PERFORMANCE MODELS OF BINARY SYNCHRONOUS COMMUNICATIONS MULTIPOINT CIRCUITS OVER VIRTUAL PRIVATE-LINE FRAME RELAY NETWORKS

Robert G. Cole

Robert G. Cole *is a member of technical staff in the Data Network Analysis Department at AT&T Bell Laboratories in Holmdel, New Jersey. He is responsible for end-to-end performance analysis of AT&T's and competitors' data network products. He received a B.S. in chemistry from the University of Miami and a Ph.D. in physical chemistry from Iowa State University. He joined AT&T in 1985.*

This paper discusses performance models of 3270 binary synchronous communications (BSC) multipoint circuits over virtual private-line frame-relay networks. The models are used to identify the key network functionality needed to minimize end-to-end delays perceived by users. We derive new models that determine the mean total transaction time between the end systems (a front-end processor and cluster controllers on multipoint circuits). The models capture queueing delays—an important component of transaction time. The models also incorporate important BSC protocol parameters, including system parameters of the front-end processor. A performance analysis based on the models demonstrates the sensitivity of the BSC protocol to the functionality at the network edge. For a simple virtual private-line replacement, egress pipelining is a necessity to maintain acceptably low end-to-end delays. For more complex replacements, where access line speeds can be economically increased and front-end-processor system generation parameters changed, the analysis demonstrates the value of edge polling.

Introduction

Currently, most data communications needs are met through private lines. In a representative application, many cluster controllers located on a customer's premises multiplex data from terminals onto a single private line (Figure 1). A front-end processor (FEP) acts as an interface between a host computer and the communications line. The processor controls the flow of data over the multipoint communications line according to a line protocol. An example is the binary synchronous communications (BSC) protocol, a half-duplex, synchronous, character-oriented protocol.[1,2] IBM Corporation developed BSC in the 1960s to

Panel 1. Acronyms and Abbreviations in This Paper

ACK	acknowledgment signal
BSC	binary sychnonous communications
CC	cluster controller
DDS	digital data service
ENQ	inquiry message
EOT	end-of-transmission signal
FEP	front-end processor
LAN	local-area network
LAPD	link-access procedures for D channel
LEC	local exchange carrier
NEGPOLP	parameter determining FEP idle time
POLL	polling signal
RESP	response message
RTS/CTS	request-to-send/clear-to-send
SEL	select signal
SERVLIM	parameter for maximum number of polls
SYSGEN	system generation parameter
TA_A	terminal adapter on access side
TA_F	terminal adapter on FEP side
VCDN	virtual-circuit data network
VDM	voice-data modem
VPL	virtual private line

extend data communications over large distances. It soon became a de facto communications standard.

A virtual circuit data network (VCDN) can handle this type of private-line traffic through various virtual private-line (VPL) replacements. The VCDN environment addressed here is based on statistical multiplexing for economic transport and on virtual circuit technology with high-speed frame-relay switching for high throughput and minimal network delays.[3] The VCDN consists of terminal adapters (for network access), frame-relay switches, and high-speed trunks.[4] The frame-relay switches perform packet switching on the basis of virtual circuit identity and error checking. The terminal adapters (TAs) accept the BSC native frames and convert the frames to be transmitted through the network into the TA-to-TA protocol, which we assume is the link-access procedures for D channel (LAPD). (The term *native frame* refers to a customer's link layer protocol frames.)

This paper presents new models for evaluating the performance of 3270 BSC, a representative standard binary synchronous communications operation, in a VCDN environment. The model's primary performance measure is the total transaction time, defined here as the time between the instant when the last bit of the inquiry (ENQ) message to the host enters the cluster controller and the instant when the last bit of the response (RESP) message to the terminal enters the cluster controller, minus the host processing time. A typical transaction consists of an inquiry message from a terminal to the host, which in turn sends a response message back to the terminal. Results of analyses with the models identify the functions and placement of terminal adapters to minimize the performance impact of VCDN virtual private-line replacements (Figures 1 to 3). Customer-premises and service-node placements of terminal adapters are examined.

