

# Explicit Rate Flow Control A 100 Fold Improvement over TCP

Dr. Lawrence G. Roberts

## The History of Flow Control in the Internet

The ARPANET: The first packet switched data network, the ARPANET or initial Internet, was built in 1969. Dr. Lawrence G. Roberts was the architect and manager. The flow control was a very simple window permitting the host computer to send a few packets before getting permission to send more. This added at least 10% to the transmission requirements and was quite slow in responding to changing network congestion, but with 50-kilobit lines, packet delays far exceeded any other delays and the window control time was not an issue. The network also used datagrams (packets with included address headers), even though that halved the transmission efficiency, because there was too little memory in the first mini-computers for the storage of virtual circuit information. As can be seen in figure 1, this simple structure was sufficient for about 15 years.

TCP/IP Era: As the Internet grew, the flow control was moved out of the network, into the host computers or workstations so that it could span many different base networks (ARPANET, packet radio, satellite net) which might have different protocols. This was based on the belief that the switches and networks would stay around, unable to adapt to changing requirements and that a separate end-to-end protocol (TCP/IP) would be more flexible to adjust. The reverse turned out to be true, the switches get replaced every 3-5 years due to the high rate of improvement in semiconductors, and TCP/IP has stayed virtually frozen for 15 years. Thus, both TCP and IPv4 now lag the needs of the Internet considerably. An upgrade to IP, IPv6, has been agreed upon, but still lags seriously in deployment. No upgrade to TCP has even been proposed, even though it is now the most serious problem in the Internet. A new flow control protocol, like ATM's explicit rate, could reduce network costs 2:1 and decrease network delays by 100:1.

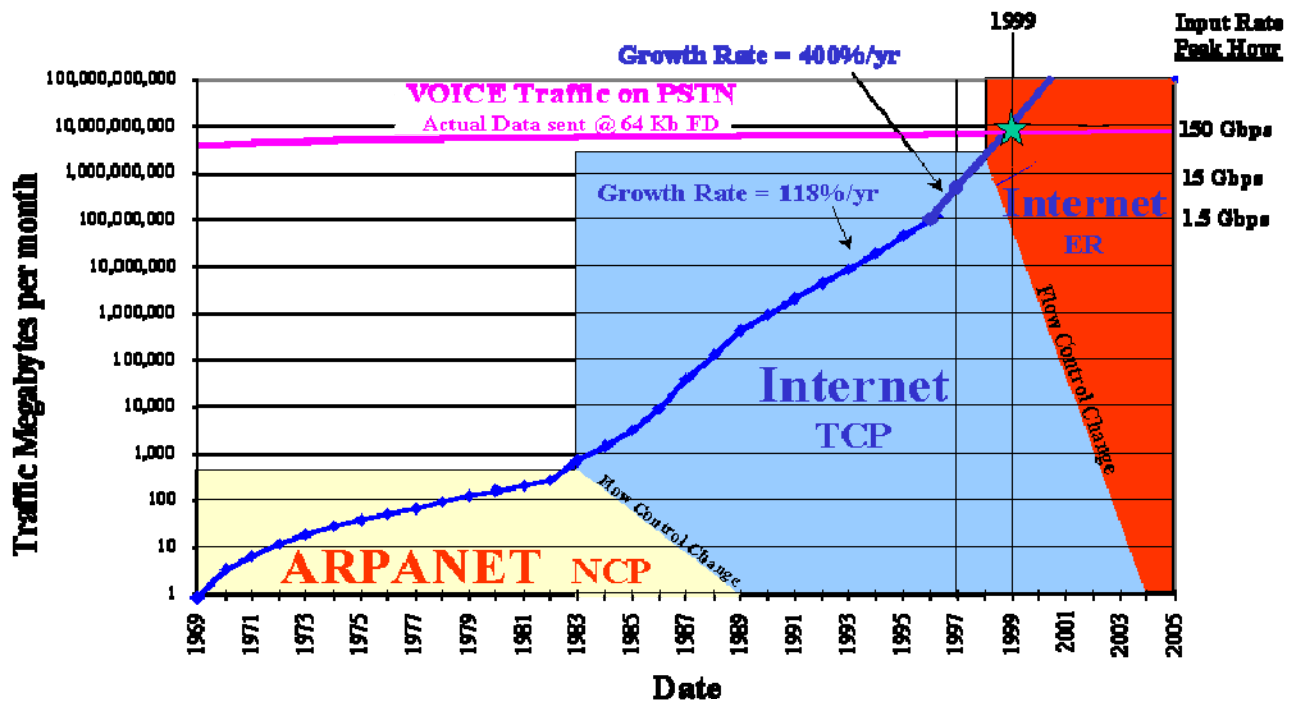


Figure 1. Internet Traffic Growth Rate, Flow Control Eras, & Telephone Traffic Growth

The Late 90's Period: Traffic growth in the Internet has doubled annually on the average since it started in 1969. This has led to an increase in traffic of one billion to one. Recently, in 1996, the rate

# Explicit Rate Flow Control

## A 100 Fold Improvement over TCP

Dr. Lawrence G. Roberts

of growth of traffic has increased to 5:1 each year, reflecting the enormous acceptance the World Wide Web application has had. Within two years the traffic will become larger than voice traffic, a critical event because this means that the pricing for transmission will have to start changing from voice based to data based. Also, since the traffic growth rate considerably exceeds that of semiconductor improvement, we cannot depend any more on scaling up the old technology, we must start to make major technological innovations in the switching technique in order to keep up with the traffic demands. There are two areas where this is necessary: switching speed and flow control speed:

Virtual Circuit (VC) Switching: from the early packet switches in the ARPANET to today's gigabit routers, the switching concept has stayed the same in the Internet, routing datagrams. To route a packet from a full worldwide address requires, after the first packet of a connection, a complex hash to turn a 32-bit address into a cached route. IPv6 with its longer 128-bit address will in fact make this even worse. There has been a raging battle for 25 years (starting with X.25 vs. IP) about whether to route using full addresses or switch using Virtual Circuit (VC) numbers. There are many pros and cons to each but now that the switch speed is being stressed extremely hard, it is becoming clear that VC switching is required in order to achieve the required switching speed. VC switching, as done in ATM, requires only a single memory lookup, not a hash, to switch. Even the router vendors are now planning to convert to ATM style VC switching for routers under the name of Tag Switching or now, Multi-protocol Label Switching (MPLS). Thus, the 25-year battle will be settled in favor of VC switching within a few years. This performance improvement through VC switching is easy to obtain since ATM is fully developed and can provide at least a factor of 10 increase in throughput over router technology. Flow Control: The largest innovation required in the Internet or large WAN networks is to improve the control time of the flow control. TCP's life has already been overextended and a major improvement is required in flow control in order to reduce network delay, improve Web access time, and reduce network cost. Explicit rate flow control does this and has already been specified for ATM. The comparison of explicit rate with TCP is set out below. Incorporating explicit rate into all networks, end-to-end is a critical next step resulting in a hundred-fold improvement.

