White Paper

July 1999

www.wwgsolutions.com

Frame Relay Flow Control and Data Transmission Part 1

At its core, Frame Relay is a simple Layer 2 protocol. A single physical interface allows connection to multiple remote locations through a series of Frame Relay switches. Each end-to-end connection is a called a Virtual Circuit (VC). VCs are available in two varieties, permanent or switched.

The Permanent Virtual Circuit, or PVC, is set up by the service provider and is always connected between endpoints. The Switched Virtual Circuit, or SVC, is set up by the user when needed, much like a dial modem connection. Once the SVC is connected, the operation is similar to that of the PVC with regard to throughput and congestion. Once data transmission is complete, the SVC is disconnected.

For identification, each VC is assigned a Data Link Connection Identifier (DLCI). The DLCI is a number, generally between 17 and 1022, contained in the Frame Relay header. The router will encapsulate data in Frame Relay with the appropriate DLCI and transmit the frame onto the Frame Relay network.

For users more familiar with LAN networks, there are analogies between Frame Relay and switched Ethernet networks. Frame Relay operates under the same general concept as switched Ethernet. The Frame Relay network directs packets between ingress and egress ports based upon the DLCI, much as an Ethernet switch directs packets between switch ports based upon the MAC address. Mixing of speeds between ports is supported, much like some Ethernet switches support both 10 and 100 Mbit/s to allow high-speed backbone ports to connect to lower-speed client ports.

Frame Relay Terms and Acronyms

When discussing Frame Relay, terms such as CIR, B_c and B_e are encountered. In practice, service providers will configure Frame Relay circuit parameters differently depending upon tariffs, corporate policy and customer-specific issues. The following information is intended to explain both how the specifications intend Frame Relay to work as well as what the typical user can expect.

Following is a description of acronyms and terms common to Frame Relay.

- **CIR** (Committed Information Rate) is the average throughput rate, in bits per second, that the user can expect from a PVC. In theory, the user should be able to transmit data continually without problems over a Frame Relay PVC at this *"average bits per second"* rate. The CIR is uniquely configurable for each PVC in the provider's Frame Relay switches as well as the CPE equipment.

- $\mathbf{B_c}$ (Committed Burst) is the total number of bits (*not bits per second*) the user is allowed to transmit onto the Frame Relay circuit in a set time period (T_c). B_c is uniquely configurable for each PVC in the provider's Frame Relay switches as well as the CPE equipment.

- $\mathbf{B_e}$ (Excess Burst) is an amount of data in bits (*not bits per second*) above B_c that if transmitted by the user within T_c, the network will *attempt* to deliver. B_e is uniquely configurable for each PVC in switches and CPE equipment. Both ANSI and the ITU consider anything above B_e to be data that *will* be discarded.

Wavetek Wandel Goltermann
Communications Test Solutions

- **T_c** is the time period calculated by dividing B_c by CIR. It is used to determine the time period from which data will be measured, in total bits, to determine if the user is within his agreement. T_c is generally not a directly configurable parameter, but is derived from the values entered for B_c and CIR.

- **VC** (Virtual Circuit) is a data path between two points in the Frame Relay network. A single interface may support multiple PVCs with each connecting two points in the Frame Relay network. Generally, these will be permanent virtual circuits or PVC. In the event the user is using switched virtual circuits, the acronym SVC is used.

- **DLCI** (Data Link Connection Identifier) is the number assigned to identify a VC. By placing this number within the Frame Relay header it identifies both the source and destination of a user data frame.

- **Access Rate** is the actual interface speed to which the user's equipment connects. While any rate is technically possible, most users will find rates based on standard digital interface speeds. In North America, these will be 56 or 64 KBPS, T1/FT1 or T3 rates. In other parts of the world, these rates will be 64 KBPS, E1/FE1 or E3 rates.