The 3270 BSC models reported in this paper expand on previous BSC models[5-8] in that the effects of the FEP system parameters—for example, SERVLIM and NEGPOLP—are incorporated into the models (see Panel 2). This is accomplished by modeling two pseudo-independent queueing systems—one representing the actions of the FEP while polling for ENQs and the other the FEP actions while transmitting RESPs. These effects are important when the performance of VPL replacements is modeled, as discussed in the section "Configuration 3."

Further, the models incorporate different terminal adapter functions from a range of functionality that includes

- LAPD wrapping
- Edge-of-the-network polling
- Pipelining.

In *LAPD wrapping*, the access terminal adapter breaks up incoming native protocol frames into smaller LAPD frames for transport through the network. (We use the terms *access TA* for the terminal adapter on the origination side of the network and *egress TA* for the terminal adapter on the destination side of the network.) The egress TA is then responsible for reconstructing the original BSC frame for transmission on the egress line. All traffic, including BSC control frames, is passed through the network in this case. This function is also referred to as *remote polling* (the FEP polls remote cluster controllers)

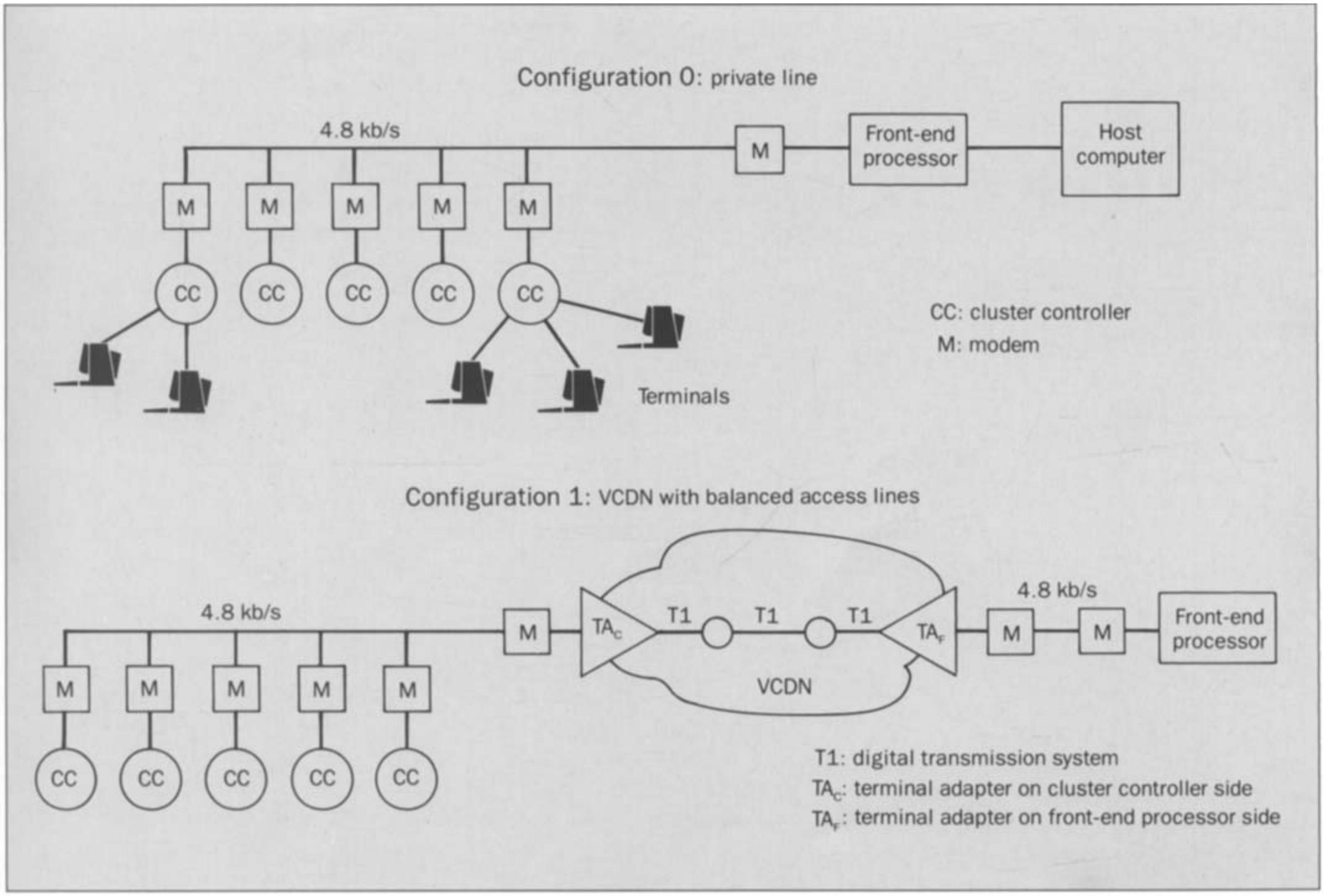


Figure 1. Multipoint private-line configuration and VCDN virtual private-line replacement.

