



## **Connexan Enterprise iSCSI Target (CiT™) Best Practices**

Connexan Inc.

Published: June 28, 2007

Authors: Connexan and IBM

### **Abstract**

This white paper describes the best known practices and recommendations for planning, installing, and configuring the Connexan Enterprise iSCSI Target (CiT™). The paper is a work in progress and is updated as new knowledge is acquired and new products and updates are released to the market.

# Contents

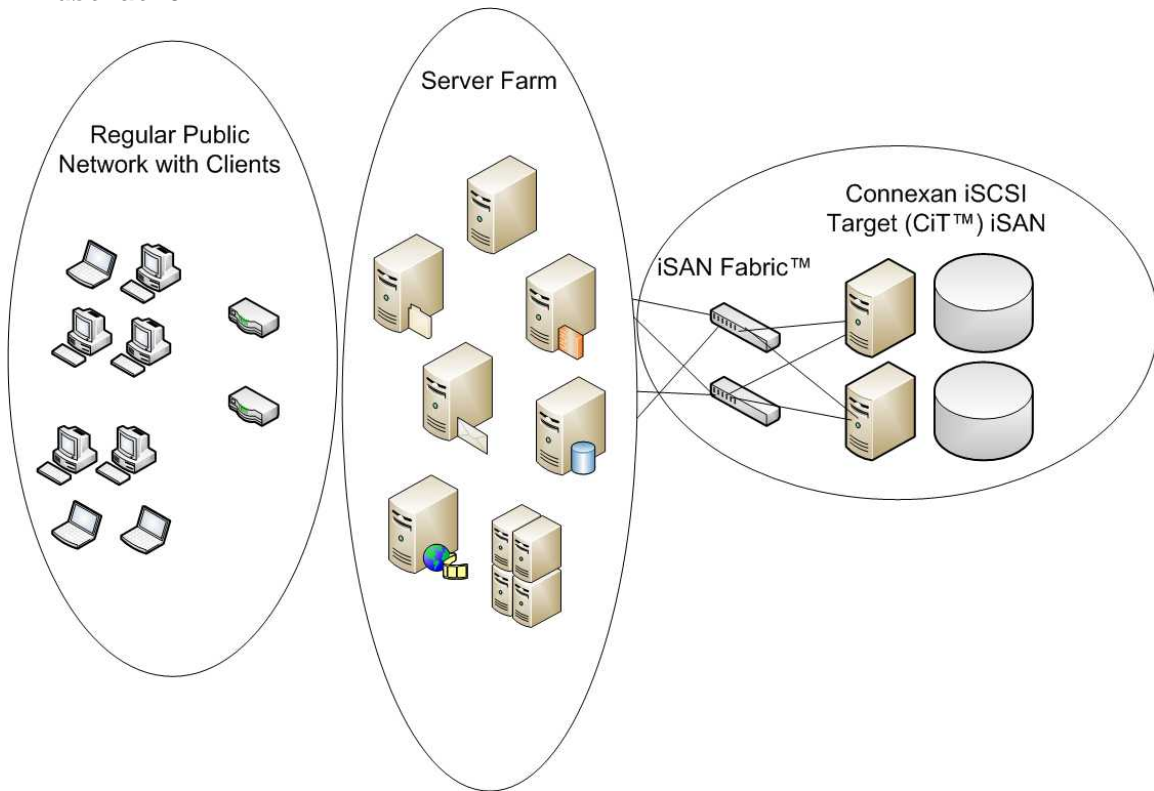
Connexan Enterprise iSCSI Target (CiT™) Best Practices .....	1
Abstract .....	1
Contents .....	2
Introduction .....	3
Best Practices.....	4
Microsoft Windows Server x64 Enterprise or Datacenter Edition should be used for the CiT™ Storage Nodes .....	4
A dedicated domain should be assigned to the CiT™ Storage Cluster .....	4
Microsoft Scalable Networking Pack should be installed .....	4
Modern network adapters should be used .....	5
CiT™ Storage Node network adapters should be teamed .....	5
Dedicated Gigabit or 10Gb Ethernet Switches should be used .....	5
iSCSI HBAs are for Initiators, not Targets.....	5
CiT™ Storage Nodes should be directly attached to redundant UPS .....	6
CiT™ Storage Node cache should be committed on a regular basis .....	6

# Introduction

The Connexan™ Enterprise iSCSI Target (CiT™) is a robust, high-performance, iSAN-enablement technology that is designed for Windows. The target is capable of providing shared, iSCSI-based, block-level access disk to a multitude of client initiators. The CiT™ is not limited to providing disk only to Windows-based systems, for example, many Linux client initiators are supported (and in use at customers today).

