AS2 Disaster Recovery

Implementation Guide

Issue 1, Approved, 18-Nov-2010

Document Summary

Document Item	Current Value
Document Title	AS2 Disaster Recovery Implementation Guide
Date Last Modified	18-Nov-2010
Current Document Issue	Issue 1
Status	Approved
Document Description	The purpose of this AS2 Disaster Recovery Guideline is to provide information about Disaster Recovery Planning (DRP) as it relates to Applicability Statement 2 (AS2) an industry standard for Internet-based data exchange.

Contributors

Name	Organization

Log of Changes in Issue 1

Issue No.	Date of Change	Changed By	Summary of Change

Disclaimer

Whilst every effort has been made to ensure that the guidelines to use the GS1 standards contained in the document are correct, GS1 and any other party involved in the creation of the document HEREBY STATE that the document is provided without warranty, either expressed or implied, of accuracy or fitness for purpose, AND HEREBY DISCLAIM any liability, direct or indirect, for damages or loss relating to the use of the document. The document may be modified, subject to developments in technology, changes to the standards, or new legal requirements. Several products and company names mentioned herein may be trademarks and/or registered trademarks of their respective companies.

Table of Contents

1. Introduction

1.1. Purpose of this Document

The purpose of this AS2 Disaster Recovery Guideline is to provide information about Disaster Recovery Planning (DRP) as it relates to Applicability Statement 2 (AS2) an industry standard for Internet-based data exchange.

1.2. Who Will Use this Document?