or *through-the-network polling* (polls travel through the network, as opposed to edge-of-the-network polling discussed below).

Breaking up the native frames into smaller LAPD packets allows through-the-network pipelining. The origination TA accomplishes through-the-network pipelining by transmitting the first 40 bytes of the native frame as a single LAPD packet as soon as they arrive on the access line. It then accumulates the next 40 bytes and immediately transmits them. This speeds up transport through the network, but its real advantage is in conjunction with out-of-the-network pipelining. We always assume through-the-network pipelining and imply out-of-the-network pipelining when saying merely "pipelining."

In *edge-of-the-network polling* or *local polling,* the terminal adapter on the cluster controller (access) side of the network (TA_C) emulates the FEP by polling over the cluster controller access line, and the terminal adapter on the FEP side of the network (TA_F) emulates the cluster controllers acting as virtual cluster controllers. In this case, when a cluster controller sends an ENQ message,

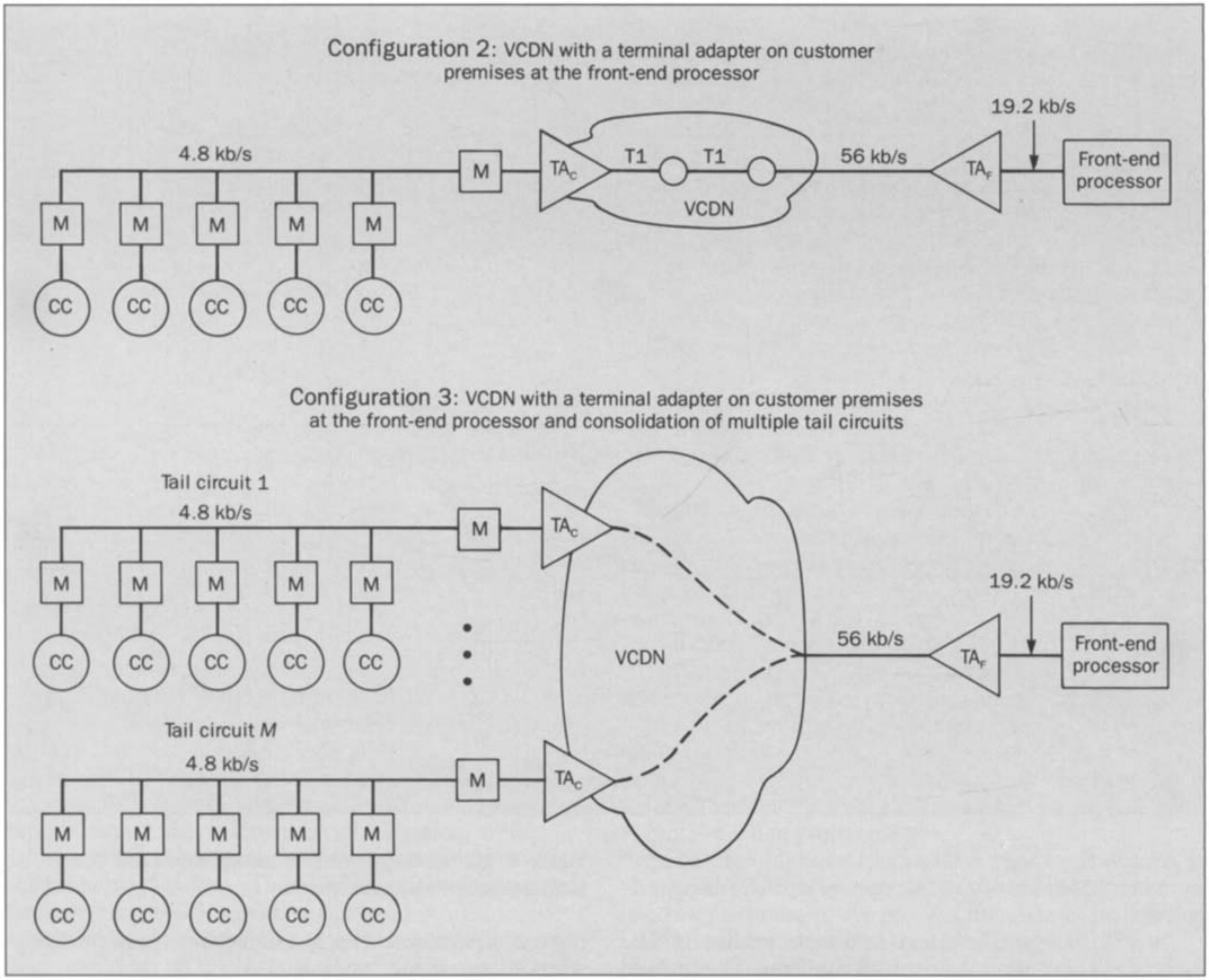


Figure 2. VCDN virtual private-line replacements with high-speed FEP access and tail circuit consolidation.

the TA_C breaks the message into smaller LAPD packets for transport through the network. However, the TA_C intercepts the BSC control frames sent by the cluster controllers, terminating the layer 2 protocol to emulate the FEP. The TA_F reconstructs the ENQ message and queues it as a cluster controller would while waiting for the FEP

