FUJITSU

WHITE PAPER

BS2000/OSD
S series Global Storage

Issue March 2010

Pages 4

The BS2000/OSD answer to data bottlenecks

The overall performance of an IT system is determined not just by the processor performance but also, crucially, by the speed of access to data and by the data throughput rate. The resulting response times and transaction rates are key factors in the successful and economical operation of an IT installation.

IT systems must observe given service-levels for different applications and operation types like OLTP, Batch, Data Warehouse and SOA applications.

When, with peak loads there come up slow response times or when in spite of high CPU power the time window for business critical batch runs becomes too small, then the reason usually may be found with data bottlenecks.

Data access bottlenecks are caused by the still existing performance gap between the CPU and the IO to the non volatile storage medium disk.

Our answer to data bottlenecks is the non-volatile semiconductor memory global storage.

Using global storage allows to increase - about factor 100 - the throughput and access to performance-critical data over disk systems even with cache facilities.

The global storage (GS) of the S-series BS2000/OSD Business Servers offers considerable additional benefits in terms of data-access speed, data throughput and overall system performance. It is ideally suited to high-intensity access to performance-critical data, e.g. for transaction processing and for databases, but also for any other frequently accessed files.

It is quite simple to use. Global storage can be used with volume emulation GSVOL like one or several disks and/or with the intelligent caching function AutoDAB from DAB (disk access buffer) as a CPU-like disk cache.

Contents

What is global storage

Global storage is a system extension to the S-series BS2000/OSD Business Servers.
It consists of a non-volatile semiconductor memory which is connected to the processor system via a special high-performance interface.
Rather than being just an "extension of the main memory", it is an "external device" to which data is written at enormous speeds, and from which it can be read just as quickly.
There are no mechanical rotary movements as with magnetic disks, no mechanical positioning times for read/write heads, and no wait times like those incurred with disk I/Os through the serialization of parallel accesses.
The connection to the processor system is not via classic I/O systems with channels, SCSI interfaces or Fibre Channel, but via a special high-performance interface directly at the system boards.
This interface allows an effective data throughput rate of up to 1.8 Gbyte per second. It's some 100 times higher than the maximum throughput rate of a disk on a classic ESCON fiber-optic channel and still 20 times higher than throughput rates of fibre channel attached disk systems at S servers.
Global storage is an ideal storage medium for writing and reading frequently accessed data.

Notable technical highlights of global storage

The global storage holds the customer's most important application data. Every precaution has been taken to ensure that global storage provides the highest possible security, incorruptibility and availability for this data.

Fail-safety

The Global Storage is a non-volatile data storage medium. GS data persist as long as power is supplied even when the server fails. To avoid loss of data in case of power outage a dual power supply is available as an option.
In addition, in the event of power failures an optional battery is activated, guaranteeing preservation of the data in the Global Storage for up to 24 hours.

Redundant data storage and disaster protection

If two global storage units are used, the data can be mirrored, i.e. kept in duplicate, by way of the "hardware duplication" function. The data will then continue to be available in the event of a fault.
The second global storage unit is built in another cabinet. This separation provides additional protection for GS data.

Recoverability after server failure

All metadata of the files is kept in the global storage so that, following a server failure and restart, you can continue working with the data in the global storage immediately.

System embedding as device

In the global storage GS volumes are configured as normal devices, i.e. applications can work unchanged with the GS volumes in the same way as with normal disks.

Expandability in addition to main memory

The global storage consists of one or two independent hardware units, each currently expandable to 64 Gbytes.

