

Datasheet

Fujitsu BS2000/OSD HIPLEX MSCF V7.0 Software

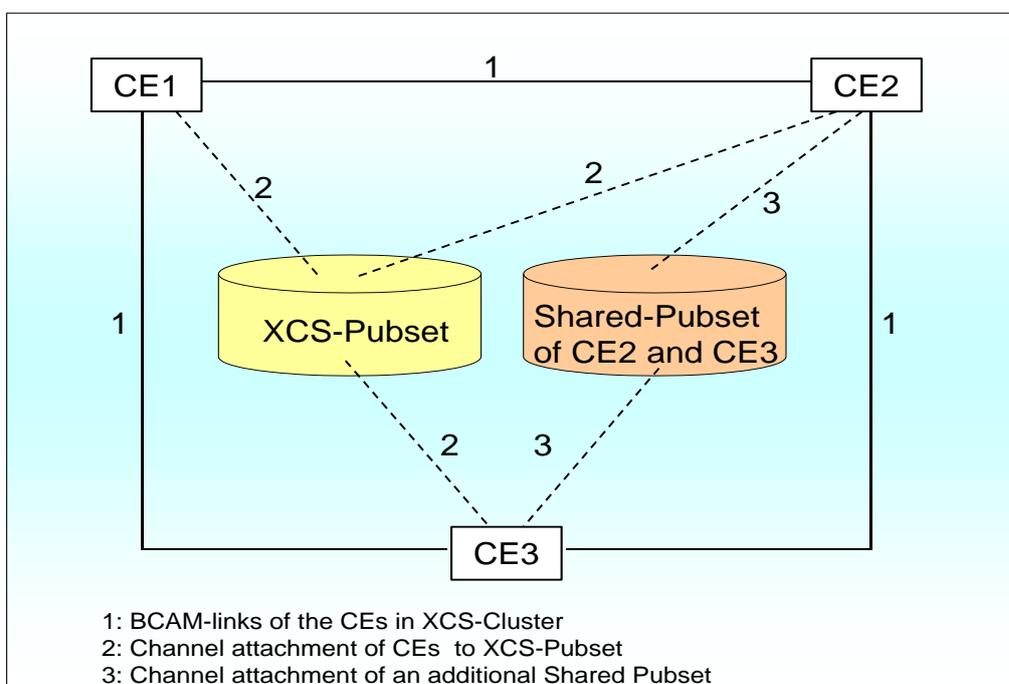
Highly Integrated System Complex

Abstract

HIPLEX (Highly Integrated System Complex) is the system designed by Fujitsu Technology Solutions to implement a load sharing cluster and an availability cluster of multiple BS2000/OSD business servers. Each server attached to the HIPLEX is referred to as a cluster element (CE). HIPLEX MSCF establishes the basic structures for load sharing and availability clusters with up to 164 BS2000/OSD business servers, as well as for distributed applications. The HIPLEX servers attached to the HIPLEX MSCF network must be networked together via openNet Server.

HIPLEX MSCF can implement four kinds of coupled systems which build upon one another:

- **Loosely Coupled System (LCS)** with up to 164 BS2000/OSD servers for job control, catalog access and job monitoring using JVs in the network.
- **Closely Coupled System (CCS)** The CCS network includes LCS functionality and adds password protection as well as basic functions for SPVS – and XCS-Coupled Systems.
- **Shared Pubset – SPVS Coupled System** with shared pubsets and master reconfiguration for up to 16 BS2000/OSD Servers per pubset. The SPVS Coupled System includes CCS- functionality.
- **Cross Coupled System (XCS)** with up to 16 BS2000/OSD servers and XCS pubset for consistent cluster administration and functions for distributed applications. The XCS Coupled System includes SPVS- functionality.



Topics

Functional Description

In the **LCS network** HIPLEX MSCF extends the initially CE-local basic functions of

- catalogue access,
- file access,
- starting of ENTER jobs,

to the HIPLEX MSCF network. Using job variables a load-sharing cluster can be configured in which it is possible with JVs to define dependencies between jobs and/or programs and to achieve synchronous and asynchronous event control of the production process.

In the **SPVS Coupled System** some or all the cluster elements of a HIPLEX operate at least one shared pubset. In the SPVS Coupled System there is a master CE (master sharer), which carries out the meta operations (catalog accesses) for accessing the shared pubset, for the other CEs (slave sharers) in the shared pubset. Data access to the shared pubset on the other hand is carried out by each CE. Once a backup master has been defined out of the group of slave sharers, then automatic changeover of master takes place in the event of a master failure.