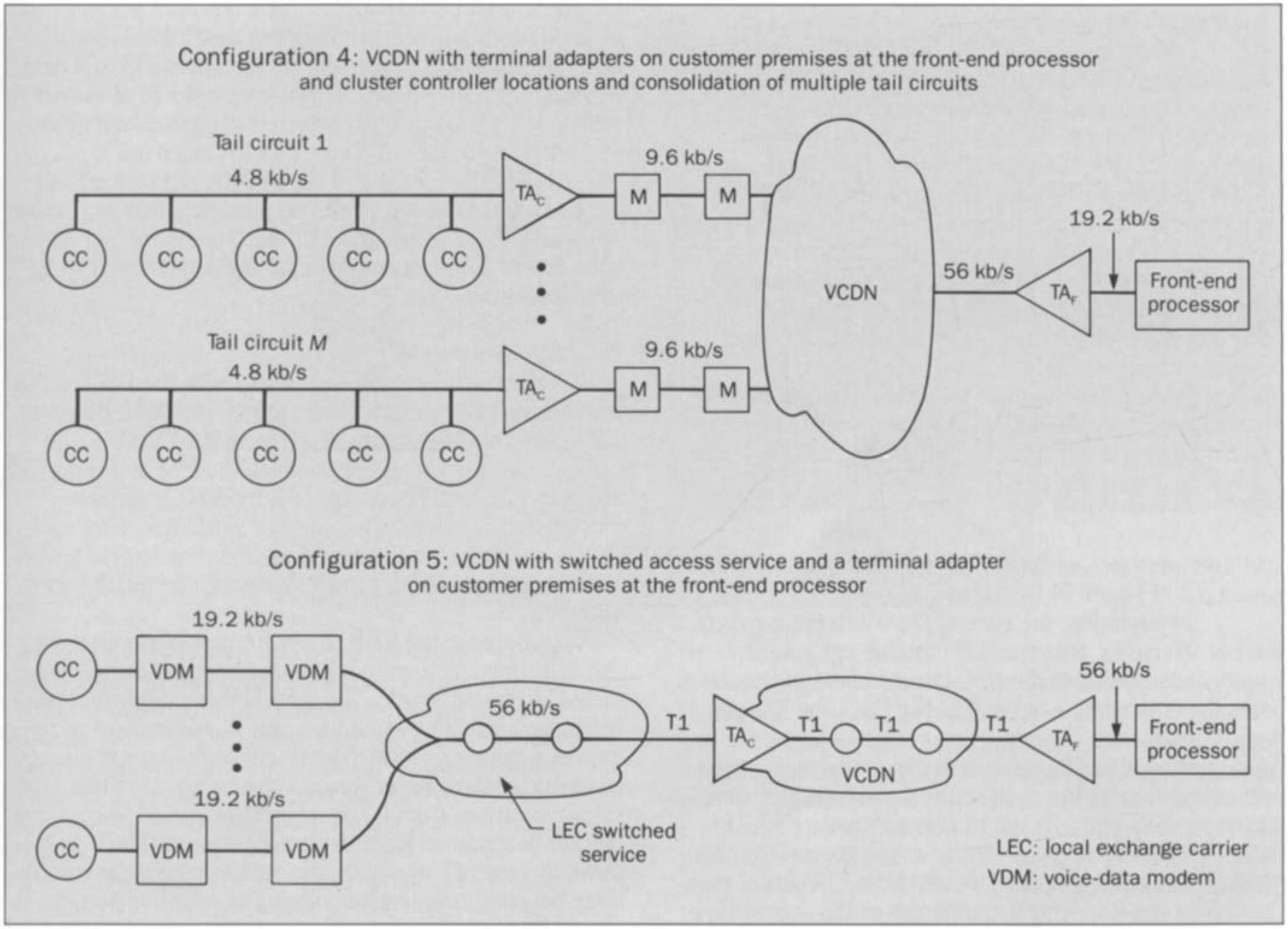


Figure 3. VCDN virtual private-line replacements with both TA_F and TA_C located on customer premises or in central office with LAN access.

to poll the particular cluster controller. The TA_F intercepts the polls from the FEP, terminating the layer 2 protocol to emulate the cluster controller, and transmits the ENQ message to the FEP. These functions have the advantage of reducing the polling cycle time—at the expense of introducing queueing at an additional device, however.

We distinguish between *edge-of-the-network polling* and *local polling* solely by the placement of the TAs; the TA functionalities are identical for both polling cases. *Edge-of-the-network polling* refers to configurations with TAs located at service nodes on the network edge (for example, see configuration 1 of Figure 1). *Local polling* refers to configurations with the TAs located on the customer premises (for example, see configuration 4 of Figure 3). Mixed configurations are allowed; for example, the TA_C may be located at a service node while the TA_F is collocated on

Panel 2. SYSGEN Parameters

Many system generation (SYSGEN) parameters are initialized when a front-end processor is set up. The parameters are used to optimize FEP performance for a given FEP and cluster controller configuration. Two SYSGEN parameters relevant to this article are SERVLIM and NEGPOLP. The parameter SERVLIM indicates the maximum number of consecutive polls sent out on the multipoint line by the FEP during any one polling period. The parameter NEGPOLP controls the length of time the FEP goes idle on the multidrop line after it receives an end-of-transmission frame in response to a poll. Such a frame indicates that the polled cluster controller had no ENQ messages queued for delivery to the host. The idle-line period serves as a load-balancing mechanism that gives the FEP more time to poll busy lines and less to poll idle lines.

