QoS and Frame Relay

Vesa Kosonen
Helsinki University of Technology
Laboratory of Telecommunications Technology
P.O.Box 3000
02015 HUT, Finland
Email: vkosonen@tct.hut.fi

Abstract

The purpose of this paper is to study quality of service properties of frame relay. In this paper we first concentrate on explaining the protocol architecture and related matters. The frame format is studied in detail. The behavior of frame relay network is under consideration, too. How does the frame find its way to the destination and what happens when there are errors? Then we will study some properties unique to frame relay. Finally we take a look how congestion control is realized in frame relay to keep up with high quality of service.

Introduction

The rapid growth of Internet sets new requirements to the networks used for the services of Internet. Originally best-effort type of Internet faces today demands for quality of service. As new applications are coming into existance the capacity of the existing networks has to be developed, too. There needs to be less errors, bandwidth needs to grow manifold, delay and jitter should be as small as possible, they should be able to endure congestion, etc. In other words, transmission properties of networks should be of high quality as they try to serve the growing demands of the customers.

As an excample lets consider how to transmit voice over Internet Protocol (VoIP), which is one of the hottest topics of research today. Since voice is very sensitive to delay and jitter these properties of Internet has to be developed before VoIP will gain support from the big audience. The new technology needs to offer something new and better to the existing telephone system and at the same time try to maintain the good old level of quality of the voice.

Frame relay is an old technology and it was initially developed to function as the transmission media for narrowband ISDN. Since that time it has much developed and has today also many independent applications outside ISDN. It e.g. supports SONET/SDH enabling higher bandwidth capabilities in the frame relay network. No longer will users, service providers and equipment vendors have to turn to other technologies just to access greater bandwidth [1]. As another example it can be mentioned that there are many research projects going on where frame relay is involved. As we can see frame relay is still in the headlines. What makes frame relay as an interesting alternative for transport technology? How is quality of service maintained in case of congestion? These are some of the questions that will be discussed in this study.

1. A Short History of Frame Relay

In 1988 came out I.122 Recommendation entitled as "Framework for Providing Additional Packet Mode Bearer Services". It introduced a new form of packet transmission that has become one of the most significant contributions of narrowband ISDN work. The title of this recommendation was changed with the release of the 1993 version to "Framework for Frame Mode Bearer Services", which we nowadays know as *frame relay*. The former term emphasizes the service being offered to the user, while the latter emphasizes the protocol that implements the service. In Table 1 can be found the ITU-T documents related to frame relay [2].

Table 1. ITU-T Recommendations on Frame Relay

Number	Title	Date
I.122	Framework for Frame Mode Bearer Services	1993
I.233	Frame Mode Bearer Services	1992
I.370	Congestion Management for the ISDN Frame Relaying Bearer Service	1991
I.372	Frame Relay Bearer Service Network-to-Network Interface Requirements	1993
I.555	Frame Mode Bearer Service Interworking	1993
Q.922	ISDN Data Link Layer Specification for Frame Mode Bearer Services	1992
Q.933	Signaling Specifications for Frame Mode Call Control	1995

Packet switching technology was developed at a time when digital transmission facilities exhibited a relatively high error rate compared with today's facilities. As a result there is a considerable amount of overhead built into packet switching schemes (e.g. X.25) in order to compensate errors [2].

With modern, high-speed telecommunications systems, this overhead is unnecessary since the amount of errors has been dramatically lowered. Frame relay was developed to take advantage of the high data rates and low error rates [2]. Frame relay was originally designed to be used for speeds up to T1/E1 speeds (2 Mbps), but has already been demonstrated to be practical for speeds up to 50 Mbps and even higher.

Today the main usage of frame relay is to connect LAN/WAN networks. Other services include image transfer, private line replacement and Internet access [3]. Frame relay has become very popular and many applications are being developed based on it.

2. Protocol Architecture

The protocol stack of frame relay is simple, only one and a half layers. The protocol layers used are the Physical Layer and a subset of the Data Link Layer, called *LAPF core* (LAPF = Link Access Procedures to Frame Mode Bearer Services) which is defined in Q.922. LAPF is based on and is an extension of LAPD, which is used in ISDN (Figure 1).

