172.17.0.1	icinga	SNMPv2c	ANY	ANY			

Total: 1

Mail configuration

Create mail configuration

SMTP server	Return address		
fujitsu.com	se-alarm-mgmt@reply.no		

Mail receivers

Add new mail receiver

Mail receivers	Component	Weight			
john	All	All			
john.doe@mycompany.com	Sys-Mgmt	>= WARNING			

Total: 1 of 5

Further details can be found in the manual “Administration and Operation“ [1] and in the online help of the SE Manager.

## 2.4 Functions to be synchronized - a choice

In this chapter the most important functions, which must be synchronized between the servers or the MUs, are described briefly:

- [LDAP](#)
- [IP-based access rights](#)
- [NTP](#)
- [Location](#)
- [Add-ons](#)

## 2.4.1 LDAP

The LDAP configuration is server-specific.

Different LDAP configurations can be used in different locations, if needed.

The superordinate “Change LDAP configuration” dialog box (button above the table) offers also the convenience function that allows to manage all SE Servers together and to supply identical configuration data.

The LDAP configurations can be set differently with the server-specific dialog box (pencil icon).

Example with two servers with the same LDAP configurations:

The screenshot displays the LDAP configuration interface. The sidebar on the left contains the following menu items: Dashboard, Systems, Applications, Devices, Hardware, Cluster, Authorizations, Users, Roles, Configuration, Certificates, Logging, and Service. The main content area is titled "LDAP IP-based access rights" and contains three sections:

- All servers: LDAP configuration**: Contains a "Change LDAP configuration" button.
- Server SE1-Schweiz: LDAP configuration**: Contains a table with the following configuration details:
 

IP address / host name	g02.fujitsu.local
Communication	Secured with TLS
Port	389
Server type	ActiveDirectory
Domain	g02.fujitsu.local
Base DN	DC=g02,DC=fujitsu,DC=local
Account	EPSIdap
Status	Active
- Server SE4-Baden-Wuerttemberg: LDAP configuration**: Contains a table with the following configuration details:
 

IP address / host name	g02.fujitsu.local
Communication	Secured with TLS
Port	389
Server type	ActiveDirectory
Domain	g02.fujitsu.local
Base DN	DC=g02,DC=fujitsu,DC=local
Account	EPSIdap
Status	Active

Further details can be found in the manual “Administration and Operation“ [1] and in the online help of the SE Manager.

## 2.4.2 IP-based access rights

The IP-based access rights are specific to a server.

Different IP-based access rights can be used in different locations if needed.

The superordinate “Allow IP address” dialog box offers also the convenience function that allows to add entries for all SE Servers at the same time and to manage all SE Servers together this way. Also the other superordinate dialogs can be executed server-specific or for all servers together.

Example:

The screenshot displays the 'IP-based access rights' configuration page in the SE Manager. The left sidebar contains navigation links for various system components. The main area shows a table of existing IP-based access rights and a form to add new ones. The 'Allowed IP addresses' table is as follows:

Server	IP address	Description	Status
se2	Filter	Filter	All
SE2-Biel	10.172.102.17/16	Office network	Active
SE2-Biel	172.17.20.8	nera	Active
SE2-Biel	172.26.0.0/16	Developers and friends	Active

Below the table, the 'Allow new IP address' form is visible. The 'Server' dropdown menu is open, showing the following options:

- All
- SE2-Biel
- SE3-Wald
- SE5-Suisse
- SE6-Svizzera

Further details can be found in the manual “Administration and Operation“ [1] and in the online help of the SE Manager.

### 2.4.3 NTP

When using external time servers (NTP servers) it must be taken into account that they have to be configured in all MUs.

This must be done in SE Manager under

*Hardware -> Units -> <server> (SE<model>) -> <unit> (MU) -> Management, tab System time.*

In case of different locations, the NTP servers in use can be different, but for the MUs of one SE Server they should always be identical.

Further details can be found in the manual "Administration and Operation" [1] and in the online help of the SE Manager.

## 2.4.4 Location

The location is set up by the service engineer when setting up the Management Cluster and is displayed in the main window

*Information (Hardware -> Units -> <server> (SE<model>), tab Information), on the one hand.*

On the other hand, the user can enter the so-called SYSLOCATION in the main window *SNMP* for every MU.

*(Hardware -> Units -> <server> (SE<model>) -> <unit> (MU) -> Management, tab SNMP)*

It is displayed in the SEM header next to the MU name.

Example:

The screenshot shows the SE Manager interface for a Management Unit named 'abgsilver'. The breadcrumb trail at the top indicates the path: Management Unit (abgsilver) > Location 1. The left sidebar shows a tree view of units, with 'Management' selected. The main content area is divided into three sections:

- Management Unit abgsilver: Configuration of local system data**: Shows 'SYSLOCATION' set to 'Location 1' and 'SYSCONTACT' as an empty field.
- Management Unit abgsilver: Allowed read accesses**: Shows a table with one entry: 'Read community' is 'icinga' and 'Restricted to' is empty. Total: 1.
- Management Unit abgsilver: Trap receiver**: Shows a table with two entries:
 

Trap receiver	Trap community	SNMP version
Filter	Filter	All
	icinga	SNMPv1
	icinga	SNMPv2c

 Total: 2.

### **i** Recommendation

It should be assured that all three entries (1 SE Server + 2 MU) are the same for the one SE server resp. for the whole location.

Further details can be found in the manual "Administration and Operation" [1] and in the online help of the SE Manager.

## 2.4.5 Add-ons

Every add-on can be installed in principle on every MU of the Management Cluster.

All installed add-ons with a GUI are linked in SE Manager and can be called directly and without further logins from there, if the current role allows this.

Further details can be found in the manual “Administration and Operation“ [1].

### **i** Configuration and usage of add-ons installed multiple times

Some add-ons must be installed on all MUs of the SE configuration to ensure correct functioning, in the case of other add-ons this is optional.

Please refer to the respective add-on-specific documentation for the recommended configurations and scenarios.

## 2.5 Dealing with an inhomogeneous Management Cluster - Notes

An inhomogeneous Management Cluster is a Management Cluster consisting of servers with different SW versions, e.g. of an SE server with all units (MU, HNC, SU x86) with version V6.5 and another SE server with all units (MU, HNC, SU x86) with version V6.4. This situation can occur if the servers of a Management Cluster are not upgraded at the same time but in steps.

This chapter deals with specific topics and limitations that must be considered in an inhomogeneous management cluster during normal operation.

### General

In general, you should log in to a V6.5 MU to use it to manage V6.5 units, while logging in to a V6.4 MU to use it to manage V6.4 units.

### Roles and accounts

The role concept in V6.5 is not compatible with that in V6.4.

Therefore the following applies:

- V6.4 accounts should only be used on V6.4 MUs, V6.5 accounts only on V6.5 MUs.
- The service will advise you and help you find the right approach when moving to V6.5, so that you can use the accounts you want in the homogeneous V6.5 end state.

### SE operation state

The SE operation state (including the maintenance state) can be edited only in V6.5 and it is not visible in V6.4. The server information (predecessor of the SE operation state) can be edited only in V6.4 and it is not visible in V6.5.

### New menu structure and the main windows

In V6.5 the structure of the Units and Service menus has been changed:

The submenus Hardware -> <MU|HNC|SUx86> -> Service (Update | CSR | Diagnostics | Remote Service) have been moved from the Hardware menu to the Service menu.

When changing the MU in SEM (e.g. using the links in the header) from a MU with V6.5 to one with V6.4, with one of the above main windows open in SEM, this is not found and the browser displays "Not Found / The requested URL ... was not found on this server."

With an analogous change from a MU with V6.4 to one with V6.5, the Service menu is changed correctly, whereas all calls of the FC menu are redirected to the Overview window.

### Hardware properties

Call the "Hardware details for unit" dialog for a unit from a MU of the same server, otherwise it will fail.

Call the "Change scheduled power on/off" dialog for a unit from a MU of the same server, otherwise it will fail.

## **Add-ons**

Version dependency

With V6.5 the add-ons had to adapt to this new version.

Please refer to the respective add-on-specific documentation for the recommended versions and recommended procedures.

## 3 SU Cluster

An SU Cluster consists of two Server Units of the same type (SU /390 or SU x86) and offers Live Migration (LM) of BS2000 systems as a cluster function.

To differentiate (in description, not in SE Manager), it is spoken also of SU /390 Clusters and SU x86 Clusters.

An SU Cluster can be configured by the customer service if requested.

The configuration of an SU Cluster doesn't require an own inspection window or downtime.

In this chapter the following subjects are discussed in regard to SU Clusters:

- [General information about SU Clusters](#)
- [Functionality of an SU Cluster](#)
- [SEM main window for SU Cluster](#)

The Live Migration is described in detail in the following chapter -> [Live Migration \(LM\) for BS2000 systems](#).

## 3.1 General information about SU Clusters

### Prerequisites

- An SU Cluster between two Server Units in different SE Servers can be set up only after a Management Cluster is configured.
- An SU Cluster between two SU x86 in an SE Server does not require a Management Cluster. However, external configuration disks (CRD) are required, which must also be configured on the Management Units of the server.
- A SAN-X configuration of the RAID peripheral is needed for SU Clusters. (Every external configuration disk must be reachable from every unit participating in the SU Cluster.)
- The same data and administration networks (DANPU and MANPU/MONPU) must be available on the participating SE Servers for an SU Cluster (or for the LM functionality) based on a Management Cluster. The network communication between the participating SE Servers must be realized in the same network segment or without routers/gateways.  
Only in such a way it can be assured that the network connections of the BS2000 systems still function correctly after the LM.  
Further details about the networks of the SE Server can be found in the manual “Administration and Operation“ [\[1\]](#).

### **i** Recommendations and notes

- After a Live Migration the migrated BS2000 system accesses the storage systems from the remote location.  
Thus for performance reasons a maximal distance of 15 km between the locations is recommended.
- Longer distances and lower data transfer rates are conceivable but the realization must be tested in practice in the customer's specific configuration.
- Note:  
The cluster software does not check whether the above conditions are met. This has to be ensured by the customer through a suitable configuration of the infrastructure.

### Configuration

- The configuration of the SU Cluster consists of a cluster name and a selection of both Server Units between which the Live Migration function should be made possible.
- Multiple SU Clusters can be configured within a Management Cluster. However, an SU can belong to only one SU Cluster.

