# Compaq SANworks™

## Application Notes
## Data Replication Manager
## Over a WDM-enabled Intersite Link

---

*This document details an extended configuration that is available for Data Replication Manager. This Disaster-Tolerant (DT) solution provides controller-based mirroring across an extended Fibre Channel link. For complete details on Data Replication Manager, refer to the Data Replication Manager HSG80 ACS V8.5P Operations Guide.*

---

First Edition (March 2000)
Part Number: EK–DRMWD–AA. A01
**Compaq Computer Corporation**

# Notice



Printed in the U.S.A.

Application Notes - Data Replication Manager Over a WDM-Enabled Intersite Link
First Edition (March 2000)
Part Number EK–DRMWD–AA. A01

# Application Notes Contents

These application notes contain information on the following topics:

# Disaster Tolerance and Data Replication Manager

During normal data processing, data is simultaneously written to both initiator (local) and target (remote) sites. While copies of data reside at both sites, all host data access occurs through the initiator site. The initiator site system performs all data acquisitions unless a failure or catastrophe occurs that disables processing at that site.

With Disaster Tolerance (DT), your operating system can detect hardware and software failures that affect the initiator site storage. In the event of an initiator site failure, another site can continue processing the data in the interim.

The Data Replication Manager provides rapid data access recovery and continued data processing after the loss of one or more components. Data Replication Manager, in conjunction with the HSG80 Array Controller, is used on a controller host-port-to-controller-host-port basis. This allows data to be migrated synchronously from one storage subsystem to another—even if the subsystems are located at different physical sites.

The standard disaster-tolerant (DT) solution requires two different types of fiber optic cables and the addition of Gigabit Interface Converters (GBICs; depending on the connection). Each site houses Fibre Channel switches, hosts, and controllers. Host bus adapters establish an interconnect between the controllers and host, while short-wave Gigabit Interface Converters (GBICs) connect the host and controllers to the switches.

The controller, the switch, and the host are cabled with 50 micron multi-mode fiber optic cables. These cables support a maximum length of 500 meters.

If the initiator and target sites are more than 500 meters apart, a single-mode fiber (SMF) cable is used to connect the two sites together. SMF cable works with long-wave GBICs and can typically span distances of up to 10 kilometers.

Figure 1 illustrates the traditional Data Replication Manager setup.

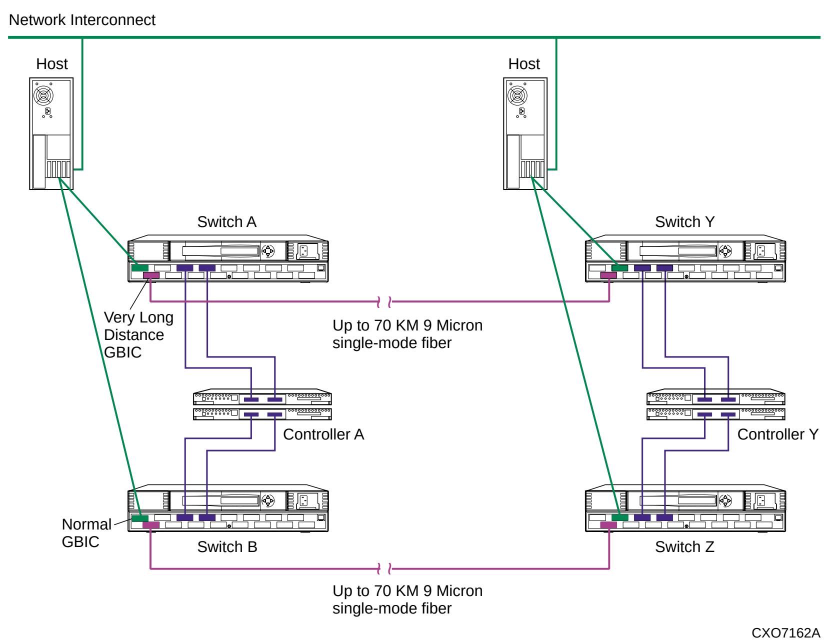


Figure 1. Data replication manager (typical configuration without WDM)

For distances of greater than 10 Km, system limitations become defined primarily by: data latency issues, signal attenuation, temporal jitter, and noise.