Transporting Frames Through the Network

As frames arrive at the Frame Relay switch, the DLCI will be examined to determine where to send the frame next. Each switch will forward the frame until it reaches the final destination, usually a router or FRAD. No acknowledgment of received frames is provided between Frame Relay devices; once transmitted, a frame is forgotten at the Frame Relay transport layer. This means that if a frame is lost or damaged in the network, it is up to the host protocol stack to identify and retransmit user data that has been lost. In Figure 1, identification and retransmission of lost data is the responsibility of the client or server protocol stack.

In Figure 1, there are four PVCs, one for each connection between the server and each of the remote hosts. The link connecting router 1 to the Frame Relay network is operating at 1,536 KBPS, while the sum of all of the remote link speeds is 2368 KBPS. The important fact is not the link speeds, but the sum of the CIRs. Most service providers allow the sum of the CIRs to equal more than the link speed supporting them. This is based on the idea that the user data will burst at intervals and not be constant, so that when data needs to be sent the bandwidth will be available.

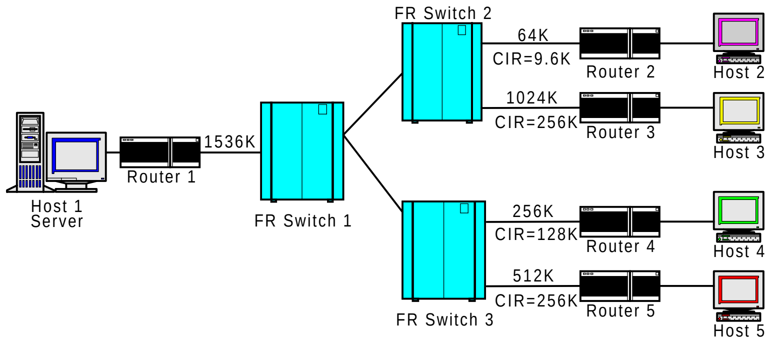

Figure 1

Understanding how a Frame Relay switch is configured gives some insight into the function of DLCI numbers. When configuring a Frame Relay switch, a Port/PVC mapping structure is used. The service provider selects two physical ports on a Frame Relay switch and assigns a DLCI number to each port. The DLCI number may be the same or different, but the number must be unique for the port. Once the ports are selected, parameters are set for CIR, B_c and B_e and then the switch calculates T_c. The result of the PVC mapping can be seen in Figure 2.

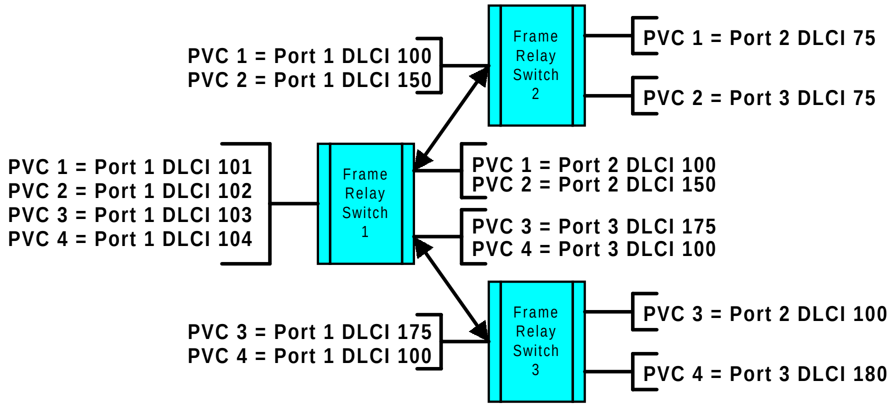

Figure 2

Frame Relay switch ports are interconnected in different ways, but more and more the switches are interconnected using ATM. Regardless of how the switches are connected, the user will see only Frame Relay at his location. Each switch in the network will have a PVC-in-PVC-out mapping. The end switches will have WAN circuits connecting the port circuit that terminates at the user site. A router or a FRAD will generally terminate this circuit at the user location.