This document is for anyone that is responsible for preventing, planning for, or recovering from an event that could negatively impact the integrity of electronic data.

2. Implementation Procedures

The following sections provide an introduction to Disaster Recovery Planning as it relates to AS2, and an overview of the activities recommended for successful disaster recovery.

In this section we review the various types of “disaster” level failures that must be planned for, some of which are:

- Software failure
- Localized hardware failure
- Network failure
- Full scale disaster involving a hot site

2.1. Software Failure

A software failure (crash or bomb) is when a program is not able to continue processing due to incorrect programming logic. Examples are when the web site is down or there is a corrupted database.

2.2. Localized Hardware Failure

A localized hardware failure is when there is a malfunction within the electronic circuits or electromechanical components (disks, drives, cables, circuit boards, etc.) of a computer system at the production or local site. (The production or local site is abbreviated as CoLoProd in diagrams later in this document.) Recovery requires troubleshooting to find the component and repair or replace the malfunctioning part.

2.3. Network Failure

A network failure is a malfunction of any of the following components

- Network operating system in a client or server machine
- Cables connecting machines
- Hardware in between, such as bridges, routers, and switches
- In a wireless system, antennas and towers

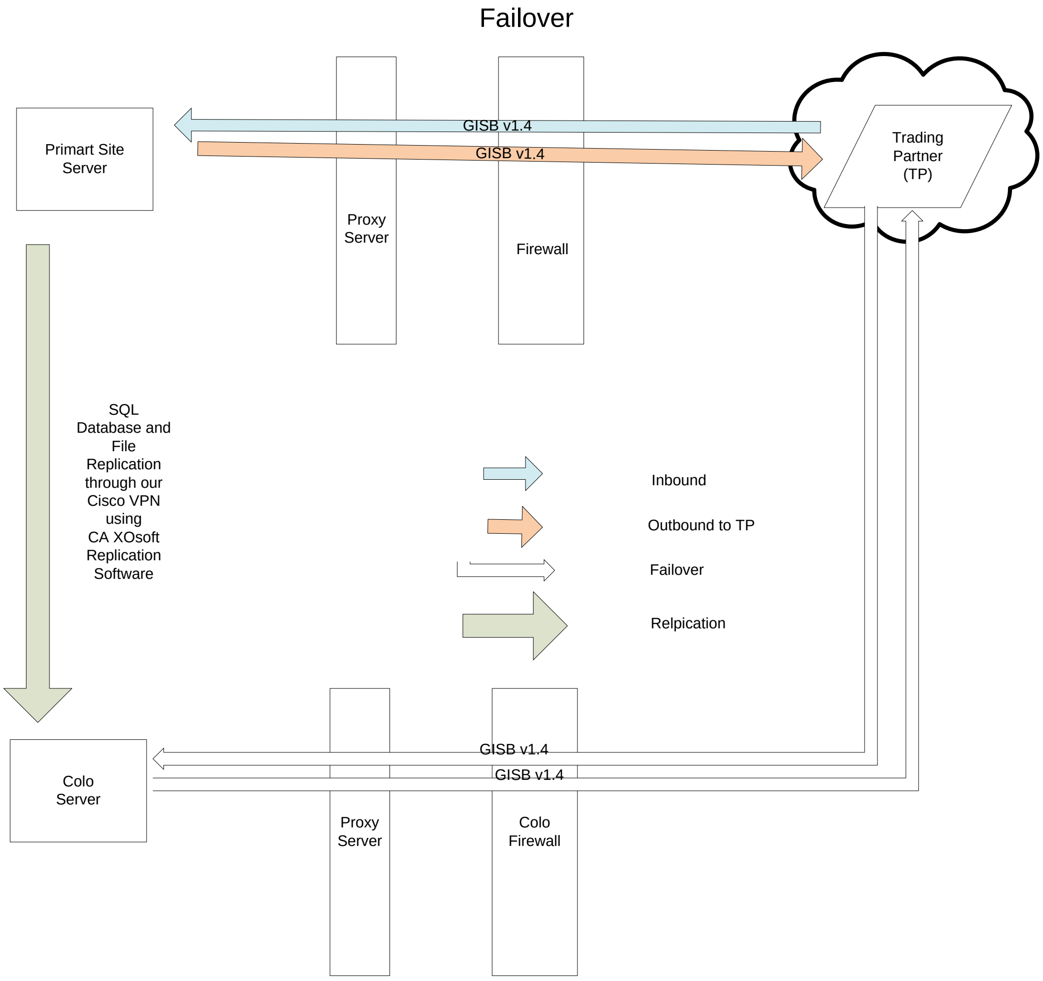
2.4. Full Scale Disaster Involving a Hot, Warm, Cold or “Live-Live” Site