Usage scenarios of global storage in BS2000/OSD

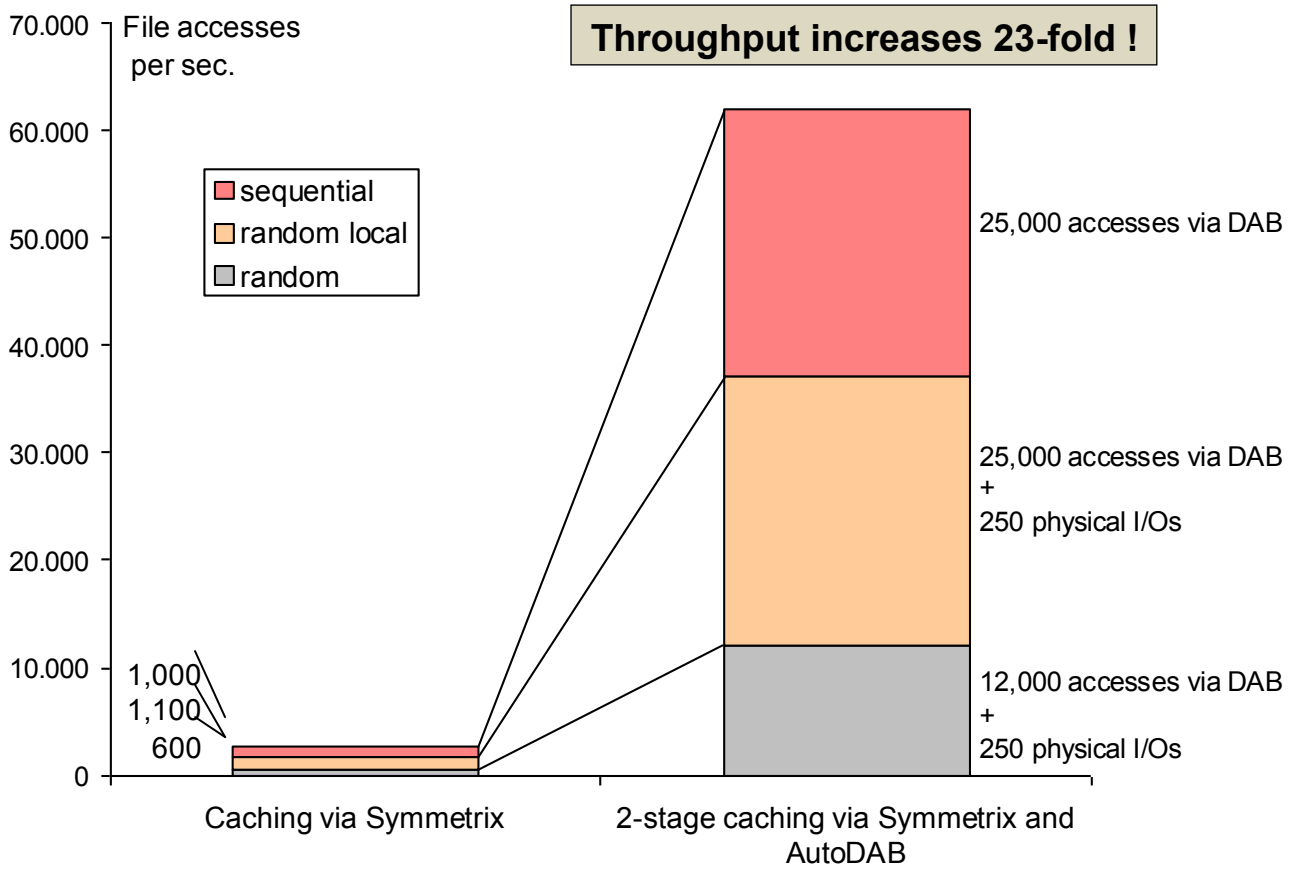
To allow users to get the full benefit of global storage, this storage medium is supported by the BS2000/OSD operating system as follows:

Using global storage as read/write cache

With the product DAB (disk access buffer), the global storage can be used as a CPU-like read/write cache for files. The data throughput is many times faster than with conventional disk caching.
With AutoDAB, the files suitable for caching are determined automatically through cyclical monitoring, i.e. files with bad cache hit rates or low access frequency are displaced from the global storage to make room for those which are frequently accessed.

Performance measurement results that speak for self

The diagram below shows a comparison using a disk benchmark between a Symmetrix 8430 Disk Storage System with 8 Gbytes disk cache on the one hand, and additional CPU-like caching, also with 8 Gbytes global storage, via AutoDAB (automatic disk access buffer) on the other.
The benchmark runs were performed on an S170-80 BS2000/OSD Server with a mixed access profile: one part sequential access, one part random access without locality, and one part random access with good locality. The data volume was twice the cache size (of which 6% seq. / 47% random / 47% random with temporal locality).