# Wavelength Division Multiplexing

Wavelength Division Multiplexing (WDM) is an optical technology used to add connection capacity over an existing fiber optic network. It works by combining and transmitting multiple optical signals simultaneously at different wavelengths down a single fiber. In effect, one fiber is transformed into multiple virtual fibers. WDM then, multiplies the effective bandwidth capacity of the optical fiber.

For example, by multiplexing eight 1 Gb/s signals into a single fiber, you would increase the data-carrying capacity of that fiber from 1 Gb/s to 8 Gb/s in aggregate. Single fibers are currently able to transmit data at speeds up to 400 Gb/s. Research is progressing toward adding more and more channels to a single fiber, increasing capacity accordingly.

The maximum data-carrying capacity that can be designed into a WDM system is strongly dependent on the spacing between the wavelengths being used. For silica fiber, the range of usable wavelengths lies within two spectral windows, nominally centered around 1300 and 1550 nm and approximately 30 nm wide.

The spectral windows occur within the infrared spectrum where the optical scattering and absorption losses associated with silica fiber are minimal and transmission is maximized. Fiber optic systems in general are currently designed to operate within these two transmission windows. WDM systems typically combine four to eight channels within either window where channel separations are on the order of the window width divided by the number of channels.The following is an example:

(30 nm window width) / (8 channels) = 3.75 nm channel spacing

Thus, the effects of upgrading a WDM system to provide more channels are:

- To decreases the spacing (and the optical isolation) between adjacent channels
- To increase the potential for channel-to-channel crosstalk

The most demanding of WDM configurations combine up to 80 channels within a single transmission window and are known as dense-WDM (DWDM) systems, where adjacent wavelengths are separated by sub-nanometer distances.

A key advantage to WDM is that most architectures are protocol and bit-rate independent. WDM-based networks can simultaneously transmit data in Fibre Channel, IP, ATM, SONET /SDH and Ethernet protocols, and can simultaneously handle multiple bit-rates.

## Commercial Applications

From a Quality of Service (QOS) viewpoint, WDM-based networks create a lower-cost way to quickly respond to customers' bandwidth demands and protocol changes.

Once WDM has been implemented, service providers can establish a grow-as-you-go infrastructure, allowing them to add current and next-generation time division multiplexing (TDM) systems for virtually endless capacity expansion. WDM also gives service providers the flexibility to expand capacity in any portion of their networks—an advantage no other technology can offer. Carriers can address specific problem areas that are congested because of high capacity demands. This is especially helpful where multiple rings intersect between two nodes, resulting in fiber exhaust.[1]

By partitioning and maintaining different dedicated wavelengths for different customers, for example, service providers can lease individual wavelengths—as opposed to an entire fiber—to their high-use business customers.

## System Characteristics

Acceptable and optimal WDM systems have certain key characteristics. These characteristics are recommended for any WDM system in order for carriers to realize the full potential of this technology.

The following characteristics are desired in a WDM system:

- WDM systems should use the full capacity of the customers' existing dark fiber.
- WDM systems should offer component reliability, 24x7 availability and expandability.
- Optical signal amplification and attenuation. Desirable on the client side and long-haul side interfaces to increase the transmitted/received signal-to-noise ratio.
- Signal conditioning (that is, the retiming and reshaping of the optical data-carrying signal), for optimization of the bit error rate.
- Channel add/drop capability (the ability to change the number of data channels by adding or dropping optical wavelengths at any network node location).
- Compensation of power levels (preferably automatic – without manual intervention), especially to facilitate adding (or dropping) channels.
- Upgradable channel capacity and/or bit rate. Each time the number of channels or the bit rate is doubled, 3 dB of additional signal-to-noise margin is needed.

