Network Working Group R. Dixon
Request for Comments: 1434 D. Kushi
IBM
March 1993

Data Link Switching: Switch-to-Switch Protocol

Status of this Memo

This memo provides information for the Internet community. It does not specify an Internet standard. Distribution of this memo is unlimited.

Abstract

This RFC describes IBM's support of Data Link Switching over TCP/IP. The RFC is being distributed to members of the Internet community in order to solicit their reactions to the proposals contained in it. While the issues discussed may not be directly relevant to the research problems of the Internet, they may be interesting to a number of researchers and implementors.

Any questions or comments relative to the contents of this RFC should be sent to the following Internet address: **dlsw@ralvma.vnet.ibm.com**.

Table of Contents

1. Introduction

Data Link Switching (DLS) is a forwarding mechanism for the IBM SNA and IBM NetBIOS protocols. It does not provide full routing, but instead provides switching at the Data Link layer and encapsulation in TCP/IP for transport over the Internet. This memo documents the Switch-to-Switch Protocol (SSP) that is used between IBM 6611 Network Processors.

Today, the IBM 6611 supports SNA (PU 2 and PU 4) systems and NetBIOS systems attached to token-ring networks, as well as SNA (PU 2) systems attached to SDLC links. For the later case, the SDLC attached systems are provided with a LAN appearance within the IBM 6611. For the LAN attached systems, the IBM 6611 appears as a source-routing bridge. Remote systems that are accessed through the IBM 6611 appear as systems attached to an adjacent ring. This ring is a virtual ring that is manifested within each IBM 6611.

2. Overview

Data Link Switching was developed to provide support for SNA and NetBIOS in multi-protocol routers. Since SNA and NetBIOS are basically connection oriented protocols, the Data Link Control procedure that they use on the LAN is IEEE 802.2 Logical Link Control (LLC) Type 2. Data Link Switching also accommodates SNA protocols over WAN links via the SDLC protocol.

IEEE 802.2 LLC Type 2 was designed with the assumption that the network transit delay would be small and predictable (i.e., a local LAN). Therefore the LLC elements of procedure use a fixed timer for detecting lost frames. When bridging is used over wide area lines (especially at lower speeds), the network delay is larger and it can vary greatly based upon congestion. When the delay exceeds the time-out value LLC attempts to retransmit. If the frame is not actually lost, only delayed, it is possible for the LLC Type 2 procedures to become confused. And as a result, the link is eventually taken down.

Given the use of LLC Type 2 services, Data Link Switching addresses the following bridging problems:

- DLC Time-outs
- DLC Acknowledgments over the WAN
- Flow and Congestion Control
- Broadcast Control of Search Packets
- Source-Route Bridging Hop Count Limits

NetBIOS also makes extensive use of datagram services that use LLC Type 1. In this case, Data Link Switching addresses the last two problems in the above list.

The principal difference between Data Link Switching and bridging is that DLS terminates the Data Link Control whereas bridging does not. The following figure illustrates this difference based upon two end systems operating with LLC Type 2 services.

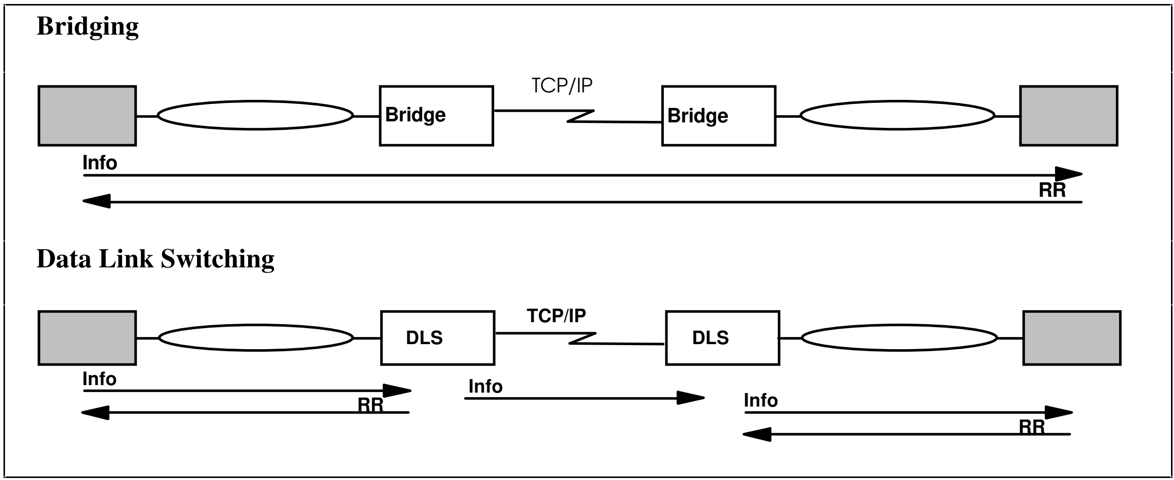

Figure 1. Data Link Switching Contrasted to Bridging

In traditional bridging, the Data Link Control is end-to-end. Data Link Switching terminates the LLC Type 2 connection at the switch. This means that the LLC Type 2 connections do not cross the wide area network. The DLS multiplexes LLC connections onto a TCP connection to another DLS. Therefore, the LLC connections at each end are totally independent of each other. It is the responsibility of the Data Link Switch to deliver frames that it has received from a LLC connection to the other end. TCP is used between the Data Link Switches to guarantee delivery of frames.

As a result of this design, LLC time-outs are limited to the local LAN (i.e., they do not traverse the wide area). Also, the LLC Type 2 acknowledgments (RR's) do not traverse the WAN, thereby reducing traffic across the wide area links. For SDLC links, polling and poll response occurs locally, not over the WAN. Broadcast of search frames is controlled by the Data Link Switches once the location of a target system is discovered. Finally, the switches can now apply back pressure to the end systems to provide flow and congestion control.

Data Link Switching uses LAN addressing to set up connections between SNA systems. SDLC attached devices are defined with MAC addresses to enable them to communicate with LAN attached devices. For NetBIOS systems, Data Link Switching uses the NetBIOS name to forward datagrams and to set up connections for NetBIOS sessions. For circuit establishment, SNA systems send TEST (or in some cases, XID) frames to the null (x'00') SAP. NetBIOS systems have an address resolution procedure, based upon the Name Query and Name Recognized frames, that is used to establish an end-to-end circuit.

Since Data Link Switching may be implemented in multi-protocol routers, there may be situations where both bridging and switching are enabled. SNA frames can be identified by their link SAP. Typical SAP values for SNA are x'04', x'08', and x'0C'. NetBIOS always uses a link SAP value of x'F0'.

3. Transport Connection

Data Link Switches can be in used in pairs or by themselves. A Single DLS internally switches one data link to another without using TCP (DLC(1) to DLC(2) in the figure below). A paired DLS multiplexes data links over a reliable transport using a Switch-to-Switch Protocol (SSP). This RFC will document the frame formats and protocols for this multiplexing between Data Link Switches. The initial implementation of SSP uses TCP as the reliable transport between Data Link Switches. However, other transport connections such as OSI TP4 could be used.