When mapping the PVC between ports and switches, each "leg" of the VC may have a unique DLCI. In Figure 2, router 1 recognizes DLCI address 101 for all packets to and from router 2. From the perspective of router 2, DLCI address 75 is used for all packets to and from router 1. Each router has its own idea of what the DLCI number represents to a given VC. In this respect, the DLCI address has local significance only.

As a frame is followed from its source to its destination, it goes through some changes. In Figures 1 and 2, we can follow a packet from host 1 to host 2 over PVC 1.

1. A packet is generated by host 1 with a destination IP address of host 2
2. Router 1 receives the packet and examines its routing table to determine the proper DLCI
3. Router 1 encapsulates the packet in Frame Relay with the proper DLCI of 101
4. Router 1 transmits the packet to switch 1
5. Switch 1, using its mapping table, replaces the DLCI with 100 and sends the packet to switch 2.
6. Switch 2, using its mapping table, replaces the DLCI with 75 and sends the packet to router 2.
7. Router 2 strips the Frame Relay header and passes the packet to host 2.
8. The return path is the reverse, starting at router 2 with a DLCI of 75 and arriving at router 1 with a DLCI of 101.

Each of the throughput parameters (CIR, B_c, B_e) has an effect on the others. For example, if the CIR = 1000 BPS and B_c = 2000 bits and B_e = 0 bits, then T_c = 2000/1000 or 2 seconds. Given these numbers, imagine a router sends 2000 bits of data in the first .1 second. In order to stay within the defined service parameters, the router must not send any more data until T_c expires, in this case 1.9 seconds. If more data is sent within the 1.9 seconds, it will be in excess of the subscribed service and *may* be discarded. In this case, the subscribed service states the user has the right to transmit 2000 bits onto the Frame Relay network in any two-second period. If the user exceeds 2000 bits in two seconds, the excess data *may* be discarded by the network.

One way of increasing the throughput would be to increase the B_e. If the B_e were to be increased to 1000 bits, the user would have the right to send 3000 bits within the same two-second period. The 1000 bits above the B_c will be delivered by the network on a best-effort basis. This means that if there is bandwidth available and no congestion, the network will deliver the data, but there is no guarantee. This allows the user to have a good probability of delivering a burst of 1000 bits above the 2000 bits guaranteed. Data sent above B_c and B_e will *generally* be discarded.

Some service providers mandate that the CIR equal the B_c, resulting in a T_c of one second. The B_e may be configured for any rate between 0 and the access rate – B_c. Switches will treat data in excess of the subscribed CIR in one of two ways. The first is to discard anything in excess of the B_e. If the B_e is configured for zero, then every frame above B_c may be discarded, depending upon other aspects of the switch configuration. Some service providers use this method to “force” the user to stay within the service level agreement. This is the least favorable method for the user, but it allows the service provider to more accurately allocate bandwidth and control traffic rates over the Frame Relay network to guarantee some level of service to all users.

The second method of treating excess data is to always attempt a best effort at delivery and resort to dropping frames only when the switch begins to suffer from congestion related to excessive traffic. In this mode of operation, short bursts above B_c will be delivered if possible. However, if the burst above B_c is large enough to cause congestion, the switch will discard the packets. This is a common method of switch configuration.

- *Expect anything above B_e to be discarded by the Frame Relay network. However, predicting how any given switch will react to excess traffic levels is difficult due to individual switch design and option settings. Always discuss with your service provider how this excess data will be handled.*

Because the user PVC can pass through a number of Frame Relay switches, it is important that the CIR, B_c and B_e configuration be the same for your PVC in every switch. If there are differences, the entire PVC will function at the configuration with the lowest throughput settings. In theory, it is possible to configure different values for each direction of a PVC, but this is almost never done.