### **SU Cluster in deactivated state**

An SU Cluster can be deactivated (status summary INACTIVE). No LM can be performed in the deactivated state. We recommend deactivating an SU Cluster in maintenance situations, for example, to avoid sending unnecessary alarms.

## 3.2 Functionality of an SU Cluster

### Live Migration of BS000 systems

Live Migration (abbrev. LM) of BS2000 systems is the central function of an SU Cluster.

LM is not only understood as relocating a BS2000 system in a state RUNNING, but also in the states DOWN and INIT\_ONLY in which resources are also allocated.

### Migrating of (deactivated) VM definitions

Deactivated VM definitions (VMs in the state DEFINED\_ONLY) can be relocated in the same way as in case of LM of an SU.

As the resources are not allocated, certain constraints are omitted.

However, as the process is essentially analogous to LM, both this document and SE Manager also use the term LM in this case. If it's necessary the (small) differences are indicated.

### Restoring (recovering) (deactivated) VM definitions after an outage

Restoring (recovering) (deactivated) VM definitions makes a simple transfer of BS2000 systems from an SU to a different SU in an SU Cluster after an outage (or power down) of the SU or of the whole SE Server possible. The command /RECOVER-VM-DEFINITION can be used for this in every SU cluster (SU /390 and SU x68).

The LM is described in detail in chapter -> [Live Migration \(LM\) for BS2000 systems](#).

Restoring deactivated VM definitions after an outage is described in detail in the following chapter -> [Restoring VM definitions](#).

### Role Concept

- The SU Cluster main windows are visible for all users except add-on administrators.
- Activating/deactivating of SU Clusters is reserved for administrators.
- LM is not only allowed for administrators but also for BS2000 administrators and BS2000 operators (if their individual access rights regarding the BS2000 systems to be migrated are set).

### 3.3 SEM main window for SU Cluster

The SEM main window SU Cluster shows the summary status of the SU Cluster in the first field.

The “Change the status of the SU Cluster” dialog box allows to activate/deactivate the SU Cluster.

The second field contains a table with the SUs and their values that are relevant for the cluster functionality.

Note regarding the names of the SUs:

- The names of the SUs of type SU /390 are freely configurable in SEM and are used in all SEM displays and actions.  
In VM2000, these names are not known; instead, the BCAM host names of the monitor systems are used.
- The names of SUs of type SU x86 are the X2000 hostnames.  
These names are used both in SEM and in VM2000 for displays and actions.

Example of an SU x86 Cluster within an SE Server and the dialog for deactivating the SU Cluster:

The screenshot shows the SE Manager interface with a sidebar on the left containing navigation options like Dashboard, Systems, Applications, Performance, Devices, Hardware, Cluster, Overview, Management cluster, orngbrwn, Authorizations, Logging, and Service. The main content area displays the 'SU cluster orngbrwn: Status' window with a dropdown menu open, and the 'SU cluster orngbrwn: Units' table below it. A dialog box titled 'SE Manager :: Action - Mozilla Firefox' is overlaid on top, asking for confirmation to deactivate the SU cluster 'orngrbrwn'.

**SU cluster orngbrwn: Status**

Cluster name	orngrbrwn
Function	Live migration
Status summary	✓ NORMAL

**SU cluster orngbrwn: Units**

Unit	Model	Server	Operation mode	Power status	CRD disks	Network heartbeat
abgorange	RX300	SE-Server-2	✓ VM2000	▶ ON	✓ NORMAL	✓ NORMAL
abgbrown	RX300	SE-Server-2	✓ VM2000	▶ ON	✓ NORMAL	✓ NORMAL

**SE Manager :: Action - Mozilla Firefox**

Change status of SU cluster

Do you really want to deactivate SU cluster orngbrwn?

⚠ After deactivating the SU cluster no live migration can be processed.

Deactivate Cancel

Example of an SU /390 Cluster:

In case of SU /390, the MUs and their resources are also important for the cluster functionality and are displayed because of this.

The screenshot shows the SE Manager interface displaying the 'SU cluster SU390-SE2-SE4: Status' window and the 'SU cluster SU390-SE2-SE4: Units' table. The status window shows the cluster name 'SU390-SE2-SE4', function 'Live migration', and status 'NORMAL'. The units table lists two main units, SU390-SE2 and SU390-SE4, each with multiple sub-units (MUs) and their respective resources.

**SU cluster SU390-SE2-SE4: Status**

Cluster name	SU390-SE2-SE4
Function	Live migration
Status summary	✓ NORMAL

**SU cluster SU390-SE2-SE4: Units**

Unit	Model	Server	Operation mode	Power status	CRD disks	Network heartbeat
SU390-SE2	SU700	SRV-SE2	✓ VM2000	▶ ON	✓ NORMAL	✓ NORMAL
abgse2mu2				▶ ON	✓ NORMAL	✓ NORMAL
abgse2mu1				▶ ON	✓ NORMAL	✓ NORMAL
SU390-SE4	SU710	SRV-SE4	✓ VM2000	▶ ON	✓ NORMAL	✓ NORMAL
abgse4mu2-1				▶ ON	✓ NORMAL	✓ NORMAL
abgse4mu1-1				▶ ON	✓ NORMAL	✓ NORMAL

Total: 2

## 4 Live Migration (LM) for BS2000 systems

The details of Live Migration (LM) for BS2000 systems are described in this chapter in the following order:

- Usage scenarios and preconditions for LM
- Operating the LM through the SE Manager
- Operating the LM through VM2000 commands
- LM procedure
- Effect of the LM within the migrated BS2000 system
- Messages during the LM
  - Example with console messages on SU /390
  - Example with console messages on SU x86
  - Messages in the Event logging of the SE Manager
- Rejection and abortion of an LM
- Further notes about LM

## 4.1 Usage scenarios and preconditions for LM

During Live Migration (LM) a BS2000 system is moved from a Server Unit (source SU) to a different Server Unit with the same architecture (target SU), without affecting the availability of the BS2000 and the running applications for the user.

The operating system and the running applications continue running during relocation, which means that the user will not notice the relocation.

LM is possible both for SUs with /390 architecture and VM2000 as well as for SUs x86 with Native BS2000 or with VM2000.

### Usage scenarios for Live Migration

- Before planned maintenance, upgrade, reconfiguration and hardware or firmware update. The running BS2000 system is migrated to a different SU before the pending date.
- To avoid or solve a resource bottleneck (workload management)
- Reversion: The running BS2000 system can later be relocated back again without interruption

#### **i** Recommendation

It is recommended that the LM is done during a low-peak time.

For reasons see the following chapter with the frame conditions and details for the sequence of events of an LM.

### Prerequisites for Live Migration

- SU Cluster
  - The source SU and the target SU must belong to the same active SU Cluster.
  - For an LM, two servers of the same type (with the same architecture) are needed. Mixed configurations within the /390 family are allowed, for example SU320 with SU330.
  - If the Server Units are on different SE Servers then these SE Servers must be connected to a Management Cluster (see chapter -> [Management Cluster - central management of SE Servers](#) ).
- Name resolution on source and target SU

To allow internal name resolution in the SENET on source and target SUs, at least one MU of the respective SE server must be available. If this is not always the case, it is recommended in a cluster to extend the DNS name server configuration of the SUs by the SENET address of an MU of the other SE server; this makes live migration possible even if no MU of your own SE server is available.
- BS2000 operating mode

The source SU and target SU must run in the same BS2000 operating mode (LM of native BS2000 is only possible for SU x86).
- CPU

The physical CPU expansion of the source and target SUs can be different. The number of active CPUs of the migrated BS2000 system is automatically adjusted for an LM if necessary (see also "General" and "Notes on VM2000 Operation" in chapter -> [Further notes about LM](#)).
- Main memory

The physical main memory expansion of the source and target SUs can be different. There must be enough main memory on the target SU to take over the BS2000 system that needs to be migrated – at least as much as the BS2000 system has allocated on the source SU.

- VM name  
The VM name is not used on the target SU yet.
- VM index  
The VM index of the guest system that needs to be transferred must be free during VM2000 operation, when the VM was created with a permanent VM index.  
If the VM was created with VM-INDEX = \*ANY then it's enough that a random VM index is free on the target SU.
- Maximal number of VM definitions  
The maximal number of VM definitions was not yet reached on the target SU.
- Network connections  
The network connections configured in BS2000 must be implemented on the participating SE Servers in the same network segment. Only in this way can it be ensured that the network connections of the BS2000 systems work correctly after LM.  
Note: This condition is not checked before the LM is carried out.
- BS2000 devices/ I/O configuration  
The configuration of the peripherals used in BS2000 (IORSF/X2000) must be compatible for source SU and target SU, i.e. the BS2000 MNs must be configured the same on both server units (same host connector and unit ID) and assigned to the same devices. (see also the notes on I/O configuration in chapter -> [Further notes about LM](#) ).
- Availability of the BS2000 devices  
All of the devices assigned to the guest system that is to be relocated must be available on the target SU.
- Global VM2000 resources (CPU pools, VM groups, assignment sets)  
If the VM2000 guest system is assigned to a VM group or to a specific CPU pool, then they must also exist on the target SU (the CPU pool must have at least one CPU attached).
- Status of the BS2000 system  
The status of the guest system that needs to be migrated must be RUNNING, INIT\_ONLY, DOWN or DEFINED\_ONLY (i.e. the status cannot be WAIT).  
The status is not changed by the LM.
- Special case Native BS2000 for SU x86  
The target SU must also be in the BS2000 operation mode Native BS2000 and there must not exist any BS2000 there. If necessary, it must be deleted in the SE Manager.
- Subsystem REWAS  
The subsystem REWAS (at least V2.0C) must be active in both monitor systems of the SU /390 Cluster.  
However, for a full functionality, it is recommended that the subsystem REWAS is active in every BS2000 system of a SE Server.
- Product version  
The product versions of VM2000, BS2000 and ONETSERV/BCAM necessary for the LM are listed in the release documentation.

## 4.2 Operating the LM through the SE Manager

The function Live Migration as well as checking the configuration for an LM are called in the main window “Operation” of the system to be migrated.

