Environmental Considerations
White Paper

VERSION 2020

Contents

Environmental Considerations for Deploying JADE

The objective of this white paper is to outline environmental considerations required to deploy JADE systems to achieve a high degree of confidence in the levels of data integrity and availability that can be maintained in the event of catastrophic failures.

Owing to the vast range of commodity hardware available, the information in this document does not attempt to provide complete environmental specifications but focuses on the critical components of memory and secondary storage. You should therefore read it in conjunction with a number of other available documents.

For minimum hardware requirements and a link to example production hardware configurations, see "Hardware Requirements" under "Operational Requirements" at:

- https://secure.jadeworld.com/JADETech/JADE2020/Docs/InstallConfig.pdf

See also operating system requirements, at https://www.jadeworld.com/jade-platform/developer-centre/learn/documentation/operational-requirements.

Note   As these recommended configurations are the minimum specifications required to deploy JADE, they are not sufficient for high-availability systems.

Types of Failure

An unfortunate aspect of every database system is the possibility of failures due to hardware faults or environmental issues that result in data corruption or the complete loss of data.

Some common types of failure include:

- Media failure, when one or more of the disk drives holding a database fail and you are faced with a complete loss of data unless you can restore and recover from a backup of the database.

  Note   It is essential that you understand that in certain redundant disk configurations, a common controller, cable, or I/O bus problem can corrupt the data on multiple physical drives.

- Application errors, when an application makes a large number of invalid modifications to data and the best way to deal with the problem may be to restore and recover the database from a backup to a specific date and time before the modifications were made.

- Software failure, when the operating system, device drivers, or environmental software incurs a problem that leads to file system or database corruption.

- Permanent loss of a server, when a server is disabled permanently or a site is lost due to a natural disaster.

Machine Profiles

The machine processing profiles whose environmental considerations are described in this document are:

- Machines not Hosting a JADE Database Server Node

- Machines Hosting a JADE Database Server Node

Machines not Hosting a JADE Database Server Node

The minimum requirements to satisfy the maintenance of data integrity on machines that update persistent objects (in particular, application server and standard client nodes) are:

- Server-class hardware with Error Checking and Correcting (ECC) memory

- A suitable application server or standard client operating system, as listed at
  https://www.jadeworld.com/jade-platform/developer-centre/learn/documentation/operational-requirements
  (for example, Windows 2008 Server)

Machines running thin client software only are not affected by these requirements, as a thin client does not directly update objects.

Machines Hosting a JADE Database Server Node

When a machine hosts a JADE database, minimum requirements must be met to satisfy the maintenance of data integrity in the case of hardware or environmental failures.

There are further requirements that must be met to satisfactorily provide for data recoverability. In this document, we consider these two processing profiles as:

- Non-Business-Critical Systems
- Business-Critical Systems

Non-Business-Critical Systems

If it is acceptable to your business processing to lose all transactions since your last full database backup in the event of a catastrophic failure, you can operate your JADE database with archival recovery disabled and the following base configuration.

- Mandatory

  - Server-class hardware with ECC memory

  - A suitable database server operating system, as listed at https://www.jadeworld.com/jade-platform/developer-centre/learn/documentation/operational-requirements (for example, Windows 2008 Server)

  - Disk subsystem guarantees that acknowledged writes are non-volatile (for details about non-volatile disk subsystems, see "Storage Requirements", later in this document)

- Recommended

  - Configure a separate disk subsystem for disk-resident database backups

Business-Critical Systems

If it is not acceptable to your business processing to lose committed transactions in the event of a catastrophic failure, you must operate the database with archival recovery enabled, by using the base configuration documented under "Non-Business-Critical Systems", earlier in this document, and with the following additional requirements.