Data Link Layer	LAPD	
		LAPF core
Physical Layer	Physical Layer	Physical Layer
OSI	ISDN	Frame Relay

Figure 1. Comparing Protocol Stack of OSI, ISDN and Frame Relay

The Physical Layer is no different from any Physical Layer with definitions of how bits are transmitted. The Data Link Layer on the other hand provides some of the same functions as defined in OSI model such as framing, addressing and bit error detection. The difference is that there is no sequencing or no acknowledgements. In case of errors the frame is simply discarded. Since frame relay has no error correction it assumes that the network infrastructure is relatively error-free (e.g. fiberglass). Multiplexing is also performed at the Data Link Layer [3].

2.1 LAPF Frame Format

The operation of frame relay for user data transfer is best explained by beginning with the frame format (Figure 2).

Flag	Address	Information	FSC	Flag
1 oct.	2-4 oct.	variable	2 oct.	1 oct.

Figure 2. LAPF Core Formats

- *Flag (1 oct.):* At the beginning and in the end there is a bit pattern 01111110 to tell the limits of the frame.
- *Address (2-4 oct.):* Contains the Virtual Circuit address and extra bits for congestion control.
- *Information (variable):* Contains higher layer protocol data. This field can contain any integral number of octets. The frame relay standards suggest that the minimum length should be 1600 octets in connection of LAN interconnection applications. Most frame relay services support a maximum length of 4096 octets.
- *FCS (2 oct.):* The frame check sequence [3]

Note the absence of a Control field in this frame. Since frame relay treats all frames the same regardless of their type and does not provide any guarantee of sequentially, there is no reason for frame relay to ever examine a Control field. If one is present, LAPF's core protocol would consider it to be part of a higher-layer protocol [3].

The address field contains information necessary for the operation of the frame relay service. This field contains addressing information (DLCI) as well as congestion and fairness indicators (C/R, EA, FECN, BECN and DE) (Figure 3). The default length of this field is two octets. Three or four octet address field may be employed, too [3].

6 oct.			1 oct.	1 oct.
DLCI (high order)			C/R	EA0
DLCI (low order)	FECN	BECN	DE	EA1
4 oct.	1 oct.	1 oct.	1 oct.	1 oct.

Figure 3. Address Field Format – 2 octets (default)

Explanations to Figure 3:

- DLCI = Data Link Connection Identifier (10 bits)
- C/R = Command/Response bit (1 bit)
- EA = Address Field Extension bit (1 bit)
- FECN = Forward Explicit Congestion Notification(1 bit)
- BECN = Backward Explicit Congestion Notification (1 bit)
- DE = Discard Eligibility (1 bit)

The Address Extension (EA) bits are used in accordance with extended address fields. If EA is set to 0 it means that more octets follow in this field. EA value 1 on the other hand indicates that this is the final octet [3].

3. The Frame Relay Network

Frame relay provides an *unreliable, connection-oriented* service to the user. It means that a virtual circuit (VC) has to be established before any data can be sent, but network doesn't give any guarantee that the data will be delivered to the destination [3]. There are two options for

VC's to choose from: Permanent Virtual Circuit (PVC) or Switched Virtual Circuit (SVC). PVC is the original VC, but SVC is gaining more popularity due to its flexibility.

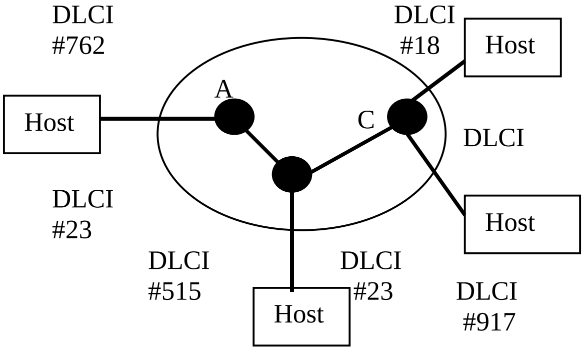

Figure 4. Frame Relay Network

The frame relay network consists of a group of interconnected nodes (switches), which relay the frame relay data across the network on the appropriate VC (Figure 4). A frame relay switch uses only the DLCI (only local value) information contained in the frame relay header to forward the frame across the network to its destination. The path through the network is transparent to the user. The DLCI does not include any description of how the connection transverses the network, or routing topology of the network [4].