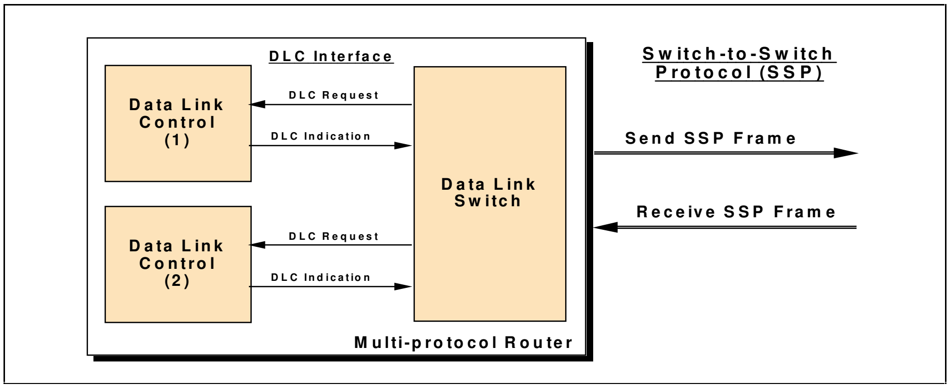

Figure 2. DLS System Diagram

Before Data Link Switching can occur between two routers, they must establish a TCP connection between them. Each DLS will maintain a list of DLS capable routers and their status (active/inactive). Once this connection is established, the DLS will employ SSP to establish end-to-end circuits over the transport connection. Within the transport connection is a specific set of DLS message units. The message formats and types for these PDUs are documented in the following sections.

The default parameters associated with the TCP connections between Data Link Switches are as follows:

• Socket Family	AF_INET	(Internet protocols)
• Socket Type	SOCK_STREAM	(stream socket)
• Read Port Number	2065	
• Write Port Number	2067	

Two or more Data Link Switches may be attached to the same LAN, consisting of a number of token-ring segments interconnected by source-routing bridges. In this case, a TCP connection is not defined between bridges attached to the same LAN. This will allow using systems to select one of the possible Data Link Switches in a similar manner to the selection of a bridge path through a source-routed bridged network. The virtual ring segment in each Data Link Switch attached to a common LAN must be configured with the same ring number. This will prevent LAN frames sent by one Data Link Switch from being propagated through the other Data Link Switches.

3.1. SSP Frame Formats

The following diagrams show the two message headers for traffic between Data Link Switches. The control message header is used for all messages except information messages. The information message header is 16 bytes long, and the control message header is 72 bytes long. The first sixteen bytes of the control message header are identical to the information message header.

CONTROL MESSAGES

Version Number	Reserved Field
Message Length	
Remote Data Link Correlator	
Remote DLC Port ID	
Reserved Field	
Message Type	Reserved Field
Protocol ID	Header Number
Header Length	
Reserved Field	
Reserved Field	Message Type
Target MAC Address	
Origin MAC Address	
Origin Link SAP	Target Link SAP
Frame Direction	Reserved Field
Message Length	
DLC Header Length	
Origin DLC Port ID	
Origin Data Link Correlator	
Origin Transport ID	
Target DLC Port ID	
Target Data Link Correlator	
Target Transport ID	
Reserved Field	

INFORMATION MESSAGE

Version	Reserved Field
Message Length	
Remote Data Link Correlator	
Remote DLC Port ID	
Reserved Field	
Message Type	Reserved Field

The Version Number is set to x'4B' ,indicating a numeric value of 75.

The Header Length is x'00 48', indicating a numeric value of 72 bytes.

The Header Number is x'01', indicating a value of one.

The Frame Direction field is set to x'01' for frames sent from the origin DLS to the target DLS, and is set to x'02' for frames sent from the target DLS to the origin DLS.

> Note: The Remote Data Link Correlator and Remote DLC Port ID are set equal to the Target Data Link Correlator and Target DLC Port ID if the Frame Direction field is set to x'01', and are set equal to the Origin Data Link Correlator and Origin DLC Port ID if the Direction Field is set to x'02'.

The Protocol ID field is set to x'42', indicating a numeric value of 66.

The Message Length field defines the number of bytes within the data field following the header. Note that this value is specified in two different fields of the message header.

The DLC Header Length is set to zero for SNA and is set to x'23' for NetBIOS datagrams, indicating a length of 35 bytes. This includes the Access Control (AC) field, the Frame Control (FC) field, Destination MAC Address (DA), the Source MAC Address (SA), the Routing Information (RI) field (padded to 18 bytes), the Destination link SAP (DSAP), the Source link SAP (SSAP), and the LLC control field (UI).

The values for the Message Type field are defined in a later section. Note that this value is specified in two different fields of the message header.

Reserved fields are set to zero upon transmission and should be ignored upon receipt.

3.2. Address Parameters

A data link is defined as a logical association between the two end stations using Data Link Switching. It is identified by a Data Link ID (14 bytes) consisting of the pair of attachment addresses associated with each end system. Each attachment address is represented by the concatenation of the MAC address (6 bytes) and the LLC address (1 byte).

DATA LINK ID

Target MAC Address	
Origin MAC Address	
Origin Link SAP	Target Link SAP

An end-to-end circuit is identified by a pair of Circuit ID's. A Circuit ID is a 64 bit number that identifies the DLC circuit within a single DLS. It consists of a DLC Port ID (4 bytes), and a Data Link Correlator (4 bytes). This value is unique in a single DLS and is assigned locally. The pair of Circuit ID's along with the identifiers of the Data Link Switches, uniquely identify a single end-to-end circuit .

Each DLS must keep a table of these Circuit ID pairs, one for the local end of the circuit and the other for the remote end of the circuit. In order to identify which Data Link Switch originated the establishment of a circuit, the terms, origin DLS and target DLS, will be employed in this document.

CIRCUIT ID

DLC Port ID	
Data Link Correlator	

The Origin Transport ID and the Target Transport ID fields in the message header are used to identify the individual TCP/IP port on a Data Link Switch. The values have only local significance. However, each Data Link Switch is required to reflect the values contained in these two fields, along with the associated values for DLC Port ID and the Data Link Correlator, when returning a message to the other Data Link Switch.

The following figure shows the use of the addressing parameters during the establishment of an end-to-end connection. The CANUREACH, ICANREACH, and REACH_ACK messages all carry the Data Link ID, consisting of the MAC and Link SAP addresses associated with the two end stations. Upon receipt of a CANUREACH message, the target DLS starts a data link for each port, thereby obtaining a Data Link Correlator. If the target station can be reached, an ICANREACH message is returned to the origin DLS containing the Target Circuit ID parameter. Upon receipt, the origin DLS starts a data link and returns the Origin Circuit ID to the target DLS within the REACH_ACK message. If the REACH_ACK message is not successfully received, the target Data Link Switch can obtain the Origin Circuit ID from a subsequent message (i.e., CONTACT, XIDFRAME, or DGRMFRAME).