customer premises with the FEP (for example, see configuration 2 of Figure 2).

In *pipelining*, the egress TA, which reconstructs the BSC frame for transmission onto the egress line, begins transmission of the BSC frame before fully accumulating the entire frame at the TA. For BSC, the TA may begin transmission onto the egress line as soon as the first LAPD packet of the BSC frame arrives, minimizing transport delays. In pipelining, the end-to-end transport time for a large BSC frame is the time to accumulate a LAPD packet's worth of the native frame at the access TA, plus network delay, plus the time to insert the BSC frame on the egress line (the largest component of the delay). However, without pipelining, the end-to-end transport time is at least *twice* the time to insert the BSC frame on the egress line, regardless of the speed of the network components.

Inevitably, because of variance in the network delays, there will be occasions during pipelining when the TA finishes sending an LAPD packet's worth of the native frame before the next LAPD packet of the BSC frame arrives. In this case, we assume that the TA will insert filler characters into the BSC frame in order to maintain synchronization while waiting for the next packet. When the end systems do not allow this, a build-out scheme must be used, as in synchronous data link control.[3,9] When a build-out scheme is implemented, the egress TA will delay the insertion of the first LAPD packet's worth of the native frame for a period of time, referred to as *build-out delay*, to allow subsequent LAPD packets to "catch up."

Remote polling and edge-of-the-network polling are mutually exclusive. However, pipelining may or may not be used with either. In general, then, there are four distinct sets of TA functionality to be studied for each VPL configuration.