The **XCS network** is a tighter coupled system than the SPVS Coupled System. In the SPVS Coupled System a number of cluster elements of the MSCF network are interconnected, but usually not all. Consequently in the SPVS Coupled System there is no view shared by each network subscriber of all the other subscribers in the network. Each network subscriber sees only the CEs to which he has an MSCF link. In the XCS network, by contrast, HIPLEX MSCF combines up to 16 cluster elements, which are interconnected via a BCAM subnetwork, very closely together: the cluster elements in the network have BCAM links **to every other CE in the XCS network**. Via a task on each subscriber in the XCS network, HIPLEX MSCF ensures coordinated operation of the network. The administration of the XCS Coupled System is carried out via a shared pubset, which all the CEs in the XCS Coupled System operate jointly, the XCS pubset.

Management of the XCS network

Prior to building up the XCS Coupled System, the remote catalogs of network subscribers, which should later be accessible from a particular CE are recorded in the MRSCATs in the specified CE. Then the initiative to subscribe to a network can be taken by each CE individually. As soon as HIPLEX MSCF starts with a modified configuration file (XCS configuration manager) the new network subscriber is made known to all the other existing network subscribers. Conversely, the new network member learns which other CEs are subscribing to the network. The exit of a CE from the network is similarly communicated to all the remaining network subscribers. For the users, subscription to the XCS network means the extension of services on to the network.

In addition to the basic functions these include

- the Distributed Lock Manager,
- Shared File System (UPAM, FASTPAM, DIV),
- shared PLAM libraries.

These functions are no longer available to a CE after it has quit the XCS network.

Catalog access (all network types)

HIPLEX MSCF permits access to system catalogs (TSOSCATs) from cluster elements subscribing to the network. This is achieved by unique catalog IDs for the system catalogs which apply across the MSCF network. In the event of a CE failure, its system catalog can be imported to a CE and thus remains available to the remainder of the MSCF network.

File access (SPVS-, XCS networks)

The Shared Pubset function enables up to 16 servers to operate a shared pubset. The sharing granule is the file. One of the CEs becomes the master which administers the meta data of the shared pubset and performs operations on meta data, e.g. OPEN/CLOSE operations, for all the other CEs (slaves). These operations are initiated via the BCAM links between the CEs. Access to the data of the shared pubset, by contrast, takes place via the direct hardware path between CE and shared pubset.

In comparison to the procedure with shared private disk (SPD), the analogous function in relation to private disks, with shared pubsets the lockout protocol is not applicable to disk. This small communication overhead leads to a high transmission rate, which is largely determined by the data transfer rate of the disk subsystem and which is not diminished even by a high OPEN/CLOSE rate.

The **XCS pubset** is the central shared pubset, which contains the meta data of the XCS network. The XCM manages the meta data and thereby also the network. In contrast to other pubsets, the XCS pubset has to be attached to all the BS2000/OSD servers which subscribe to the HIPLEX MSCF network.

Reconfiguration of shared pubsets

To cover the possibility of a master failure, a slave CE is predefined as a backup master. Via live monitoring, which is based on a query of specific counters in the SPVS pubset and on a periodic check of the BCAM link, the backup master initiates an automatic recovery and takes on the new role of master CE. Via XCM all the network subscribers which are still active are informed about the configuration change. Where a slave failure occurs, this is similarly configured by XCM out of the network and communicated to the rest of the network (consistent display). In addition, on the master all the resource-locks held by the slave are replaced, i.e. the resources are released for the rest of the network.

Job variables

The software products HIPLEX MSCF and JV make job variables and conditional job control available across the MSCF network. Their uses include job distribution and monitoring in the MSCF network.

Static Load sharing

HIPLEX MSCF provides functions relating to job distribution and monitoring within the HIPLEX MSCF network. After the network has been set up, jobs can be initiated and distributed to any of the network partners and their status monitored across the network. For distributing jobs the target CE can be addressed directly via a BCAM name or indirectly via job variables. This enables user-driven load sharing in accordance with the distribution of resources in the network. The job variables allow cross-CE dependencies between jobs to be taken into consideration. Job distribution can also be automated using products like AVAS (job networks covering several CEs) or HIPLEX AF (planned relocation of applications).

Dynamic Load Sharing (XCS network with HIPLEX JMS)

HIPLEX JMS extends the highly efficient Job-Management System (JMS) on each OSD server to OSD multiservers. In an XCS network, this allows an optimizing and system-driven distribution of batch loads. The basis is formed by the server-local JMSs with their TU schedulers, job classes, streams and a class scheduler per server. Before the start of a job, HIPLEX JMS does a search for the server with the lightest load. It does this by performing a network-wide job class analysis across the XCS systems and then, with the aid of the local JMS, starts the job on the server that currently has the lightest load. The request for dynamic distribution for a job is made by means of the new host parameter "*ANY" in the ENTER-JOB, ENTER-PROCEDURE commands. Dynamic and static job distribution are of course possible concurrently in an XCS network (for different jobs).