- Mandatory

  - Transaction journals must be located on a separate physical disk subsystem or volume to that hosting database files

- ▫ Database journal volume must be mirrored, and ideally duplexed (that is, have a dual controller)

- ■ Recommended

    - ▫ Database files should be located on a fault-tolerant disk subsystem

    - ▫ Configure separate physical disk subsystem for disk-resident transaction journals and database backups

    - ▫ Use a local tape backup unit or network backup strategy to allow database and journal backups to be moved off the machine that is hosting the database

        Note   Separate physical disk subsystem implies a separate I/O channel, disk controller, and at least one disk drive

See also "Storage Requirements", later in this document.

Virtualized Operating System Environments

Jade Software Corporation (JSC) provides support to customers who run the JADE product on supported native operating systems, subject to conditions described later in this section. JSC also supports the JADE product running on enterprise-class virtualized environments using specified operating systems and hardware platforms. For details, see at https://www.jadeworld.com/jade-platform/developer-centre/learn/documentation/operational-requirements .

However, you should be aware that JADE database recovery cannot be guaranteed when using virtualized environments with a disk subsystem that has write caching enabled, unless acknowledged writes are guaranteed to be non-volatile under all failure conditions. (For details, see "Storage Requirements", later in this document.) For example, at least one virtualized platform provider, VMware, supports a published set of certified operating systems and hardware. You and VMware are responsible for any interactions or issues that arise at the hardware or operating system layer as a result of your use of VMware.

See http://www.vmware.com for a list of VMware certified platforms.

Although JSC will not require you to recreate and troubleshoot every issue in a virtualized environment, we *do* reserve the right to ask you to diagnose certain issues in a native certified operating system environment operating without the virtual environment.

JSC will make this request only when there is reason to believe that the virtual environment is a contributing factor to the issue. Any time spent on investigation of problems that may, in the sole opinion of JSC, be related to virtualization will be handled in the following fashion.

1.   JSC will provide standard support for the JADE product.

2.   If a problem is encountered while JADE is running in a virtualized environment, you may be required to recreate the problem on a non-virtualized server unit, at which time JSC will provide regular support.

3.   You can authorize JSC to investigate the virtualization-related items at an agreed time-and-materials rate. If such investigation shows that the problem is related to virtualization, you can contract JSC to provide a software change to JADE in order resolve the issue, assuming such a resolution is possible.

4.   Regardless of the problem type or source, if the problem is determined not to be related to platform virtualization, time spent on investigation and resolution will be covered as part of regular maintenance, and support will be provided as usual.

Environments Using Windows Clusters

The JADE product is not currently "cluster-aware". As a result, a cluster fail-over when a JADE database is open will most likely result in a database recovery when the system restarts after fail-over. When planning clustered configurations, you should take special care to meet the storage requirements documented under "Storage Requirements", later in this document.

You may need to conduct your own verification tests for a range of scenarios, to gain confidence that database recovery is possible after typical failure modes.

Disk Subsystem Allocation Strategies

Redundant Array of Inexpensive Disks (RAID) is an important component for servers used for business-critical data processing. RAID systems have several configurations (levels) that have different performance profiles and fault tolerance capabilities.

In general, we recommend the use of mirroring (RAID 1) to achieve redundancy combined with striping (RAID 0) for performance.

The RAID levels 3 through 6, which implement different forms of parity, typically provide poor write performance and are not recommended for hosting database and transaction log files. On the other hand, a RAID 5 or RAID 6 subsystem may be suitable for use for database backups where capacity is more of a consideration than performance.

The following allocation strategy employs a combination of RAID levels, depending on the file types.

RAID Level 1, Level 0+1, or Level 1+0 (RAID 10)

- Database files
- Transaction journals
- Operating system files
- Database backups (performance)

RAID Level 0

- Transient database files
- Operating system page files
- Software code files

RAID Level 5 or Level 6

- Database backups (capacity)
- Transaction journal backups

Storage Requirements

The following subsections cover storage requirements. When planning, deploying, and maintaining a JADE database server installation, you should ensure that the I/O system supports all of the factors outlined in this article.