A frame relay network operates at an OSI Layer 2 router network. Each frame relay access node puts the routing information (destination DLCI) in the data link layer (frame relay header) of the frame. In other words, the frame relay network nodes look only at the header and the FCS. This is important because the network will not be effected if the data is eg. encrypted [4].

The frame relay switch (node) uses a two step review process to forward frames across the network (Figure 5):

1. Integrity of the frame is checked using the Frame Checker Sequence (FCS), and if an error is indicated, the frame is discarded.
2. The destination DLCI address is validated, and if it is invalidated the frame is discarded [4].

All frames that are not discarded as a result of the FCS or DLCI checks are forwarded. The frame relay node makes no attempt to correct the frame or to request a re-transmission of the frame. This makes for an efficient network, but requires that the user end-stations assume responsibility for error recovery, message sequencing and flow control [4].

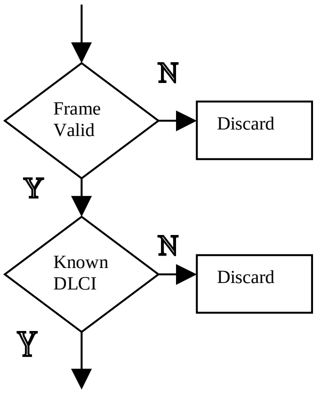

Figure 5. Nodal Processing in Frame Relay

Thus, frame relay switches do not look at the user data packets, which makes the network transparent to all protocols operating at levels above OSI level 2 [4].

4. Some unique aspects of Frame Relay

While frame relay protocol is relatively simple, it is very different from the bit-oriented protocols employed in X.25 and on the ISDN D-channel. This section explains some of the unique aspects of the frame relay protocols [3].

4.1 Committed Information Rate (CIR) and Class of Service Parameters

One of frame relay's main benefits is that it makes a pool of bandwidth available to many VCs. However, there is a danger that some applications consume all of the available bandwidth leaving nothing or only small amount of bandwidth to other applications. To prevent this to happen frame relay has a value called Committed Information Rate (CIR). The purpose of CIR is to provide fair access to the network's bandwidth by all user applications [3].

Imagine that the available bandwidth is 64 kbit/s. Application A needs much of the bandwidth most of the time and application B needs only a small amount of the bandwidth every now and then. Application A is assigned 48 kbit/s CIR and application B is assigned 16 kbit/s CIR. This arrangement provides guaranteed access for both of the applications [3].

But is this good use of resources? Application A has only 48 kbit/s available even it would need a lot more while application B's 16 kbit/s stays unused. The solution to this problem is to allow application A to exceed its CIR. In our example it would mean that application A is allowed to send traffic up to 64 kbit/s. To protect the application B, however, all the traffic that exceeds application A's CIR would be marked as *discard* *eligible* by setting the DE-bit in the frame address field (Figure 3). In this way the applications that need to send traffic only now and then is guaranteed bandwidth while the applications that need to send much traffic are allowed to do so. Bandwidth across the access line never goes unused as long as some applications have data to send (Figure 6) [3].

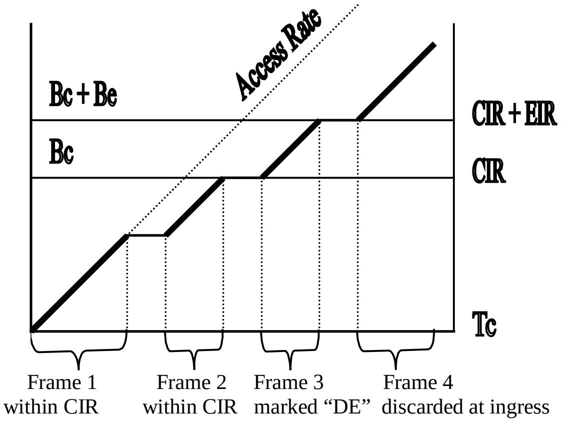

Figure 6. Relationship between CIR, DE and other Class of Service Parameters.