### LM of Native BS2000 (only SU x86)

The main window “Operation” is called in Native BS2000 mode in the following way:

*Systems -> <SE name> (SE...) -> <SU name> (SU...) -> BS2000 -> Operation.*

To check the preconditions or/and to start the LM, the action “Migrate BS2000” needs to be chosen and “Execute” be clicked in the action field of the main window:



The action or the wizard “Migrate BS2000” is only offered when the SU is a part of an SU Cluster and an LM is possible in principle.

The procedure is analogous to the procedure shown further for a BS2000 VM.

#### Note:

- Before the migration of the Native BS2000 system, the Native BS2000 system on the target SU must be shut down and deleted (action “Delete BS2000”).
- The Native BS2000 can be restored using the action "Restore BS2000" if needed.

### LM of a BS2000 VM

The main window “Operation” is called in VM2000 mode in the SEM menu in the following way:

*Systems -> <SE name> (SE...) -> <SU name> (SU...) -> <VM name> -> Operation.*

To start checking the preconditions and/or to start the LM, the action “Migrate BS2000 VM” needs to be chosen and “Execute” be clicked in the action field of the main window:



After calling the LM function a new window is opened, in which the VM and the status of the system that needs to be migrated and the names of the source SU and target SU are displayed.

The target SU – for two Server Units in an SU Cluster – is automatically identified:

Migrate BS2000 VM ?

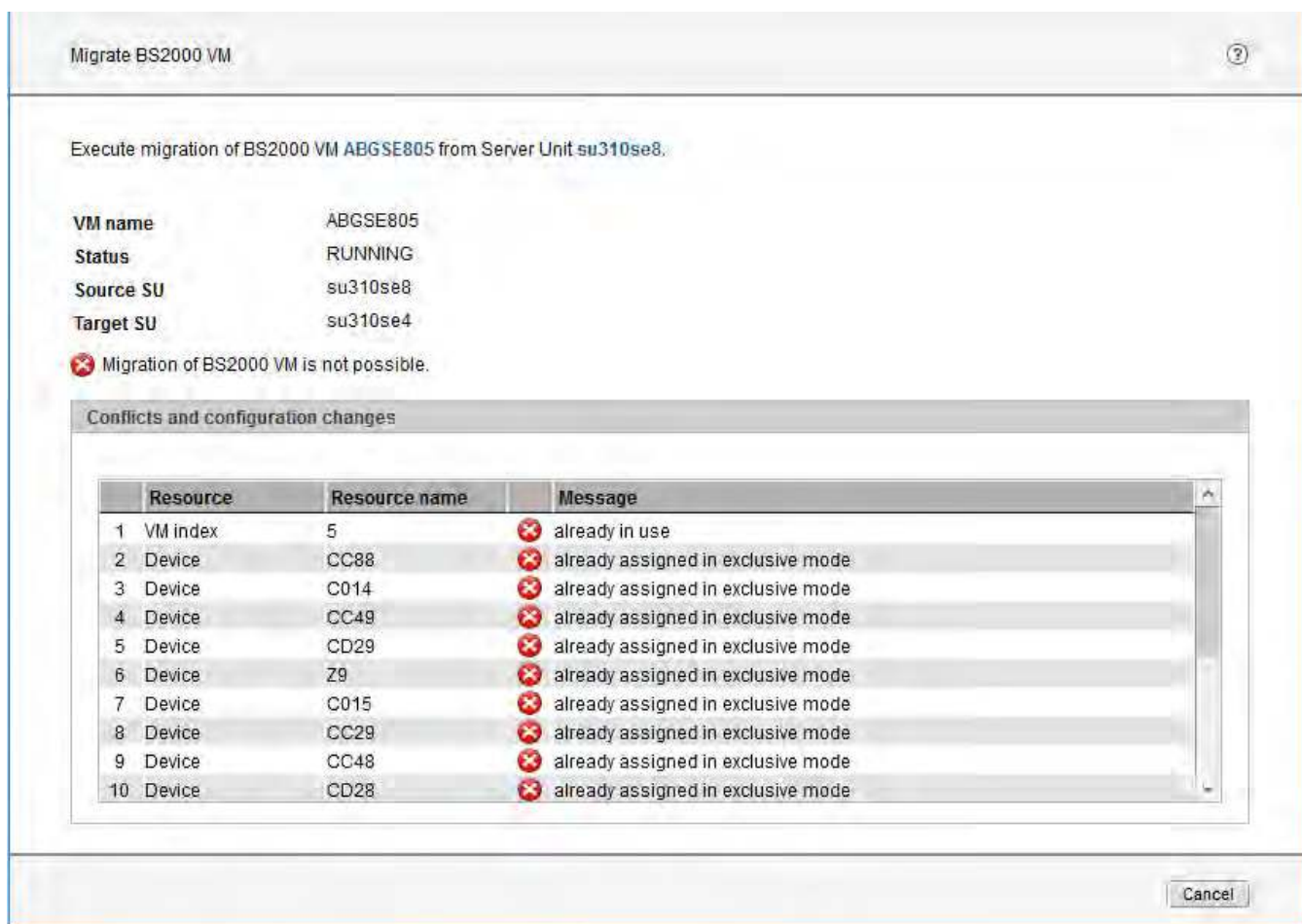
Execute migration of BS2000 VM **ABGSE805** from Server Unit **su310se8**.

<b>VM name</b>	ABGSE805
<b>Status</b>	RUNNING
<b>Source SU</b>	su310se8
<b>Target SU</b>	<input type="text" value="su310se4"/>

Only after clicking „Check preconditions“, the real LM can be started or the action can be aborted depending on the result of this check.

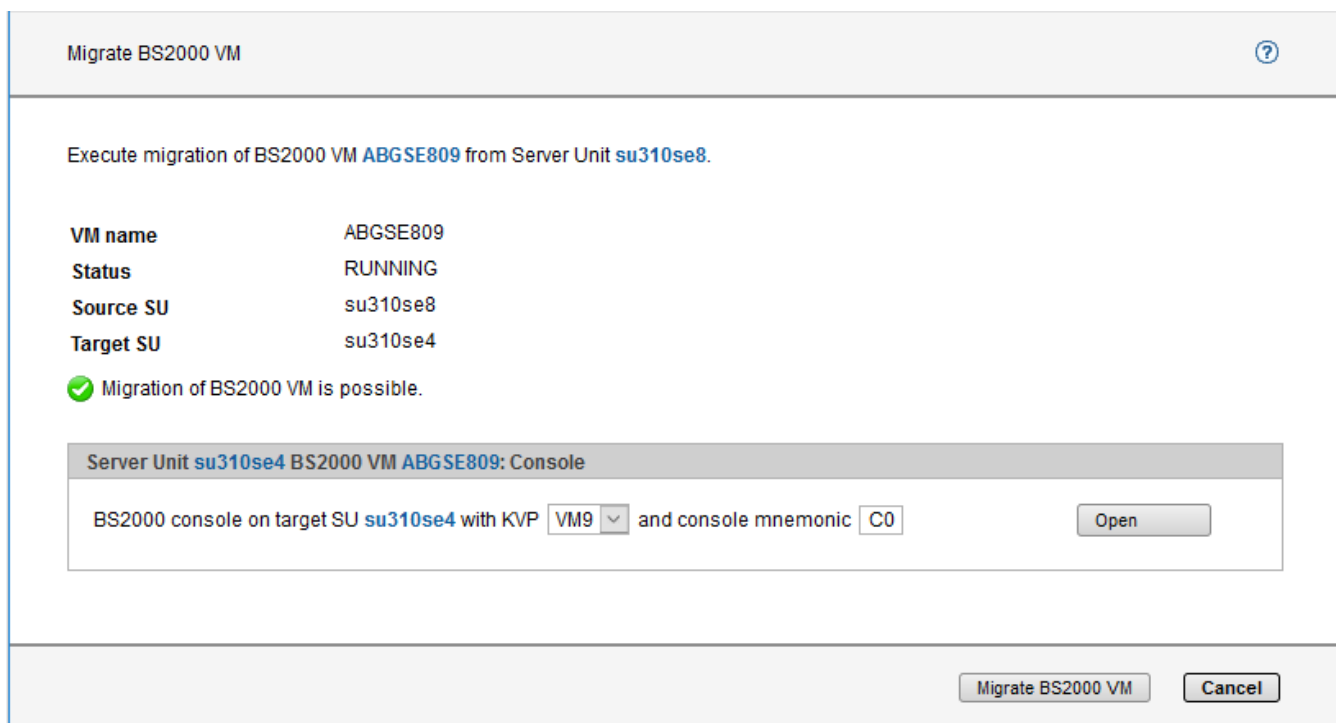
It is checked if the resources and prerequisites needed for the BS2000 system to be migrated are fulfilled on the target SU (devices, size of the needed main memory, free VM index, etc. See chapter -> [Usage scenarios and preconditions for LM](#)). The discovered issues or conflicts are reported directly to the caller in the wizard. The LM is rejected if one or more conflicts are discovered.

Example for the rejection of LM because of insufficient main memory on the target SU:



All reported conflicts must be resolved, otherwise the LM cannot be carried out.

If no conflicts were discovered, the button “Migrate BS2000 VM” appears in the wizard so that the LM can be started:



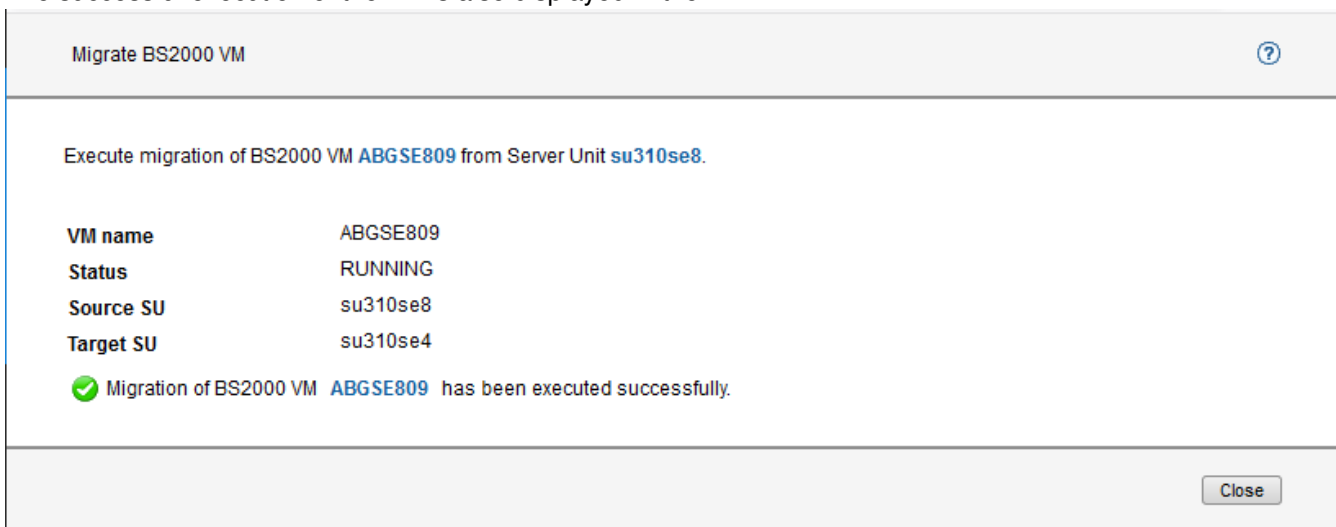
After the LM, console messages of the migrated system are no longer displayed on the KVP console of the source SU but they are displayed on the KVP console of the target SU. The wizard offers the possibility to open a KVP console on the target SU with a mouse click already before starting the migration so that the administrator can see all the messages during the LM.