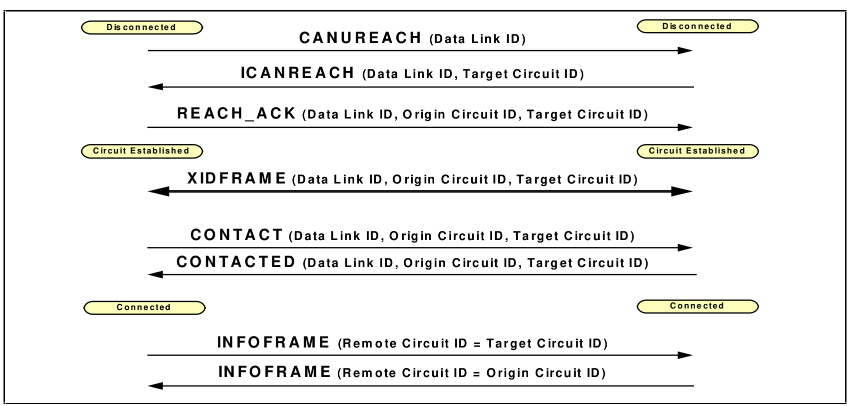

Figure 3. DLS Circuits and Connections

During the exchange of the XIDFRAME, CONTACT, and CONTACTED messages, the pair of Circuit ID parameters is included in the message format along with the DATA LINK ID parameter. Once the connection has been established, the INFOFRAME messages are exchanged with the shorter header. This header contains only the Circuit ID associated with the remote DLS. The Remote Data Link Correlator and the Remote DLC Port ID are set equal to the Data Link Correlator and the DLC Port ID that are associated with the origin or target Data Link Switch, dependent upon the direction of the packet.

3.3. Message Types

The following table lists the protocol data units that are exchanged between Data Link Switches. All values not listed are reserved for potential use in follow-on releases.

Command	Function	Hex Value
CANUREACH	Can U Reach Station	x'03'
ICANREACH	I Can Reach Station	x'04'
REACH_ACK	Reach Acknowledgment	x'05'
DGRMFRAME	Datagram Frame (See note)	x'06'
XIDFRAME	XID Frame	x'07'
CONTACT	Contact Remote Station	x'08'
CONTACTED	Remote Station Contacted	x'09'
RESTART_DL	Restart Data Link	x'10'
DL_RESTARTED	Data Link Restarted	x'11'
INFOFRAME	Information (I) Frame	x'0A'
HALT_DL	Halt Data Link	x'0E'
DL_HALTED	Data Link Halted	x'0F'
NETBIOS_NQ	NetBIOS Name Query	x'12'
NETBIOS_NR	NetBIOS Name Recognized	x'13'
DATAFRAME	Data Frame (See note)	x'14'
NETBIOS_ANQ	NetBIOS Add Name Query	x'1A'
NETBIOS_ANR	NetBIOS Add Name Response	x'1B'

Table 1. SSP Message Types

Note: Both the DGRMFRAME and DATAFRAME messages are used to carry information received by the DLC entity within UI frames. As will be explained below, the DGRMFRAME message is addressed according to a pair of Circuit IDs, while the DATAFRAME message is addressed according to a Data Link ID, being composed of a pair of MAC addresses and a pair of link SAP addresses. The latter is employed prior to the establishment of an end-to-end circuit when Circuit IDs have yet to be established.

For the exchange of NetBIOS control messages, the entire DLC header is carried as part of the message unit. This includes the MAC header, with the routing information field padded to 18 bytes, and the LLC header. The following message types are affected: NETBIOS_NQ, NETBIOS_NR, NETBIOS_ANQ, NETBIOS_ANR, and DATAFRAME when being used by NetBIOS systems. The routing information in the DLC header is not used by the remote Data Link Switch upon receiving the above five messages.

4. Protocol Specification

This section provides a description of the Switch-to-Switch Protocols. Included is a set of high-level protocol flows and a detail set of state transition tables. The states and the protocols are described in terms that are intended to be generic to different platforms. Emphasis of the technical details is to ensure operability of the IBM 6611 with another vendor's implementation. Notes are inserted at points where the IBM 6611 performs local actions that are specific to the AIX platform upon which it operates.

4.1. Protocol Flow Diagrams

The switch-to-switch protocols are used to setup and take down circuits between a pair of Data Link Switches. Once a circuit is established, the end stations on the local networks can employ LLC Type 1 (connectionless) protocols. In addition, the end systems can establish an end-to-end connection for support of LLC Type 2 (connection oriented) protocols.

The term, Data Link, is used in this document to refer to both a "logical data link" when supporting Type 1 LLC services, and a "data link connection" when supporting Type 2 LLC services. In both cases, the Data Link in defined by the concatenation of the destination MAC address (DA), the source MAC address (SA), the destination link SAP (DSAP) and source link SAP (SSAP).

4.1.1. Connect Protocols

The following figure depicts the protocol flows that are used for the establishment of a circuit between a pair of Data Link Switches, followed by the establishment of a connection between the pair of end systems. The figure is drawn assuming that the two end systems are SNA (the protocol flow for NetBIOS systems is described in a later paragraph).

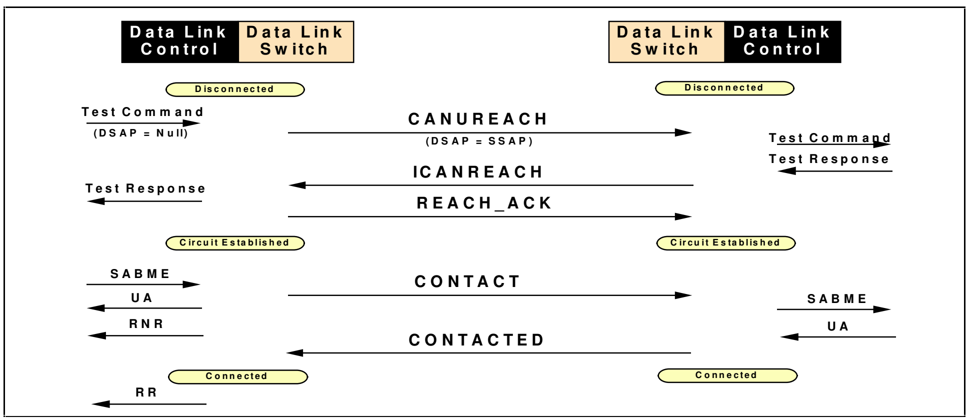

Figure 4. DLS Connect Message Protocols

Upon receipt of a Test command from the origin station, the origin DLS will send a CANUREACH (i.e., can you reach) message to the target DLS. If the target DLS is not known to the origin DLS, the CANUREACH message is sent to all remote Data Link Switches defined to the origin DLS. The receipt of the CANUREACH message causes the target DLS to send a Test command searching for the target station. The target station will return a Test response, causing the target DLS to return an ICANREACH (i.e., I can reach) message to the origin DLS. If multiple Data Link Switches can reach the target station, the origin DLS will receive multiple ICANREACH messages. The origin DLS will select the first message and send a REACH_ACK (i.e., reach acknowledgment) message to the selected Data Link Switch. During this exchange of messages, both Data Link Switches change states from the Disconnected state to the Circuit Established state. Once the circuit is established, Type-1 frames, such as XID, may be exchanged between the origin and target stations.

To establish a connection, the origin station sends a SABME command. Upon receipt of this command, the origin DLS will send a CONTACT message to the target DLS and return a UA response to the origin station. To inhibit traffic flow until the connection is established to the remote station, a RNR supervisory frame is sent to the origin station. The CONTACT message will cause the target DLS to send a SABME command to the target station, which in return will reply with a UA response. Upon receipt of the UA response, the target DLS will send a CONTACTED message to the origin DLS. The origin DLS will now send an RR supervisory frame to the origin station. During this exchange of messages, both Data Link Switches change states from the Circuit Established state to the Connected state.