- Stable Media

- I-O Basics

- Forced Unit Access (FUA) and Write-Through

- Write Ordering

- Compression

- Network File Systems

Stable Media

JADE uses the Write-Ahead Logging (WAL) protocol to maintain the Atomicity, Consistency, Isolation, and Durability (ACID) properties of the database and to ensure data integrity. WAL requires stable media capabilities. A storage solution must comply with this stable media intention.

In the event of system failures, it is vital that storage solutions used to host JADE database files and transaction journals guarantee that acknowledged write operations are non-volatile, no matter what fails.

If this requirement is not met, data can be lost in the event of power failures or the resetting of the server. Such data loss can result in an irrecoverable file system or database corruption.

I/O Basics

In general, I/O requirements for JADE systems related to stable media are very similar to those published for Microsoft SQL Server products. These requirements are discussed at length at http://technet.microsoft.com/en-us/library/cc966500.aspx.

Microsoft publishes a utility, SQLIOSim, which simulates SQL Server activity on a disk subsystem. This utility is discussed at and can be downloaded from https://support.microsoft.com/en-us/kb/231619. We recommend that you use this utility as part of verification that your operating system environment and hardware configuration meets minimum basic requirements. More information about the utility is available at https://techcommunity.microsoft.com/t5/sql-server-support/how-it-works-sqliosim-checksums/ba-p/315463.

Note that a successful result using the SQLIOSim utility with a specific configuration does not provide any assurance that writes will be persisted on disk under all conditions such as system reset and power interruption. For critical applications, you may need to conduct additional testing to include such conditions, in order to establish a high level of confidence in your configuration.

Forced Unit Access (FUA) and Write-Through

All components in a storage solution must honor the write-to-stable media intent.

A solution that has write-back caching enabled must guarantee acknowledged writes are non-volatile under all failure conditions. JADE uses two mechanisms to achieve the stable media requirement.

Note Using an external Uninterruptible Power Supply (UPS) to achieve this on its own is not sufficient, because failure modes that are unrelated to power *do* occur.

- A **Write Through** command, which is available only for Small Computer System Interface (SCSI) disk devices, is implemented by issuing a **WRITE** command to the disk with the Force Unit Access (FUA) bit set. This command is used by JADE database journaling, and it instructs the disk to write the current packet of data to the disk immediately, bypassing the onboard write cache.

  Storage solutions that do not support FUA semantics (for example, Integrated Drive Electronics (IDE)/ATA Packet Interface (ATAPI) disks) must be configured with write-back caching disabled.

  Legacy IDE drives provide no support for the equivalent of SCSI FUA in drives based on any version of the Advanced Technology Attachment (ATA) specification prior to version 7. That means that if you were to host journals on an IDE drive, controller write caching must be disabled, which is not usually the default. This issue was addressed in ATA-7, on which most-current Serial ATA (SATA) drives are based.

  If the write journal drive is an ATA-type device, you cannot rely on data being stable on the media for the **WRITE** command with FUA. Although it has been the standard since ATA-7 (2005), it is not always supported by commodity drives or device drivers. From JADE releases 7.1.09 and 16.0.02, writes to the audit files are now synchronized on such devices.

- A **Flush Buffers** command, which is available for SCSI and IDE/ATAPI disk devices, instructs the disk to write all cached data to the disk immediately. For SCSI disks, this is implemented by issuing the **SYNCHRONIZE CACHE** command to the disk. For IDE/ATAPI disks, the **FLUSH CACHE** command is sent to the disk.

Caching storage solutions designed for use in data-critical transactional environments may not honor these commands immediately. Such solutions *must* guarantee that, after a failure such as power loss or system reset, cached data is eventually persisted on the physical media.

As each new volume is encountered when opening a writable file, the JADE database logs the device write-cache settings. If the device reports that write-cache is enabled and the write-cache type is not write-through, warning messages are logged if the device does not support flush-cache operations or you have configured the device as power protected (suppressing flush-cache operations).