Some carriers offer a 0 CIR service. In this method, every frame is DE marked and the provider will attempt to make a best-effort delivery, but with no guarantee. Some less critical applications can run acceptably over low-cost 0 CIR circuits. Real-time applications may experience a somewhat erratic performance, with 0 CIR causing them to perform unacceptably at times.

All of the Frame Relay configuration options related to throughput in a router or FRAD apply only to the data being transmitted by the device. Frames received from the Frame Relay network by the router have already made it through the network and are no longer subject to the Frame Relay network restrictions of CIR and burst.

The fact that 0 CIR is offered by some service providers provides insight about CIR. The switch uses the value of CIR and B_c to determine T_c. It is the values B_c, T_c and B_e that are of real importance to the operation of a Frame Relay PVC. In fact, the main use of CIR, and the reason I use it in this document, is that it is easy to understand. The true subscribed service could more accurately be defined as *"B_c bits within T_c seconds not exceeding a limit of B_e bits of burst above B_c within T_c".*

- *CIR is the common term used in Frame Relay discussions, but B_c , T_c , and B_e are the settings that actually determine how the network will operate and what the throughput will be.*

The following shows the valid relationships between the VC configuration parameters.

If CIR > 0 then the following applies:
- $B_c \geq$ CIR
- $0 \leq B_e \leq$ (access rate - B_c)
- $T_c = B_c$ / CIR

If CIR = 0 then the following applies:
- $B_c = 0$ (prevents divide by 0 error)
- $0 < B_e \leq$ access rate in BPS (If $B_e = 0$ then all frames would be dropped)
- $T_c = B_e$ / access rate in BPS

Frame Relay Parameters

Figure 3 is an example of how frames transmitted onto the Frame Relay network are treated. The graph displays a 64 KBPS interface with specific settings for CIR, B_c, and B_e. One important point to note is the function of the value for CIR. With the settings in the graph, 16 Kbits can be transmitted every two seconds (T_c). If the CIR were raised to 16 KBPS, the time interval (T_c) would be reduced to one second. This would result in doubling the throughput. This change in the T_c will have an effect on the B_e as well by allowing a 16 Kbit burst above B_c every one second instead of every two seconds.