### **ABR (Available Bit Rate) traffic service**

The newest functional addition to ATM is ABR traffic service, specified by the ATM Forum in May 1996 and soon to be released in ATM products. ABR joins several other traffic types in ATM including CBR, UBR, and VBR. To understand the new function ABR brings to ATM, an understanding of the previous options is necessary. The following definitions are provided by the ATM Forum Glossary:

- CBR (Constant Bit Rate) is an ATM service category which supports a constant or guaranteed rate to transport services such as video or voice as well as circuit emulation which requires rigorous timing control and performance parameters. Constant bit rate also means fixed bandwidth, and when voice and video traffic is not active, overall bandwidth is wasted.
- VBR (Variable Bit Rate) is an ATM Forum defined service category which supports variable bit rate data traffic with average and peak traffic parameters. VBR provides some relief from the strong demands made by CBR, but still typically wastes a factor of two in bandwidth in order to insure that there is low cell loss.
- UBR (Unspecified Bit Rate) is an ATM service category which does not specify traffic related service guarantees. Specifically, UBR makes no commitments about the cell loss or the bandwidth, other than the peak bandwidth. However, the nature of Internet TCP/IP traffic is such that UBR is next to useless for such data traffic in wide area applications due to excessive data losses.

# Explicit Rate Flow Control

## A 100 Fold Improvement over TCP

Dr. Lawrence G. Roberts

- ABR (Available Bit Rate) is an ATM layer service category for which the limiting ATM layer transfer characteristics provided by the network may change subsequent to connection establishment. A flow control mechanism is specified which supports several types of feedback to control the source rate in response to changing ATM layer transfer characteristics. It is expected that an end-system that adapts its traffic in accordance with the feedback will experience a low cell loss ratio and obtain a fair share of the available bandwidth according to a network specific allocation policy. ABR supports two types of flow control, a binary rate technique called EFCI, for backward compatibility, and Explicit Rate (ER), a major improvement in flow control.

### Explicit Rate Flow Control

The most important addition that is supported in ATM is Explicit Rate Flow Control. Flow control is used to control the traffic sources so that they do not send too much data into the network at any moment. If a trunk in the network is overloading, all the sources using that link must be told to slow down. This is absolutely necessary for a data network because the self-similar characteristics of data traffic are such that simply under-loading the network will not work. The critical difference between flow control techniques is the time it takes to tell the source about congestion and to get it under control. Today, TCP is typically used in LANs to control the data flow. TCP was created 15 years ago when networks were much slower. It operates a binary rate flow control algorithm between the two end-stations, slowing down the data flow if the return path indicates that data was lost. If data is not being lost, it speeds up the flow. It always oscillates with a period of 1-2 seconds on the Internet, losing data each cycle and under-utilizing the network on the other part of the cycle. Its characteristic time is about one second, that is the time it takes TCP to stop the data flow due to increased congestion.

### Simulation Comparison of Explicit Rate and Binary Rate Flow Control

#### Basic Response to Network Rate Change

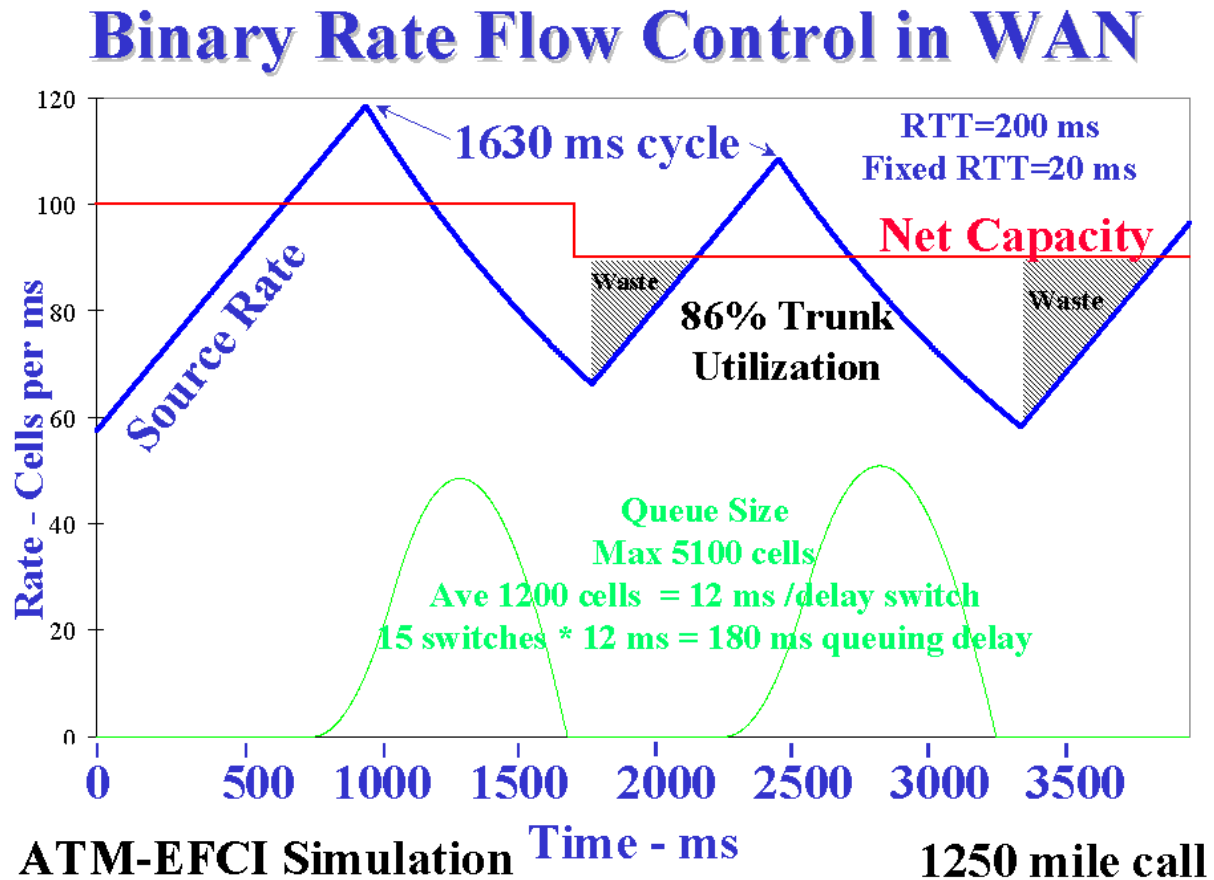
The first simulation examples will be to explore the effect of a minor (10%) decrease in the available network bandwidth, but with an otherwise stable traffic capacity. The source is persistent and has already adjusted its operating rate to the network capacity. The network is modeled after the typical Internet call, 15 switches in the forward path and a path length of 1250 miles. This distance results in a fixed round trip time of 20 milliseconds assuming no queuing delays.