For NetBIOS end systems, the protocol flows are similar but employ different frames and SSP messages. Instead of using a Test command frame to initiate the circuit, a NetBIOS system will use a Name Query frame. Receipt of a Name Query frame will cause the Data Link Switch to issue a NETBIOS_NQ message instead of the CANUREACH message. In a like fashion, the Test response is replaced with a Name Recognized frame and the ICANREACH message is replaced with a NETBIOS_NR message. As with the SNA protocol flows, the receipt of a NETBIOS_NR message causes the origin Data Link Switch to respond with a REACH_ACK message.

4.1.2. Link Restart Protocols

The following figure depicts the protocol flows that result from restarting the end-to-end connection. This causes the Data Link Switches to terminate the existing connection and to enter the Circuit Established state awaiting the start of a new connection.

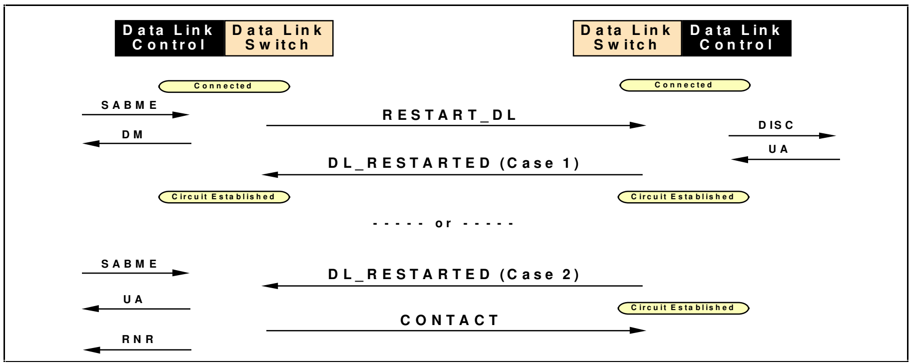

Figure 5. DLS Link Restart Message Protocols

Upon receipt of a SABME command from the origin station, the origin DLS will send a RESTART_DL message to the target DLS. A DM response is also returned to the origin station and the data link is restarted.

Upon receipt of the RESTART_DL message, the target DLS will issue a DISC command to the target station. The target station is expected to return a UA response. The target DLS will then restart its data link and send an DL_RESTARTED message back to the origin DLS. During this exchange of messages, both Data Link Switches change states from Connected state to Circuit Established state.

If the origin station now resends the SABME command, the origin DLS will send a CONTACT message to the target DLS. If the SABME command is received prior to the receipt of the DL_RESTARTED message (case 2 in the figure), the CONNECT message is delayed until the DL_RESTARTED message is received. The resulting protocol flows at this point parallel those given above for the connect sequence.

4.1.3. Disconnect Protocols

The following figure depicts the protocol flows that result from the end system terminating an existing connection. Not only is the connection terminated, but the circuit between the Data Link Switches is taken down.

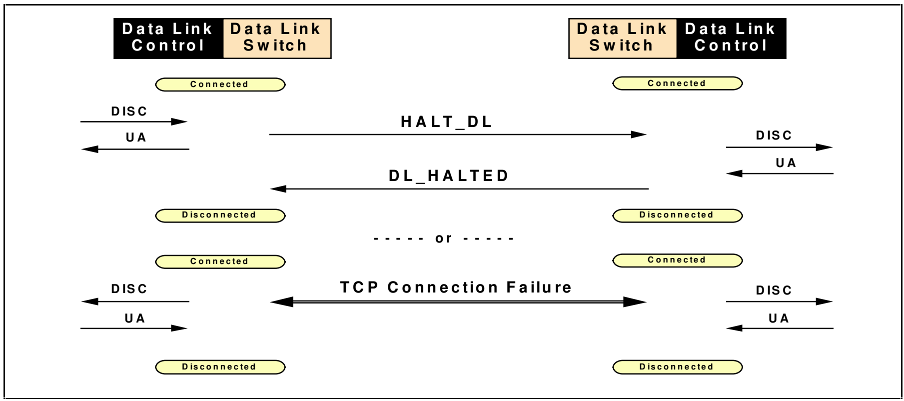

Figure 6. DLS Disconnect Message Protocols

Upon receipt of a DISC command from the origin station, the origin DLS will reply with a UA response and issue a HALT_DL message to the target DLS. Upon receipt of the HALT_DL message, the target DLS will send a DISC command to the target station. The target station will then respond with a UA response, causing the target DLS to return a DL_HALTED message to the origin DLS. During this exchange of messages, both Data Link Switches change states from the Connected state to the Disconnected state.

If the TCP connection between two Data Link Switches fails, all connections that are currently multiplexed on the failed TCP connection will be taken down. This implies that both Data Link Switches will send DISC commands to all the local systems that are associated with the failed connections. Upon sending the DISC command, the Data Link Switch will enter the DISCONNECT state for each circuit.

4.2. DLS State Machine

The following state tables describe the states for a single connection through the Data Link Switch. State information is kept for each connection. The initial state for a connection is DISCONNECT. The steady state is either CIRCUIT_ESTABLISHED or CONNECTED. In the former state, an end-to-end circuit has been established allowing the support of Type 1 LLC between the end systems. The latter state exists when an end-to-end connection has been established for the support of Type 2 LLC services between the end systems.

For SNA, circuit establishment is via the use of IEEE 802.2 Test or XID frames. SNA devices send these frames to the null SAP in order to determine the source route information in support of bridging. Normally SNA devices use SAP x'04', x'08', or x'0C'. Typically the SAP would be used to determine if the Test frames should be sent to the DLS code in the router. If both bridging and DLS are enabled, this allows the product to ensure that SNA frames are not both bridged and switched.

For NetBIOS, circuit establishment is via the Name Query and Name Recognized frames. These frames are used for both address resolution and source route determination. Normally NetBIOS devices use SAP x'F0'.

4.2.1. Data Link Switch States

The Switch-to-Switch Protocols will be formally defined through a single state machine. The following table lists the eleven possible states. A separate state machine is employed for each end-to-end circuit that is maintained by the Data Link Switch. The three steady states are DISCONNECTED, CIRCUIT_ESTABLISHED, and CONNECTED.

State Name	Description
CIRCUIT_ESTABLISHED	The end-to-end circuit has been established. At this time LLC Type 1 services are available from end-to-end.
CIRCUIT_PENDING	The target DLS is awaiting a REACH_ACK response to an ICANREACH message.
CIRCUIT_RESTART	The DLS that originated the reset is awaiting the restart of the data link and the DL_RESTARTED response to a RESTART_DL message.
CONNECTED	The end-to-end connection has been established, thereby allowing LLC Type 2 services from end-to-end in addition to LLC Type 1 services.
CONNECT_PENDING	The origin DLS is awaiting the CONTACTED response to a CONTACT message.
CONTACT_PENDING	The target DLS is awaiting the DLC_CONTACTED confirmation to a DLC_CONTACT signal (i.e., DLC is waiting for a UA response to an SABME command).
DISCONNECTED	The initial state with no circuit or connection established, the DLS is awaiting either a CANUREACH, an ICANREACH, a NETBIOS_NQ, or a NETBIOS_NR message.
DISCONNECT_PENDING	The DLS that originated the disconnect is awaiting the DL_HALTED response to a HALT_DL message.
HALT_PENDING	The remote DLS is awaiting the DLC_DL_HALTED indication following the DLC_HALT_DL request (i.e., DLC is waiting for a UA response to a DISC command).
RESTART_PENDING	The remote DLS is awaiting the DLC_DL_HALTED indication following the DLC_HALT_DL request (i.e., DLC is waiting for a UA response to a DISC command), and the restart of the data link.
RESOLVE_PENDING	The target DLS is awaiting either the DLC_DL_STARTED indication following the DLC_START_DL request (i.e., DLC is waiting for a Test response as a result of sending a Test command), or a NB_Name_Recognized frame in response to a NB_Name_Query frame.