Availability cluster (SPVS-, XCS networks)

The failure of a CE in the MSCF network is recognized by HIPLEX MSCF. Subsequently, a reconfiguration of the network takes place. Furthermore, HIPLEX AF can use the failure data coming from HIPLEX MSCF and fully automatically relocate applications from the failed CE to another CE which is still working. In order to do this, HIPLEX AF switches the pubsets required by the applications and the network

access over to the new CE (fail-over functionality). Shared pubsets do not need to be switched over.

Distributed Lock Manager (DLM, XCS network)

DLM is the basic structure implemented in HIPLEX MSCF to manage competition for access to resources in distributed applications, e.g. in ORACLE PARALLEL SERVER. DLM possesses various types of interconvertible locks. Using these, finely granular locks can first be defined and then very flexibly managed. By distributing to all the CEs the DLM is fail-safe, i.e. using XCM it can successfully be reconfigured in the event of a CE failure.

Shared file system (XCS network)

Based on DLM and HIPLEX MSCF in a shared file system (UPAM, FASTPAM, DIV) a file in an XCS pubset can be shared at data record level, i.e. the lock granule is a data block: Two or more CEs can simultaneously gain read and/or write access to various data records in one and the same file.

Shared library (XCS network)

With HIPLEX MSCF, DLM and LMS as of V3.4, all PLAM-libraries on shared pubsets are automatically "shared libraries". All CEs, which are members of a XCS Coupled System, can open them. Library elements are the locked objects.

Fast crash-monitoring with HIPLEX AF V3.3

Usually a HIPLEX-cluster is controlled via shared pubsets and/or MSCF-links. If HIPLEX AF V3.3 is installed partner systems within a HIPLEX are additionally capable of controlling the cluster via the Live-Monitor provided by HIPLEX AF V3.3. This means that a partner-crash is noticed within seconds. The failover-procedure can start immediately. This failover-speedup reduces overall system downtime.

Technical details

Technical data

Hardware	BS2000/OSD Business Server S and SX series
Software	For S-Server: BS2000/OSD-BC V9.0 openNetServer V3.5 JV V15.1 (optional) For SQ-Server: OSD/XC V9.0
Operating mode	Timesharing operation
Implementation language	SPL, Assembler
User interface	Commands: English Message texts: either English or German
Installation	By the user, on the basis of the Release Notice
Documentation	HIPLEX MSCF V6.0 (BS2000/OSD) user manual (also applies to V7.0)
Demands on the user	Knowledge of multi-systems / HIPLEX
Conditions	This software product can be purchased by the customer against a single payment or leased in accordance with our conditions for the use of software products.
Ordering and delivery	This software product may be obtained from your local Fujitsu Technology Solutions regional office.

More information

Fujitsu platform solutions

In addition to Fujitsu BS2000/OSD, Fujitsu provides a range of platform solutions. They combine reliable Fujitsu products with the best in services, know-how and worldwide partnerships.

Dynamic Infrastructures

With the Fujitsu Dynamic Infrastructures approach, Fujitsu offers a full portfolio of IT products, solutions and services, ranging from clients to datacenter solutions, Managed Infrastructure and Infrastructure as a Service. How much you benefit from Fujitsu technologies and services depends on the level of cooperation you choose. This takes IT flexibility and efficiency to the next level.

Computing products

www.fujitsu.com/global/services/computing/

- PRIMERGY: Industrial standard server
- SPARC Enterprise: UNIX server
- PRIMEQUEST: Mission-critical IA server
- ETERNUS: Storage system
- BS2000/OSD mainframes

Software

www.fujitsu.com/software/

- Interstage: Application infrastructure software
- Systemwalker: System management software

More information

Learn more about Fujitsu BS2000/OSD HIPLEX MSCF V7.0, please contact your Fujitsu sales representative, or visit our website.

http://ts.fujitsu.com/products/bs2000/software/cluster_technology/hiplexmscf.html

Fujitsu green policy innovation

Fujitsu Green Policy Innovation is our worldwide project for reducing burdens on the environment. Using our global know-how, we aim to resolve issues of environmental energy efficiency through IT. Please find further information at:

<http://www.fujitsu.com/global/about/environment/>



Copyright

© Copyright 2011 Fujitsu Technology Solutions GmbH

Fujitsu and the Fujitsu logo are trademarks or registered trademarks of Fujitsu Limited in Japan and other countries.

Other company, product and service names may be trademarks or registered trademarks of their respective owners.

Disclaimer

Technical data subject to modification and delivery subject to availability. Any liability that the data and illustrations are complete, actual or correct is excluded. Designations may be trademarks and/or copyrights of the respective manufacturer, the use of which by third parties for their own purposes may infringe the rights of such owner.

Contact

FUJITSU Technology Solutions GmbH
Address: Domagkstraße 28, 80807 Munich, Germany
E-mail: bs2marketing@ts.fujitsu.com
Website: <http://ts.fujitsu.com/bs2000>
2011-10-18 EM EN