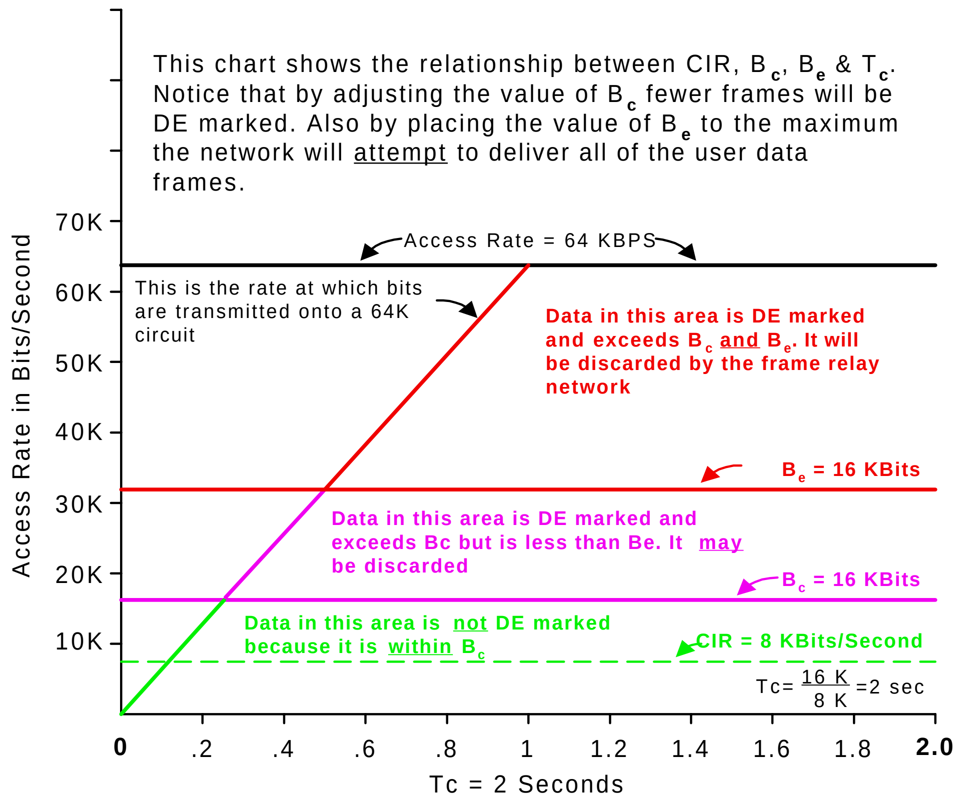

Figure 3

Router-Network Interaction

From the point of view of a server, the connection type, in this case Frame Relay, to the remote client is unknown. The server has no way of knowing the access rate or CIR for its own router much less the access rate of the remote routers. Because it has no knowledge of the access rate linking it to the destination, the server's protocol stack is incapable of adjusting flow based upon access rate or CIR.

Another issue is the fact that router 1 in Figure 1 has no knowledge of the access rate of the remote routers. When frames arrive from the server destined for any remote router, the frames will be placed on the Frame Relay network using the total available bandwidth. In the case above, router 1 will place packets on the Frame Relay interface at 1.536 Mbit/s. While it is true that the router is configured for the CIR of each PVC on the Frame Relay circuit, generally it does not use the CIR to adjust the transmit rate onto the network. Data that needs to be routed onto the Frame Relay network is placed there as fast as possible. If the CIR is exceeded, so be it. The idea is, *"send them all and let the network figure out what to pass and what to drop."*

As can be seen in Figure 1, if the server is sending a file to remote host 2, there are going to be problems related to the Frame Relay buffer size. As frames enter the network from router 1, at 1.536 Mbit/s, they can only exit switch 2 at 64 KBPS, the access rate of router 2. Mandating that the Frame Relay switches buffer all frames transmitted by router 1 results in problems.

The primary problem with buffering all of the data has to do with timing. In our example above, the data enters PVC 1 of the network 24 times faster that it can leave. This means that buffering one second of input means buffering 24 seconds of output data. Protocol stacks will time out and retransmit unacknowledged data before 24 seconds have elapsed. The result would be the original packet as well as the retransmitted packets being received by the end station eventually. This would result in more problems than simply dropping packets when the network is congested.

Many routers offer some form of enforcing or adjusting traffic levels for each PVC. This may be done by limiting the amount of data placed on a PVC based upon the throughput parameters configured for that PVC. While this can be useful in reducing congestion, it reduces the Frame Relay ability to provide burst capabilities when bandwidth is available while reserving bandwidth for a PVC that currently has no traffic to transmit.

Another feature of some routers is the ability to set transmission priority based upon user-defined criteria. This can be a useful way of ensuring delay-sensitive protocols receive transmission-queuing priority over less sensitive data. Priority selection may be in the form of protocol priority or IP destination address priority. The best systems allow a variety of selections, such as TCP/IP port number or encapsulated protocol type. This becomes important when real-time data, such as VoIP or VoFR, is mixed with less delay-sensitive data on the same physical interface.

Congestion in Frame Relay

While Frame Relay does not employ a *mandated* method of flow control, it does contain mechanisms intended to notify routers of congestion conditions. The router *may* reduce transmission rate and thus alleviate the congestion condition by using these notification bits.

- **FECN** (Forward Explicit Congestion Notification) is a bit set in a frame to indicate to the receiving router that this frame passed through one or more congested nodes on the way to its destination.

- **BECN** (Backward Explicit Congestion Notification) is a bit set in a frame to indicate to the router receiving the BECN that the frames it transmits are going to encounter congestion before they get to their destination.