If the configuration should be checked for a later LM, the function can be aborted.

The LM is started by clicking the button "Migrate BS2000 VM".

It is waited for the end of the action.

The successful execution of the LM is also displayed in the LM:



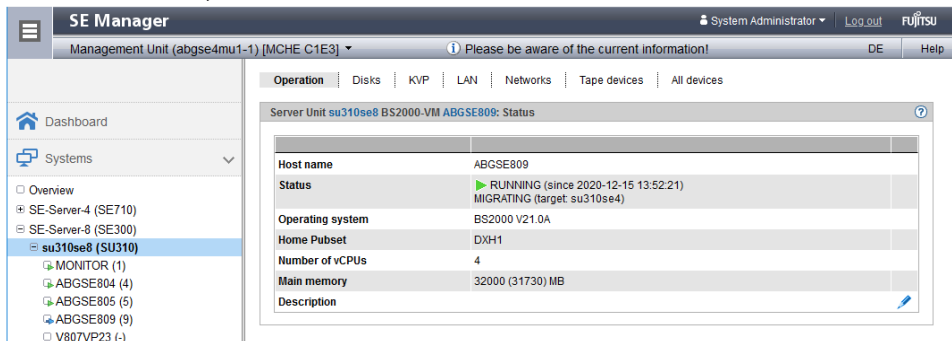
As soon as the wizard is closed after a successful migration, the SE Manager redirects automatically to the "Operation" window of the migrated VM on the target SU.

### Further messages during an LM

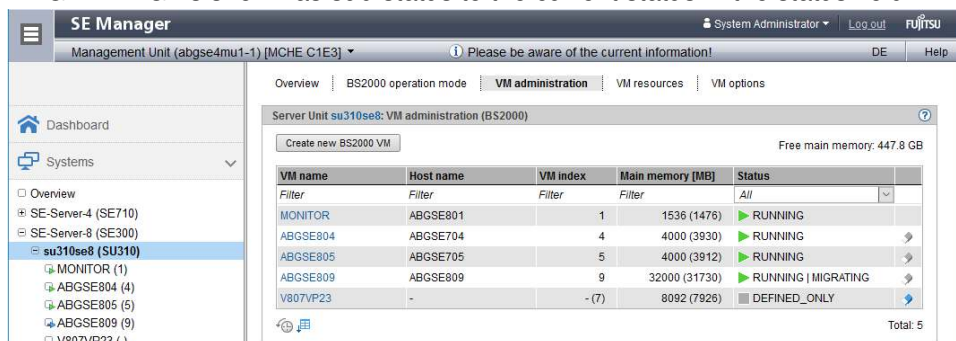
While the LM is running, this is displayed in SEM as follows:

- Main window "Operation"

In the status field, "MIGRATING" is shown as sub-status to the current status as well as the target SU.



- Main window with system overviews for example "VM administration"  
"MIGRATING" is shown as sub-status to the current status in the status field.



- Navigation

The LM of the BS2000 VM is represented by a blue arrow in the primary navigation.

### 4.3 Operating the LM through VM2000 commands

The LM can be controlled also with VM2000 administrator commands in VM2000 operation mode.

The command is entered in the monitor system of the source SU.

For SU /390 the SU name is the same as the BCAM hostname of the monitor system, for SU 86 the SU name is the X2000 hostname.

(Note regarding SU /390: In SEM the names of the SU /390 are freely configurable. These names are used in all SEM displays and actions.)

In the following, the commands /MIGRATE-VM, /CHECK-VM-MIGRATION and /MIGRATE-VM-DEFINITION are explained.

These commands are described in detail in the VM2000 user manual [3].

#### Command /MIGRATE-VM

/MIGRATE-VM initiates the Live Migration.

The system that needs to be migrated is specified through the VM index or VM name.

Example of an LM of the guest system with VM index 12 for the target SU ABGSE701:

```
/MIGRATE-VM VM-ID=12, SERVER-UNIT-NAME=ABGSE701
```

The progress of the LM is protocolled on the console with messages.

Additionally, the progress can also be monitored with a MONJV.

Example of an LM with specification of the VM name of the guest system that needs to be migrated and of a MONJV to monitor the LM, for example from an SDF-P procedure:

```
/MIGRATE-VM VM-ID=VM12PROD, SERVER-UNIT-NAME=ABGSE701, MONJV=MONJV.LM.VM12PROD
```

After sending the command, the status of the LM is displayed in one of the first three bytes of the MONJV. Further information follows it (see current handbook for VM2000).

Possible values of the status of the MONJV and their meanings are:

Status	Meaning
\$S	Task accepted for migration
\$R	Migration in progress
\$T	Migration terminated normally
\$TW	Migration terminated with warnings
\$AE	Migration ended abnormally because of errors
\$AF	Migration task rejected

**Command /CHECK-VM-MIGRATION**

The command checks the preconditions for carrying out an LM.

Checking serves the purpose of finding possible issues in the configuration in the preliminary stages of a planned LM.

The command /MIGRATE-VM also executes the function /CHECK-VM-MIGRATION (first) internally.

The result of the check is returned through VM2000 messages.

Example:

```
/CHECK-VM-MIGRATION VM-ID=12, SERVER-UNIT-NAME=ABGSE501
% VMS2404 LIVE MIGRATION OF VM (12,VM12PROD) REJECTED BY THE TARGET SU 'ABGSE501'.
REASON:
VMS2437,4683
% VMS2461 LIVE MIGRATION OF VM (12,VM12PROD) TO TARGET SU 'ABGSE501' NOT POSSIBLE
/HELP-MSG-INFORMATION MSG-IDENTIFICATION=VMS2437, LANGUAGE=E
% VMS2437 DEVICE '(&00)' NOT ASSIGNABLE TO VIRTUAL MACHINE ON THE TARGET SU
% ? (&00): Mnemonic device name
%   The device is already assigned (exclusively) to another VM.
```

LM cannot be done in this situation, because the system to be migrated is using a device with MN 4683 on the source SU, but on the target SU this device is assigned exclusively to a different VM.

As measure the device must either be made available on the target SU or removed from the system to be migrated on the source SU.

After removing the inconsistency, the command delivers a positive result:

```
/CHECK-VM-MIGRATION VM-ID=12, SERVER-UNIT-NAME=ABGSE501
% VMS2460 LIVE MIGRATION OF VM (12,VM12PROD) TO TARGET SU 'ABGSE501' POSSIBLE
```

The LM is now possible.

**Command /MIGRATE-VM-DEFINITION**

Persistent VM definitions of not initialized VMs (state DEFINED\_ONLY in SE Manager) are transferred with this VM2000 administration command within an SU Cluster from the local SU of the cluster that needs to be removed.

In the VM2000 monitor system one receives the message DEFINED\_ONLY as status ACT[IVE]=NO:

```
/SHOW-VM-DEFINITION
%                               CPU-POOL (P)
% VM-ID          ACT PER HA VM-GROUP (G)  PROC  MEM-SIZE MAX-IO
...
%   - OSD11TST  NO  YES NO *STDPOOL (P)   1     512 100
...
%   12 VM12PROD YES YES NO *STDPOOL (P)   1     2048 100
```

The VM with the VM name OSD11TST isn't active (ACT = NO), which means that its VM definition can be transferred with the command /MIGRATE-VM-DEFINITION to the SU ABGSE501:

```
/MIGRATE-VM-DEFINITION VM-NAME=OSD11TST, SERVER-UNIT-NAME=ABGSE501  
% VMS2320 DEFINITION OF VM 'OSD11TST' MIGRATED TO TARGET SU 'ABGSE501'
```

The VM with the VM name VM12PROD is active (ACT = YES), which means that its VM definition cannot be transferred with the command /MIGRATE-VM-DEFINITION:

```
/MIGRATE-VM-DEFINITION VM-NAME=VM12PROD, SERVER-UNIT-NAME=ABGSE501  
% VMS4015 VIRTUAL MACHINE (12,VM12PROD) ALREADY INITIALIZED. /MIGRATE-VM-DEFINITION  
REJECTED
```

## 4.4 LM procedure

The following short presentation of the LM procedure serves to create a better understanding of connections and limits itself to the aspects essential for it.

### Management of BS2000 devices

At the start of the LM, the attached tape devices (also emulated tape devices) are detached by the BS2000 system that needs to be migrated (/DETACH-DEVICE).

This is why, it's recommended to stop processing the tapes before the LM to avoid jobs being aborted and so as to not hinder the DETACH command.

After the relocation the tape devices remain detached and must be attached if needed (/ATTACH-DEVICE).

Disks emulated by the local SU/MU are also detached before the LM, which is why their usage should be stopped before the migration.

PAV alias devices on SU /390 are also automatically detached before the LM, but are attached again after the LM on the target SU.

Pay attention to the hints in chapter -> [Further notes about LM](#).

An LM in BS2000 consists of multiple phases that follow one another:

### Phase 1 – Setup phase

In the setup phase the preconditions for the LM are checked (see chapter -> [Usage scenarios and preconditions for LM](#)). If all prerequisites are fulfilled then the necessary process environment is provided for the system under migration to the target SU (the VM is furnished with the necessary main memory and assigned devices on the target SU).