The following features are required in a WDM system for inter-operability with DRM:

- Standards-compliant interfaces such as Fibre Channel, SONET, ATM, and so on.

1. Fiber "exhaust" is the situation that exists when the traffic volume on the Internet and other networks have exhausted collective bandwidth available through installed optical fiber lines.

# Configuration

The following configuration issues have been defined and are described in the following sections:

- Test Configuration: DRM over WDM
- Interfaces
- WDM Product Variations
- Extended Fabric
- Future Availability
- Configuration Notes and Recommendations

## Test Configuration: DRM over WDM

The configuration addressed in this application note and pictured in Figure 2 can be described as a disaster-tolerant data storage solution operating over extended WDM-enabled intersite links. The components of this DRM solution consist of:

- Four HSG80 Array Controllers running ACS V8.5P with remote copy functionality
- Four Brocade Silkworm optical switches
- Microsoft Windows NT V4.0 operating system running on two Proliant 5500R Servers, Secure Path (Raidisk)
- WDM hardware

The HSG80 Array Controller running ACS V8.5P along with the Brocade Silkworm switches, provides a DRM solution for distributed environments. With an array controller connected to a host at a primary data center, configurations can be created with local critical data replicated to remote storage site(s).

The addition of WDM hardware allows for multiplexing of multiple optical signals that can be amplified as a group and transmitted long distances over a single fiber. The configuration shown in Figure 2 has the WDM hardware integrated into the intersite link and is configured to simultaneously transmit four optical signals:

- Send and receive for fabric 1
- Send and receive for fabric 2

The four signals are optically multiplexed and transmitted down the duplex fiber linking the primary and remote data centers.

The performance of this configuration was certified by imposing a requirement to demonstrate disaster-tolerant operation over a series of tests designed to simulate possible failure scenarios specific to the WDM enabled intersite links.

When designing the desired DRM solution for a given application, another key consideration is the inter-site link that is used to connect the local and remote storage sites. A trade-off exists between the higher cost of providing redundant links (to lower the risk of link failure) and the cost reduction associated with a simpler system having a single point-of-failure.

Figure 2 represents a comparison of three possibilities for DRM systems. As one would expect, a single fiber link would offer a lower cost solution for applications that can tolerate periods of downtime while the higher cost of link redundancy would be recommended for applications requiring high availability.