A full scale disaster is when a whole site is down malfunctions are extensive enough to where the whole local production system is duplicated and switched over to another geographic location.

A hot, warm, or cold site is a site that is switched over to in the case of failure of a whole system. There are 4 different categories of disaster recovery scenarios with different recovery times.

- **Cold site** –Software and hardware is assembled and built and switched over to replace the production site. Typically it is turned over to in a 48 hour recovery period.
- **Warm site** – A duplicate system is built (software and hardware) and standing by running but not quite ready for turnover. Typically it is turned over to in a 24 hour recovery period.
- **Hot site** – A complete duplicate system is already installed, assembled and built. It is running and standing by to be turned over to. Typically, it is turned over within 15 minutes.
- **Live-Live site** – A complete duplicate system at another site is always running and processing a portion of the total load. Typically AS2 network transmissions are variably directed to one site or the other. Control information about the status of transmissions, partner information, etc., is continually synchronized between the two sites. Received files are directed to the appropriate internal application.

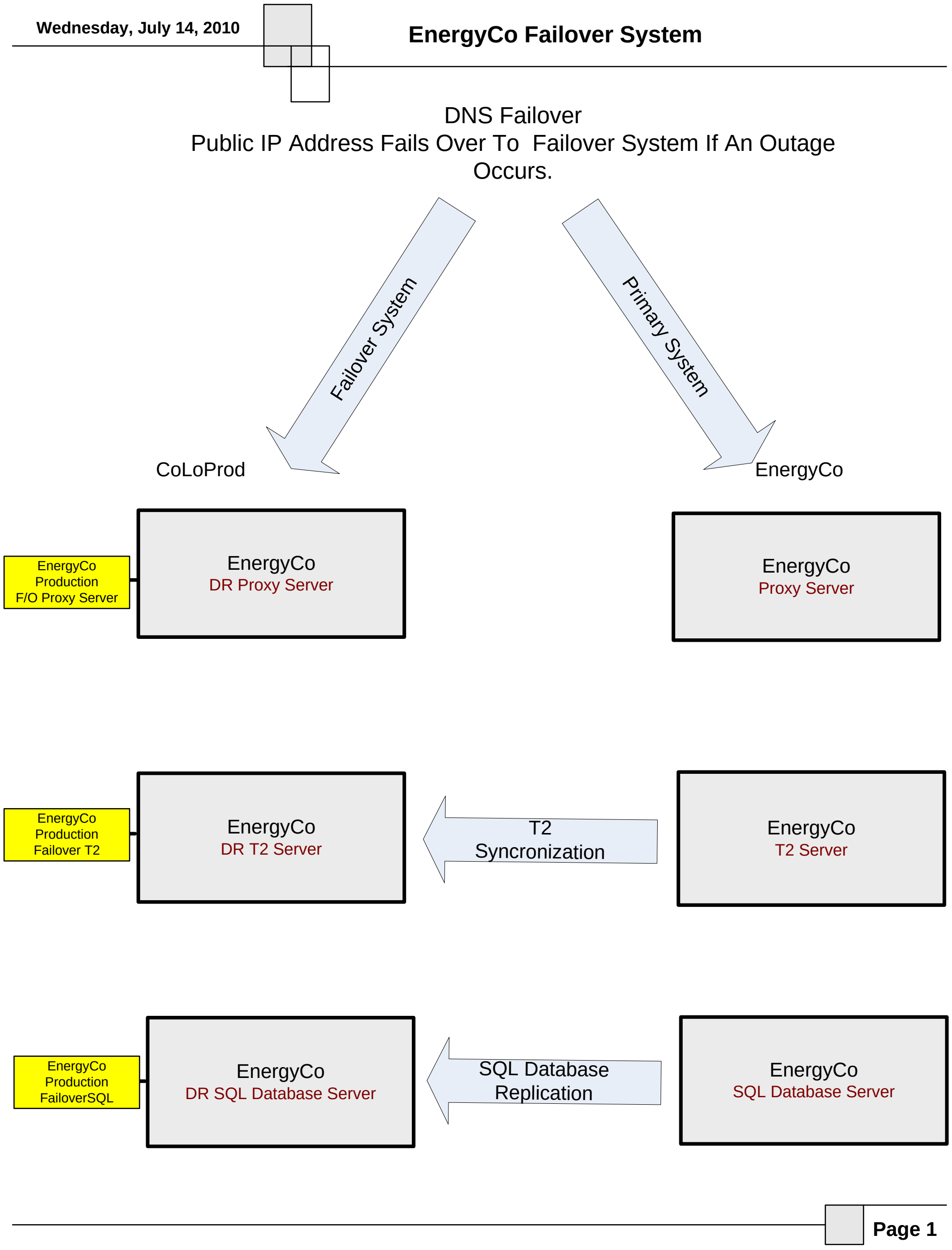
3. Generalized Diagram of AS2 Disaster Recovery