Binary Rate Protocol This simulation example is for ATM using EFCI binary rate protocol. The source rate oscillates around the network capacity. If no switch in the network is congested, the rate is allowed to increase. If one of the switches is congested, it marks the EFCI bit in the data cells as they go by and the destination converts this EFCI mark into a CI bit mark in the RM cells that return to the source. The source must then decrease its rate until it receives unmarked RM cells. Since the data cells transverse the network slowly, waiting at each switch behind all the queued data, the Round Trip Time (RTT) for each basic step is 200 milliseconds, 10 times slower than the speed of light. However, since all that is conveyed is a simple up/down signal, it then takes about eight RTTs for one oscillation.

The data queue in each switch builds up to 5100 cells with an average queue depth of 1200 cells. This means that data crossing 15 switches, the typical number of switches in an Internet path, will experience 180 ms of queuing delay (12 ms per switch times 15 switches) plus the 20 ms for the speed of light around the network. Thus the total RTT is 200 ms which is the measured Internet RTT during peak hour. In this case the protocol is EFCI, not TCP like the Internet, but both of these protocols are binary rate with similar characteristics and control times. The control-time is the time it

**Explicit Rate Flow Control**  
**A 100 Fold Improvement over TCP**  
 Dr. Lawrence G. Roberts

takes for the EFCI source to adjust to the new rate after the network capacity decrease. The binary rate control-time is around one second.



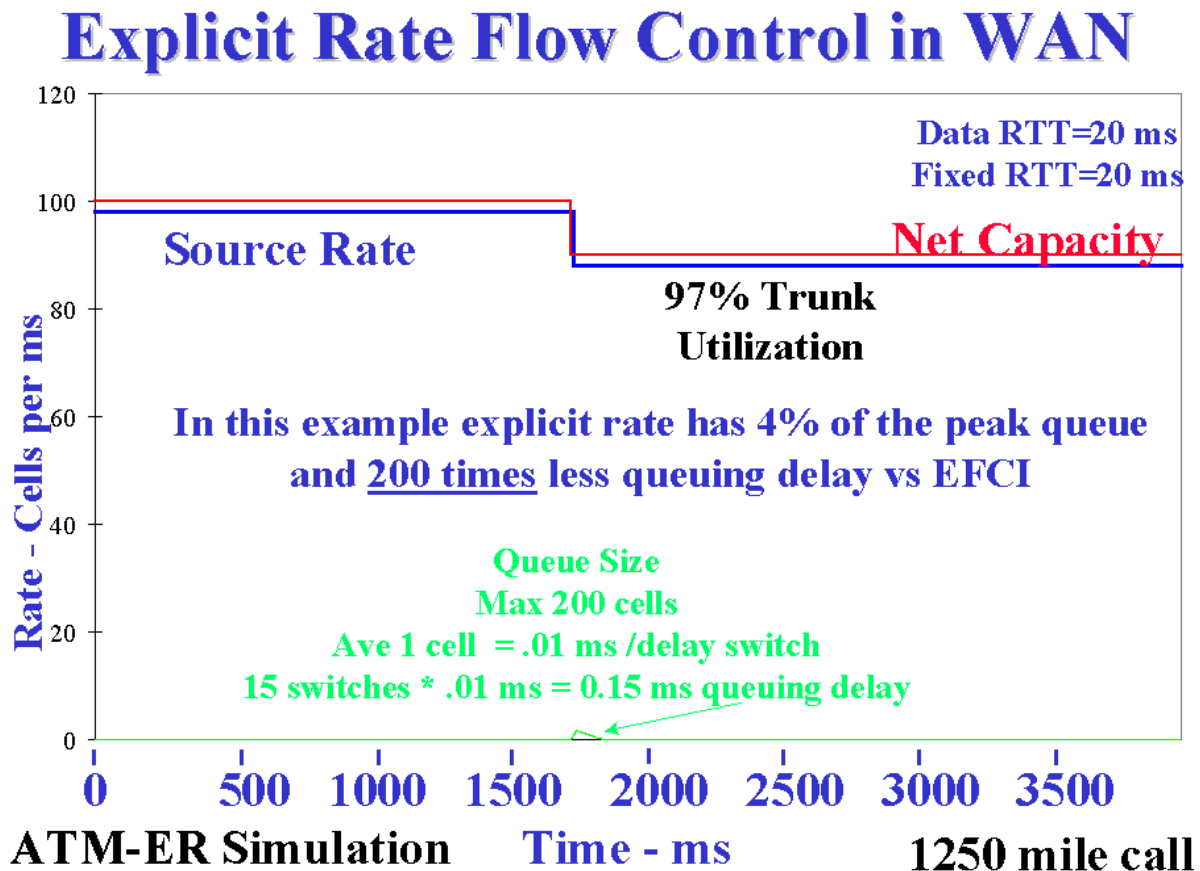
**Figure 2. Simple Rate Change with Binary Rate Flow Control**

Explicit Rate Protocol This simulation example is for ATM using Explicit Rate (ER). The source sends Resource Management (RM) cells every 32 cells along with the forward data, all marked with the explicit rate, which the source desires. The RM cells are given RM cell priority so that they move through the network at about the speed of light (20 ms round trip). The destination returns them in reverse direction and the switches in the path all mark down the explicit rate field if it is greater than that switch can support. Thus when the RM cells arrive at the source they specify the maximum rate that the source can transmit without congesting the network and it adjusts its rate to the specified rate.

As can be seen, Explicit Rate controls the source rate to track the network capacity extremely closely. The switches in this example opted to set the rate at 97% of their capacity since that permits them to quickly clear the queue that occurs when a capacity change occurs. The source learns about the capacity change within about 6 milliseconds and thus a small queue of 200 cells builds up at the switch before the source responds. This results in an average delay of one cell per switch or 150 microseconds for 15 switches,. Thus the RTT for data is 20 milliseconds, virtually the same as the fixed RTT. In this case, the control-time for Explicit Rate is 6 milliseconds. This results in the average

**Explicit Rate Flow Control**  
**A 100 Fold Improvement over TCP**  
 Dr. Lawrence G. Roberts

delay being 200 times less than for binary rate, the peak queue size being 25 times less, and the network utilization being 13% better.



**Figure 3. Simple Rate Change with Explicit Rate Flow Control**

**More Typical Network Traffic Variation – 65% Random Network Capacity Variation**

In a typical loaded network, the network capacity is continually changing, sometimes due to new CBR calls, sometime due to changing VBR rates, and sometime due to the number of other ABR calls active. In this example, the variation has a maximum range of 65% of the total capacity and varies randomly. This environment is very typical of an actual network. The network has 15 switches in the forward path. In this example, the network distance has been reduced to Metropolitan Area Network size, 375 miles, which has a 6 millisecond fixed round trip time.