Table 2. Data Link Switch States

The following figure depicts the events that cause a transition to a new state, along with the principal protocol action(s) that occurs. On each transition, the event that caused the transition is shown above the line, and the resulting action is shown below the line. SSP messages that are sent and received are shown in plain text, while the DLC requests and indications are prefaced by the letters, DLC, and are shown in italics (i.e., *DLC_CONTACT*). The DISCONNECT_PENDING state will be entered whenever a DLC error condition occurs in any of the other states (except RESOLVE_PENDING), or when a DISC command is received by the DLC.

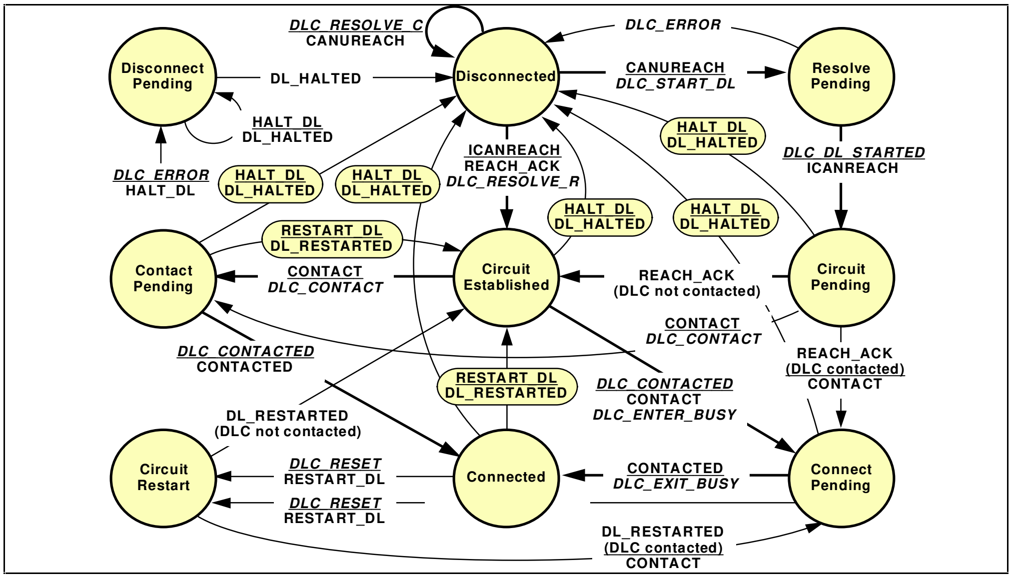

Figure 7. DLS State Transitions

The DISCONNECT state is the initial state for a new circuit. One end station starts the connection via a TEST or XID command (i.e., *DLC_RESOLVE_C*) or a NetBIOS Name Query command (not explicitly shown). Upon receipt, the Data Link Switches exchange a set of CANUREACH, ICANREACH and REACH_ACK messages. Upon completion of this three-legged exchange, both Data Link Switches will be in the CIRCUIT_ESTABLISHED state. Two pending states also exist during this exchange. The RESOLVE_PENDING state is entered by the target Data Link Switch awaiting a Test response to a Test Command and the CIRCUIT_PENDING state is entered by the target DLS awaiting the REACH_ACK reply to an ICANREACH message.

The CIRCUIT_ESTABLISHED state allows for the exchange of LLC Type 1 frames such as the XID exchanges between SNA stations that occurs prior to the establishment of a connection. Also, datagram traffic (i.e., UI frames) may be sent and received between the end stations. These exchanges use the XIDFRAME and DGRMFRAME messages sent between the Data Link Switches.

In the CIRCUIT_ESTABLISHED state, the receipt of a SABME command (i.e., *DLC_CONTACTED*) causes the origin DLS to issue a CONTACT message, to send an RNR supervisory frame (i.e., *DLC_ENTER_BUSY*) to the origin station, and to enter the CONNECT_PENDING state awaiting a CONTACTED message. The target DLS, upon the receipt of a CONTACT message, will issue a SABME command (i.e., *DLC_CONTACT)* and enter the Contact Pending state. Once the UA response is received (i.e., *DLC_CONTACTED*), the target DLS sends a CONTACTED message and enters the CONNECTED state. When received, the origin DLS enters the Connected state and sends an RR supervisory frame (i.e., *DLC_EXIT_BUSY*).

The CONNECTED state is the steady state for normal data flow once a connection has been established. Information frames (i.e., INFOFRAME messages) are simply sent back and forth between the end points of the connection. This is the path that should be optimized for performance.

The connection is terminated upon the receipt of a DISC frame or under some other error condition detected by DLC (i.e., *DLC_ERROR*). Upon receipt of this indication, the DLS will halt the local data link, send a HALT_DL message to the remote DLS, and enter the DISCONNECT_PENDING State. When the HALT_DL frame is received by the other DLS, the local DLC is halted for this data link, a DL_HALTED message is returned, and the DISCONNECTED state is entered. Receipt of this DL_HALTED message causes the other DLS to also enter the DISCONNECTED state.

The CIRCUIT_RESTART state is entered if one of the Data Link Switches receives a SABME command (i.e., *DLC_RESET*) while in the CONNECTED state. This causes a DM command to be returned to the origin station and a RESTART_DL message to be sent to the remote Data Link Switch. This causes the remote data link to be halted and then restarted. The remote DLS will then send a DL_RESTARTED message back to the first DLS. The receipt of the DL_RESTARTED message causes the first DLS to issue a new CONTACT message, assuming that the local DLC has been contacted (i.e., the origin station has resent the SABME command). This is eventually responded to by a CONTACTED message. Following this exchange, both Data Link Switches will return to the CONNECTED state. If the local DLC has not been contacted, the receipt of a DL_RESTARTED command causes the Data Link Switch to enter the CIRCUIT_ESTABLISHED state awaiting the receipt of a SABME command (i.e., *DLC_CONTACTED* signal).

In the figure, the oval state containing (HALT_DL and DL_HALTED) represents the HALT_PENDING state and the oval state containing (RESTART_DL and DL_RESTARTED) represents the RESTART_PENDING state. Both these oval states have only one input event and only one output action. While in these states, the Data Link Switch is awaiting responses from the local station on the adjacent LAN (i.e., a UA response to a DISC command). Also in the RESTART_PENDING state, the Data Link Switch will attempt to restart the data link prior to sending a DL_RESTARTED message. For the IBM 6611 implementation, the start of a data link involves the exchange of a Test command/response on the adjacent LAN (i.e., *DLC_START_DL*). For other implementations, this additional exchange may not be required.

