Dynamic Path Failover Driver for AIX FibreQuik Host Bus Adapter Installation and User Guide

Cambex Corporation
115 Flanders Road
Westborough, MA 01581

Customer support
support@cambex.com

Document: 081-468-025
Date: 4/15/05
Rev.:   E

Table of Contents

Dynamic Path Failover for AIX

Introduction

The Cambex FibreQuik™ Fibre Channel host bus adapters provide attachment of Fibre Channel devices to PCI or Micro Channel based servers. They support the SCSI-FCP protocol.

When used with the Cambex DPF (Dynamic Path Failover) software, two adapters in a system provide greater throughput and total protection from failure of a fibre channel component.

In addition, the Cambex AutoRecover daemon, cbxdpfar, will automatically detect when a failed path has been repaired and make it available as a standby path. For arrays that support preferred paths, AutoRecover will automatically resume operation on the repaired path, otherwise, an operator may manually change paths to do load balancing.

Storage (SAN) and communications (LAN) networks have resulted in a physical and logical separation between the storage device (typically a SAN-connected RAID array), the storage user (typically an Enterprise Server, which we will refer to below as the server) and the administration station (typically a remote workstation).

As storage systems have evolved from JBOD (just bunch of disks) to complex shared RAID systems, the management software has also grown in complexity. The following sections describe the different levels of management software.

SAN Management

SAN Management is the top layer. It is responsible for assigning the required access for each server to each storage device. The unit of storage is the LUN (Logical Unit Number) which acts the same as a single disk directly attached to the server. In reality, a LUN is made up of portions of a number of disks under the direction of a RAID controller in the storage device. SAN Management typically requires a process running on each managed server communicating with the manager running on the administering station.

Storage Management

The Storage Management layer, or RAID manager, is the layer that communicates with the RAID controllers on the storage device. This is the layer used to create, modify or delete LUNs.

The Storage Management layer may provide SAN Management functions, but only for the specific storage device being managed. Storage management typically runs on a server attached to the storage device (in-band) but may also run over a network (out-of-band).

Path Management

Path Management includes a variety of software for managing the various fibre channel components such as HBAs (Host Bus Adapters), hubs and switches. It is used to display path status and allow control over multiple paths between servers and storage. To avoid having to access a large number of unrelated software interfaces, integration of the various Path Management components is usually done using a management tool

communicating via a network connection to each managed device.

The Cambex Dynamic Path Failover driver is a Path Management component that allows multiple paths to a storage device and provides automatic failover between available paths. Cambex Dynamic Path Failover software provides both a command line interface and access through smit, the standard integrated management tool for AIX.

SCSI and FCP Drivers

AIX makes two types of device drivers available. SCSI device drivers support SCSI-2 disk and tape devices. FCP device drivers support newer SCSI-3 disk and tape devices, which provides a larger address space for target ID and LUN. Cambex supports adapter drivers to work with each type of device driver. When using the SCSI adapter driver, adapter devices will be configured with a name like **scsi2**. When using the FCP device driver, the adapter will be configured with a name like **fscsi2**. The choice of which driver to use should depend on which devices are best supported by both the hardware and software in the system.

Load Balancing

Load balancing is a term which has been in use for many years in the High Perfomance Computing area to mean sharing tasks between multiple processors so that no single processor is required to do all of the work. In the storage area we do this by allowing each LUN of each target device to be accessed over a specific path. By monitoring I/O activity and allocating LUNs to appropriate paths, high utilization of the aggregate fibre channel bandwidth may be achieved.

Adapter Specifications

PC2000LC-HPSP

Bus	PCI	
Data rate (each dir.)	200	MB/s
Cable length - copper	30	m
- multimode optical	500	m
- single mode optical	10	Km
Voltage	3.3 or 5	V
Power	7.2	W
Shipping weight	10	oz

PC2000LC-A

Bus	PCI	
Data rate (each dir.)	200	MB/s
Cable length - copper	30	m
- multimode optical	500	m
- single mode optical	10	Km
Voltage	3.3 or 5	V
Power	7.2	W
Shipping weight	10	oz

Trademarks

Dynamic Path Failover and FibreQuik are trademarks of
Cambex Corporation.