The 3270 BSC Models

The primary goal of analysis with the BSC models is to derive the mean transaction time on a 3270 BSC multidrop line. In the diagram of the model in Figure 4, the circles represent the states in which the FEP is polling (S_e) and selecting (S_r). For polling, the model is a multiple-queue, cyclic-server system, with a single queue for each cluster controller on the multidrop line. For selecting, the model is a single-queue, single-server system with server vacations.

Because the FEP shares its time between polling and selecting, each system sees a server that is only partially available for work. The models therefore treat the polling and selecting systems as quasi-independent queueing systems in which interactions between the two queueing systems occur through the moments of the server vacations (that is, the server's vacation characteristics are determined from the server sojourn time). To make the analysis tractable, the models ignore the correlations between consecutive polling and selecting periods. A closed-form expression for the mean transaction time results from conservation laws relating the two quasi-independent queueing systems.

The total response time, if host processing time is ignored, is

$$E(T) = [E(W_e) - h_e^{(1)}] + [E(W_r) - h_r^{(1)}] + t_e + t_r \quad (1)$$

where $E(W_e)$ and $E(W_r)$ are the mean waiting times (queueing plus service), $h_e^{(1)}$ and $h_r^{(1)}$ are first moments of the service time distribution functions for ENQ and RESP

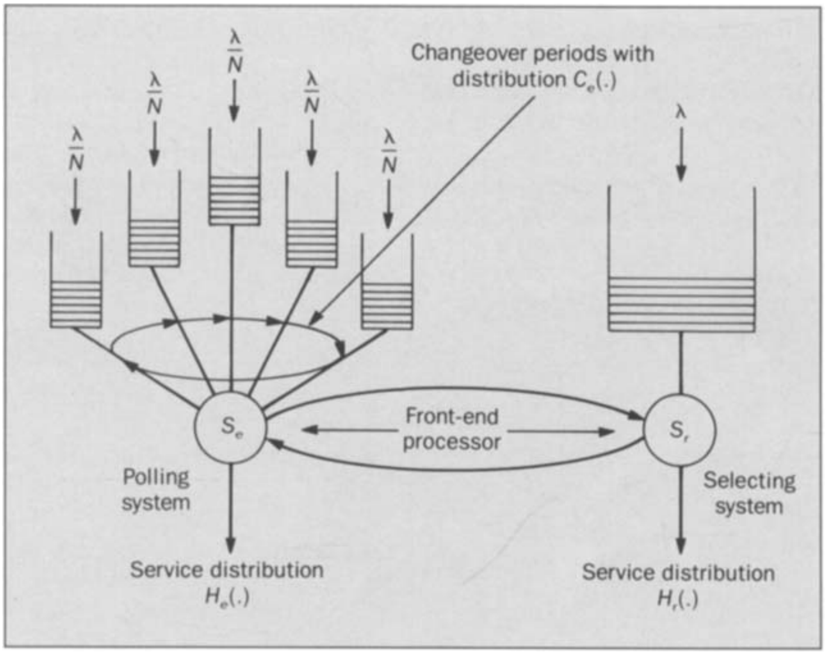


Figure 4. 3270 BSC models as two pseudo-independent queueing systems.

messages respectively, and t_e and t_r are mean transport times for ENQ and RESP messages, respectively. For simplicity, we assume a fixed message size for ENQs and RESPs. Most quantities (for example, t_e and $h_e^{(1)}$) can be derived directly from a timing diagram (see Figure 5). Remaining quantities [for example, $E(W_e)$] can be derived by analyzing the polling and selecting systems discussed below.

Polling System. The waiting time for ENQs is obtained from the expression for the mean waiting time in a multiple-queue, cyclic-server system with independent changeover times and Poisson arrivals, first derived by Hashida:[10]

$$E(W_e) = \frac{(N-1)c_e^{(1)}}{2(1-\rho_e)} + \frac{c_e^{(2)}}{2c_e^{(1)}} + \frac{\lambda h_e^{(2)}}{2(1-\rho_e)} + h_e^{(1)} \quad (2)$$

where $c_e^{(1)}$ and $c_e^{(2)}$ are the moments of the changeover, or *walk-time*, distribution; ρ_e is the polling system load; λ is the system total arrival rate; and N is the number of cluster controllers or queues. We need only determine $c_e^{(1)}$ and $c_e^{(2)}$, since all other quantities are model inputs (for example, N and λ) or can be obtained directly from Figure 5.