#### Hints:

- The necessary main memory is allocated on the target SU and this means it's no longer free. However in the system overviews of SEM, the inbound BS2000 system is not visible on the target SU. This creates a hole in SEM to the end of the LM in the storage display – the free main memory is smaller than the sum of main memory values of the displayed BS2000 systems would make one expect. In VM2000 the inbound BS2000 system is visible on the target SU.
- The same is valid in SEM also for the devices assigned to the BS2000 system on the target SU: though they aren't free, they are also not shown as assigned.

### Phase 2 – Copying the main memory from the source SU to the target SU

In the first step all the pages of the main memory of the system to be migrated are transferred from source SU to target SU. Additionally, VM200/X2000 starts monitoring the working memory of the system to be migrated on the target SU and notes every subsequent change.

In the next step, the memory pages that have been changed in the meantime are transferred again.

This step is repeated multiple times depending on necessity, where usually the number of the pages to be copied becomes smaller with every run.

At the end of phase 2, the I/O system is stopped for the BS2000 system under migration. While the programs can start I/Os, they are not executed, but instead the I/O tasks are noted (by VM2000/X2000) and started only after migration on the target SU. This way the number of changing memory pages is diminished further.

At the very end of Phase 2, a last storage transfer is carried out.

The transfer of the main memory is done through the internal network -> [MCNPR](#).

This phase of the LM lasts longest and depends on the size of the main memory, the transfer speed in the MCNPR and on the activity within the BS2000 system (how often the individual sites are changed during LM).

The BS2000 and the running applications are running in this phase without impairment, however, the CPUs are loaded additionally on SU /390 in monitor system by transferring the main memory.

### **Phase 3 – Transfer of the system to the target SU**

If the whole memory (or at least a large part) is successfully transferred, then the system is stopped on the source SU and the possibly not transferred storage sites and current register content are copied to the target SU.

Afterwards, the already migrated system is continued on the target SU at the point when it was interrupted and the I/O processing is also restarted.

The time between stopping the system on the source SU and restarting it on the target SU is usually only a few seconds.

If restarting the system on the target SU fails, then the LM is aborted and the system is continued on the source SU. All preparations made on the target SU are reversed.

### **Phase 4 – Post-treatment**

After restarting the system on the target SU, resources reserved for this system are released on the source SU (devices, main memory, etc.).

Important BS2000 components such BCAM, I/O control and MSCF are integrated into the LM to avoid automatic MSCF error recognition and to adjust the own I/O configuration on the target SU.

## 4.5 Effect of the LM within the migrated BS2000 system

The LM is transparent for customer application, i.e. they aren't affected by the transfer to the target SU. If necessary, it gives the possibility to inform oneself if an LM took place, which happens through the BS2000 commands /SHOW-SYSTEM-INFORMATION and /SHOW-LIVE-MIGRATION-HISTORY.

In the following chapter, the usage of the above commands is described in the context of LM. Further details can be found in the handbook "BS2000 Commands" [8] with the description of the corresponding commands.

Notice:

The SU500 and SU700 shown in the examples below are no longer supported. However, this does not change the validity of the LM sequences and displays shown.

### Command /SHOW-SYSTEM-INFORMATION

After a BS2000 has been migrated, the command /SHOW-SYSTEM-INFORMATION issues additional the message NSI6231 or NSI6232 when called with standard parameters, depending on the value of the system parameter MIGHOST:

```
% NSI6231 LIVE MIGRATION HAS OCCURED; DATA OF THE IPL HOST WILL BE DISPLAYED
```

or

```
% NSI6232 LIVE MIGRATION HAS OCCURED; DATA OF THE CURRENTHOST WILL BE DISPLAYED
```

Example:

```
/SHOW-SYSTEM-INFORMATION
```

```

/SHOW-SYSTEM-INFORMATION
% NSI6231 LIVE MIGRATION HAS ALREADY TAKEN PLACE; DATA OF IPL HOST WILL BE
DISPLAYED
%CONFIGURATION = 390SU- 500-10B
%CURRENT-CONFIGURATION = 390SU- 700-70
%CPU-ID-LIST : ADR 0 = 0602210424000000
% ADR 1 = 0612210424000000
%HSI-ATT : TYPE = IX
% ASF = YES
% OPERATION-MODE = VM2000
%MEMORY-SIZE = 4096 MB
%MINIMAL-MEMORY-SIZE = 4096 MB
%BS2000-ID : NAME = O10BXS
% VERSION = V20.0B0000
% OSD-BC-VERSION = V11.0B0000
% CREATED = 2018-01-19,14:22:52
%IOCONF-ID : NAME = SU500SE1
% VERSION = V19.0A01
% CREATED = 2019-08-16,13:19:25
% FORMAT = IORSF01
%IPL-TIME = 2019-09-26,14:15:40
%SYSTEM-CONF : SYSID = 127
% HOME-PUBSET = SBZ1
% HOST-NAME = BS2PROD
% VM-INDEX = 12
% VM-NAME = VM12PROD
% SYSTEM-NAME = *NONE
% SYSPAR-BS2-SEL = *STD
% LIVE-MIG-COUNT = 3
%VM2000-VERSION = V11.5A
%VM2000-MONITOR- OSD-BC-VERSION = V11.0B0000
% SYSTEM: HOST-NAME = ABGSE501
%SYSTEM-TIME- ZONE = +01:00
% PARAMETER: SEASON = S
% SEASON-DIFFERENCE = 01:00
% PREV-CHANGE-DATE = 2019-03-27,02:00:00
% NEXT-CHANGE-DATE = 2019-10-30,03:00:00
% SYNCHRONIZATION = SERVER-CONN-EXT-REF
% EPOCH = 00

```

The parameter LIVE-MIG-COUNT shows how often this system has been migrated.

From the parameter CONFIGURATION one can see that this VM has been started initially on a SU500 and the parameter CURRENT-CONFIGURATION shows that it is currently running on a SU700.

The following CPU-ID-LIST shows by default the CPU IDs of the installation on which the system has been loaded (in this example, the CPU ID of the SU500).

The CPU IDs of the installation on which the system is currently running are displayed if the command is called with the parameter SERVER-UNIT=\*CURRENT:

```
/SHOW-SYSTEM-INFORMATION SERVER-UNIT=*CURRENT
%CONFIGURATION                = 390SU- 700-70
%CURRENT-CONFIGURATION        = 390SU- 700-70
%CPU-ID-LIST :   ADR  0       = 6002900126000000
%                           ADR  1       = 6012900126000000
%HSI-ATT :   TYPE           = IX
...
```

The corresponding user-macro is called NSINF and is described in the "Executive Macros" [10] manual.

The standard output of the command is influenced by the setup of the system parameter MIGHOST.

### System parameter MIGHOST

System data changes because of an LM to a different SU, for example the name of the VM2000 monitor system or the VM index.

To avoid compatibility problems when an application uses such data (for example as part of a file name), all information functions of the systems deliver also after the LM the data valid at IPL time.

If such compatibility requirements are not given, we recommend to change the system parameter MIGHOST in such a way that current data is output instead of the IPL data.

The parameter can be changed while the system is running.

Further details can be found in the handbook "Introductory Guide to Systems Support" [9] in the description of the system parameter MIGHOST.

### Command /SHOW-LIVE-MIGRATION-HISTORY

The command /SHOW-LIVE-MIGRATION-HISTORY delivers additionally also the date and the time of every performed Live Migration.

Example of an output after the first migration of the VM 12 in a SU /390 Cluster:

```
/SHOW-LIVE-MIGRATION-HISTORY
LM-COUNT    DATE        TIME          CONFIGURATION  VM
00000000    2019-12-06   10:07:02     390SU- 500-10B  12
00000001    2019-12-07   09:57:38     390SU- 700-70   12
```

The system has been loaded on 06.12.2019 at 10:07 on SU500 and migrated to a SU700 on 07.12.2019 at 9:57.

## 4.6 Messages during the LM

The sequence of events during an LM is protocolled both on the consoles of the monitor systems that are involved, as well as on the console of the systems under migration.

The following examples are presented in this chapter:

- [Example with console messages on SU /390](#)
- [Example with console messages on SU x86](#)
- [Messages in the Event logging of the SE Manager](#)

## 4.6.1 Example with console messages on SU /390

The following example shows the most important messages during an LM between two SU /390 with VM2000. The messages are output on consoles of three different systems. To differentiate between those three better, they outputs are written in different colours. The console header of the messages was cut off to the timestamp, as it has no role here:

Console messages of the monitor system from the protocol file on the target SU:

```
000659 MIGRATE-VM VM-ID=12, SERVER-UNIT-
NAME=ABGSE701                               Start of LM
000700 % VMS2450 LIVE MIGRATION OF VM (12,VM12PROD) TO TARGET SU 'ABGSE701'
STARTED                                     1
000719 % VMS2452 VM (12,VM12PROD) SUSPENDED AT LIVE MIGRATION TO TARGET SU
'ABGSE701'                                  7
000723 % VMS4009 DEVICE 'CS' REMOVED FROM VIRTUAL MACHINE (12,
VM12PROD)                                  13
...
000723 % VMS4009 DEVICE 'FAA4' REMOVED FROM VIRTUAL MACHINE (12,VM12PROD)
000723 % VMS4009 DEVICE 'FAA5' REMOVED FROM VIRTUAL MACHINE (12,VM12PROD)
000723 % VMS4024 VIRTUAL MACHINE (12,VM12PROD) TERMINATED
14
000724 % VMS2451 LIVE MIGRATION OF VM (12,VM12PROD) TO TARGET SU 'ABGSE701'
COMPLETED                                  15
```

Console messages of the monitor system from the protocol file on the target SU:

```
000701 % VMS2430 LIVE MIGRATION OF VM (12,VM12PROD) FROM SOURCE SU 'ABGSE501'
STARTED                                     4
000701 % VMS3050 VIRTUAL MACHINE (12,VM12PROD) CREATED
5
000703 % VMS4005 DEVICE 'CS' ASSIGNED TO VIRTUAL MACHINE (12,
VM12PROD)                                  6
...
000703 % VMS4005 DEVICE 'FAA4' ASSIGNED TO VIRTUAL MACHINE (12,VM12PROD)
000703 % VMS4005 DEVICE 'FAA5' ASSIGNED TO VIRTUAL MACHINE (12,VM12PROD) (TYPE=SD)
000722 % VMS2432 VM (12,VM12PROD) RESUMED AT LIVE MIGRATION FROM SOURCE SU
'ABGSE501'                                  9
000723 % VMS2431 LIVE MIGRATION OF VM (12,VM12PROD) FROM SOURCE SU 'ABGSE501'
COMPLETED                                  12
```