Software and Hardware Installation

Driver Installation (command line)

The examples assume that the driver CD is mounted at /mnt and that you are installing on AIX 5.1. The name and revision level of the uninstalled driver image may be examined by issuing the command (as root):

```
installp -ld <device>
```

where <device> is the path to the drivers, such as /mnt/AIX4/ for the CD or the name of a driver directory on the hard disk. Example output would be:

```
Fileset Name            Level      I/U Q Content
=============================================
PC1000.driver.obj       2.5.10.25      I  b usr
# PC1000 Fibre Channel Adapter Driver (AIX 5L)

cbxdpf.driver.obj       5.1.0.23       I  b usr
# Cambex Dynamic Path Failover Driver
```

Before installing a new version of the driver, delete the previous version of the driver. Note that any file systems must be unmounted, volume groups exported and SCSI devices must be removed before removing the driver. Be sure to have a complete backup of all data before updating the driver. The command to remove the driver is:

```
installp -u <name>.driver.obj
```

where names are cbxdpf and PC1000 as shown above.

The new driver may be installed on the system with the command:

```
        installp -acd <device> all
```

where <device> is as described above.

example:

```
        installp -acd /mnt/AIX5.1 all
```

After installing the driver, see the README file for additional
information and release notes.

```
        /usr/lpp/cbxfc/READ_ME
        /usr/lpp/cbxdpf/READ_ME
```

Driver Installation (smit)

The smit path is
>   Software Installation and Maintenance
>   Install and Update Software
>   Install and Update Software by Package Name

For item " INPUT device / directory" use **/mnt/AIX5.1** or the directory where the image files are located. A list should pop up with the name of the drivers. Select and press enter.

Also make sure that the following drivers are loaded from the AIX installation media.

```
devices.fcp.disk.rte
devices.fcp.tape.rte
```

The recommended maintenance levels for AIX are specified in the image file *prereq* section. If not at the correct maintenance level, the driver installation will fail. The current level may be found from the command:

```
instfix -i |grep ML
```

Hardware Installation

Power down the system and install the adapter in a free slot. The PCI cards may be installed in either 32 or 64-bit PCI card slots, however the 32-bit slots may limit performance at high data rates. Make sure that any cables do not block airflow to the card. Observe proper static precautions. Detailed procedures for opening the case and for suggested adapter placement may be found in the *Operator Guide* for the specific RS/6000 model.

When using copper cable, it is extremely important that the building wiring present a low impedance ground path to the frame of the device and the host with which it will communicate. Noise due to copiers or large machinery sharing the electrical circuits is a common cause of signal integrity problems.

AutoRecover

cbxdpfar, the AutoRecover daemon, is automatically installed and enabled when the DPF driver is installed. Control of the daemon is done using SRC.

stopsrc -s cbxdpfar Stops AutoRecover
startsrc -s cbxdpfar Starts AutoRecover

AutoRecover works by periodically polling the standby paths to each LUN. If the polls fail, the path is marked as failed. If the polls succeed, the path is marked as available for use. If path preferences are set up, the AutoRecover daemon will move I/O to the preferred paths when they become available. AutoRecover also deletes paths to hdisks that have been removed from the system. See the section on SMIT for changing the default AutoRecover timing parameters.

AIX Configuration

During power-up initialization, AIX will find and configure the new adapter(s) and any supported devices using the driver's configuration method. Devices are assigned a SCSI ID as they are discovered starting with 0. SCSI ID is not related to loop ID, port ID or to SCSI ID assignment on other hosts. It is a logical value bound to the fibre channel WWN of the device. Configuration may also be run from the command line:

```
cfgmgr
```

Once disks have been configured, the standard AIX utilities may be used to build volume groups, logical volumes and file systems. See the *AIX System Management Guide* for a detailed explanation.

Find out the names of the installed adapters with the command:

```
lsdev -Cc adapter
```

In arbitrated loop, the ID of each device must be unique. For disk drives, the ID is typically set by switches on the drive enclosure. The adapter ID is soft assigned and may be examined after it is participating with the *mbtest* command. If the adapter name found by the *lsdev* command were *fscsi1*, the command would be:

```
/usr/lpp/cbxfc/mbtest fscsi1 id
```

The requested ID may be changed in the database to any non-conflicting value in the range 0 to 126 with the command:

```
chdev -l fscsi1 -a id=6 -P
```

The ID attribute has no effect in a fabric environment.