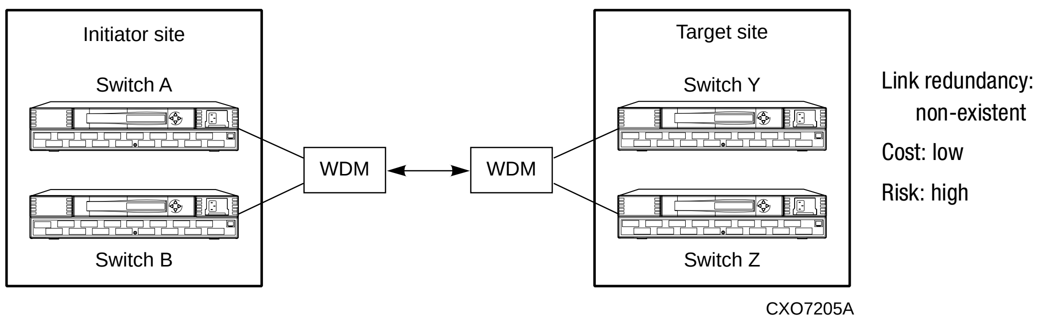


Figure 2a. Simple Point-to-Point configuration using one long-haul fiber link

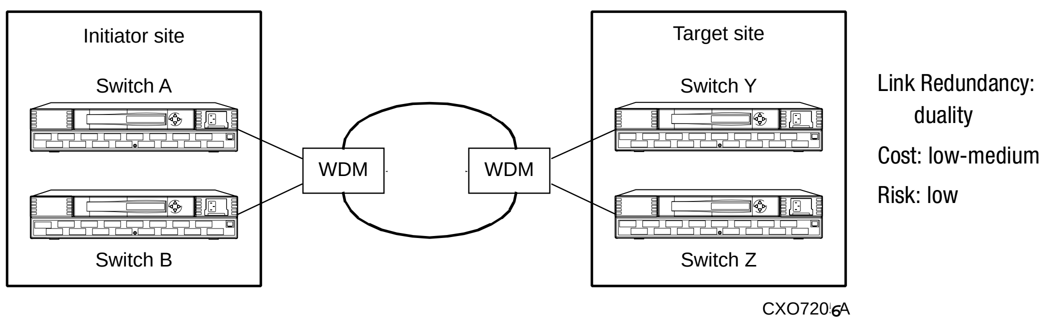


Figure 2b. Point-to-Point Loop configuration using two long-haul fiber links

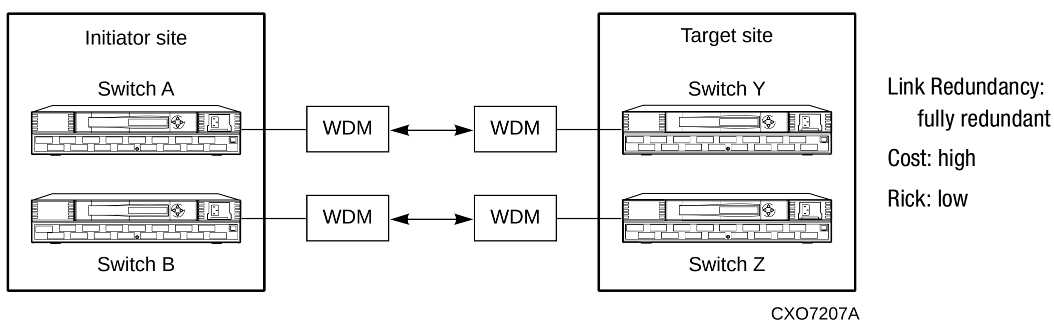


Figure 2c. Redundant WDM Loop configuration using two (or more) long-haul fiber links

Figure 2. Comparison of three inter-site link options

## Interfaces

The physical interfaces that connect the DRM system to the WDM-enabled intersite link are straightforward: there are two duplexed fiber optic cables, each with SC connectors, that make up the fabric links between the optical switches of each DRM site and the input modules to the WDM hardware.

Each Fibre Channel switch contains multiple I/O ports with gigabit interface converters (GBICs). The switch has the function of relaying or routing incoming signals from one switch port to another. The GBIC interface serves as an optical transceiver that provides high-speed serial links by:

1. Conversion of incoming optical signals from the fiber optic link to equivalent electrical signals for the switch
2. Conversion of electrical signals from the switch to optical signals that are to be launched down the fiber.

The specifications of the GBIC-to-fiber interface conform to the American National Standards Institute's (ANSI) Fibre Channel, FC-0 Standard for long wavelength operation.[2]

These long-wave GBICs (1310 nm) enable data transmission over distances up to 10 Km on single mode 9/125 (9 micron core diameter / 125 micron cladding diameter) optical fiber.

The interface at the WDM hardware similarly converts the incoming optical signals to electrical signals, then uses those signals to modulate a narrowband laser. The resulting optical signal is then multiplexed with other such signals before being launched over the long-haul fiber link. As before, the fiber-to-WDM hardware interface must support the ANSI Fibre Channel, FC-0 Standard.

## WDM Product Variations

WDM vendors are currently offering at least three variations of the product described in this application note. These variations are significant and warrant a separate discussion because of their potential for compatibility issues with DRM systems. WDM architectures can be delineated in the most general sense as being either:

- Passive
- Active (with respect to signal amplification)
- Sensitive or insensitive (with respect to channel protocol)

DWM products on the market today fit uniquely into one or possibly two of these categories.