Console messages of the system VM12PROD that needs to be migrated before the migration:

```
000701 % ETMDR01 DYNAMIC RECONFIGURATION WITH FUNCTION 'MIGRATE' STARTED
2
000701 % MCS1002 GLOBAL MSCF CONFIGURATION PARAMETER 'RECOVERY-START' CHANGED FROM
'*AUTOMATIC' TO 3
'LOCKED'
```

Console messages of the system VM12PROD that needs to be migrated after the migration:

```
000721 % ETMDR03 LIVE MIGRATION FROM SOURCE CONFIGURATION '390SU- 500-10B' TO TARGET
      8
      CONFIGURATION '390SU- 700-70' FINISHED. INITIAL CONFIGURATION WAS '390SU-
700-70'
000723 % BCA07DC LINE "L#MANLO1" active
      10
000723 % BCA05EC VLAN-LINE "L1#DANPU" active
000723 % BCA05EC VLAN-LINE "L1#MANPU" active
000723 % BCA05EC VLAN-LINE "L1#MCNPR" active
000723 % MCS1002 GLOBAL MSCF CONFIGURATION PARAMETER 'RECOVERY-START' CHANGED FROM
'LOCKED' TO      11
      '*AUTOMATIC'
```

Explanation of the messages in chronological sequence of appearance on the respective consoles: (Indentations and colours mean: **target SU** | **migrating VM** | **target SU**.)

1: After entering the command /MIRATE-VM the preconditions for an LM on the source SU and target SU were checked, and the LM was started.

2: The guest system was informed about the upcoming transfer.

3: The automatic MSCF outage recognition was deactivated locally and on all partners.

4: The LM is announced in monitor system of the target SU.

5: The VM configuration of the guest system that needs to be transferred was transferred to the monitor of the target SU.

With the information from the VM configuration the VM is set up on the target SU.

Then, transfer of the main memory of the VM from the source SU to target SU starts.

6: The devices assigned VM 12 on the source SU are also assigned on the target SU.

7: The guest system was stopped on the source SU and the remaining part of the main memory and the register contents were transferred from the source SU to target SU.

8: The guest system is started on the target SU and I/O system was started.

9: In the monitor system of the target SU the successful restart of the guest system was protocolled.

10: In the guest system BCAM has activated the IP addresses of the system on the new LAN devices of the target SU and made the changed addresses known in the network.

11: The MSCF outage recognition was activated.

12: The successful end of the LM was protocolled on the console of the monitor system of the target SU.

13: The system is running now on the target SU again, which also means that the no longer needed resources on the source SU are released.

14: The VM with VM ID 12 was terminated on the source SU and the VM definition was deleted.

15: Final message VMS2541 after which all actions of the LM were executed.

## 4.6.2 Example with console messages on SU x86

The LM is executed by X2000 on SU x86.

In operation mode VM2000 other message numbers are issued in part than on SU /390.

The following example shows the most important messages during an LM between two SU x86 with VM2000.

The messages are output on the source SU and target SU. To differentiate them better they are shown in different colours (source SU | target SU).

The console header of the messages was cut off to the timestamp as it plays no role:

Console messages of the monitor system from the protocol file on the source SU:

```
162955 MIGRATE-VM VM-ID=3, SERVER-UNIT-NAME=ABGSE30B, MONJV=JV.VM.MIGVM.3
162958 % VMS2450 LIVE MIGRATION OF VM (3,VM03PROD) TO TARGET SU 'ABGSE30B' STARTED
...
163014 % VMS4009 DEVICE 'CC44' REMOVED FROM VIRTUAL MACHINE (3,VM03PROD) (BY SYSADM)
163014 % VMS4009 DEVICE 'CC45' REMOVED FROM VIRTUAL MACHINE (3,VM03PROD) (BY SYSADM)
...
163014 % VMS4009 DEVICE 'FAAE' REMOVED FROM VIRTUAL MACHINE (3,VM03PROD) (BY SYSADM)
163014 % VMS4009 DEVICE 'FAAF' REMOVED FROM VIRTUAL MACHINE (3,VM03PROD) (BY SYSADM)
163014 % VMS6021 VIRTUAL MACHINE (3,VM03PROD) TERMINATED BY SYSTEM ADMINISTRATION
163020 % VMS2451 LIVE MIGRATION OF VM (3,VM03PROD) TO TARGET SU 'ABGSE30B' COMPLETED
```

In contrast to SU /390 no message VMS2452 is output on the source SU x86 when the migrating BS2000 system is stopped.

Instead, the message HAL0331 is output on the KVP console of the system under migration on the source SU:

```
HAL0331 BS2000 SUSPEND
```

Console messages of the monitor system from the protocol file of the target SU:

```
163002 % VMS6020 VIRTUAL MACHINE (3,VM03PROD) CREATED BY SYSTEM ADMINISTRATION
...
163002 % VMS4005 DEVICE 'FAAF' ASSIGNED TO VIRTUAL MACHINE (3,VM03PROD) (BY SYSADM)
163002 % VMS4005 DEVICE 'CC44' ASSIGNED TO VIRTUAL MACHINE (3,VM03PROD) (BY SYSADM)
163002 % VMS4005 DEVICE 'CC45' ASSIGNED TO VIRTUAL MACHINE (3,VM03PROD) (BY SYSADM)
163002 % VMS4005 DEVICE 'FAAE' ASSIGNED TO VIRTUAL MACHINE (3,VM03PROD) (BY SYSADM)
163032 % VMS2050 GUEST SYSTEM ON VM (3,VM03PROD) READY
```

In contrast to SU /390 no message VMS2432 is output on the target SU 86 when restarting the migrating BS2000 system.

Instead, the message HAL0332 is output on the KVP console of the migrating system on the target SU:

```
HAL0332 BS2000 RESUME, CONTINUING
```

The BS2000 messages of the migrating guest system are for SU x86 analogous to the ones for SU /390.

### 4.6.3 Messages in the Event logging of the SE Manager

In the event logging of SEM it is logged when a BS2000 system is started, shut down and deleted, among others. Those also are events during an LM that are protocolled: shutting down the BS2000 system on the source SU, creating the BS2000 system on the target SU and deleting the BS2000 system on the source SU (For further details see the chapter -> [LM procedure](#)).

With the help of the filter function the user can have those events displayed for a specific BS2000 system to check if and when they were migrated.

In the following example, two migrations can be seen of the guest system "V213WUTH" on an SU /390. Three events are created for every migration.

The messages are sorted by timestamp, the newest event is displayed above.

Example:

- The first LM took place on 17.12.2019 at 11:05 o'clock from the source SU "SU390-SE4" to the target SU "SU390-SE2".
- The LM in reverse direction was done on 18.12.2019 at 13:38 o'clock.

The screenshot shows the SE Manager interface with the 'Event logging' section selected in the left sidebar. The main area displays a table of events for the VM 'V213WUTH'. The table is filtered to show events from 2019-12-17 11:00:00 onwards. The events are sorted by timestamp, with the newest event at the top. The events are:

Date	Weight	Unit	Component	Message
2019-12-18 13:38:46	NOTICE	abgse2mu2	Sys-Mgmt	VM 'V213WUTH' on Server Unit 'SU390-SE2' deleted
2019-12-18 13:38:43	NOTICE	abgse2mu2	Sys-Mgmt	VM 'V213WUTH' on Server Unit 'SU390-SE2' paused
2019-12-18 13:38:41	NOTICE	abgse4mu2-1	Sys-Mgmt	VM 'V213WUTH' on Server Unit 'SU390-SE4' created
2019-12-17 11:05:36	NOTICE	abgse4mu1-1	Sys-Mgmt	VM 'V213WUTH' on Server Unit 'SU390-SE-4' deleted
2019-12-17 11:05:34	NOTICE	abgse2mu1	Sys-Mgmt	VM 'V213WUTH' on Server Unit 'SU390-SE-2' created
2019-12-17 11:05:33	NOTICE	abgse4mu1-1	Sys-Mgmt	VM 'V213WUTH' on Server Unit 'SU390-SE-4' paused

The table also shows a 'Total: 6 of 10743' at the bottom right.

The messages in the events table have the following meaning during Live Migration:

VM ... paused „BS2000 Suspend“ on the source SU. The VM was paused.

VM ... created „BS2000 Resume“ on the target SU. The VM was started.

VM ... deleted The VM was deleted on the source SU.

Additionally, the name of the affected VM and on which SU the event happened is displayed in the message.

In case of an LM between SU x86 four events will be created for every migration - the VM is being deactivated before deletion - apart from that the notifications are the same.

## 4.7 Rejection and abortion of an LM

If an LM cannot be done then it is rejected or aborted and the BS2000 system that needs migrating continues running on the source SU.

All measures undertaken by the LM are automatically reversed.

### Rejecting an LM

Before the migration it is first checked if all preconditions for a successful LM are fulfilled.

If the preconditions are not fulfilled, then the task is rejected and the reasons are provided (as conflicts) to the caller in individual messages. These can be, for example:

- There is not enough free main memory on the target SU.
- A device that is assigned to the guest system that needs to be transferred is not generated on the target SU or is assigned exclusively to another system.

As a measure, the reported issues or conflicts must be solved. The process can be repeated then.

Details about the prerequisites are described in chapter -> [Usage scenarios and preconditions for LM](#).

### Aborting an already running LM

Also after the start of the LM problems can occur that force the abortion of the migration.

In this case, all already finished action are reversed (this is valid also for the devices that were automatically detached – they are attached again automatically).

The reason for an abortion is reported to the caller in additional messages.

The following reasons can cause an abortion:

- Problems when assigning the necessary resources to the target SU.
- The networks connection between the SE Servers was interrupted.
- The always changing main memory pages of the guest system cannot be transferred from the source SU to target SU within the intended time (SU /390).
- The attached disks in the guest system are not in operable state on the target SU (SU /390).

### Abortion of the Live Migration because of issues when transferring the main memory

On SU /390:

VM2000 checks the necessary time for memory transfer to keep the interruptions of the system under migration between SUSPEND and RESUME as short as possible.