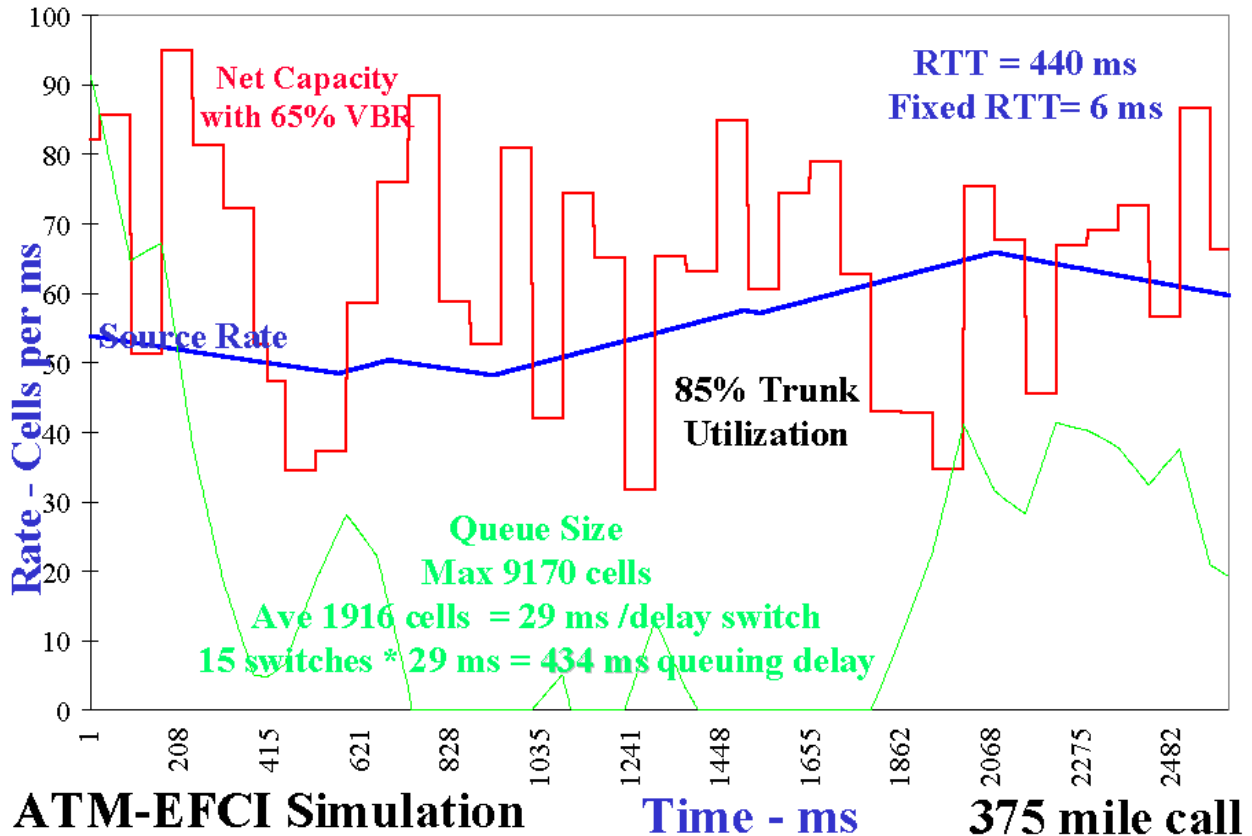
Binary Rate Protocol With such a highly varying network capacity, binary rate protocol has a very difficult time tracking the capacity changes given its long control-time. The best it can do is to achieve about the same rate average over one of its cycles. As a result, the peak queue size (see figure 4) nearly doubles from 5100 cells in the simple case due primarily to the oscillation, to 9170 cells. The average delay also increases from 180 ms to 434 ms. This occurred even though there was a major reduction in the network path distance from 1250 to 375 miles. Thus it is clear that binary rate can build up major queuing delays even at MAN distances. The buildup is more related to the number of switches in the path and the time constant of the network capacity variations than the actual path

# Explicit Rate Flow Control A 100 Fold Improvement over TCP

Dr. Lawrence G. Roberts

distance. Although this simulation is for ATM-EFCI, all binary rate flow control protocols (like TCP) operate in much the same way, if their oscillation period is larger than the network capacity variation period, the flow control cannot track the capacity and major queuing delays result, even at short distances.