Explanation of the benchmark:
The total throughput of the benchmark is significantly increased by virtue of the two-stage caching.
Why does this increase occur?
The better the hit rate in a DAB cache, the more existing CPU performance can be converted into I/O throughput.

- Because of the very large prefetch (data areas are read into the cache in advance), the sequentially processed files register a very high hit rate and are then kept practically resident in the DAB cache on account of the low data volume
- Because of the intelligent algorithms of AutoDAB, the random files with temporal locality also have very high hit rates and occupy only a small proportion of the cache storage
- Because DAB uses the cache storage so efficiently, another large proportion is available for caching the files with a random access pattern, which also helps increase the throughput.

Even in case of high performance disk systems connected to the server via fibre channel, the Global Storage provides significant higher performance (more than factor 5).

What can you achieve with a smaller global storage?

Because of the AutoDAB algorithms, with a smaller DAB cache the random files without locality are first "deselected" from the cache, i.e. no longer cached until all accesses to these files only occur on the disk subsystem.
If the cache is further reduced in size, random files with temporal locality are (partially) displaced from the cache. The increase in throughput here will remain constantly high.

High availability in database operation

In addition to the acceleration of database write I/Os and the associated improvement in user response times, GS can also be used as a precaution in case of errors.
Our measurements have shown that, when GS is used, recovery of a database following errors can be **5 times** faster.
This means a significant reduction in downtime after unscheduled stoppages

Using global storage for file storage - emulated volumes

On the global storage, so-called GS volumes can be set up. This involves the emulation of disk volumes on which the applications and their data can work in the same way as on real disk storage, but significantly more efficiently.
In a HIPLEX coupled via global storage (parallel HIPLEX) the data can be accessed simultaneously from up to four real and 15 virtual systems.
As with conventional disk operation, backup tools such as HSMS and ARCHIVE are available for the data backup and archiving of GS volumes.
The fast GS access ensures that, during data backup and reconstruction, the magnetic tape cartridge periphery used for backup can work at the maximum data rate.

The usage of global storage depends essentially on the temporal and geographical distribution of the file accesses.

- Performance-critical files which fit fully into the global storage should be stored on **GS volumes**. This produces the shortest access times and constantly high access rates.
  Examples of this are KDCFILE from UTM and database log files.
- For performance-critical files that are too large to fit fully into global storage, the GS should be used as **read/write cache with AutoDAB**. By exploiting temporal and geographical localities, you can thus achieve very short access times and high throughput rates.

VM2000 and virtual global storage

Global storage is also supported in the virtual machine concept VM2000.
Here, global storage partitions can be allocated to the guest systems as virtual storage either exclusively or for shared access.

Using global storage in a HIPLEX

In a parallel HIPLEX the global storage can be exclusively allocated to the real or virtual systems and can be used for shared access to shared data by several systems.
Moreover, the signaling function of global storage can be used for very fast communication between multiple systems, e.g. for a lock protocol which coordinates the accesses to shared resources or data.

System embedding of global storage

Read/write operations on devices such as disks occur asynchronously, i.e. they cause the processor to switch tasks with a considerable administration overhead. Writing on and reading from global storage, on the other hand, occurs synchronously, i.e. I/O operations do not cause a change of task and therefore a loss of context. This results in an higher I/O speed and improvement in overall system performance.

Customer application of global storage

Global Storage has been used in numerous installations for years.
With its efficient functions and operating-system support in BS2000/OSD, it optimizes customer operations in terms of I/O performance, data throughput and overall operating behavior, thus making a vital contribution to the successful and economical running of customer installations and to the satisfaction of their users.

Published by department:

Product Marketing BS2000/OSD
Phone: ++49 89 62060 1977
Fax: ++49 89 62060 3291977
bs2marketing@ts.fujitsu.com

Fujitsu Technology Solutions GmbH
ts.fujitsu.com