4. Scenarios for AS2 Disaster Recovery

The following section contains various scenarios for AS2 Disaster Recovery

4.1. Energy Company Failover System – Mid-Level Transaction Volume with Very Low Tolerance for Recovery Delays



4.1.1. Operational Description

Requires two identical sets of hardware and software at two remote locations both responding to the same URL. Both sites are capable of processing all AS2 communications. Trading partners will send data to the common URL. The AS2 transmission is routed to the appropriate Content Smart Switch by best path determination.

After being directed to the Content Smart Switch the transmission is sent to the Web Server cluster completing the HTTP AS2 session. AS2 messages are then passed through a firewall (if allowed) to the High Availability Cluster of AS2 servers.

Data and configurations are synchronized live and continuously between the two sites. Within each site a failure of a server would cause the transmission to be directed to another within the cluster. If the entire site were to experience an outage the transmissions would be sent to the alternate site for processing.

4.1.2. Software Requirements

Software licensing will be required for both sites. Software supporting synchronization and High Availability over a WAN connection will have to be purchased for both sites.

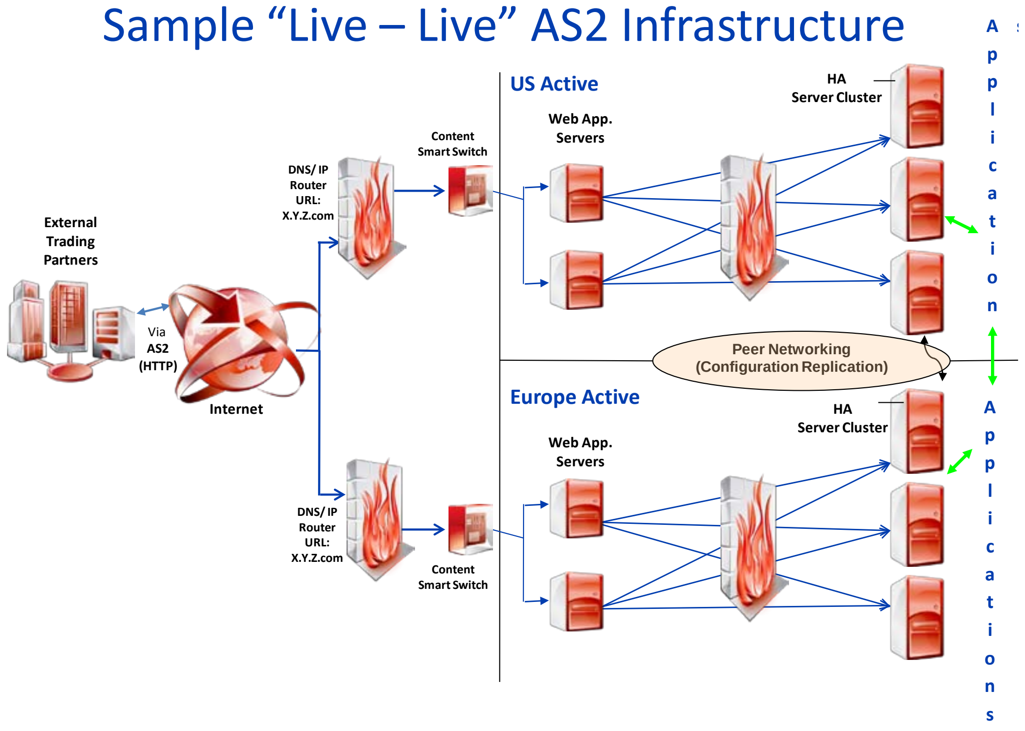
4.1.3. Hardware Requirements

Duplicate hardware will be required at the sites. A reliable high bandwidth connection between the sites will be necessary also. A point–to-point VPN connection between the sites is highly recommended.

4.1.4. Cost Guidelines

The costs for this scenario will be a little more than twice that of a single site.

4.2. Live-Live Failover System – High-Level Transaction Volume with Very Low Tolerance for Recovery Delays



4.2.1. Operational Description

Two complete duplicate systems at different sites have the same domain name (URL) and are always processing a portion of the total AS2 message load. Trading partners send AS2 messages which are variably directed to one site or the other. The AS2 transmission is processed by the DNS/IP Router at the site which responds first (based on internet latency).

At either site, the AS2 message goes through a firewall into the DMZ and is passed to a Content Smart Switch which directs it to one of two Web Application Servers. The Web Application Server completes the HTTP AS2 session in the DMZ. The AS2 message is transferred through a second firewall (out of the DMZ) to one of several High Availability AS2 servers based on load balancing.

Control information about the status of transmissions, partner information, etc. is continually synchronized between the two sites. If one site fails, all transmissions automatically go to the other site. This approach can handle both planned and unplanned outages. This approach is useful if virtually no AS2 downtime is acceptable

4.2.2. Software Requirements

A second AS2 software license is required, as well as an AS2 software solution that supports synchronization of control information (configuration replication) across a wide area network. A software application to transfer payload files to and from internal applications across a wide area network is also required.