The *walk-time*, the time from one cluster controller relinquishing line control to the next cluster controller gaining line control, plays a critical role in models of polled protocols. The minimum, or bare, walk-time is simply the time necessary for the cluster controller to transmit a short control frame to the FEP, plus the time for the FEP to transmit a poll to the next cluster controller, plus propagation delays due to distance and modem processing. In addition, walk-times may be enlarged by (1) FEP pauses of duration equal to the NEGPOLP setting and (2) interruption of the polling period so that the FEP can transmit RESP messages before the next poll. The latter interruption occurs whenever the maximum number of consecutive polls, equal to the value of the SERVLIM parameter, is sent or a polled cluster controller sends ENQ messages to the FEP.

Let x be the random variable representing the assumed independently distributed walk-times. We further

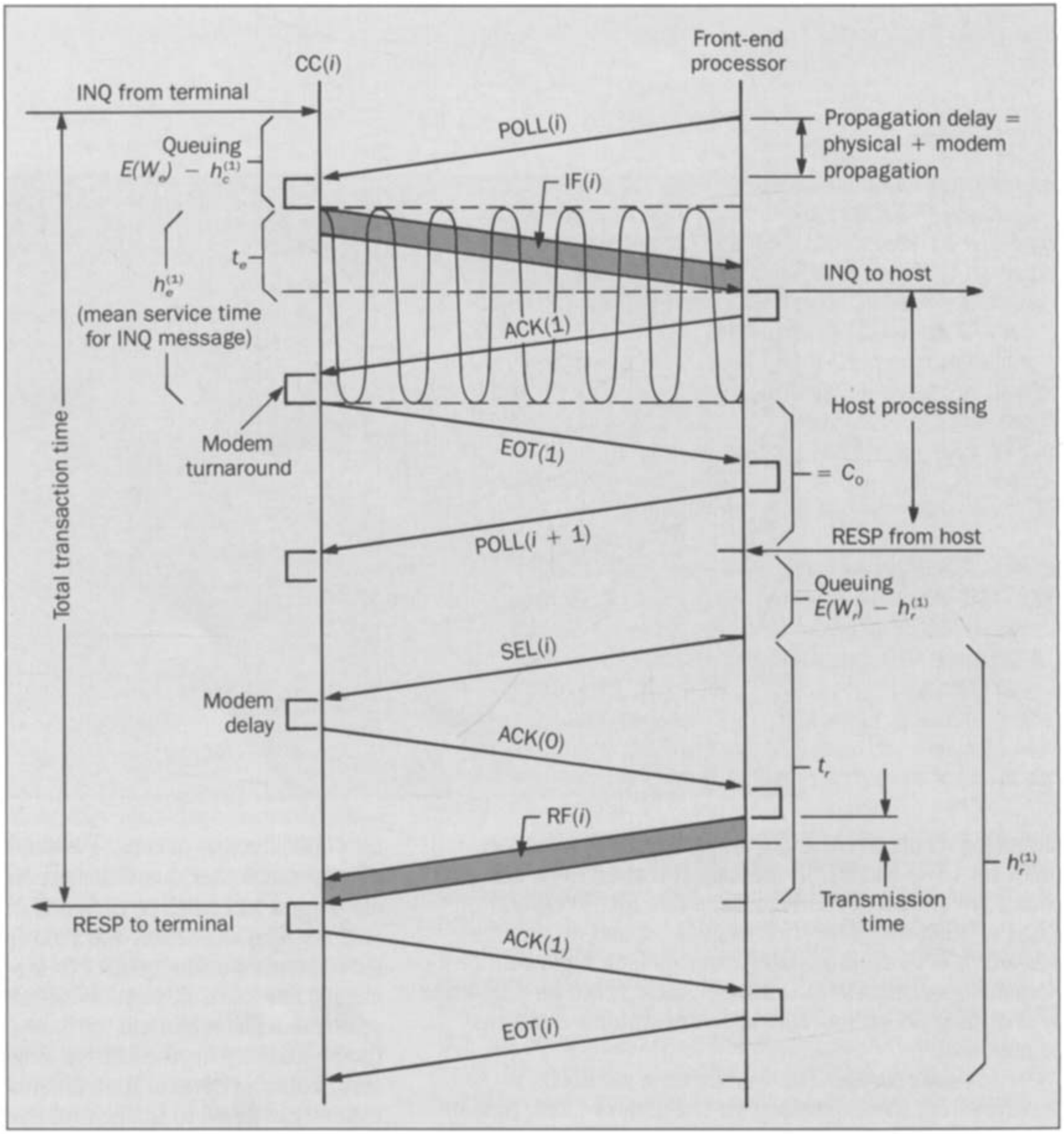


Figure 5. 3270 BSC private-line timing diagram for a typical transaction, consisting of a single ENQ frame and a single RESP frame with companion BSC control frames.

assume x to be the sum of three independently distributed variables:

$$x = x_0 + x_n + x_s$$

where x_0 = bare walk-time
x_n = contribution of FEP pause times
x_s = contribution of polling interruptions to transmit RESP messages

with distributions

$$D(x_0) = H(x_0 - c_0)$$

$$D(x_n) = (1 - p_0)H(x_n) + p_0 H(x_n - f_{\text{neg}})$$

$$D(x_s) = (1 - \beta)H(x_s) + \beta V(x_s)$$

Here c_0 is obtained from Figure 5, $H(x)$ is the Heavyside function, p_0 is the probability that a cluster controller is empty when it receives a poll, f_{neg} is the value of the NEGPOLP parameter, β is the probability that the FEP interrupts polling during an arbitrary changeover, and $V(.)$ is the distribution for the FEP sojourn time in an arbitrary selecting period with moments $d_e^{(1)}$ and $c_e^{(2)}$. By equating the mean number of ENQ arrivals to the mean number of RESPs served during a mean polling cycle, we can relate the mean of x, $c_e^{(1)}$, to the means of x_0 and x_n.[11] We get