## Binary Rate Flow Control with VBR

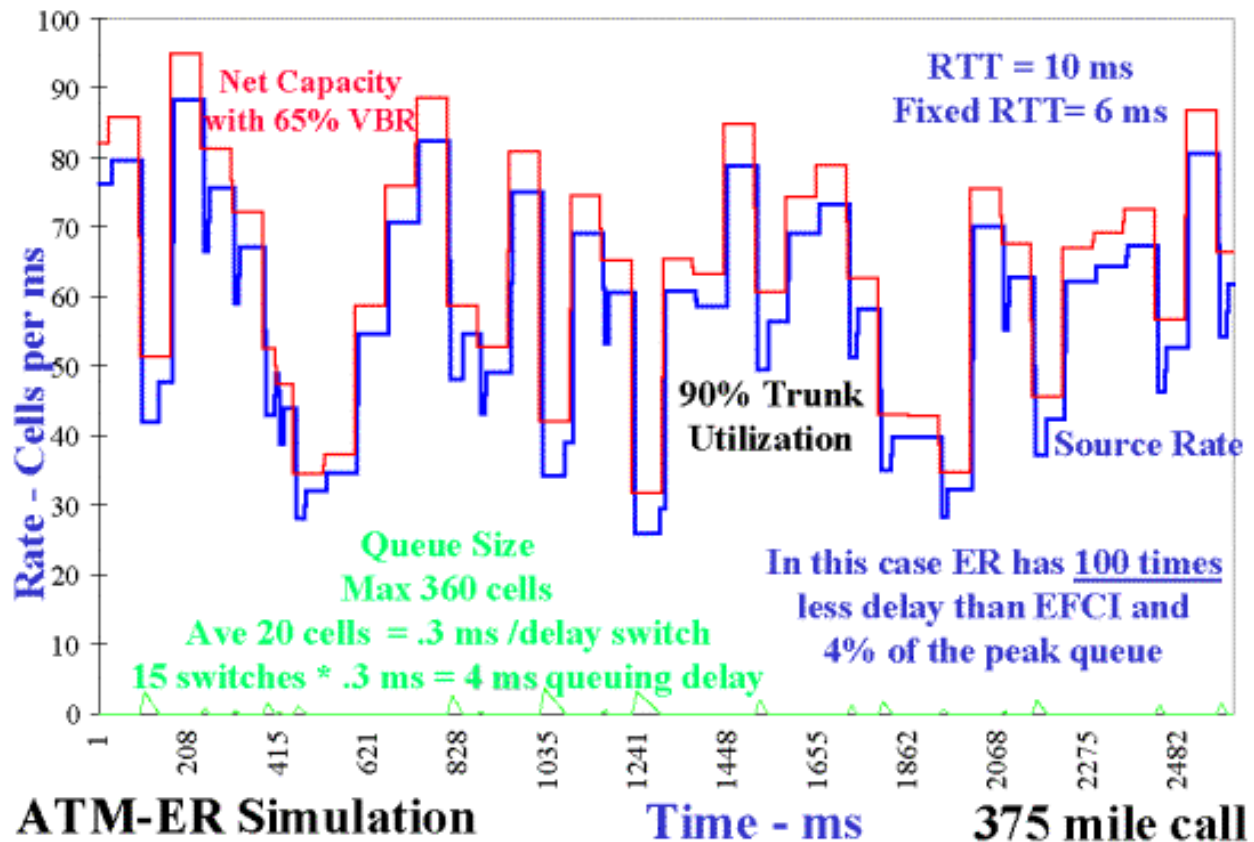


**Figure 4. Random 65% Rate Change with Binary Rate Flow Control**

Explicit Rate Protocol With Explicit Rate flow control, the source can be controlled to track the network capacity almost exactly. As shown in figure 5, with Explicit Rate the maximum queue size is again 25 times less than for binary rate and the average delay is only 4 ms, 100 times less than for binary rate. This is typical of the real difference Explicit Rate can make; two order of magnitude improvements in delay, delay variance, and control-time over all binary rate protocols.

**Explicit Rate Flow Control**  
**A 100 Fold Improvement over TCP**  
 Dr. Lawrence G. Roberts

## Explicit Rate Flow Control with VBR



**Figure 5. Random 65% Rate Change with Explicit Rate Flow Control**

### The IETF Alternative: TCP/IP

The Internet Engineering Task Force (IETF) supports the TCP/IP protocol and is currently drafting a signaling protocol for QOS based traffic, RSVP. TCP/IP is used in the Internet and many LAN's. It consists of several parts. IP addressing is a good addressing plan and will continue to be used with or without ATM or CIF. IP addresses can easily be translated into ATM addresses through an Address Resolution Protocol (ARP) request that can be handled by a router incorporated in the CIF switch.

### TCP Flow Control

TCP operates in the following way. It starts sending data at a very slow rate (slow start) and watches the acknowledgements from the destination to see if any data is lost. If none is lost, it speeds up. It keeps speeding up until data is lost at which time it decreases its rate. It keeps decreasing its rate until no data is lost. It then increases its rate again, oscillating like this continually.

The most critical problem today in the Internet is the slow control-time of the TCP flow control. TCP ships about one second of data into the network before congestion can be stopped. This one-second time is more or less fixed by the speed of light, the operating system delay and the queuing delay multiple. So long as TCP operates only in the end-stations its operation cannot be substantially improved. When the router or switch memories in the data path all have sufficient memory, TCP

# Explicit Rate Flow Control

## A 100 Fold Improvement over TCP

Dr. Lawrence G. Roberts

works, although slowly. However, the memory required is the delay-bandwidth product of TCP's cycle time and the full input bandwidth of the router or switch. As the new switching hardware like ATM is added to the network, the switch memory is typically not large enough for TCP's one-second cycle. ATM switches have perhaps 1/50 of the memory required for TCP operation. Therefore, if TCP is not replaced, TCP will cause major overloads and outages on long haul networks like the Internet.

Secondly, the user is severely impacted by the slow start-up rate and high delay variance inherent with TCP.

The IETF has not even considered revising TCP. In fact no study has been done by the IETF on flow control because everyone seems to believe, "if it worked in the past, it will continue to work". Nothing can be further from the truth in an environment like the Internet where growth is a factor of 5 per year. TCP must be replaced with a new flow control as good as explicit rate flow control as soon as possible.

### IETF Protocol Future

The IETF protocol stack is today grossly deficient in both flow control and QOS. The IETF has responded by working on a draft of RSVP, which is billed as fixing all the problems but in fact will only address the QOS signaling issue. Both flow control and QOS routing have been ignored. Support for Major Functional Requirements by Protocol

Functional Requirement	ATM Protocol	TCP/IP
QOS Signaling	UNI-SIG 4.0	RSVP (draft)
QOS Routing	PNNI 1.0	None
Explicit Rate Flow Control	TM 4.0	None

There is no easy way for the IETF to come up with new, unique solutions to these protocol issues in less than four years and then by the time all the network equipment is replaced to incorporate the new protocols, it will be eight years. Thus, the logical conclusion is for the now complete ATM protocol suite to be used, either on ATM switches or with CIF on the older legacy protocols. This can either be done by the IETF adopting the ATM protocols (unlikely) or the user and Internet Service Providers (ISP's) making the decision to use CIF and ATM.

### Control Time Analysis of TCP and Explicit Rate

Simulations only let one observe individual cases and it is very hard to conclude general results from such special cases. They are good to get the feel for what the behavior is like, but for a true comparison, analysis is far preferable where possible. Thus, the following is a far more precise and complete comparison of TCP and ER.

There are three major factors involved in determining the control-time of a flow control loop: The propagation speed of the control signal. With explicit rate there is the option to give the RM cells priority so that they can move near the speed of light, not the speed of the data flow. With TCP this is impossible because the control signal is the loss of data from the data path. For the Internet this factor results in, on the average, a factor of 10 improvement in control-time.