4.2.2. State Transition Tables

This section provides a detail representation of the Data Link Switch, as document by a set of state machines. The first state machine documents the sending and receiving of SSP messages. Many of the transitions are dependent upon local signals between the Data Link Switch entity and one of the DLC entities. These signals and their definitions are given in the following tables.

Event Name	Description
DLC_CONTACTED	Contact Indication: DLC has received an SABME command and will send a UA response, or DLC has received a UA response as a result of sending an SABME command.
DLC_DGRM	Datagram Indication: DLC has received a UI frame.
DLC_ERROR	Error condition indicated by DLC: Such a condition occurs when a DISC command is received or when DLC experiences an unrecoverable error.
DLC_INFO	Information Indication: DLC has received an Information (I) frame.
DLC_DL_HALTED	Data Link Halted Indication: DLC has received a UA response to a DISC command.
DLC_DL_STARTED	Data Link Started Indication: DLC has received a Test response from the null SAP.
DLC_RESET	Reset Indication: DLC has received an SABME command during the time a connection is currently active and has responded with DM.
DLC_RESOLVE_C	Resolve Command Indication: DLC has received a Test command addressed to the null SAP, or an XID command addressed to the null SAP.
DLC_XID	XID Indication: DLC has received an XID command or response to a non-null SAP.

Table 3. Local DLC Events

Action Name	Description
DLC_CONTACT	Contact Station Request: DLC will send a SABME command.
DLC_DGRM	Datagram Request: DLC will send a UI frame.
DLC_ENTER_BUSY	Enter Link Station Busy: DLC will send an RNR supervisory frame.
DLC_EXIT_BUSY	Exit Link Station Busy: DLC will send an RR supervisory frame.
DLC_HALT_DL	Halt Data Link Request: DLC will send a DISC command.
DLC_INFO	Information Request: DLC will send an I frame.
DLC_RESOLVE_R	Resolve Response Request: DLC will send a Test response or XID response from the null SAP.
DLC_START_DL	Start Data Link Request: DLC will send a Test command to the null SAP.
DLC_XID	XID Request: DLC will send an XID command or an XID response.

Table 4. Local DLC Actions

The Data Link Switch may be described by a state transition table consisting of eleven states. Each of these states is described below in terms of the events, actions, and next state for each transition. If a particular event is not listed for a given state, no action and no state transition should occur for that event. Any significant comments concerning the transitions within a given state are given immediately following the table representing the state.

A separate state machine is maintained by the Data Link Switch for each end-to-end circuit. The number of circuits that may be supported by each Data Link Switch is a local implementation option.

State: DISCONNECTED

Event	Action(s)	Next State
Receive CANUREACH	DLC_START_DL	RESOLVE_PENDING
Receive ICANREACH	Send REACH_ACK, DLC_RESOLVE_R (See note 1)	CIRCUIT_ESTABLISHED
Receive DATAFRAME	DLC_DGRM	
Receive NETBIOS_NQ	DLC_DGRM	RESOLVE_PENDING
Receive NETBIOS_NR	Send REACH_ACK, DLC_DGRM (See note 2)	CIRCUIT_ESTABLISHED
DLC_RESOLVE_C	Send CANUREACH	
DLC_DGRM	If NB_Name_Query: Send NETBIOS_NQ, Else: Send DATAFRAME	

It is assumed that each Data Link Switch will build a set of topology tables giving the identity of each Data Link Switch that can reach a specific MAC address or a specific NetBIOS name. This table can be built based upon the origin address information received within the CANUREACH message or

NETBIOS_NQ message and the target address information within the ICANREACH message or NETBIOS_NR message. As a consequence, the amount of search traffic can be kept to a minimum.

Upon receipt of a Test command or XID command to the null SAP (i.e., DLC_RESOLVE_C signal from DLC), the Data Link Switch will check the topology table prior to sending the CANUREACH message. If the target MAC address is in the table, the CANUREACH message will be sent to only those Data Link Switches that are known to be able to reach the given MAC address. If the MAC address is not in the table, the CANUREACH message will be sent to all known Data Link Switches. Since the destination link SAP (DSAP) value is null, the Target Link SAP field in the CANUREACH message header is set equal to the Origin Link SAP value.

Upon receipt of a NB_Name_Query the Data Link Switch will check the topology table prior to sending the NETBIOS_NQ message. If the target NetBIOS name is in the table, the NETBIOS_NQ message will be sent to only those Data Link Switches that are known to be able to reach the given NetBIOS name. If the NetBIOS name is not in the table, the NETBIOS_NQ message will be sent to all known Data Link Switches.

For SNA, the DISCONNECTED state is exited upon receipt of a CANUREACH message by a prospective target Data Link Switch, or upon receipt of an ICANREACH message by the origin Data Link Switch. In the former case, the Data Link Switch will issue a Test command to the target station (i.e., DLC_START_DL signal is presented to DLC). In the later case, a Test response is sent to the origin station (i.e., DLC_RESOLVE_R will be issued) and a REACH_ACK message will be returned to the target Data Link Switch.

> Note 1- The IBM 6611 will not send a Test response, but will send a Test command to the station that originated the resolve procedure (i.e., a DLC_START_DL will be issued) in order to start the data link.

For NetBIOS, the DISCONNECTED state is exited upon the receipt of a NETBIOS_NQ message by the prospective target Data Link Switch, or upon the receipt of a NETBIOS_NR message by the origin Data Link Switch. In the former case, the Data Link Switch will send a NB_Name_Query frame. In the later case, the Data Link Switch will send a NB_Name_Recognized frame to the origin station and a REACH_ACK message will be returned to the target Data Link Switch.

> Note 2- The IBM 6611 will also send a Test command (i.e., a DLC_START_DL will be issued) to the station that originated the name resolution in order to start the data link.

State: RESOLVE_PENDING

Event	Action(s)	Next State
Receive DATAFRAME	DLC_DGRM	
DLC_DL_STARTED	Send ICANREACH	CIRCUIT_PENDING
DLC_ERROR		DISCONNECTED
DLC_DGRM	If NB_Name_Recognized: Send NETBIOS_NR (See note), Else: Send DATAFRAME	If NB_Name_Recognized: CIRCUIT_PENDING

The RESOLVE_PENDING state is entered upon receipt of a CANUREACH message or a NETBIOS_NQ message by the target DLS. In the former case, a data link is started, causing a Test command to be sent by the DLC. In the later case, a NB_Name_Query frame is sent. This frame is directed to a group MAC address.

Any CANUREACH messages received in the RESOLVE_PENDING state will be responded to if a DLC_DL_STARTED signal is received. The Data Link Switch may also update its topology information based upon the origin MAC address information in each CANUREACH message.

Upon the receipt of a DLC_DL_STARTED signal in the RESOLVE_PENDING state, the Data Link Switch may update its topology table base upon the remote MAC address information. The ICANREACH message should be returned to all Data Link Switches that had sent a CANUREACH message. In a similar fashion, the Data Link Switch may update its topology table upon the receipt of a NB_Name_Recognized frame and a NETBIOS_NR message will be returned to all Data Link Switches that have sent a NETBIOS_NQ message.

The RESOLVE_PENDING state is exited once the data link has been started (i.e., a DLC_DL_STARTED signal is received as a result of a Test response received by the DLC) or a NB_Name_Recognized frame is received (i.e., a DLC_DGRM signal is received). The target Data Link Switch will then enter the CIRCUIT_PENDING state.