### Active Signal Conditioning Systems

Certain WDM products are offered with line amplifiers and attenuators. These features are included primarily to facilitate interfacing via fiber optic links to other telecom hardware.

2. American National Standards Institute Inc. (ANSI), T11, Fibre Channel-Physical and Signaling Interface (100-SM-LC-L, rev. 3.0).

Line amplifiers allow the boosting of weak signals that are received form peripheral network components, as well as the signals that are being transmitted that might otherwise fall below the threshold sensitivities of receiving equipment.

Similarly, incoming and outgoing signals can be attenuated if they are sensed as being above receiver saturation levels. A typical active system monitors power levels to ensure operation is maintained within the power budget of the hardware. Power monitoring capability is usually accomplished with hardware/software control-loops, which can add significantly to the cost of the product.

### Protocol Insensitive Systems

These WDM systems are transparent to transmission protocol and data rate. As such, they establish open interfaces that give operators the flexibility to provide the following protocols over the same fiber optic cable:

- Fibre Channel
- SONET/SDH
- Asynchronous/PDH
- ATM
- Frame Relay
- Others

A truly passive optical system also passes the optical signal without any form of signal conditioning (such as amplification or attenuation beyond the levels inherent in the system components).

### Protocol Sensitive Systems

While most WDM systems are designed to be protocol-independent, at least one vendor is known to offer a system with protocol-specific capabilities for Fibre Channel. This design enables digital time division multiplexing (TDM) on top of existing optical multiplexing to support multiple channels per wavelength.

This design also allows for network monitoring, digital re-timing (to reduce jitter), link integrity monitoring, and distance buffering. At the time of the publication of this application note, this architecture was available for supporting a mix of Fibre Channel or Gigabit Ethernet protocols only. Considering the added sensitivity to protocols, this WDM variant seems straightforward in point-to-point configurations, but may require additional and potentially costly transmission hardware when deployed in meshed networks.

### Network Management System

A fourth category, which may be described as more of an add-on feature than an architectural variant, is known as a Network Management System (NMS). This feature is worthy of note (and is included in this section) because of its utility as a diagnostic tool.

NMS is simply a subsystem responsible for managing at least part of a WDM-enabled network. NMSs allow communication between users or between network nodes to enable tracking of the network statistics, resources, and performance.

Current NMSs use an optical service channel that is independent of the WDM channels to create a standards-based data communications network allowing service provides to remotely monitor and control system performance and use.

## Extended Fabric

Fibre Channel fabrics are made up of one or more electro-optical switches connected to each other with fiber optic cables. These fabric lengths are capable of being extended without the need for protocol translation to long distances (up to approximately 120 Km) while maintaining efficient link utilization.

To extend a Fibre Channel fabric, it is necessary to account for the time it takes for a single Fibre Channel data frame to travel the round trip distance over the fiber optic link.

If the extended ports (E-ports) on all the switches all have one E-port buffer,[3] then a single frame can be sent over the fabric link at a time. For long distances, sending one frame at a time would be very inefficient. It makes more sense to send frames end-to-end, and in succession at a rate that would “fill up” the fiber link with fibre channel frames.

To fill up a fiber link, one must calculate the number of frames that can exist end-to-end on the fiber link at a time. This calculation is performed in Appendix A.