Persistent Bindings

Persistent Bindings is the term used to indicate the mapping of SCSI ID to FC Worldwide Name (WWN). By keeping a persistent binding, the loop ID or port ID of a device may change and the device will still be properly addressed.

In the Cambex AIX driver, persistent binding is automatic. At configuration time, all devices that are found are automatically assigned to a SCSI ID and a persistent binding is created. Persistent bindings are by portname and are kept in the ODM database. They are preserved and restored at reboot for all FC devices.

To display the SCSI ID to WWN mapping for all adapters and devices, issue the command:

```
/usr/lpp/cbxfc/fctest
```

When replacing a failed device on a SAN, we want to keep the same SCSI ID but change the WWN binding to the replacement. A perl script is provided for doing this:

```
/usr/lpp/cbxfc/FRU_replace
```

SMIT Commands

Adapter Management
The adapter management commands available in SMIT are in the *devices* menu under *Cambex Fibre Channel Adapter*:

List all Cambex FC adapters
This provides the name and location of each adapter.

Change / Show Characteristics of an Adapter
Allows the user to display and modify attributes. Attributes (and default values) include:

exec_throttle (256) number of simultaneous I/O operations.

dma_window (64) PCI DMA memory window in MB.

retry_delay (3) Delay in seconds after bus reset or other error.

logout_delay (7) Time in seconds allowed for ports to be logged out before failing I/O.

topology_list (LOOP_P2P) Automatically detect topology. May also be forced to LOOP_ONLY or P2P_ONLY.

frame_size (2048) Maximum FC frame size. Choices are 1024 or 2048.

infinity (300) Maximum SCSI command timeout in seconds. Set to 0 to disable.

GHz (auto) Auto-detect fibre channel rate in GHz. May also be forced to 1 or 2. Ignored on 1GHz cards.

sw_fc_class (3) Fibre channel class.

devtype_exclude ( ) List of SCSI device types to ignore during configuration. For example, specifying "1,5" would exclude both tape and CD devices.

Attributes only used in loop mode include:

id (0) Default L_ID for adapter. See table of L_ID to ALPA.

login_on_lip (yes) Full login if LIP occurs.

full_duplex (no) Use full duplex loop open, OPNyx.

Generate Error Report
Accesses the AIX error logging facility.

Trace Cambex FC Adapters
Runs the AIX trace facility, looking only for the entries specific to the adapter driver.

Path Management

The path management commands available in SMIT are in the *devices* menu under *Cambex Dynamic Path Management*:

List all Cambex Dynamic Paths

This provides the information on all paths and devices known to the DPF driver. Use this when installing to verify that all paths have been found and configured.

Change / Show Characteristics of a Dynamic Path

Displays the current path status and allows the user to place paths online/offline or to change path preferences.

The autorecover daemon will automatically move operation to preferred paths. The preferred path is the combination of path preference values for HBA, device and path. For some arrays, target device preference is read automatically from the array at configuration time. Consult the array user manuals for more information.

Trace Cambex FC Adapters

Runs the trace facility, looking only for the entries specific to the DPF driver and adapter driver.

Manage AutoRecovery Subsystem

Provides the following functions:
  Show Status of AutoRecovery Subsystem
  Start AutoRecovery Subsystem
  Stop AutoRecovery Subsystem
  Change / Show AutoRecovery Subsystem Parameters

Parameters are:
  -n   Number of successful inquiries (3 - 9)

-t   Seconds between inquiries (3 - 3599)
-r   Seconds between ODM updates (automatic)

Manage Licenses

Allows the user to enter a new license for the DPF software, overwriting the existing license, or to display the expiration date of the current license. To enable a one-time-use, 30-day license for use until the purchased license is received. Issue the commands:

```
cd /usr/lpp/cbxdpf
mv license.30day license
cfgmgr
```

 Licenses for OEM versions of the driver are obtained from the storage vendor who provided the software.