If the device settings cannot be determined, the following warning is logged.

```
<<WARNING>> Unable to determine write-cache settings for volume <volume>, drive
letter=<drive>, label=<label>
```

If write-cache is enabled and it is not write-through but the device reported flush-cache as not supported, the following warnings are logged.

```
<<WARNING>> Unsafe write-cache settings for volume <volume>, drive letter=<drive>,
label=<label>

<<WARNING>> Device does not support host software forced flush of device
write-cache
```

If write-cache is enabled and it is not write-through but you have configured the device as power protected on the **Policies** sheet in device properties, the following warnings are logged;

```
<<WARNING>> Possibly unsafe write-cache settings for volume <volume>, drive
letter=<drive>, label=<label>

<<WARNING>> Host software forced flush of device write-cache is disabled
```

If a warning is logged for a device, it means that the reported device configuration does not guarantee that acknowledged writes are non-volatile and therefore the device configuration does not comply with the database stable media requirement.

If a warning was logged, an indication of the implications of a failure for the system is also logged.

If archival recovery is enabled, the following message is logged.

```
A power outage might result in data loss or corruption necessitating roll-forward
recovery from a checked backup
```

If archival recovery is disabled, the following message is logged.

```
A power outage might result in irrecoverable data loss or corruption (archival
recovery is disabled)
```

Write Ordering

Database journaling and recovery requires write-order preservation. This write-ordering requirement applies to both local and remote I/O destinations.

When a database is split across physical devices, all devices must honor the ordering across all database files. Specifically, when a write **w1** is issued to a device **d1** and is acknowledged as completed before a write **w2** to device **d2** is initiated, the data written by **w2** cannot be present on physical media if the data written by **w1** is not present.

Compression

Locating JADE database files or journals on compressed drives or directories is not supported.

The issue is that when compression is used, the actual file data is handled by the operating system in large chunks. When the database engine updates a record or block, therefore, the system is really working on a large chunk of data and rewriting it. The data rewrite breaks the WAL protocol because data written to disk will be rewritten, thereby breaking write ordering rules. The rewrite can lead to caching and other activities that are not appropriate for database ACID properties.

All database journal and data files that can be updated must remain in an uncompressed state.

Network File Systems

Owing to the complexity and range of performance configuration options in network file servers combined with the issues of distributed cache and write-back caching on the file server, we strongly recommend that you do *not* locate JADE database files and transaction logs on remote network file systems. This includes most vendors' implementation of a Network Attached Storage (NAS) system; that is, a disk system that clients attach to through a network redirector using a network protocol such as Transmission Control Protocol/Internet Protocol (TCP/IP).

In contrast, a Storage Area Network (SAN) operates using access methods and technologies such as SCSI and fiber-channel, which extend the reach and capacity of the local disk subsystem. Block storage protocols such as Internet Small Computer System Interface (iSCSI) that operate over TCP/IP are an alternative to fiber-channel.

When considering an iSCSI storage solution on a Windows platform, we recommend that you verify that the solution implements all mandatory requirements documented in the latest version of *Microsoft Windows Logo Program System and Device Requirements* on the Microsoft web site.

SAN drives associated with a specific server appear as local disk drives to the operating system and support the same file access methods as any other locally attached storage. This type of technology is often referred to as *channel attached storage*, and doesn't suffer from the reliability issues associated with network file systems.

Note   Locating database transaction journals on network file systems is explicitly disallowed.

Backup, Recovery, and Restore Strategy

Even when you have taken into account other items within this document, to help ensure the overall integrity and availability of the system, a well-tested backup, recovery, and restore process is required.

When developing the strategy for your backup, recovery, and restore process, see the *Developing a Backup Strategy* white paper .

Virtualization Snapshots

A plethora of solutions and add-ons exist in the marketplace, intended to provide alternative backup facilities using virtual machine snapshot technology. JSC has not qualified any of these backup solutions for use with the JADE product, and does not support their use.

This is a similar position to that taken by Microsoft Corporation with respect to their SQL Server product.

The recommended approach to backing up JADE systems is documented in the *Developing a Backup Strategy* white paper. For more details, see the "Non-JADE Backups" section in that white paper.