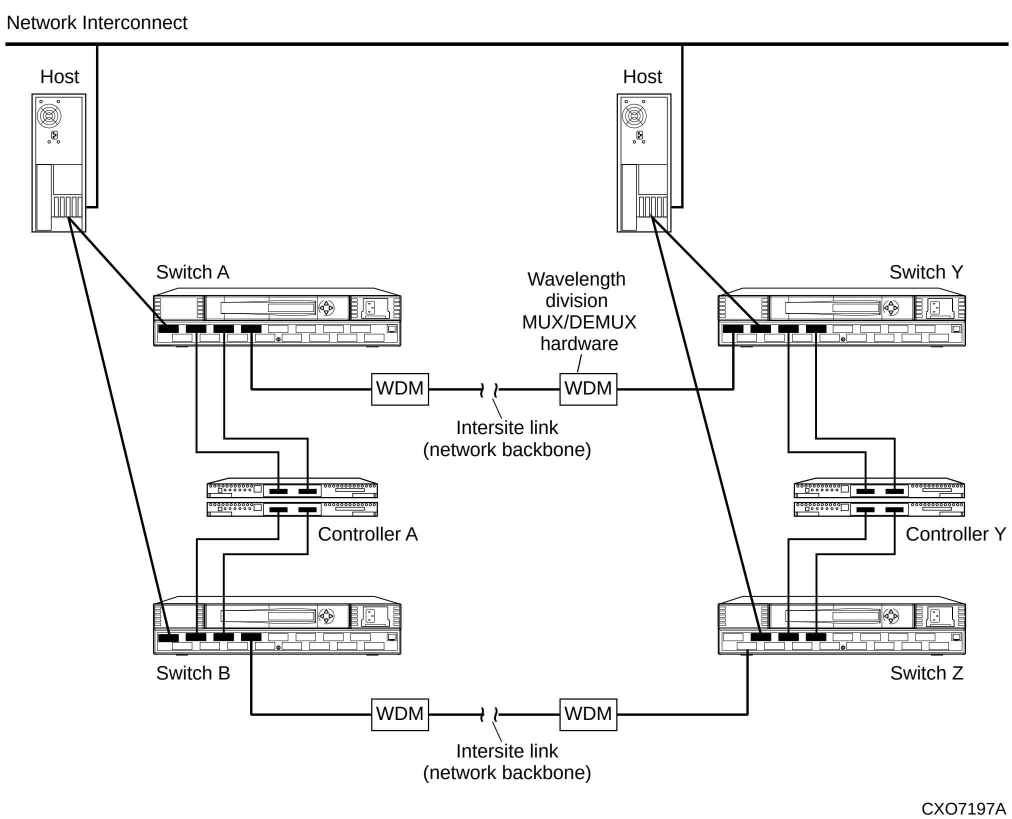


Figure 3. Data Replication manager configured with a WDM-enabled inter-site link

3. Also known as a buffer-to-buffer “credit”

## Future Availability

The factory configuration for the Brocade 2400 and 2800 switch currently provides for eight buffer-to-buffer credits per E-port. These credits cannot be adjusted unless licenses are purchased from the switch vendor and installed to enable the Extended Fabric feature. Once enabled, Extended Fabric allows the user to increase the number of E-port credits to either 27 or 55. These settings will then allow the maximum data transmission rate to be maintained over distances of up to 110 Km. This feature is expected to be available by late 2000.

## Configuration Notes and Recommendations

The following section makes recommendations about the following issues:

- Data-Path Protection
- Hardware Switching on WDM Systems Configured with Line Amplifiers

### Data-Path Protection

Most WDM systems can be configured to provide protection against loss of service. Path switching and equipment switching are offered by most WDM vendors to provide protection of data traffic from fiber cuts and equipment failures:

- Path switching protects the signal carried between the local and remote WDM hardware site locations.
- Equipment switching protects equipment that is not otherwise protected by path switching.

The idea of data-path protection is to provide an appropriate failover mechanism to automatically transfer functionality from one circuit to another.

The price the user pays for this protection feature is connection capacity:

- One protected data path requires two identical end-to-end connections (two complete data channels).
- A 16-channel WDM system provides either 16 unprotected data channels, or a maximum of 8 protected channels.

### Hardware Switching on WDM Systems Configured with Line Amplifiers

It should be noted that during the qualification testing of the DRM system over the WDM-enabled inter-site link, at least one WDM system failed to re-establish connections after performing the switch-failure test. The switch-failure test is performed by disabling one of the two established DRM fabrics, either by simply unplugging the fiber connector or by powering down the entire switch that supports one of the fabrics. In order to pass this test, the fabric components, including the WDM hardware, must reestablish the logical connection once the physical connection is restored.