- **DE** (Discard Eligibility) is a bit that is set by routers to designate some traffic as more appropriate for discarding than other traffic. The idea is that when a Frame Relay switch encounters congestion and must begin to drop frames, the first to go should be those with the DE bit set. Setting the DE bit is generally assumed to be a router function, although the switch may also set the DE bit. When no specific parameters are configured, it can be assumed that any frame exceeding the Bc will have the DE bit set.

When a Frame Relay switch begins to experience congestion, it notifies the offending routers, in this case router 1. Notification is achieved by sending a BECN. The BECN rides in a frame generated from the receiving station. This means that the receiving station, in this case host 2, must have a packet to send to the server. No packet, no place for the switch to put the BECN. Note that it is not the router that sends the BECN, but the switch that sets the BECN bit in a packet generated from host 2 and destined for the server. Only Frame Relay switches set the BECN; routers and FRADs cannot.

The other thing the switch does in response to congestion is send a FECN. The FECN bit is set by the switch in a packet from the server to host 2. The purpose is to inform the router for host 2 that there may be delays in receiving information from host 1. The idea is that the Frame Relay device receiving the FECN may have a method to inform the far end-user device that it should reduce the transmission rate. This could be accomplished by manipulating the acknowledgments of received data it sends back to the server. In practice, routers generally ignore the BECN and FECN notifications.

If the device notified with a BECN continues to transmit at unacceptably high rates, congestion will worsen. Frame Relay switches address congestion problems in a very straightforward way – they discard frames.

- *When congested, a Frame Relay switch will discard user data. This is done without any notification to any node or station on the network.*

In order to implement some control into packet loss, DE is used. Any Frame Relay device in the network can set the DE bit. In Figure 1 this includes all of the routers and Frame Relay switches. The DE bit will be set any time the amount of data transmitted onto the PVC within T_c exceeds B_c. When the Frame Relay switch must drop frames due to congestion, it *should* drop DE-marked frames first. This provides a way for the router to set a user-configurable discard priority to different types of traffic. In practice, many routers are not configured to distinguish different types of data and select particular frames for DE; it is often a pure rate-based function.

In order to set the DE bit appropriately, routers and switches must be configured correctly. When the Frame Relay parameters are being configured for a router or switch, the CIR, B_c and B_e must be defined by the user. It is the B_c setting in conjunction with the T_c (B_c /CIR) that the router will use to determine if a packet should have the DE bit set in the frame. Many routers have options to place priority on some protocols over others. This priority can cause some packets to be placed in the front of the transmit queue so that the lower priority traffic will have the DE bit set.

If a user is configuring a router and does not know the CIR, B_c and B_e of the PVC, always guess high, even up to the access rate. This will prevent any frames from being DE marked by the router, but will not prevent the network from discarding frames in the event of congestion.

- *BECN and FECN bits are set in user data packets by switches only. Any Frame Relay device such as routers and switches can set the DE bit. Once set the DE must remain set until the packet reaches it destination.*

Routers and FRADs that support Frame Relay have the ability to configure the CIR and therefore when the DE is set on transmitted frames. Most also have the ability to control the amount of data they place on the Frame Relay network based upon their configured CIR. Other configuration options reduce the amount of data placed on the Frame Relay PVC when a BECN is received. These configuration options are often not implemented and the problems of congestion and flow control are left up to the higher layer protocols.