> Note- The IBM 6611 will also send a Test command in order to start the data link to the station that responded to the Name Query frame (i.e., a DLC_START_DL will be issued).

State: CIRCUIT_PENDING

Event	Action(s)	Next State
Receive CONTACT	DLC_CONTACT	CONTACT_PENDING
Receive HALT_DL	DLC_HALT_DL	HALT_PENDING
Receive REACH_ACK	If Connected: Send CONTACT	If Connected: CONNECT_PENDING, else: CIRCUIT_ESTABLISHED
Receive XIDFRAME	DLC_XID	
Receive DGRMFRAME	DLC_DGRM	
Receive DATAFRAME	DLC_DGRM	
DLC_CONTACTED	DLC_ENTER_BUSY	
DLC_ERROR	Send HALT_DL	DISCONNECT_PENDING
DLC_XID	Send XIDFRAME	
DLC_DGRM	Send DGRMFRAME	

The CIRCUIT_PENDING state is entered by the target Data Link Switch following the sending of an ICANREACH message or the sending of a NETBIOS_NR message. In this state it is awaiting the reception of a REACH_ACK message from the origin Data Link Switch. If a connection does not exist with the target station (i.e., the normal case), the Data Link Switch will enter the CIRCUIT_ESTABLISHED state.

If the target Data Link Switch happens to receive a SABME command from the target station while in the CIRCUIT_PENDING state (i.e., a DLC_CONTACTED signal received from the DLC), the reception of the REACH_ACK message will cause the Data Link Switch to enter the CONNECT_PENDING state and to send a CONTACT message to the other Data Link Switch. Thus the target Data Link Switch has assumed the role of the origin Data Link Switch.

State: CONNECT_PENDING		
Event	**Action(s)**	**Next State**
Receive CONTACTED	DLC_EXIT_BUSY	CONNECTED
Receive HALT_DL	DLC_HALT_DL	HALT_PENDING
Receive DGRMFRAME	DLC_DGRM	
Receive DATAFRAME	DLC_DGRM	
DLC_RESET	Send RESTART_DL (See note)	CIRCUIT_RESTART
DLC_ERROR	Send HALT_DL	DISCONNECT_PENDING
DLC_DGRM	Send DGRMFRAME	

The CONNECT_PENDING state is entered by the origin Data Link Switch when a DLC_CONTACTED signal has been received from the DLC (i.e., a SABME command has been received). A CONTACT message is then issued. The state is exited upon the receipt of a CONTACTED message from the target Data Link Switch. If a DLC_RESET signal is received, the local data link is restarted and a RESTART_DL message is sent to the remote DLS.

> Note- The IBM 6611 will also send a Test command in order to restart the data link to the station that sent the SABME command (i.e., a DLC_START_DL will be issued).

State: CIRCUIT_ESTABLISHED		
Event	**Action(s)**	**Next State**
Receive CONTACT	DLC_CONTACT	CONTACT_PENDING
Receive HALT_DL	DLC_HALT_DL	HALT_PENDING
Receive XIDFRAME	DLC_XID	
Receive DGRMFRAME	DLC_DGRM	
Receive DATAFRAME	DLC_DGRM	
DLC_CONTACTED	Send CONTACT DLC_ENTER_BUSY	CONNECT_PENDING
DLC_ERROR	Send HALT_DL	DISCONNECT_PENDING
DLC_DGRM	Send DGRMFRAME	
DLC_XID	Send XIDFRAME	

The CIRCUIT_ESTABLISHED state is entered by the origin Data Link Switch from the DISCONNECTED state, and by the target Data Link Switch from the CIRCUIT_PENDING state. The state is exited when a connection is started (i.e., DLC receives a SABME command). The next state is CONTACT_PENDING for the target Data Link Switch and CONNECT_PENDING for the origin Data Link Switch.

State: CONTACT_PENDING		
Event	**Action(s)**	**Next State**
Receive HALT_DL	DLC_HALT_DL	HALT_PENDING
Receive RESTART_DL	DLC_HALT_DL	RESTART_PENDING
Receive DGRMFRAME	DLC_DGRM	
Receive DATAFRAME	DLC_DGRM	
DLC_CONTACTED	Send CONTACTED	CONNECTED
DLC_ERROR	Send HALT_DL	DISCONNECT_PENDING
DLC_DGRM	Send DGRMFRAME	

The CONTACT_PENDING state is entered by the target Data Link Switch upon the receipt of a CONTACT message. This causes the Data Link Switch to issue a DLC_CONTACT signal to the DLC (i.e., DLC sends a SABME command). This state is then exited upon the receipt of a DLC_CONTACTED signal from the DLC (i.e., a UA response received).

If a RESTART_DL message is received, indicating that the remote Data Link Switch has received a DLC_RESET signal, the local Data Link Switch will send a DISC command frame on the adjacent LAN (i.e., DLC_HALT_DL signal) and enter the RESTART_PENDING state.

State: CONNECTED

Event	Action(s)	Next State
Receive HALT_DL	DLC_HALT_DL	HALT_PENDING
Receive RESTART_DL	DLC_HALT_DL	RESTART_PENDING
Receive DGRMFRAME	DLC_DGRM	
Receive INFOFRAME	DLC_INFO	
Receive DATAFRAME	DLC_DGRM	
DLC_RESET	Send RESTART_DL (See note)	CIRCUIT_RESTART
DLC_ERROR	Send HALT_DL	DISCONNECT_PENDING
DLC_DGRM	Send DGRMFRAME	
DLC_INFO	Send INFOFRAME	

The CONNECTED state is entered by the origin Data Link Switch from the CONNECT_PENDING state upon the receipt of a CONTACTED message. The CONNECTED state is entered by the target Data Link Switch from the CONTACT_PENDING state upon the receipt of a DLC_CONTACTED signal. At this time, the target Data Link Switch will return a CONTACTED message to the origin Data Link Switch.

The CONNECTED state is exited usually under one of two conditions: a DLC_ERROR signal received from the DLC (e.g., a DISC command received by the local DLC), or a HALT_DL message received from the other Data Link Switch (e.g., a DISC command received by the remote DLC).

A SABME command (i.e., a DLC_RESET signal) received by either Data Link Switch will also cause the two Data Link Switches to leave the CONNECTED state and eventually restart a new circuit.

> Note- The IBM 6611 will also send a Test command in order to restart the data link to the station that sent the SABME command (i.e., a DLC_START_DL will be issued).

Following the receipt of a reset signal, the Data Link Switch will send a RESTART_DL message to the other Data Link Switch and will enter the CIRCUIT_RESTART state. Upon the receipt of the RESTART_DL message, the remote Data Link Switch will send a DISC command (i.e., DLC_HALT_DL signal) and enter the RESTART_PENDING state.

State: CIRCUIT_RESTART

Event	Action(s)	Next State
Receive DL_RESTARTED	If Connected: Send CONTACT	If Connected: CONNECT_PENDING, else: CIRCUIT_ESTABLISHED
Receive DATAFRAME	DLC_DGRM	
DLC_ERROR	Send HALT_DL	DISCONNECT_PENDING
DLC_DGRM	Send DATAFRAME	

The CIRCUIT_RESTART state is entered if a DLC_RESET signal is received from the local DLC. This was caused by the receipt of a SABME command while a connection was currently active. A DM response will be issued to the SABME command and the Data Link Switch will attempt to restart the end-to-end circuit.