The decisive factor here are the number of the changes of the main memory, the available CPU capacity of the monitor system and the transfer rate of the internal network MCNPR (the recommended LAN configuration is described in the chapter -> [Network connections](#)).

If it cannot be ensured that the guest system under migration can be transferred to the target SU without noticeable delays, then the LM is interrupted with the message VMS2425 and the system continues running on the source SU.

Example for output of the message VMS2425:

```
% VMS2425 CHANGING MEMORY PAGES CANNOT BE TRANSFERRED IN TIME. LIVE MIGRATION OF VM  
(12,VM12PROD) TO TARGET SU 'ABGSE701' ABORTED
```

On SU x86:

X2000 executes the memory transfer with own CPUs. This means that the LM has no influence on the CPU capacity in BS2000. However, LM causes the BS2000 to become slower because of the memory monitoring and memory transfer.

A slow ISL connection and a high memory change rate can lead to long SUSPEND/RESUME times here, but the LM is executed in every case.

## 4.8 Further notes about LM

### General

- Only one LM can be started on a source SU. Every further start is rejected as long as the LM is running. However, a further LM for which this SU functions as target SU is possible.
- The LM is rejected when a dynamic I/O configuration change is running (initiated with /START-CONFIGURATION-UPDATE in the monitor system – see the VM2000 manual [3]).
- The command /START-CONFIGURATION-UPDATE is rejected during the LM.
- A Live Migration between Server Units with different CPU performance influences the performance of the running applications.
  - If a migration is done to a target SU with more processors than on the source SU, then the number of running BS2000 CPUs remains the same.
  - If a migration is done to a target SU with less processors than the source SU, then the redundant BS2000 CPUs are detached and blocked with /DETACH-DEVICE. During a return migration the earlier blocked and detached CPUs are automatically attached.
  - If the speed of an individual CPU changes after the LM, then the data written by accounting cannot be readily used for an exact or correct calculation of used CPU times. In those cases, the accounting should be restarted after LM and the accounting files should be evaluated in a separate way.
- Tape operation is technically not possible during an LM and therefore needs to be stopped before the migration as the devices cannot be detached otherwise (/DETACH-DEVICE). Detached tape devices are not automatically connected on the target SU.
- Local disks (for example the STBY pubset which has been installed in the factory) are also detached.
- PAV alias devices on SU/390 are automatically detached before the migration and attached after the migration on the target SU. However, other PAV alias devices can be involved (DPAV) on the target SU.
- The connections started from the SE Manager to the local KVP consoles and dialogue connections through server-internal LOCLAN are disconnected by the LM. In contrast, connections from the network of the customer (also console connections through OMNIS) remain operable.

### Notes for VM2000 operation

- The monitor system of VM2000 cannot be migrated.
- The main memory size of the monitor system should be at least 512 MB.
- For SU /390 the monitor system must be equipped with enough CPU capacity (number of virtual CPUs / multiprocessor level, CPU quota, maximal CPU power consumption), otherwise performance issues can occur that in certain circumstances can lead to the LM being aborted.

Note: The CPU capacity needed for LM corresponds to approximately one real CPU on the source SU and to approximately 1.4 real CPUs on the target SU.
- Depending on the size of the CPU pool, the number of active CPUs can be increased or reduced during the migration, but it cannot be increased beyond the multiprocessor level provided by /CREATE-VM. If the target CPU pool has less real CPUs than the migrated VM has attached vCPUs, then the surplus vCPUs are automatically detached and blocked. In the reversed case possible detached vCPUs are automatically attached.
- The global VM2000 resources (CPU pools, VM groups, assignment sets) and the BS2000 devices of the migrating BS2000 system should not be changed during the LM.

- If the VM that needs to be migrated is monitored with a MONJV (defined during /CREATE-VM or /ACTIVATE-VM-DEFINITION), then on the source SU at the start of the migration the string MIGR-OUT is written at the distance 87 – 94 of the MONJV.  
If the migration is successful, then the monitoring state is set to \$T (VM terminated).
- The LM will be rejected if VM2000 shutdown has been initiated before (command /SHUTDOWN-VM VM-ID=\*VM2000).

### Notes about I/O configuration for SU /390

The I/O configurations of a BS2000 device are generated compatibly for LM if the device is generated on both server units with the same BS2000 MN and if its channel configuration or its common subset is the same (not all generated channel connections must also be physically present - "overconfiguration").

Only the I/O configuration for those BS2000 devices that are assigned to the BS2000 VM to be migrated is checked in regards to a rejection during an LM.

Nevertheless, it is recommended that the I/O configurations on the source SU and target SU are kept the same.

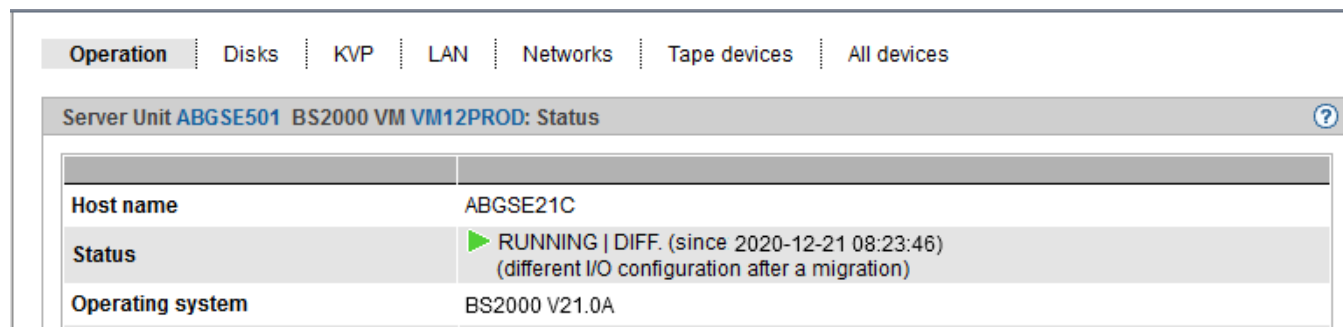
If in case of SU /390 the I/O configuration of the source SU and target SU are different (for example if additional devices have been generated on the target SU), then the VM receives in addition to its status the substatus DIFF.

For example:

```
/SHOW-VM-RESOURCES
```

```
...
%   VM-ID           STATE           VERSION  PER  ADMIN  PRIV
%   1 MONITOR      RUNNING        V21.0A   NO   YES    AS
...
%   12 VM12PROD    RUNNING (DIFF) V21.0A   NO   NO     AS
```

The status RUNNING|DIFF is displayed in SE Manager after the migration to the target SU in menu "Operation" of the migrated system:



As long as a VM has the substatus DIFF, it's excluded from the dynamic I/O configuration change and devices that were not generated on the source SU cannot be assigned to it on the target SU.

The VM keeps the substatus DIFF until it is either migrated to the source SU or it is reactivated on the target SU (new activation is done with the commands /DELETE-VM and /ACTIVATE-VM-DEFINITION and is only possible for persistent VMs.)

### **Notes for I/O configuration for SU x86**

The I/O configurations of a BS2000 device are generated compatibly for LM if the device is generated identically on both server units (i.e. same BS2000 MN, device type, host connector and Unit ID).

The I/O configuration is only checked in regards to rejection for those BS2000 devices that are assigned to the BS2000 system that needs to be migrated.

However, it is recommended that the I/O configuration on the source SU and target SU is kept the same.

The BS2000 devices that are not assigned to the BS2000 VM under migration should also be generated identically on SU x86 (sameness of device type, host connector and unit ID for the same BS2000 MN).

Otherwise, warnings will appear also for unassigned devices with different generation during LM and those devices cannot be assigned on the target SU of the migrated BS2000 VM at all.

In the case of native BS2000, all devices are assigned to the BS2000. In case of unidentical generation of BS2000 devices, an LM is possible on one side only or not at all, depending on the situation.

## 5 Restoring VM definitions after an outage

### Managing VM definitions

A VM definition that is also unambiguous in the SU Cluster exists for every BS2000 VM.

The VM definition contains a description of attributes as well as resources and devices, which are assigned to the BS2000 VM or should be assigned during their activation. A VM definition contains also the current state of the BS2000 VM.

All VM definitions are managed on the CRD (see chapter -> [CRD functionality](#); [CRD with external configuration disks](#)).

The VM definitions are identified unambiguously by their VM name and the SU to which they are assigned, and a BS2000 VM can run only on such that SU to which its VM definition is assigned.

This is why the following is valid: to start a BS2000 VM on a specific SU of the SU Cluster, its VM definition must be assigned to the desired SU.

### Command /RECOVER-VM-DEFINITION

If an SU to which there are VM definitions assigned is not available, then the VM2000 administrator can transfer the VM definitions with the command /RECOVER-VM-DEFINITION to the available SU of the cluster and start the VM there.

This situation can occur for example during deactivation caused by maintenance of an SE Server or an SU or after an SU outage or server outage.

This means that the command /RECOVER-VM-DEFINITION supports the customer's HA procedures and processes (HA means High Availability in this case) to make a quick restart of the productive BS2000 systems on another SU after an outage (failover).

The command /RECOVER-VM-DEFINITION is available in any SU Cluster (SU /390 and SU x86) with VM2000. The (apparent) status of the VM displayed by SEM does not play any role for this command, but it must be ensured that the remote SU is not running.

Example:

The VM definition of the guest system with the VM name VM12PROD is taken from the SU ABSEE701 to the local SU.

```
/RECOVER-VM-DEFINITION VM-NAME=VM12PROD
```

In a SU Cluster with two SU the name of the removed SU is unambiguous and known to the VM2000, so it can be left out.

The name can still be provided though (both commands are equivalent):

```
/RECOVER-VM-DEFINITION VM-NAME=VM12PROD, SERVER-UNIT-NAME=ABGSE701
```

This command is only performed when the following conditions are met:

- The local SU on which the command was entered and the remote SU specified in the command are part of the same SU Cluster.
- The monitor system (SU /390) resp. the X2000 (SU x86) of the remote SU has not written any alive message on the CRD in the intended time.
- The VM name is not used locally.
- The max. number of VM definitions is not reached on the local SU.

After performing the command /RECOVER-VM-DEFINITION the VM definition is in state DEFINED\_ONLY and the VM is not initialized, so it doesn't allocate any resources.