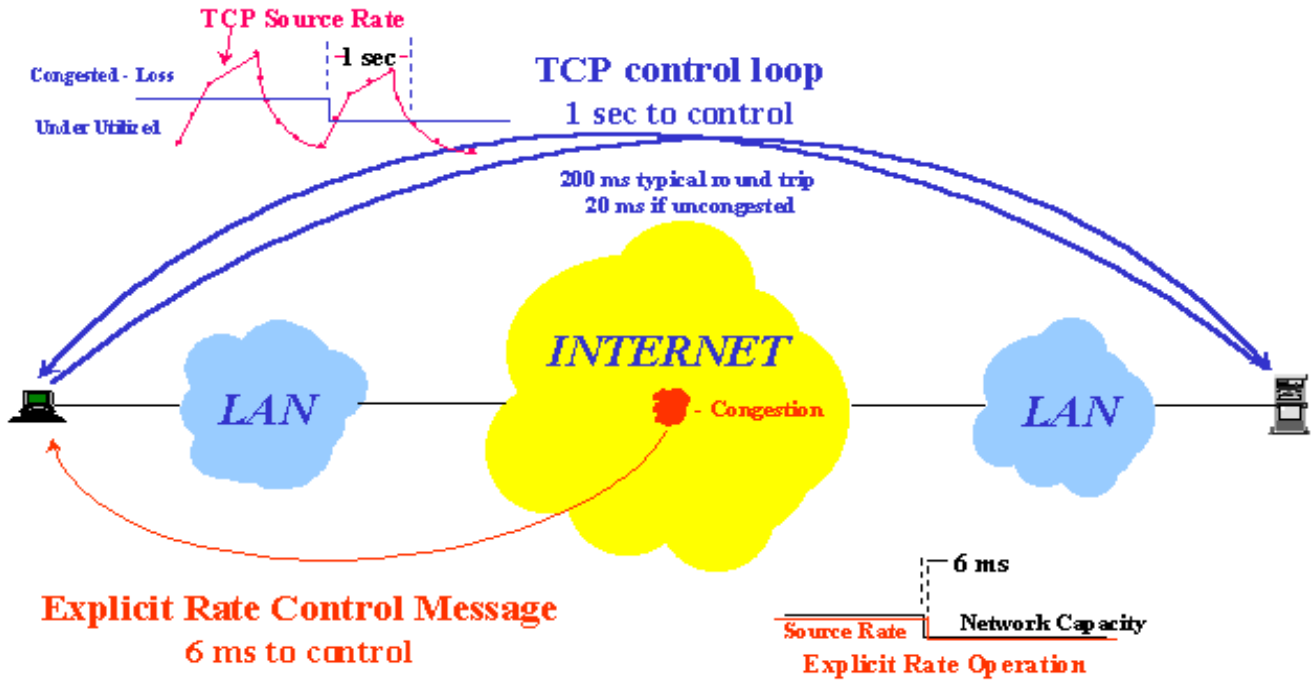
The oscillation cycle. With explicit rate there is no oscillation and the source adjusts immediately upon receiving the first RM cell. With TCP, the source adjusts over typically 5 round trips times to the network capacity. This results in another factor of 5 improvement in control-time.

# Explicit Rate Flow Control

## A 100 Fold Improvement over TCP

Dr. Lawrence G. Roberts

The distance which the control must travel. TCP must send data around the whole network to detect loss whereas Explicit Rate only needs to send the RM cell back  $\frac{1}{4}$  of the network round trip distance. This results in at least another factor of 2 improvement in control-time.



**Figure 6. Comparison of TCP and Explicit Rate Time to Control**

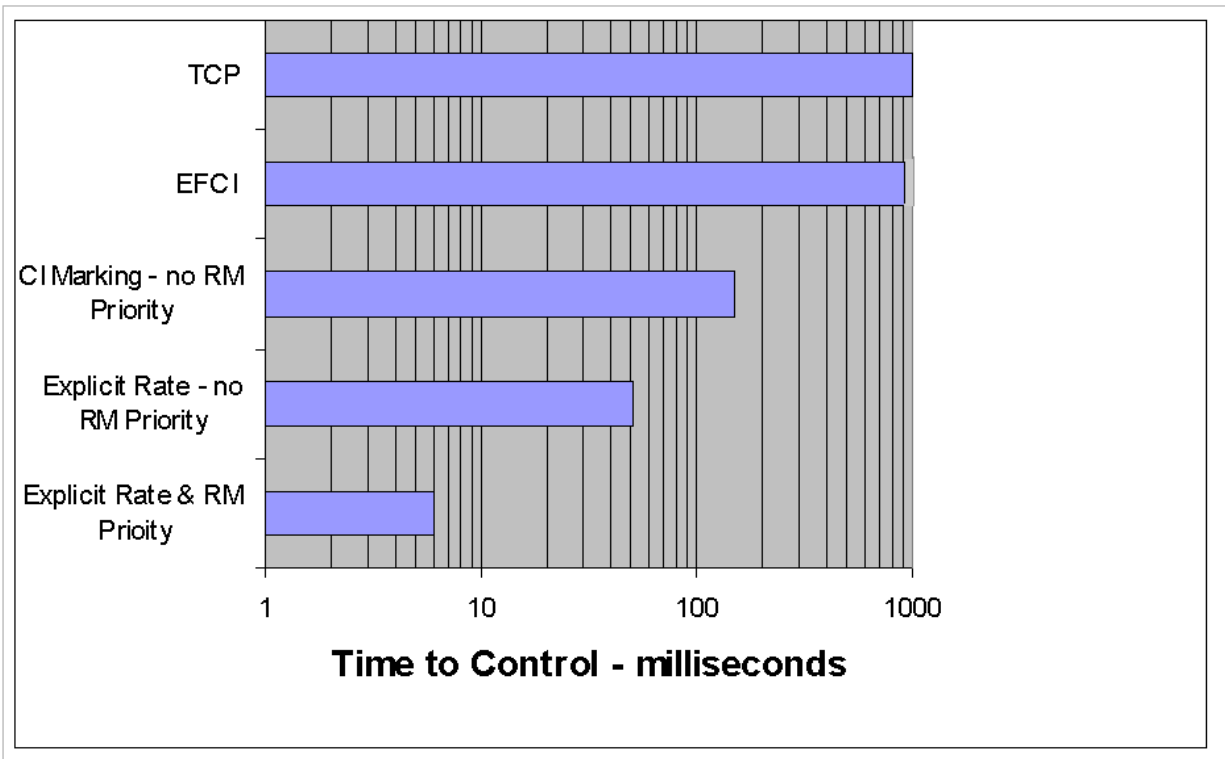
Overall these factors result in a factor of 100 in the control-time reduction from TCP to Explicit Rate.

### Comparison of Control-Times of Flow Control Options

The graph in figure 7 shows the control-time for TCP, and several ATM ABR options: EFCI, CI Marking, and Explicit Rate. CI marking is not done today since to do it fairly is just as hard as Explicit Rate, but if done is somewhat better than EFCI marking. Explicit Rate is shown with and without RM cell priority, a critical feature that reduces the time by a factor of 10.

**Explicit Rate Flow Control**  
**A 100 Fold Improvement over TCP**  
Dr. Lawrence G. Roberts

## Flow Control Time



**Figure 7. Control-times for Flow Control Options**

### **Impact of Control-Time Reduction from Explicit Rate Benefits For the Network Operator**

The network operator can gain some of the advantages of Explicit Rate, even without extending Explicit Rate end-to-end between users. If a network or sub-network is operated with Explicit Rate and the Explicit Rate is terminated at the edge of the net with edge devices, then data loss, switch memory, and transmission wastage can all be reduced for that network or subnet.

Data Loss Explicit Rate can be configured to achieve as low a cell loss ratio as desired. This is done by adjusting the Transient Buffer Exposure (TBE) parameter during signaling. Typically, it would be likely that the network operator would reduce the data loss to 10-12. Since TCP typically has loss rates of 10-1 to 10-2 this is a dramatic improvement. One major effect of this lower loss rate is that delay is dramatically reduced since retransmissions take lots of time.