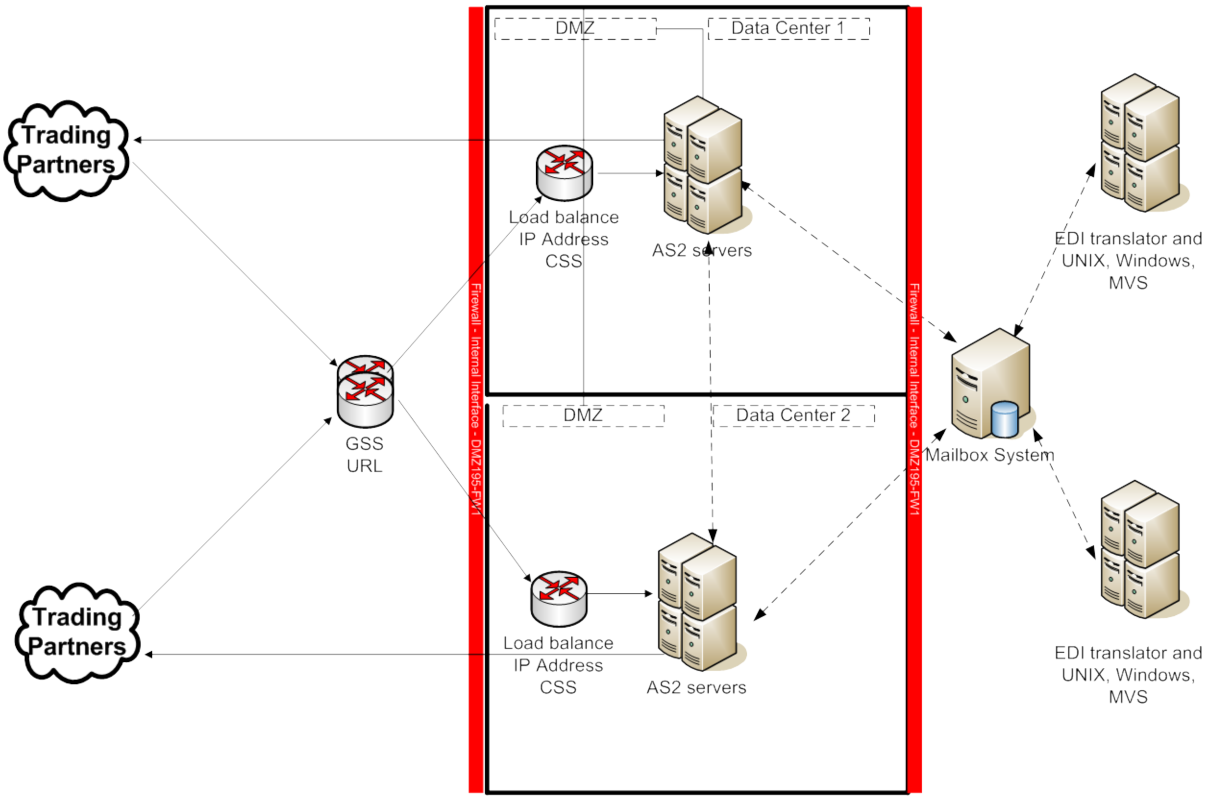
4.2.3. Hardware Requirements

Investment in duplicate hardware at a second site, as well as a High Availability server cluster at each site is required. A reliable high bandwidth network connection between the two sites is also needed.

4.2.4. Cost Guidelines

Costs will be at least two times the cost of a single site solution. However the need for a separate disaster recovery site (or contract) for AS2 connectivity is avoided. If the business cost of losing AS2 connections with customers and suppliers is high, this solution may be worthwhile.

4.3. Live-Live Failover System – High-Level Transaction Volume with Very Low Tolerance for Recovery Delays


DMZ
Data Center 1
Trading Partners
Load balance IP Address CSS
AS2 servers
EDI translator and UNIX, Windows, MVS
Firewall - Internal Interface - DMZ195-FW1
Firewall - Internal Interface - DMZ195-FW1
GSS URL
DMZ
Data Center 2
Mailbox System
Trading Partners
Load balance IP Address CSS
AS2 servers
EDI translator and UNIX, Windows, MVS

4.3.1. Operational Description

Two complete duplicate systems at different sites have the same domain name (URL). Trading partners send AS2 messages to a Global Site Selector (GSS) which then directs the traffic to one site or the other. The GSS renews the IP address from each CSS at a set interval of time. If one data center is down, all traffic will route to the other.

At either site, the AS2 message goes through a firewall into the DMZ and is passed to a Content Smart Switch which directs it to one of several AS2 Application Servers based on server and AS2 application availability. The Web Application Server completes the HTTP AS2 session in the DMZ. The AS2 message is then transferred through a second firewall (out of the DMZ) to one a High Availability server for storage and further application level routing.

Control information about the status of transmissions, partner information, etc. is continually synchronized between the two sites. If one site fails, all transmissions automatically go to the other site. This approach can handle both planned and unplanned outages. This approach is useful if virtually no AS2 downtime is acceptable.

4.3.2. Software Requirements

Depending on your companies' AS2 software contract, additional AS2 software licenses may be required. A software application to transfer payload files to and from internal applications across a wide area network is also required.