It was found, however, that certain WDM systems monitor the optical power levels being transmitted over the long-haul links. When these power levels fall below a factory-determined threshold, a fiber cut is assumed and the link is terminated in order to avoid strong back-reflections that could damage the optical amplifiers on the system.

This situation is unique to a testing environment in that there is only one user attached to the entire WDM system. In practice, there will be multiple users on line that will maintain the monitored power levels well above the shutdown threshold.

The recommendation is to be aware that this possibility exists for WDM systems that are configured with line amplifiers.

# Performance

This section describes the following performance issues:

- Data Rates
- Latency

## Data Rates

The tests to determine data rates were simplistic, and were only designed to determine the maximum data rate over varying fiber lengths. As such, 128 block (64 KB) write operations were issued to multiple units over very restricted seek distances. This resulted in all data going to cache (eliminating disk accesses). The restricted seek range prevented performance degradation due to cache locking, and the multiple units allowed several I/O operations to be in process at the same time.

Results were obtained at inter-site link distances of 25, 50, 75, and 100 Km. The extended fabric feature was enabled with:

- 27 buffer-to-buffer credits at the 25 Km and 50 Km links
- 55 credits for the 75 Km and 100 Km links.

Figure 4 shows the performance results for distances out to 100 Km for buffer-to-buffer credit settings of 8, 24, and 55.

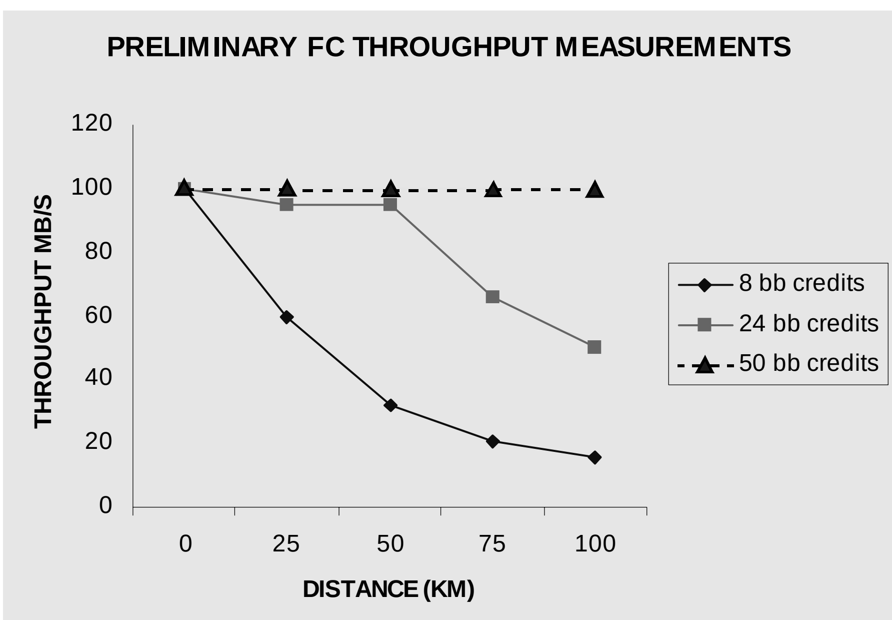


Figure 4. Performance results

### Latency

Latency tests were performed in order to isolate the effects of the WDM hardware on the total round-trip transport time. The results were obtained in the following manner:

- Measure the total round-trip transport time across the WDM-enabled long haul link.
- Measure the total round-trip transport time across the same long haul link without (bypassing) the WDM hardware.
- Determine the WDM-induced latency effort by taking the difference.

The results yielded total round-trip delay on the order of 1 microsecond, that could be solely attributed to the WDM components. Since there are four WDM "boxes" encountered during one round-trip, the delay from a single box is on the order of 0.25 microseconds which is considered negligible with respect to any effects on DRM functionality.