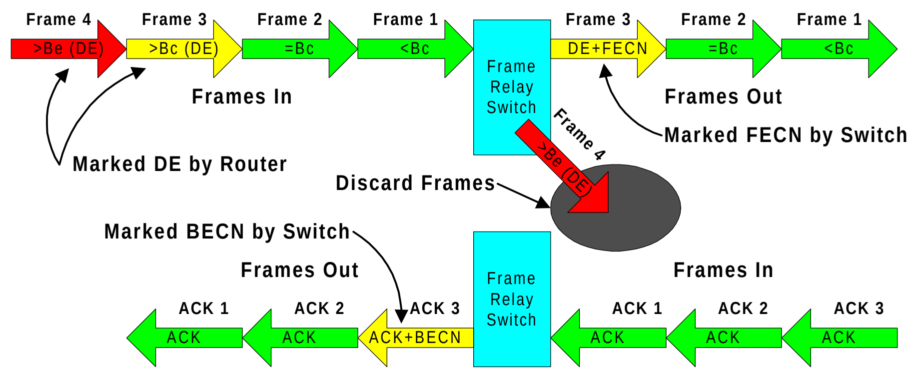

Figure 4: A Frame Relay switch experiencing congestion issuing a FECN and discarding frame. Response to the congestion is a BECN frame set in a user-data frame. Any frame can carry the BECN as long as it is on the correct PVC.

As data comes into the Frame Relay switch, it is placed in a buffer. Frames that are being transmitted to the next link are removed from the buffer. When the buffer reaches the maximum allotted level, frames will be dropped. Frame Relay assumes that upper layer protocols, like TCP, will figure out the frame was lost and do something about it. Neither retransmission nor notification of lost frames is the job of Frame Relay.

- *Acting on BECN, FECN and DE bits is optional and not required by the Frame Relay specifications. Most routers and FRADs can be configured to ignore or respond to FECN and BECN.*

Frame Size and CIR

One aspect of Frame Relay that is almost never considered is the size of user frames to be transmitted over the Frame Relay network and the values chosen for CIR and burst. The Frame Relay Forum suggests a maximum frame size of 1600 bytes. This value was chosen because it is the Ethernet Maximum Transmission Unit (MTU) with the Frame Relay header added. If you multiple 1600 by 8, the result is 12,800 bits.

Simply put, if your B_c is less than 12,800 bits, any 1600 byte frame may be marked DE. In the event your service provider is going to employ rigid enforcement, you could be faced with a network that will only pass one 1600 byte per second. If you are going to use a low CIR, be sure your excess burst is high enough to provide acceptable performance for your MTU.

In reality, very few frames achieve the maximum size. This is especially true with SNA/SDLC, which is generally less than 256 bytes and most often less than 128 bytes. Token Ring MTUs, on the other hand, can be set to 4096 bytes. Frame Relay can handle a maximum of 4096 byte frames, but anything larger renders the checksum mathematically unreliable.

Monitoring Utilization

The question of how the Frame Relay network is truly running can be difficult to answer. Any product that claims to monitor utilization related to subscribed parameters must be quite powerful to deliver accurate results.

Many routers and FRADs include some management tools to monitor utilization on the Frame Relay network. Most of these operate by querying the Management Information Base (MIB) of the device. The results returned are then placed in tables or graphs for the user to examine.

Frame Relay congestion issues are a problem that can happen in one-second intervals. Requesting one-second updates for each PVC of the router or FRAD can produce problematic levels of traffic. If the device is a remote unit, accessed over the Frame Relay network, the network may become congested with the updates. This can render the resulting statistics of questionable accuracy and value.

Routers and FRADs attempt to get around this problem by increasing sample times and averaging the results to develop the one-second statistics. Generating statistics by measuring 60-second samples and calculating the average per-second values can provide a generalized view of the network but lack the level of detail necessary to make important decisions about network performance.

For truly accurate information, a dedicated performance monitoring tool such as Wavetel Wandel Goltermann's FR-CIR must be employed. With FR-CIR, a user can monitor the network, comparing the actual traffic levels with the subscribed service parameters for CIR, B_c and B_e relative to T_c. This allows a user to quickly identify and focus on throughput and identify areas of over or under subscription.

FR-CIR also monitors FECN, BECN and DE-marked frames for each PVC. The user can compare these congestion indicators with transmitted data. This allows the user to determine if the traffic is the cause of the congestion or is they are a victim of a switch that is over subscribed.