The generic version of DPF purchased from Cambex is licensed for use only on Cambex cards. A license token is provided with the purchase. Contact cambex support by email to obtain a license. You will need to provide

The license tokens for each card
Host machine ID from 'uname -m'
The output of /usr/lpp/cbxfc/fcinfo

Troubleshooting

Use the *errpt* command to identify the cause of failure. The
lsdev command will provide the location of the component. For
example, if *errpt* identifies a loose cable on device hdisk1, *lsdev*
will give an address similar to:

```
hdisk1   Available 00-02-01    Cambex SCSI...
```

and *lsattr -El hdisk1* will show:

```
clr_q         no                   Device CLEARS it's
                                   Queue on error
location                           Location Label
lun_id        0x0                  Logical Unit Num ID
node_name     0x20000080e511b744 FC Node Name
pvid          none                 Physical Volume id
q_err         yes                  Use QERR bit
q_type        simple               Queuing TYPE
queue_depth   16                   Queue DEPTH
reassign_to   120                  REASSIGN time out
rw_timeout    30                   READ/WRITE timeout
scsi_id       0x0                  SCSI ID
start_timeout 60                   START UNIT timeout
ww_name       0x23000080e511b744 FC World Wide Name
```

Which indicates that the adapter is in card location 00-02, and the
drive is at SCSI ID 0, LUN 0.

Checklist for driver and HBA installation

The following commands may be used to isolate problems
during installation of the fibre channel driver and card. Type
the command shown in **bold** and verify the output.

lslpp -l PC1000.driver.obj   must give:
  Fileset   Path: /usr/lib/objrepos   PC1000.driver.obj
  Level   01.05.00.00 (or greater)
  State   Committed
  Description   Cambex Fibre Channel Adapter Driver

lsdev -Cc adapter should print a line for each card:
  fscsi$<x>$ $<x>$ is a unique number for each SCSI card
  Available $<SS-SS>$ $<S>$ is the location in the machine where
               the card is installed
Check that the number of SCSI cards and their slots are correct
If 'Available' then card and driver are installed.
If 'Defined' or not listed then card was not found.

lscfg -vl fscsi4   should print more information for a card:

```
  DEVICE              LOCATION              DESCRIPTION

  fscsi4              00-03                 Cambex Fibre
Channel I/O Controller

     Manufacturer................CAMBEX
     Machine Type and Model......MC1000F
     ROS Level and ID............V2.01.04
     Serial Number...............   2354
     Displayable Message.........Cambex Fibre ...
     Network Address.............500502E020000932
     Device Specific.(Z0)........Build:TMNX
```

cfgmgr -v >/tmp/cfg.log 2>&1
Should identify the card and configure it.

If cfg.log contains the ID 7710002x and no driver is found, then the adapter driver is not installed.

During boot, the number displayed when configuring the adapter is hex 0x911 (FC adapter).

`/usr/lpp/cbxfc/mbtest fscsi1`
This test code may be used to get fibre channel status from the adapter or from attached devices. Type the command as above for a list of subcommands.

`/usr/lpp/cbxfc/mbtest fscsi1 fab`
This command will list the persistent bindings between SCSI ID and WWN on adapter scsi1. It will also list the loop ID and port ID (fabric address) which can change dynamically as devices are connected or disconnected.

`/usr/lpp/cbxfc/mbtest fscsi1 getp`
This command will show the fibre channel ports which are currently logged in to adapter scsi1.

`/usr/lpp/cbxfc/sctest fscsi1`
This test code may be used to get IOCINFO information from the adapter or to see the results of SCSI Inquiry or TestUnitReady commands to SCSI devices attached to the adapter. Type the command as above for a list of subcommands.

`/usr/lpp/cbxdpf/dpfutil listall`
This will list the failover devices and paths. The active path will be shown on the left. The letter **F** indicates a failed path. The message ***no failover devices available*** indicates that the attached storage is not recognized by the DPF driver. This could mean that a platform kit needs to be installed for your storage. Contact your storage supplier for more information.

/usr/lpp/cbxfc/runtrace
This should be run to start a trace if Cambex support requests a trace of the card failure. To stop the trace and generate a trace report, type:

```
trcstop; trcrpt > /tmp/trc.log
```

Questions or problems related to Cambex fibre channel adapters may be e-mailed to: support@cambex.com

Error Messages

Various errors and notifications will be placed in the AIX error log. These can be read using the errpt command, or automatic notification may be set up.

DPF_FAILOVER
This indicates that the active path to a device has failed, and that DPF successfully transferred I/O to an alternate path.

DPF_STDBYDN
This indicates that a path other than the active path has failed to a device.

DPF_UNRECOV
This indicates that the active path to a device has failed, and that DPF was not successful in transferring I/O to an alternate path.

DPF_OFFLINE
DPF_ONLINE
These messages indicate that a path has been manually set to the offline or online state.