Switch Memory Size Even though there tends to be statistical smoothing of the data transients at a switch for both techniques, the reduction of delay is directly reflected in reduced buffer size. This is critical because high-speed ATM switches cannot afford the same memory as the older software switches and routers had. A 100 Mbps router often has 12 Megabytes of buffer storage or about 1 second of buffering at full input rate. A 10 Gbps ATM switch typically has about 10 ms of storage at

# Explicit Rate Flow Control

## A 100 Fold Improvement over TCP

Dr. Lawrence G. Roberts

full input rate and could not afford to increase it to 1 second since the cost would be unreasonably high. Explicit rate flow control makes it possible for an ATM switch to use only 10 ms of memory and still support near loss-less data transmission.

The reduction of switch memory requirements by 100:1 results in a switch cost reduction of between 3:1 and 10:1. Today, the price spread between ATM switches and routers is more like 10:1 but that is partly due to memory and partly due to the greater difficulty of routing compared to switching. For a pure ATM switch using synchronous DRAM, the price improvement due to Explicit Rate would be at least 3:1.

Line Utilization Line utilization can be controlled to be much better with a faster control-time. In the simulation examples the improvement was 6-13%. However, there is a total tradeoff possible between utilization and queuing delay. One can fill memory with data to insure high line utilization but this increases the delay and buffer size required. There are some Internet sites which run their lines at 25% utilization to reduce delay and data loss. These sites could greatly increase utilization with Explicit Rate. Alternatively, other sites run their lines at 90% load with high delay and data loss. These sites could reduce their delay, data loss and memory requirements with Explicit Rate.

### Benefits For the End User

Once Explicit Rate has been established from end-to-end the user can gain other advantages. To gain these, TCP must be avoided or modified. To avoid TCP one can run a native ATM application. But to gain the broadest benefit, TCP should be modified so that when a VC has been established end-to-end with Explicit Rate, then TCP does not use slow start and lets the data flow be controlled by ER. If the VC path includes a router without ER, then TCP should continue to use slow start. This change to TCP is minor and a revised version could easily be downloaded if the user wants the benefits.

World Wide Web Access Time Today, TCP is used for all WWW activity. Each page access starts a new connection with TCP in slow start at about 2,000 bits/second. Since the typical page (not including graphics) is small – perhaps 5,000 bytes, then TCP typically takes 10 seconds to retrieve a page. If no loss is detected the rate might be increased, but most page access is done at slow start today. The graphics, being a larger block, may move TCP out of slow start, but even so many seconds are lost up front getting TCP up to speed. Thus slow start is the largest negative impact on Internet performance today.

The reason for slow start is that TCP has no knowledge of the network capacity until it has operated for several seconds, testing the network operating point. It would be catastrophic to the network if TCP were to start faster without any ability to determine the network status faster. The network would be inundated. However, with a flow control that could determine network capacity faster, the start rate could be made proportionately faster. This is because:

$$\text{Start Rate} = \text{Buffer Size} / \text{Control-Time}$$

Thus, since explicit rate is 100 times faster than TCP, the start rate could be increased to at least 200 Kilobits/second which would result in WWW page access in 0.1 second rather than 10 seconds. Such a change would be a major increase in performance for the WWW.

# Explicit Rate Flow Control

## A 100 Fold Improvement over TCP

Dr. Lawrence G. Roberts

Reduction in Delay Variation Explicit rate switches have 100 times less delay variance the IP switches, down from 3 sec. to 30 ms. This dramatically improves the consistency of the service for all data applications and also improves the utility of the data service for highly interactive activities.

Reduction in Data Loss and Retransmission When explicit rate is complete end-to-end, there is virtually no data loss (10-12) whereas TCP/IP has 10-20% data loss (and retransmission) in peak hours today on the Internet. Thus, 20% would be saved on line cost.

### Cells in Frames (CIF) to Extend Explicit Rate End-to-End

Implementing Explicit Rate Flow Control in the core of the Internet will improve overall bandwidth utilization and performance significantly. However, the benefits of flow control extend only as far as the RM cell can travel. If it is too difficult or expensive to expand ATM from the ATM-ER core to the desktop, Cells In Frames technology can extend the ATM protocol including ABR-ER across legacy frame based media, to the desktop.

CIF is ATM with variable length packets on the lines and trunks. The CIF Alliance has specified a protocol which allows ATM to be embedded into various frame based legacy protocols (Ethernet & Token Ring), using only one ATM header for up to 31 cells from the same virtual circuit in a packet. The specification of CIF over PPP and Sonet is underway. A significant feature of CIF is that ATM can be transported to workstations without changing the legacy NIC card because the necessary processing is done in simple downloaded software "SHIM" on the workstation. CIF also permits the connection of workstations and servers with about the same expense as standard Ethernet. CIF is described in another paper, "Cells In Frames" , and the specification is available "ATM Forum 96-1104, Request for Coordination of Cells In Frames Specification, August 1996".

### ABR-ER vs. UBR Simulations

There have been numerous simulations that apparently show that Explicit Rate ABR is no better than TCP over UBR. There are several flaws in these simulations which help explain why something 100 times worse is being made to look similar:

The simulations all presume that TCP is used with BOTH ABR and UBR. This is assumed because routers without explicit rate were assumed to be in the path and people were not sure how TCP flow control would be eliminated when ABR was used. It is simple however, to remove TCP flow control by running ATM native mode, running UDP, or by modifying TCP. Either way the bad effects of TCP can be eliminated. But in all the simulations this was not done and routers without explicit rate were assumed to be in the path in addition to ATM explicit rate switches. Thus, TCP still was allowed to oscillate back and forth between slow start and major data loss. All the loss was typically in the routers. It is like comparing a leaky hose with a good hose by comparing them in all cases with a leaky hose in series with the good hose. Explicit rate must be end-to-end before a reasonable comparison can be made.

The simulations all presume no RM cell priority. Thus, a factor of 10 is lost up front. Data on a loaded network typically travels at 10% of the speed of light. This 10:1 data slow down is the actual experience on the Internet.

The test cases were not 1000% overloaded without flow control as in the real Internet. Thus, the severity of the problem with UBR was not demonstrated. It would be a disaster to really use UBR with its small buffers in the Internet with TCP unless it was carefully configured to never be overloaded.

# Explicit Rate Flow Control

## A 100 Fold Improvement over TCP

Dr. Lawrence G. Roberts

Given these flaws, the simulations are meaningless. The comparison is easy to make without simulation. One only has to evaluate time to control as done above, and the results follow fall out without a computer.

Explicit rate flow control is only useful to the user if the RM (Resource Management) cells flow end-to-end and all the congestion points arrange to mark the RM cells with the correct rate. If a current day router or bridge is in the path, the RM cells will be delayed behind data and no marking will be done. This will not have a major effect if the device is always significantly under-loaded, but if it ever becomes overloaded then TCP will need to be operated on top of the explicit rate flow control. Further, both end-stations must participate in the explicit rate flow control in order for it to work. If one or both do not support explicit rate, TCP will be required and all the user benefit will be lost.