To better understand the concept behind the CiT™, see Illustration 1 below. As opposed to using traditional, expensive, fibre-channel connectivity for high performance, centralized data storage for servers in a server farm, the CiT™ creates an iSAN Fabric™, based on standard, high-performance Gigabit or 10Gb Ethernet technologies.

Illustration 1



# Best Practices

## **Microsoft Windows Server x64 Enterprise or Datacenter Edition should be used for the CiT™ Storage Nodes**

Even for customers not planning to implement high availability initially, it is a good idea to have that capability in place. Many CiT™ customers whom have installed the CiT™ have dramatically exceeded their initial expectations of what they would implement on it. Some of those that installed using a Standard Edition of Windows have had to rebuild the CiT™ Storage Nodes in order to implement high availability.

Although the 32-bit editions of Windows Server are stable and perform extremely well, Connexan still recommends that customer begin using the x64 editions. This enables the customer to further increase CiT™ Storage Node cache in the future without modifying the underlying Operating System. It also “future-proofs” the CiT™ Storage Nodes to a certain degree, allowing integration with future Windows-based enhancements that are only available to x64-based systems. Windows x64 editions are also showing performance scalability that 32-bit systems will not be able to match as newer I/O technologies are released to the market, for example, 10Gb Ethernet, Infiniband, High-speed Virtual I/O technologies, and more.

## **A dedicated domain should be assigned to the CiT™ Storage Cluster**

The CiT™ utilizes Microsoft Cluster Server for its high availability solution. The Storage Cluster should be treated as an integrated storage controller, not as application servers. Access to the Storage Cluster should be restricted to only iSAN and SAN administrators. They should be properly trained and knowledgeable in CiT™ management practices, and have a need to modify the back-end storage solution of the CiT. Individuals that just need to create or modify iSCSI Devices in the CiT™ do not require access directly to the CiT™ Storage Cluster. Under no circumstances should general domain or application administrators have access to the CiT™ Storage Cluster console or services.

## **Microsoft Scalable Networking Pack should be installed**

To achieve the benefits of Broadcom TOE and Intel I/OAT technologies, as well as other modern network performance improvements, the Microsoft Scalable Networking Pack (SNP) must be installed on the CiT™ Storage Nodes. The SNP enables the TCP Chimney Offload (TOE) as well as Receive Side Scaling (RSS, ability to multi-thread network I/O requests) and NetDMA (Microsoft Network Direct Memory Access).

Connexan recommends that all CiT™ Storage Nodes run Windows Server 2003 Service Pack 2 or later. Microsoft has included and enabled the feature set of the SNP in Service Pack 2, so no further installation or system modification should be necessary.

A technical overview of these technologies is available from Microsoft at:

<http://www.microsoft.com/technet/community/columns/cableguy/cg0606.msp>

The Microsoft Scalable Networking Homepage is available at: <http://microsoft.com/snp>

## **Modern network adapters should be used**

Modern network adapters from Broadcom and Intel should be used on the CiT™ Storage Nodes.

Broadcom has implemented a technology called TOE (TCP Offload Engine) on its modern server Gigabit Ethernet adapters. TOE dramatically reduces the demand placed on the CiT™ Storage Node system processor for I/O. TOE also allows the CiT™ Storage Nodes to achieve line-level performance on Gigabit Ethernet. To see which Gigabit Ethernet adapter chipsets by Broadcom support TOE, see: <http://www.broadcom.com/products/technologies/Gigabit-Ethernet>