Explanations to Figure 6:

- Tc= the Committed Rate Measurement Interval, is the interval of time over which information transfer rates are measured, normally 1 s.
- Bc=Committed Burst Size, is the maximum number of bits the network guarantees to deliver during the time interval Tc under normal circumstances.
- CIR is defined as "the troughput rate that the network agrees to support under normal conditions" or CIR = Bc/Tc [bit/s]
- Be= Excess Burst Size, which is the maximum number of bits above the CIR that the network will attempt to deliver during Tc (DE=1).
- EIR= Excess Information Rate, EIR=Be/Tc
- All the frames that exceed Bc+Be will be discarded.
- To avoid frame discards: Access Rate = Bc + Be

Since most data transmission is bursty in nature, it actually means that most of the information is transmitted in a relatively short period of time.

5. Congestion Control in a Frame Relay Network

Frame Relay uses statistically multiplexed virtual circuits within a single physical bearer. As we have seen, each circuit may be configured to have a guaranteed throughput (CIR-value) while allowing for bursts of extra traffic (DE-marked frames). Frame Relay supports Congestion Management, a feature whereby the network attempts to notify end-points that the network is experiencing congestion and the volume of traffic should be reduced. Otherwise the Frame Relay nodes will start discarding frames to maintain internal buffer levels. If end-points do not reduce their transmission rate in this situation, "packet storms" can result as stations retransmit discarded data onto an already overloaded network [5].

Figure 7 shows the problem of congestion graphically. During periods of *no congestion*, network throughput is able to keep pace with the incoming traffic. At some level of incoming traffic (point A), the network may discard some frames due to localized congestion. Network throughput decreases gradually once this period of *mild congestion* begins (from A to B). As a result two responses may occur. First, end users who have lost frames will request retransmission, adding even more to the traffic load in the network. Second, congestion in the network results in delays at end-user equipment, causing timeouts and, again, requests for retransmission. As a result the incoming traffic will increase to point B, where *severe congestion* begins. If this point is reached, network throughput will degrade rapidly. To offset the negative effects of congestion, frame relay’s congestion control strategy is to *slow down the rate of incoming traffic* if the level reaches point A; in that way, incoming traffic level will never reach the potentially crippling point B [3].

Congestion control is normally the function of the Network Layer protototocol, but since frame relay doesn’t have one it is the function of the Data Link Layer. Congestion control can be implemented in a frame relay network in the following ways:

- use of the Explicit Congestion Notification bits
- the Consolidated Link Layer Management Protocol
- Implicit Congestion Notification [3].

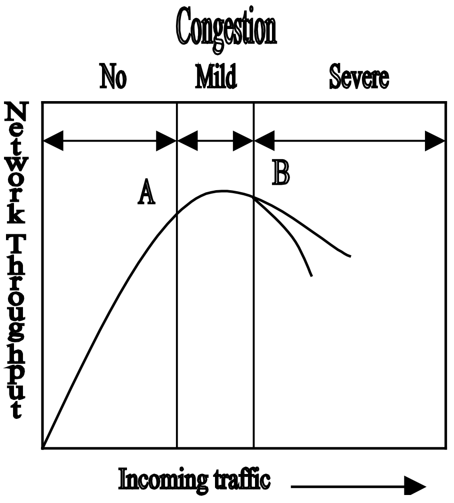

Figure 7. Throughput and Network Congestion

5.1 Explicit Congestion Notification (ECN)

In frame relay there are two congestion bits in use (Figure 3 and Figure 8):

- FECN - Forward Explicit Congestion Notification
- BECN - Backward Explicit Congestion Notification

The FECN bit is set to notify the receiving end system that the marked frame has encountered congestion. In response to this, the receiving system should try to reduce the flow of data from the sending system on this frame relay connection. The mechanism for doing so must be above the level of the frame relay bearer service, which provides no direct flow-control facilities [2].