### **Should ATM UBR (or UBR+) be used in the Internet?**

The analysis above shows that when TCP is used for flow control, the data buffers must be large enough to support on the order of one second of data from all the input lines. If the buffer is much smaller than this, the large peak demands required by binary rate flow control will cause major data losses with high probability. Since ATM-UBR switches only have 10-40 milliseconds of data storage they clearly are destined to have major data loss if used in the Internet. The only workable option to use ATM-UBR in a wide area Internet environment is to configure the network such that the trunks are always significantly underloaded (like 25% load). However, due to the fact that the trunks cost more than the switches in an ATM network this option is only attractive if no other alternative is available.

Recently UBR+ has been discussed at the ATM Forum. UBR+ is UBR with a minimum rate available. However, UBR+ still has no flow control and is expected to be controlled with TCP the same as UBR. The minimum rate has no effect on any of the discussion above, and UBR+ would fail just as badly as UBR in Internet usage.

### **RM Cell Priority**

The use of RM cell priority makes a major improvement in explicit rate. If RM cell priority is not implemented, then the gain over TCP is only 10:1 and when compared to systems with RM cell priority, one without it would be 10 times worse in time to control. This means that it would require ten times more memory in the ATM switch per port in order to achieve the same cell loss objective. This has a substantial impact on the cost of the switch. Secondly, by losing a factor of 10 in time to control, a switch without RM cell priority would have to have a startup rate for the user that is 10:1 slower than a switch which has RM cell priority.

### **Conclusion**

Flow control is the most critical factor in the performance of a data network. Since the load always seems to grow quickly to vastly exceed the available capacity, there must be some sort of traffic cop to control the load or the network will have to discard most of the incoming data and then retransmission only makes the problem worse. For the past 15 years the Internet flow control has been TCP. It was designed to work only between the end-stations so it did not need to concern itself with the network.

Today this is an unnecessary and extremely harmful restriction. The network can easily assist in determining and signaling its maximum capacity and the end-station can use this information to greatly reduce the control-time, the time from load determination to load adjustment. The ATM Forum has developed such a flow control protocol, Explicit Rate, and the completed specification was

# Explicit Rate Flow Control A 100 Fold Improvement over TCP

Dr. Lawrence G. Roberts

published in May 1996. Explicit Rate operates 100 times faster than TCP to control the rate of the traffic sources. This is near optimal given the limit of the speed of light.

The 100 times improvement in the control-time which Explicit Rate provides over TCP provides several major benefits. For the network operator, it reduces the cost of the switch by 3:1, eliminates most all data loss, and improves trunk utilization. For the user, it reduces the WWW access time from 10 seconds to 0.1 seconds, reduces the delay variance on the network from 3 sec to 30 milliseconds, and eliminates the need for almost all data retransmission.

It is critical that ATM with Explicit Rate be introduced into the Internet to reduce cost and reduce delays. To do this, it first must be introduced into the Internet backbone which will reduce its cost. Then it can be extended to the users through the use of Cells In Frames (CIF). Using CIF, an ISP can offer standard IP service to all users, but for those users who download the CIF software, the full benefits of ATM protocol can become available immediately. Thus, users can individually select to gain the benefits of Explicit Rate without requiring a flash cut to a new flow control throughout the network. Figure 8 shows how the Internet could be upgraded in the core with Explicit Rate and then at some ISP's with CIF to get the benefits to the users.

## Explicit Rate in the Internet

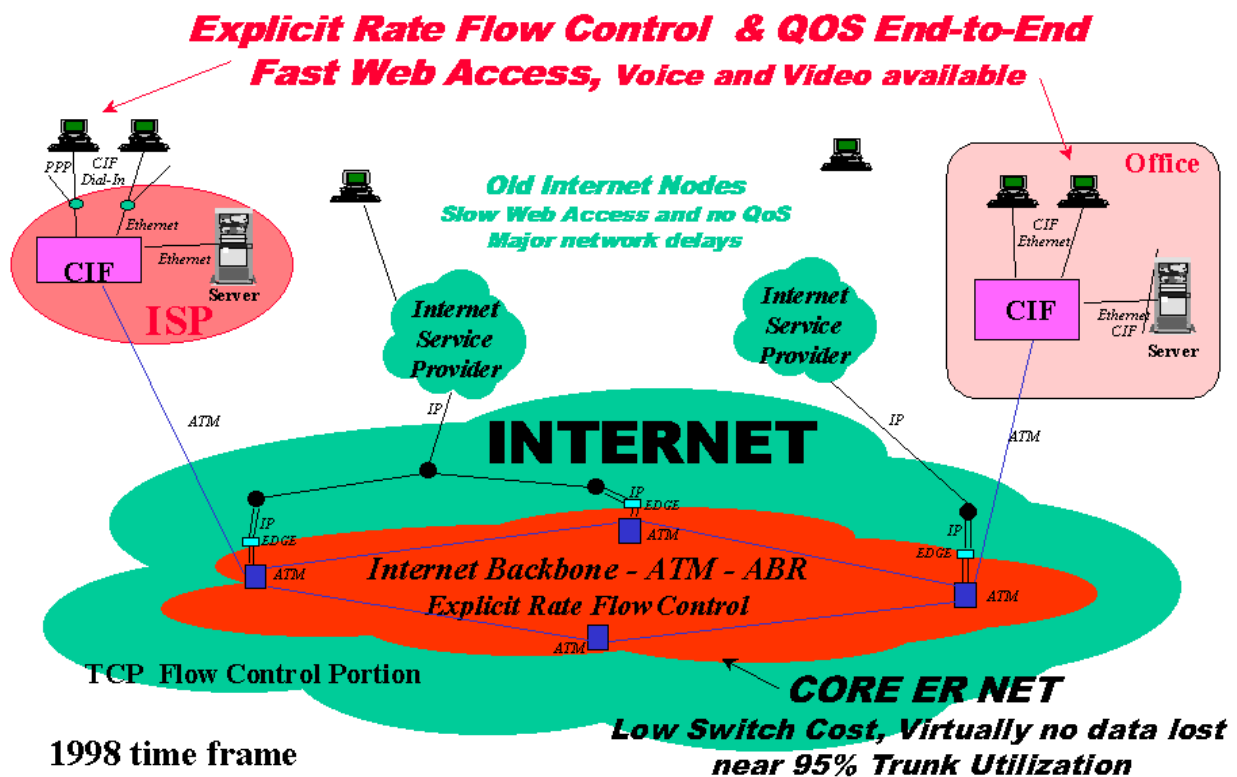


Figure 8. Upgrade Path for the Internet to Support Explicit Rate