Intel has taken an alternative approach to TOE named I/OAT (Intel® I/O Acceleration Technology). I/OAT is a TCP onloading technology, as opposed to the offloading offered by TOE. In our experience, it does not achieve the same dramatic performance improvement that TOE adapters do. I/OAT also does not significantly reduce CiT™ Storage Node processor overhead, but it does improve the manageability of the network data throughput, whereby providing a significant performance increase over traditional Gigabit Ethernet networking. I/OAT is usually a less expensive option to implement than TOE, as well. I/OAT is also supported in Linux as of the time of this paper, whereas TOE is not. I/OAT requires that the CiT™ Storage Node have an Intel Northbridge chipset, whereas TOE is chipset independent. For more information on Intel® I/OAT, see: <http://intel.com/go/ioat>

The most significant portion of data throughput overhead in pre-modern network technologies has been processor interrupts – every time data is received, the processor must be interrupted and asked if it is ok to process the data and write it to/through memory. Both TOE and I/OAT technologies are DMA (Direct Memory Access) based. This means that interrupts are reduced to a small fraction of what they previously were, and ensures that both technologies, properly configured, can significantly improve I/O performance on CiT™ Storage Nodes.

## **CiT™ Storage Node network adapters should be teamed**

For both reliability and performance, the network adapters functioning to provide access to the iSAN should be teamed. Refer to the network adapter's user guide for information on how to team the adapters in use on your CiT™ Storage Nodes.

## **Dedicated Gigabit or 10Gb Ethernet Switches should be used**

A Connexan iSAN Fabric requires a dedicated network. Although existing network infrastructure could be used by utilizing VLANs, Connexan strongly urges customers to purchase dedicated iSAN Fabric Gigabit Ethernet or 10Gb Ethernet switches.

When comparing the cost of a Connexan iSAN with that of a traditional, fibre-channel SAN, the savings are tremendous. The cost of two, dedicated switches for the iSAN Fabric is less than that of a single fibre-channel switch. The performance, security, and manageability benefits gained by dedicating a physical network to the Connexan iSAN should not be overlooked.

## **iSCSI HBAs are for Initiators, not Targets**

Connexan highly recommends iSCSI HBAs for client initiators connecting to the CiT™ iSAN Fabric™. iSCSI HBAs allow the client initiator to boot from the iSAN. They also implement a native version of TOE within their drivers. iSCSI HBAs dramatically reduce client initiator processor overhead when accessing iSAN data. They also offer increased manageability.

Connexan presently recommends and supports QLogic iSCSI HBAs. See: [http://qlogic.com/products/iscsi\\_products.asp](http://qlogic.com/products/iscsi_products.asp)

The CiT™ Storage Nodes do not presently support iSCSI HBAs. iSCSI HBAs are designed to offer iSCSI initiator support. The CiT™ Storage Node is an iSCSI target and is not able to take advantage of the benefits of iSCSI HBAs at this time. Provided modern network cards are used, HBAs are not necessary at the CiT™ Storage Node target level.

VMware ESX 3.0.1 support iSCSI, but does not support TOE or I/OAT. To achieve maximum performance as well as reduce the overhead on system processors (vital for VMs), it is highly recommended that all VMware ESX 3.0.1 implementations utilize the QLogic iSCSI HBAs.

### **CiT™ Storage Nodes should be directly attached to redundant UPS**

Each CiT™ Storage Node should be equipped with redundant power supplies. The power supplies should be connected to separate, dedicated, Uninterruptible Power Supply units, shared between the CiT™ Storage Nodes. This configuration is similar to that of integrated fibre-channel SAN controllers, but at a significantly lower price point with increased manageability and uptime.

The sync.exe program, from Microsoft's Sysinternals division, should be scheduled to be executed by the UPS management software in the event of a power loss. This will ensure that any data remaining in the CiT™ Storage Node system cache is committed to disk and not lost.

### **CiT™ Storage Node cache should be committed on a regular basis**

The Connexan iSCSI Target implements advanced caching algorithms in available in Windows. This allows the customer to purchase multiple gigabytes of RAM, as cache, for the CiT™ Storage Nodes at a fraction of what that same amount of cache would cost in proprietary fibre-channel storage controllers. The integrity of the data maintained in cache on CiT™ Storage Nodes is vital, though. Connexan is presently working on developing automated technologies for cache commitment, but as of today, cache commitment should be scheduled as a manual process utilizing Microsoft's Sysinternals sync.exe command.