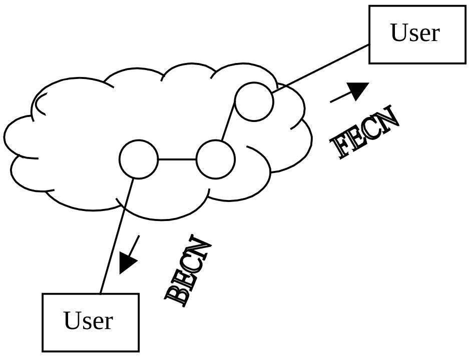

Figure 8. Use of FECN and BECN bits.

In general terms, the receiving end system should use the following strategy for each connection:

1. Compute the fraction of frames for which the FECN bit is set over some measurement interval.
2. If more frames have the FECN bit set than have an FECN bit of zero, then reduce the flow of frames from the sending system.
3. If the congestion condition persists, institute additional reductions.
4. When the congestion condition ends, gradually increase the flow of frames [2].

The BECN bit, on the other hand, is set to notify the receiving end system that the frames it transmits on this connection may encounter congestion. In response to this, the receiving end system should reduce the flow of data transmitted on that connection.

In general terms, the receiving end system should use the following strategy for each connection:

1. When the first frame with the BECN bit is received, reduce the information rate to CIR.
2. If additional consecutive frames with the BECN bit set are received, then institute additional reductions.
3. If a consecutive sequence of frames with the BECN bit set to zero are received, then gradually increase the flow of frames [2].

In general terms, for explicit congestion avoidance, the network alerts end systems to growing congestion within the network and the end systems take steps to reduce the offered load to the network [6].

5.2 The Consolidated Link Layer Management Protocol (CLLM)

There is a potential deficiency when using the ECN bits: How does the source become informed about congestion, if there are no frames being sent back to it by the destination? For this kind of situation there exists CLLM message. It is generated by the network and sent to the source host to inform to reduce the number of frames that it is sending. CLLM message is sent in LAPF frames on DLCI 1007 and indicates the list of all affected VCs as well as the cause and expected duration of the congestion [3].

5.3 Implicit Congestion Notification (ICN)

Implicit signalling occurs when the network discards a frame, and this fact is detected by the end user at a higher, end-to-end layer, such as Q.922 control protocol. When this occurs, the end-to-end software may deduce that congestion exists. Once congestion is detected, the protocol uses flow control to recover from the congestion [2].

6. Conclusions

Frame relay provides an unreliable, connection-oriented service to the user. It uses Data Link Connection Identifiers in travelling through the network. Frame relay was invented to utilize the new technology in the field of transmission facilities. Frame relay’s advantage is light protocol structure which makes fast connections possible. E.g. congestion control of frame relay operates exceptionally at the OSI layer two.

Mechanisms of maintaining the quality of service are tested especially in the situations of heavy traffic. The important question is how does the protocol resolve the congestion conditions. In case of frame relay network the method is to reduce the incoming traffic by notifying the end systems about the congestion. The notifying bits are built in the protocol structure of frame relay.

Frame relay has become an interest of many research projects and will be one of the most popular transport technologies used in the Internet in the future, too.

References

[1] http://www.frforum.com/6000/FRF_14PR.html
[2] Stallings, William: ISDN and broadband ISDN with Frame Relay and ATM, Fourth edition, Prentice-Hall, Inc. 1999.
[3] Kessler, Gary and Southwick Peter: ISDN: concepts, facilities and services, Signature (4th) edition, McGraw-Hill series, 1998.
[4] http://135.145.192.160:80/cgi-bin/auth.pl?file=framerelay_security_guide.htm&session=35f6c3c61b015f2f#2-1
[5] http://www.datacraft.com.au/whitepapers/f_relay.htm
[6] Stallings, William: High-Speed Networks, TCP/IP and ATM Design Principles, Prentice-Hall, Inc. 1998.