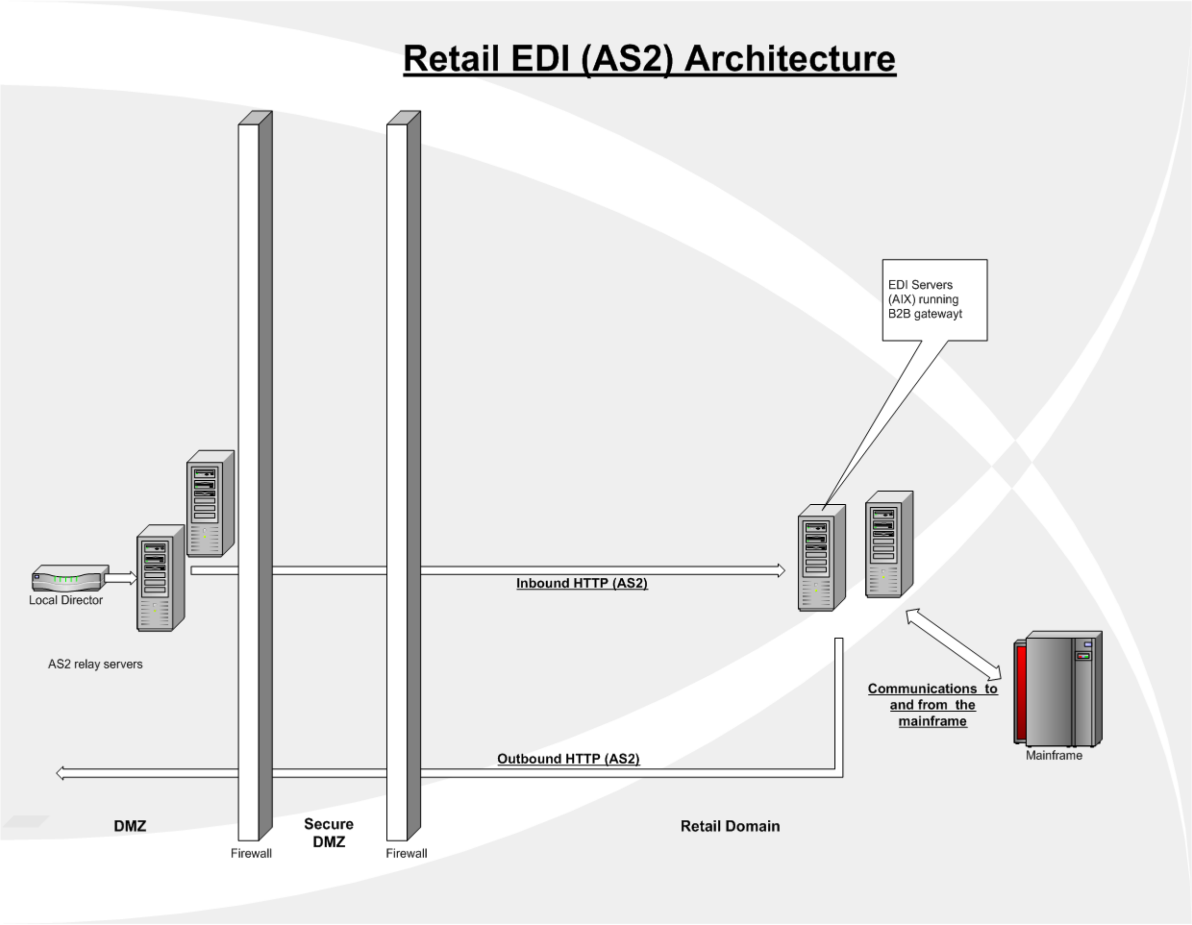
4.3.3. Hardware Requirements

Investment in duplicate hardware at a second site, as well as a High Availability server cluster at each site is required. A reliable high bandwidth network connection between the two sites is also needed. A reliable power backup for both sights is recommended. Several pieces of network routing equipment (Global Site Selector and Content Smart Switches) will be required.

4.3.4. Cost Guidelines

Costs will be at least two times the cost of a single site solution. However the need for a separate disaster recovery site (or contract) for AS2 connectivity is avoided. This solution is also useful if your business model supports units in several different time zones where it is difficult to have downtime for system maintenance. If the business cost of losing AS2 connections with customers and suppliers is high, this solution may be worthwhile.

4.4. Live-Live Failover System – High-Level Transaction Volume with Very Low Tolerance for Recovery Delays



4.4.1. Operational Description

In the primary production system, inbound calls route into the DMZ which uses a load balancer to go to multiple relay servers run off an Intel platform and bring them into the internal domain. It processes all inbound/outbound EDI transactions. The transactions run on mainframe. The gateway is doing XML as well.

There is a duplicate backup facility in a different location which runs a QA environment. Backups are done frequently with copies. There is very low loss of data. Tiered applications are provisioned by priority. By design the test platform must run at 85% of capacity of production.

If a disaster occurs, the system is brought up using a second copy at the backup facility using the same IP address. Communications start going to the secondary site. Full DR tests with simulated outages are done twice a year, isolated from production.

4.4.2. Software Requirements

Depending on your companies' AS2 software contract, additional AS2 software licenses may be required. A software application to transfer payload files to and from internal applications across a wide area network is also required.

4.4.3. Hardware Requirements

Investment in duplicate hardware at a second site, as well as a High Availability server cluster at each site is required. Duplicate hardware is not idle, but runs the QA environment except in case of disaster.

4.4.4. Cost Guidelines

Costs will be at least two times the cost of a single site solution. However the need for a separate disaster recovery site (or contract) for AS2 connectivity is avoided. This solution is also useful if your business model supports units in several different time zones where it is difficult to have downtime for system maintenance. If the business cost of losing AS2 connections with customers and suppliers is high, this solution may be worthwhile.

5. Recovery Considerations and Recommendations

Discussion of possible barriers to recovering AS2 communications and remedies/planning considerations and recommendations. Proactive and Reactive steps are included.