The VM definition can either be put in operation immediately with the command /ACTIVATE-VM-DEFINITION and /START-VM or adjustment is still needed.

### Possible scenario

A change of device assignment (/MODIFY-VM-DEFINITION) for example can be necessary if not only the SE Server but also the storage system with the source disks is damaged in a redundantly designed computer center (total outage of a site).

The system manager activates the mirror disks from the other SE Server of the SU Cluster in the other storage system (through StorMan in SE Manager of the MU or SHC-OSD in BS2000) and changes the MNs of the disk devices in the VM definition as shown in the below example.

Afterwards the VM definition is activated and the VM is initialized by this and finally the BS2000 is loaded from the mirrored IPL disk.

The reservation of the resources described in the VM definition is done with the command /ACTIVATE-VM-DEFINITION.

From this point in time, the VM attributes and device configuration can be modified with the usual VM2000 commands such as /MODIFY-VM-ATTRIBUTES or /ADD-VM-DEVICES. The changes are done then on the initialized VM and in the VM definition.

Example of changing the device configuration in the VM definition of the VM VM12PROD and the following start of the VM:

```
/MODIFY-VM-DEFINITION VM-NAME=VM12PROD, DEVICES=*REMOVE (UNITS=*DEVICERANGE (
    FROM=A000, TO=A0FF) )
/MODIFY-VM-DEFINITION VM-NAME=VM12PROD, DEVICES=*ADD (UNITS=*DEVICE-RANGE (
    FROM=B000, TO=B007, USAGE=*EXCLUSIVE) )
/MODIFY-VM-DEFINITION VM-NAME=VM12PROD, DEVICES=*ADD (UNITS=*DEVICE-RANGE (
    FROM=B008, TO=B0FF, USAGE=*SHARED) )
```

These commands cause that devices B000 –B0FF are assigned during initialization of the VM instead of the previous devices A000 – A0FF.

The devices B000 – B007 are assigned exclusively to the VM, the devices B008 – B0FF are assigned as shared.

The command for displaying the devices contained in a VM definition is done with the command /SHOWM-VM-DEFINITION, so for the above example:

```
/SHOW-VM-DEFINITION VM-NAME=VM12PROD, INFORMATION=*UNITS
```

Finally, the VM definition can be activated and the VM can be started:

```
/ACTIVATE-VM-DEFINITION VM-NAME=VM12PROD
/START-VM IPL-UNIT=B000, VM-ID=VM12PROD
```

## 6 Glossary

Glossary with terms and explanations relevant in the cluster context

### **BS2000 Operation Mode**

The BS2000 operation modes Native BS2000 and VM2000 affect the cluster status of SU Clusters insofar as Live Migration is only possible between SUs with the same BS2000 operation modes:

- SU /390: only VM2000
- SU x86: Native BS2000 or VM2000

### **Cluster**

Group with two formings in the context of the SE Server:

- Management Cluster (SE Cluster) – group between two or more SE Servers for central management (shared administration and operation)
- SU Cluster – group of two Server Units for the support of Live Migration of BS2000 systems

### **Cluster Function**

- The cluster function of the Management Cluster is the central management (or the shared administration and operation).
- The cluster function of the SU Cluster is the Live Migration of BS2000 systems.

### **Cluster Manager**

The Cluster Manager is a software component, which is installed on all MUs and is used for cluster management. At a given time, the Cluster Manager is active on exactly one MU of the cluster, on the MU which is currently the Master MU.

### **Cluster Status**

The cluster status is the general state of a cluster.

It depends on:

- the states of the involved units (for example normal operation or maintenance mode for an SU x86)
- the states of the resources (for example CRD accessible or not) and
- the redundancies available in comparison to normal operation (for example one or more active MUs in the Management Cluster)

In the state NORMAL all cluster functions are fulfilled completely.

In the state WARNING there are restrictions (for example redundant resources are missing).

In the state ERROR the cluster functions are not fulfilled (for example LM is not possible).

SU Clusters can be set into the state INACTIVE.

### **CRD / Configuration Disks**

The CRD serves for storing and managing data which must be available for all the units of the cluster (MU and SU x86), for example VM definitions and regular activity messages of the monitor systems.

The CRD (which means "Configuration Raw Device") consists of an internal and one or two external so-called configuration disks (on external storage systems).

These configuration disks are combined using software mirroring into a logical unit - the CRD resp. the CRD functionality.

Their detailed display is done in SEM per unit.

### **Events**

Events are reported and they can be redirected to users by the Alarm Management so they can react faster.

The Cluster Manager active on the Master MU creates events for every relevant change (component Cluster).

The full list of possible events is available in the online help of the SE Manager.

### **Heartbeat (Network Heartbeat)**

The Cluster Manager active on the Master MU regularly queries all units about their availability and their correct function.

Depending on the result, the states are recalculated and displayed, events are triggered and automatic measures are undertaken.

### **ISL-E**

Network connection between SE Servers respectively their Net Units.

By this, server-internal networks are transformed into shared internal networks for the Management Cluster (for example MCNPR).

### **Live Migration (LM)**

In the context of the SE Server the LM is the transfer of an active BS2000 system (state == RUNNING) from one SU to another SU without affecting the applications in BS2000 system which is migrated.

LM is only possible within an SU Cluster, i.e. between SUs of the same type and in the same BS2000 operation mode.

In extended sense, LM describes also the transfer of an inactive BS2000 system (state != RUNNING).

## **Management Cluster**

The Management Cluster is a group of two or more SE Servers for central management of the participating SE Servers (for central administration and operation).

The Management Cluster requires ISL-E connections between the SE Servers and a CRD (with external configuration disks).

An inhomogeneous Management Cluster is a Management Cluster consisting of one or more SE Servers with all units (MU, HNC, SU x86) with one version (e.g. version V6.5) and one or more SE Servers with all units with another version (e.g. version V6.4). In such a case certain restrictions have to be considered.

## **Master MU**

The Master MU designates a special role of a MU within the Management Cluster:

At a given time, the software that administrates the cluster - the Cluster Manager - is active on exactly one MU of the cluster. This MU plays currently the role of the Master MU. In case of a reboot of the Master MU another MU takes over this role.

## **MCNPR (Management Control Network Private)**

Private (internal) management network for the general SE Server communication:

The MCNPR is server-overarching in the case of a Management Cluster.

It serves for internal management of the SE Servers and also for the transfer of the BS2000 systems during LM.

## **Net Unit (NU)**

The Net Unit is a component of the SE Server and consists of LAN switches. The expansion depends on the size of the SE Server and the need for redundancy.

The Net Unit allows for server-internal communication and also connecting an SE Server to customer networks.

The Net Unit contains the HNC (High-speed Net Connect) for the LAN connection of the SU /390.

In case of a Management Cluster, the Net Units of the participating SE Servers must be connected through ISL-E. This way server-internal networks (for example MCNPR) become shared internal networks for the Management Cluster.

The Net Unit supports a transfer rate of 1 GBit/s.

## **Power Status (of a unit)**

The power status of every unit (SU or MU) affects the cluster status.

See the following examples:

- If an SU is turned off, then the function LM is not fulfilled for the SU Cluster and the cluster status is ERROR.
- If an MU is turned off, then the redundant operation of the Management Cluster is not fully given and the cluster status is WARNING.

## **Redundant SKP Functionality**

The redundant SKP functionality is only provided within an SE Server - also in a Management Cluster.

This means that a SU /390 is only operable if at least one MU of the server is running.

## **SE Cluster**

Synonymous with -> Management Cluster.

## **SE Server**

Short designation for Fujitsu Server BS2000 SE Series

- The Basic Operating Manual [4] provides information on the specific models - e.g. SE730 or SE330.
- Some models may optionally contain several SU x86 and AUs.

## **SENET**

A DNS server for the "senet" domain which provides name resolution for communication (especially for the internal communication within MCNPR) runs on every MU. The DNS server is configured in such a manner that it performs name resolutions for "senet" itself and forwards other name resolutions to external DNS servers which must be configured in addition.

## **SU /390 (Server Unit /390)**

SU type: Server Unit with /390 architecture.

A /390-based Server Unit (SU /390) makes using BS2000 systems (Native BS2000 or VM2000) possible. The Basic Operating Manual [4] provides information on the specific models - e.g. SU730.

## **SU x86 (Server Unit x86)**

SU type: Server Unit with x86-architecture.

A x86-based Server Unit (SU x86) makes using BS2000 systems (Native BS2000 or VM2000) possible. The Basic Operating Manual [4] provides information on the specific models - e.g. SU330

## **SU Cluster**

An SU Cluster is a logical unit that consists of two Server Units of the same type (SU /390 or SU x86).

The SU Cluster offers Live Migration (LM) of BS2000 systems as a basic function.

The two SUs can be located in two different SE Servers (prerequisite for this is a Management Cluster) or in the same SE Server (only for SU x86).

## **SU /390 Cluster**

An SU /390 Cluster is an SU Cluster that consists of SU /390.

## **SU x86 Cluster**

An SU x86 Cluster is an SU Cluster that consists of SU x86.

## **SU Role (in relation to LM)**

- **Source SU:** SU from which a BS2000 system is being migrated.
- **Target SU:** SU, to which a BS2000 system is being migrated.

- **Initial SU** (also IPL SU): SU, on which the IPL for a BS2000 system has been executed last.

## 7 Related publications

You can find the following BS2000 manuals on the manual server with the BS2000 documentation at <https://bs2manuals.ts.fujitsu.com>.

[1]	<b>Fujitsu Server BS2000 SE Series Administration and Operation</b> User Guide
[2]	<b>Fujitsu Server BS2000 SE Series Security Manual</b> User Guide
[3]	<b>VM2000 (BS2000)</b> Virtual Machine System User Guide
[4]	<b>Fujitsu Server BS2000 SE Series Basic Operating Manual</b>
[5]	<b>Fujitsu Server BS2000 SE Series Server Unit /390</b> Operating Manual
[6]	<b>Fujitsu Server BS2000 SE Series Server Unit x86</b> Operating Manual
[7]	<b>Fujitsu Server BS2000 SE Series Additive Components</b> Operating Manual
[8]	<b>BS2000 OSD DX Commands</b> User Guide
[9]	<b>BS2000 OSD DX Introduction to System Administration</b> User Guide
[10]	<b>Executive Macros</b> User Guide