DPF_LIC_WARNING
DPF_LIC_EXPIRE
These messages indicate that a temporary license for the DPF software is either close to or has expired. Once the DPF license has expired, you will need to uninstall the DPF driver in order to regain access to the hdisks.

Upgrading the Driver

When upgrading to a newer version of the driver there are two methods:

1) Use rmdev to remove all dependent devices, remove the existing driver(s), then install the new driver(s) and run cfgmgr to recreate the devices. This method allows a new driver to be installed without rebooting the server. File systems which are unmounted before the process may be mounted again afeterwards without affecting the data.

2) Install the new driver with devices in place. The server must be rebooted afterwards so that the driver on disk will match the driver in memory.
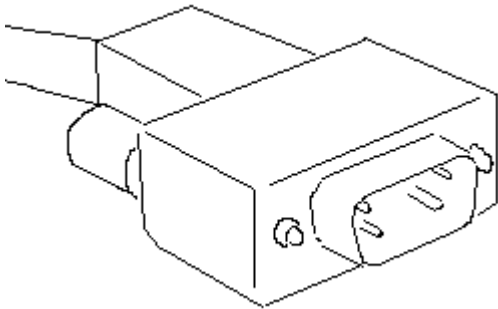
When upgrading AIX to a new major version, remove the drivers beforehand and install the new versions after completing the AIX upgrade.
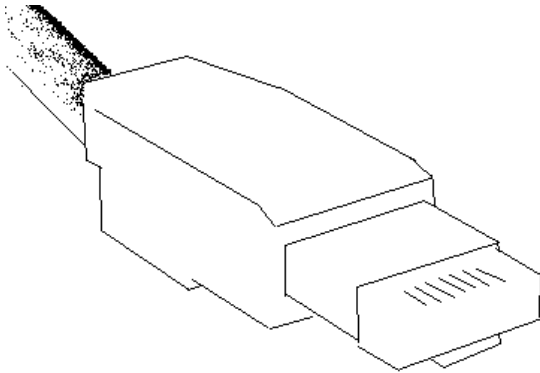
Appendix A

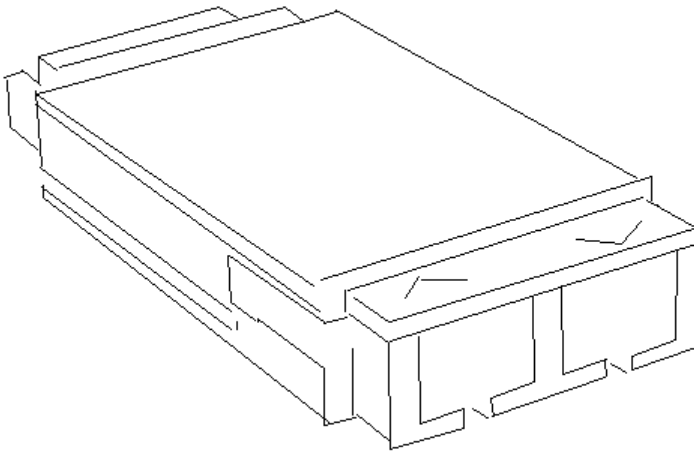## *Identifying Fibre Channel Components*
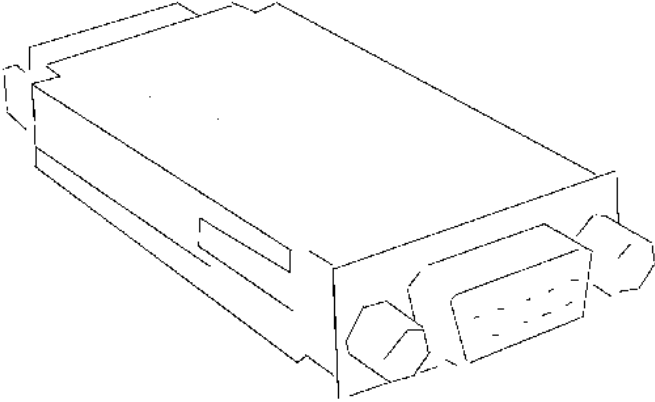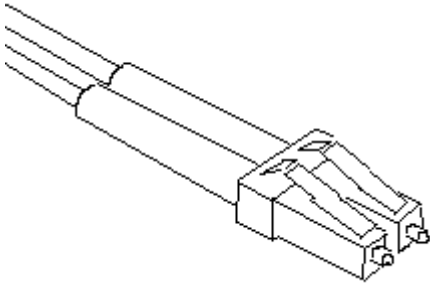