5.1. Firewall

Barrier: Firewall setup at DR site must be identical to firewall rules and port openings to production site. If not identical, trading partners will be denied access.

Recommendation: There should be careful setup of DR site firewall rules and port openings.

5.2. IP Addresses

Barrier: When DR site is activated, IP addresses of various application servers will be different.

Recommendation: Trading partners have to be informed of DR site IP addresses, as well as production site IP addresses.

5.3. Certificates

Barrier: Updating to synchronize DR and production sites can be extremely time-consuming.

Recommendation: Updating of public key certificates can be greatly improved by utilization of certificate automation standard work of GS1 eTG (reference...)

5.4. URL

Barrier: When DR site is activated DNS updates may be delayed to trading partners.

Recommendation: All trading partners should be notified to flush their DNS cache.

Appendix A. Glossary of Terms and Acronyms

The following glossary was updated for the 18-Nov-2010 publication of this document. Please refer to the GDSN glossary in the GS1 GDD (http://gdd.gs1.org/GDD/public/searchableglossary.asp) for the latest version.

5.5. Acronyms and Abbreviations

- **AS2** – Applicability Statement 2 (AS2).
- **DRP** – Disaster Recovery Planning.
- **ERP** – Enterprise Resource Planning.

5.6. Terminology

- A – B
- C – E
- G – I
- M – O
- P
- Q – S
- T –

5.6.1. A – B

Activity

An action that took place during a given time period with a defined start time and end time

AS1

Applicability Statement 1 – An Internet Request For Comment (RFC) defining how applications can securely transport EDI and XML over the Internet using SMTP. It specifies how to transport data files.

AS2

Transports business-critical data over the Internet via HTTP (Hypertext Transfer Protocol) or HTTP/S (Secure HTTP). AS2 provides additional security protection as well as responding with a message letting the sender know that the data was received.

AS3

Applicability Statement 3 – An Internet draft defining how applications can securely transport EDI and XML over the Internet using FTP. It specifies how to transport data files.

Authentication

Ensures the accurate identification of both the sender and the receiver. Is accomplished via digital signatures.

5.6.2. C – E

Ciphertext

Data that has been transformed from a ‘plaintext’ form into encrypted text (an unreadable form) via an encryption process.

Digital Certificate

A document that contains name, serial number, expiration dates & a copy of the owner’s public key; used to encrypt data & validate signatures.

Digital Signature

An electronic signature that can be used to authenticate the identity of the sender of a message, and via the encrypted document digest, to ensure that the original content of the data that has been sent is unchanged.

Document Digest

A unique "fingerprint" summary (128 or 160 bits long) of an input file. It is used to create a digital signature and to ensure that the file has not been altered. It is also called a ‘hash’ and is produced by a checksum program that processes a file.

EDI

Electronic Data Interchange – The exchange of structured business data computer to computer. EDI data format standards are developed by the EDIFACT Working Group of the United Nations and the Accredited Standards Committee (ASC) X12 of the American National Standards Institute.

EDIINT

EDI Over the Internet Working Group – A working group of the IETF that developed the AS1 and AS2 standards.

Encryption

A process that uses a mathematical algorithm and a key to transform data into an unreadable format (called ciphertext). A receiver can then use a key to restore the data to its original content.

5.6.3. G – I

HTTP

Hypertext Transport Protocol - The HyperText Transfer Protocol (HTTP) is the de facto standard for transferring World Wide Web documents.

IETF

Internet Engineering Task Force - The Internet Engineering Task Force is a large, open, international community of network designers, operators, vendors, and researchers concerned with the evolution of the Internet architecture and the smooth operation of the Internet.

Integrity

Ensures that data is neither tampered with nor corrupted in transit. Is accomplished via document digests and digital signatures.

ISP

Internet Service Provider - A company that provides end users (individuals and companies) access to the Internet.

5.6.4. M – O

MDN

Message Disposition Notification – A document, typically digitally signed, acknowledging receipt of data from the sender.

Message

An Internet message consists of header fields (collectively called "the header of the message") followed by a body. The header is a sequence of lines of characters with special syntax. The body is a sequence of characters that follows the header and is separated from the header. See RFC 2822

Message-ID

Message Identifier - A globally unique identifier for a message. The sending implementation must guarantee that the Message-ID is unique. See RFC 2822.

MIME

Multipurpose Internet Mail Extension - MIME is a specification for enhancing the capabilities of standard Internet electronic mail. It offers a simple standardized way to represent and encode a wide variety of media types for transmission via the Internet.

Non-repudiation of Receipt

Confirms that the intended party received the data. Is accomplished via digital signatures and signed MDNs.