The CIRCUIT_RESTART state is exited through one of two transitions. The next state depends upon the time the local DLC has reached the contacted state (i.e., a DLC_CONTACTED signal is presented) relative to the receipt of the DL_RESTARTED message. This signal is caused by the origin station resending the SABME command that initially caused the DATA Link Switch to enter the CIRCUIT_RESTART state. The two cases are as follows:

1) DL_RESTARTED message received before the DLC_CONTACTED signal- In this case, the CIRCUIT_ESTABLISHED state is entered.

2) DL_RESTARTED message received after the DLC_CONTACTED signal- In this case, the CONNECT_PENDING state is entered.

State: DISCONNECT_PENDING

Event	Action(s)	Next State
Receive DL_HALTED		DISCONNECTED
Receive HALT_DL	Send DL_HALTED	
Receive DATAFRAME	DLC_DGRM	
DLC_DGRM	Send DATAFRAME	

The DISCONNECT_PENDING state is entered when a DLC_ERROR signal is received from the local DLC. Upon receipt of this signal, a HALT message is sent. Once an DL_HALTED message is received, the state is exited, and the Data Link Switch enters the DISCONNECTED state.

State: RESTART_PENDING

Event	Action(s)	Next State
Receive DATAFRAME	DLC_DGRM	
DLC_DL_HALTED (See note)	Send DL_RESTARTED	CIRCUIT_ESTABLISHED
DLC_ERROR	Send HALT_DL	DISCONNECT_PENDING
DLC_DGRM	Send DATAFRAME	

The RESTART_PENDING state is entered upon the receipt of a RESTART_DL message from the remote DLS while the local Data Link Switch is in either the CONTACT_PENDING state or the CONNECTED state. These cause the local DLC to issue a DISC command. Upon the receipt of the UA response (DLC_DL_HALTED), the data link is restarted, a DL_RESTARTED message is returned to the remote DLS, and the CIRCUIT_ESTABLISHED state is entered.

Note- The IBM 6611 will send a Test command in order to restart the data link to the target station (i.e., a DLC_START_DL will be issued) prior to sending the DL_RESTARTED message.

State: HALT_PENDING		
Event	**Action(s)**	**Next State**
Receive DATAFRAME	DLC_DGRM	
DLC_DL_HALTED	Send DL_HALTED	DISCONNECTED
DLC_ERROR	Send DL_HALTED	DISCONNECTED
DLC_DGRM	Send DATAFRAME	

The HALT_PENDING state is entered upon the receipt of a HALT_DL message. This causes the local DLC to issue a DISC command. Upon the receipt of the UA response (DLC_DL_HALTED), a DL_HALTED message is returned to the remote DLS and the DISCONNECTED state is entered.

4.3. NetBIOS Datagrams

The NetBIOS protocols use a number of UI frames for directory services and the transmission of datagrams. Most of these frames are directed to a group MAC address (GA) with the routing information field indicating spanning tree explorer (STE). Two of the frames, NB_Add_Name_Response and NB_Status_Response, are directed to a specific MAC address with the routing information field indicating a specifically routed frame (SRF). The handling of these frames is summarized in the following table.

Event	Action(s)	Comment
DLC_DGRM (NB Group Address): NB_Add_Name_Query	Send NETBIOS_ANQ	Transmitted to all remote DLS
DLC_DGRM (Specific Address): NB_Add_Name_Response	Send NETBIOS_ANR	Transmitted to specific DLS
DLC_DGRM (Specific Address): NB_Status_Response	Send DATAFRAME	Transmitted to all remote DLS
DLC_DGRM (NB Group Address): NB_Name_in_Conflict, NB_Add_Group_Name_Query, NB_Datagram, NB_Datagram_Broadcast, NB_Status_Query, NB_Terminate_Trace	Send DATAFRAME	Transmitted to all remote DLS

Table 5. NetBIOS DLC Frames

The above actions do not apply in the following states: CIRCUIT_ESTABLISHED, CONTACT_PENDING, CONNECT_PENDING, CONNECTED, and CIRCUIT_PENDING. The handling of the remaining two UI frames used by NetBIOS systems, NB_Name_Query and NB_Name_Recognized, are documented as part of the DLS state machine in the previous section (i.e., DISCONNECTED and RESOLVE_PENDING states). Furthermore, the handling of NetBIOS datagrams (i.e., NB_Datagram) sent to a specific MAC address is also governed by the DLS state machine.

> Note- The IBM 6611 will also issue Test frames during the exchange of the NetBIOS, NB_Name_Query and NB_Name_Recognized. This exchange of protocol data units occurs during the start of a data link and is used to determine the routing information. Most other implementations of NetBIOS will use the NB_Name_Query/NB_Name_Recognized exchange to determine routes in conjunction with resolving the NetBIOS names. These differences are not reflected in the SSP protocols.

The handling of the NetBIOS specific SSP messages is given in the following table.

Event	Action(s)	Comment
NETBIOS_ANQ	DLC_DGRM: NB_Add_Name_Query	Routed STE (NB Group Address)
NETBIOS_ANR	DLC_DGRM: NB_Add_Name_Response	Routed SRF (Specific MAC Address)
NETBIOS_NQ	DLC_DGRM: NB_Name_Query	Routed STE (NB Group Address)
NETBIOS_NR	DLC_DGRM: NB_Name_Recognized	Routed SRF (Specific MAC Address)
DATAFRAME	DLC_DGRM	Routed STE (If NB_Status_Response: Specific MAC Address Else: NB Group Address)

Table 6. NetBIOS SSP Messages

The above actions apply to all DLS states. The handling of NetBIOS datagrams sent within DGRMFRAME messages is governed by the DLS state machine. The DGRMFRAME message type is employed instead of the DATAFRAME message type once the end-to-end circuit has been established. At that time, the message is addressed according to the pair of Circuit IDs in the message header instead of relying upon the MAC address information in the token ring header.

Acknowledgments

Randall Campbell, David Miller, Gene Cox, Ravi Periasamy, and The Ghost of Christmas Past

References

1) ISO 8802-2/IEEE Std 802.2 International Standard, *Information Processing Systems, Local Area Networks, Part 2: Logical Link Control*, December 31, 1989

2) The NETBIOS Frames Protocol, *IBM Local Area Technical Reference*, SC30-3383-03, Chapter 5, December 1990

3) ISO/IEC DIS 10038 DAM 2, *MAC Bridging, Source Routing Supplement*, December 1991

Security Considerations

Security issues are not discussed in this memo.

Authors' Addresses

Roy C. Dixon
IBM Networking Systems
Department B57, Building 060

P.O. Box 12195
Research Triangle Park, NC 27709

Phone: (919) 543-3380

EMail: rcdixon@ralvmg.vnet.ibm.com

David M. Kushi
IBM Research Division
T. J. Watson Research Center
Room H0-C07
30 Saw Mill River Road
Hawthorne, NY 10532

Phone: (914) 784-7815

EMail: kushi@watson.ibm.com

Note: Any questions or comments relative to the contents of this RFC should be sent to the following Internet address: **dlsw@ralvma.vnet.ibm.com**. This address will be used to coordinate the handling of responses.