Copper DE9 GBIC

Copper DE9 Cable

Copper HSSDC Cable
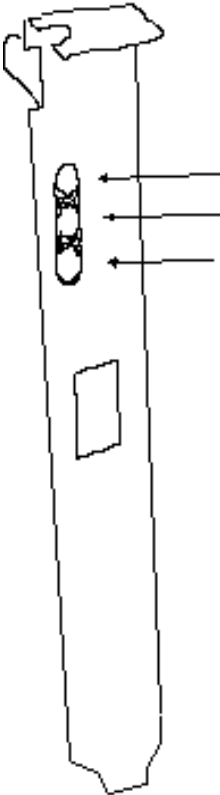
Duplex SC Optical GBIC

Small form factor (SFF) optical LC cable

PC2000 LED Meanings

Top LED – Transceiver hardware is detecting a signal.
Middle LED – Firmware is being synchronized.
Bottom LED – Flashing amber when not synchronized.

LED	NORM	BAD CABLE	NO CONN
TOP	GREEN	GREEN	OFF
MIDDLE	GREEN	OFF	OFF
BOTTOM	OFF	FLASHING	FLASHING

Declaration of conformity

We Cambex Corporation of Waltham, Massachusetts, USA, declare under our sole responsibility, that the products PC2000LC-HPSP, PC2000LC-A, and MC-1000F are in conformity with the following standards:

EN 55022:1994/A1:1995/A2:1997 Class A ITE emissions requirements
EN 50082-1:1992 EMC residential, commercial and light industrial generic
    immunity standard

Warning

This is a Class A product. In a domestic environment this product may cause radio interference in which case the user may be required to take adequate measures.

Safety Notices

FCC Warning

This equipment has been tested and found to comply with the limits for a class A computing device pursuant to Part 15 of FCC rules which are designed to provide reasonable protection against interference when operating in a commercial environment. This equipment generates, uses and can radiate radio frequency energy and if not installed and used in accordance with the instruction manual may cause interference to radio communications. Operation of this equipment in a residential area is likely to cause interference in which case the users will be required to take whatever measures may be required to correct the interference at their own expense. Changes or modifications of this device could affect the user's authority to operate this equipment.

Safety Notice

Class 1 Laser Product
Versions of this product supplied with an optical transceiver meet the safety requirements of a class 1 laser as specified by the US Food and Drug Administration Center of Drug and Radiological Health (CDRH) 21 CFR 1040.10 & 1040.11 and by IEC 825-1. The U.S. Department of Health and Human Services (DHHS) does not consider Class 1 lasers to be hazardous. Customer supplied laser GBICs must meet the same requirements. If in doubt, contact Cambex customer support before operating the product.

The International Electrotechnical Commission (IEC) 825 Laser Safety Standard requires labeling in English, German, Finnish, and French stating that the product uses Class 1 lasers. Because it is impractical to label the transceiver, the following label is provided in this manual.

CLASS 1 LASER PRODUCT
LASER KLASSE 1
LUOKAN 1 LASERLAITE
APPAREIL A LASER DE CLASSE 1

TO IEC 825 (1984) + CENELEC HD 482 S1

Dynamic Path Failover License Agreement

Cambex Dynamic Path Failover (the Software) is a licensed product. The user must consent to the provisions of this license agreement in order to use the Software. Use of the Software beyond the number of purchased licenses is prohibited.

The user may not cause or permit copying, modification, rental, sublicense, lease, transfer, reverse engineering, or disassembly or decompilation of the Software, or creation of derivative works based on the Software.

This license for the Software remains in effect until terminated by either party upon written notice. Upon termination of the license for the Software, the user must discontinue use of and destroy or delete any copies of the Software.

Except as limited by applicable law, Cambex's total liability under this agreement is limited to the purchase price of the Software. Dynamic Path Failover may not be used to provide life-critical protection such as operation of medical equipment or control of nuclear facilities. The Software is intended to provide improved availability of data but makes no claim for greater reliability or protection of data, and Cambex may not be held liable for data loss or consequential damages.

Any provision of this agreement found unenforceable will not affect other provisions. This agreement is governed by laws of the Commonwealth of Massachusetts.