5.6.5. P

Payload

The body of the message that contains a business document(s) and is protected by encryption and a digital signature.

Privacy

Ensures that only the intended receiver can view the data. Is accomplished via a combination of encryption algorithms and message packaging.

Private Key

A value known only to the owner, used to create a signature and decrypt data encrypted by its corresponding public key.

Public Key

A value, known by everyone to whom the certificate has been distributed, used to encrypt data and validate a digital signature. Although mathematically related to the private key, it is astronomically difficult to derive from the public key.

5.6.6. Q – S

Receiver of Message

The EDIINT application and/or site which receives the Message containing the business payload. The Receiver of Message sends a MDN back to the Sender of Message.

Retry

When attempting to send an AS2 message, the Sender of Message can encounter transient failures. "Retry" is the term used in this document to refer to an additional send attempt (HTTP POST) of the same message, with the same content and with the same Message-ID value. A Retry can occur whether the Sender of Message requests a Synchronous or Asynchronous MDN.

Resend

When a MDN response is not received in a timely manner, the Sender of Message may choose to resend the original message. Resend only applies when the Sender of Message requests an Asynchronous MDN. Because the message has already been sent, but has presumably not been processed according to expectation, the same message, with the same content and the same Message-ID value is sent again. This operation is referred to as a “resend” of the message. Resending ends when the MDN is received or the resend count is reached.

Resubmit

Neither Resending nor Retrying continue forever, but the data may still need to be exchanged at a later time, so a message may need to be resubmitted. When data that failed to be exchanged or was exchanged but later lost is resubmitted in a new message (with a new Message-ID value), it is called resubmission. Resubmission is normally a manual compensation.

Schema

A set of rules to which an XML document must conform in order to be considered 'valid' according to that schema. This is a specific reference to the World Wide Web Consortium's approved standard XML Schema language.

Sender of Message

The EDIINT application and/or site which transmits the Message containing the business payload to the "Receiver of Message”

S/MIME

Secure MIME - S/MIME (Secure/Multipurpose Internet Mail Extensions) provides a consistent way to send and receive secure MIME data. Based on the popular Internet MIME standard, S/MIME provides the following cryptographic security services for electronic messaging applications: authentication, message integrity and non-repudiation of origin (using digital signatures) and privacy and data security (using encryption).

SMTP

Simple Mail Transport Protocol - An Internet standard for transporting e-mail.

Symmetric Key

A single secret key used to encrypt or decrypt a file, known only by the sender and receiver. It is used in the “symmetric-key encryption” process in which each computer has a secret key (code) that it can use to encrypt information before it is transmitted to another computer. Symmetric-key encryption requires that you know which computers will be talking to each other so you can install the key on each one. Symmetric-key encryption is essentially the same as a secret code that each of the two computers must know in order to decode the information. The code provides the key to decoding the message.

5.6.7. T – Z

UN/EDIFACT

United Nations / Electronic Data Interchange for Administration, Commerce and Transport

VAN

Value Added Network

XML

eXtensible Markup Language. A widely used standard from the World Wide Web Consortium (W3C) that facilitates the interchange of data between computer applications. XML is similar to the language used for Web pages, the HyperText Markup Language (HTML), in that both use markup codes (tags). XML allows the developers create customized tags that offer greater flexibility in organizing and presenting information than is possible with HTML.

Appendix B. Frequently Asked Questions

Question	Reply
What is DR?	Disaster Recovery
What is AS2?	AS2 (Applicability Statement 2) is an Internet Engineering Task Force (IETF) standard (RFC4130) that specifies how to transport data securely and reliably over the Internet. Data can consist of Electronic Data Interchange (EDI) messages or XML messages, but may also be of any other message type or format. AS2 specifies how to connect, deliver, validate and acknowledge data. AS2 creates an envelope for a message which is then sent securely over the Internet. Security is achieved by using digital certificates and encryption. An implementation of AS2 involves two machines, a client and a server, communicating with each other over the Internet. On the operating system level, the AS2 client may be a server, too, offering its communication services to application software. The client sends data to the server, e.g. a trading partner. On receipt of the message the receiving application sends an acknowledgement or MDN (Message Disposition Notification) back to the sender.
Are AS1, AS2, and AS3 official standards?	Yes, AS1, AS2, and AS3 are official standards. The AS2, AS1, and AS3 standards were developed and by and currently governed by the Internet Engineering Task Force (IETF –www.ietf.org). AS2 is registered as IETF standard “RFC4130”, AS1 is registered as “RFC3335”, and AS3 is registered as “RFC4830”.