## Disclaimer

Compaq does not warrant Third Party Products which customers may purchase for peripheral use with their DRM solution. These products are used on an "AS IS" basis unless otherwise specified in the Warranty Attributes of the Third Party Products. Third Party Products may be warranted by the third party as specified in the documentation provided with the Third Party Products.

## Related Documentation List

**Table 1 HSG80 Array Controller Documentation List**

| Document Title | Order Number |
|---|---|
| HSG80 Array Controller ACS V8.4 Configuration and CLI Reference Guide | 118619-001 / EK-HSG84-RG |
| HSG80 Array Controller ACS V8.3/8.4 Maintenance and Service Guide | 118629-001 / EK-HSG84-SV |
| StorageWorks Fibre Channel Storage Switch Service Guide | 135268-001 / AA-RHBZA-TE |
| StorageWorks Fibre Channel Storage Switch User's Guide | 135267-001 / AA-RHBYA-TE |
| Compaq StorageWorks RA8000 and ESA12000 Storage Subsystems User's Guide | 387404-001 / EK–SMCPR–UG |
| Compaq StorageWorks RA8000 and ESA12000 Fibre Channel Cluster Solutions for Windows NT Installation Guide | 101471-001 / EK-NTC8K-IG |
| RA8000 and ESA12000 HSG80 Solution Software V8.3/V8.4 for WindowsNT Server - Intel Installation Reference Guide | 387387-002 / AA-RFA9B-TE |
| RA8000 and ESA12000 Fibre Channel Storage Subsystem for WindowsNT Server - Intel Quick Setup Guide (for ACS V8.3) | 387387-002 / AA-RFA7A-TE |
| Application Note: Compaq SANworks Data Replication Manager over an ATM Link | N/A/EK-DRMAL-AA. A01 |

**Table 1 HSG80 Array Controller Documentation List (Continued)**

| | |
|---|---|
| RA 8000 and ESA12000 Fibre Channel Storage Subsystem for WindowsNT Server - Intel Quick Setup Guide (for ACS V8.4) | 136258-001 / AA-RHH5A-TE |
| StorageWorks Secure Path for Windows NT, A High Availability MultiPath Solution Installation Guide | 123995-001 / EK-WNTMP-MH |
| KGPSA PCI-to-Fibre Channel Host Adapter | EK–KGPSA–UG |
| Compaq StorageWorks Ultra SCSI RAID Enclosure (DS-BA370-Series) User’s Guide | 387403–001 / EK–BA370–UG |
| Command Console Version 2.1 (HSG80) for RA8000/ESA12000 User’s Guide | 387405-003 / AA-RFA2C-TE |

# Appendix A - Calculation for the Optimum Number of Buffer-to-Buffer Credits

To “fill up” a fiber link, calculate the number of frames that can exist end-to-end on the fiber link at a time. Since frames are made up of a fixed number of bytes, it is a simple exercise to determine the number of bytes that can exist on a round-trip loop at one time for a given data rate.

As an example, consider data flowing at a rate of 100 MB/sec down a 10 Km fiber. Since a pulse of light travels down fiber at a speed of 5 microseconds per kilometer, the round-trip travel time for a single pulse can be calculated as:

2 (10 Km  X  5 microseconds/Km) = 100 microseconds.

In 100 microseconds, the number of bytes that can fill up the round trip distance is:

100 microseconds  X  100 MB/second =  10000 bytes.

To determine the number of frames that will fill up this same round-trip distance, divide the total number of bytes by the frame size. For a frame size of 2 KB/frame, the calculation yields:

10000 bytes / (2000 bytes per frame) = 5.

Thus, to be able to efficiently fill up the round-trip fiber distance (20 Km) with 2 KB data frames at a transmission rate of 100 MB/sec, the E-port would need to be configured with at least 5 buffer-to-buffer credits. Similar calculations using 2 KB frames would yield 25 credits for 50 Km and 50 credits for 100 Km.