$$c_e^{(1)} = (c_0 + p_0 f_{\text{neg}})\frac{(1 - \rho_e)}{(1 - \rho_e - \rho_r)}$$

where ρ_r is the load on the selecting system. By definition, we get

$$c_e^{(2)} = c_0^{(2)} + p_0 f_{\text{neg}} + \beta v_e^{(2)} + 2\,[p_0 c_0 f_{\text{neg}} + \beta c_0 v_e^{(1)} + p_0 \beta v_e^{(1)} f_{\text{neg}}]$$

Given that the arrivals are Poisson, we can approximate p_0 as

$$p_0 = \int e^{-\lambda T_B/N} d(1 - e^{-T_B/E(T_B)})$$

where T_B is the time between the server departure from an arbitrary queue to the server return to the same queue, which we have assumed to be exponentially distributed. From Reference 12, $E(T_B) = Nc_e^{(1)}(1 - \rho_e/N)/(1 - \rho_e)$ and

$$p_0 = \left[1 + \lambda c_{e,o}^{(1)} \frac{1 - \rho_e/N}{(1 - \rho_e - \rho_r)} \right]^{-1} \qquad (3)$$

By assuming that times of initiation of polling periods are renewal points, we get (see Appendix 3)

$$\beta = \frac{1}{(1 - p_0) \sum_{i=0}^{f_{\max}-1} (i + 1) p_0^i + f_{\max} p_0^{f_{max}}} \qquad (4)$$

where $f_{\max}$ is the value of the FEP parameter SERVLIM.

The only remaining quantities to be found are the two moments $v_e^{(1)}$ and $v_e^{(2)}$ of the FEP sojourn time in a selecting period. If we ignore correlations between the polling and selecting periods, the moments are simply related to $v_r^{(1)}$ and $v_r^{(2)}$, alternately interpreted as either moments of the FEP vacation distribution from serving RESPs or the FEP sojourn time in an arbitrary polling period. We get[12]

$$v_e^{(1)} = \frac{\rho_r v_r^{(1)}}{1 - \rho_r}$$

$$v_e^{(2)} = h_r^{(2)} \frac{\lambda_r v_r^{(1)}}{1 - \rho_r}$$

where ρ_r is the load on the selecting system.

Selecting System. The waiting time for RESP messages is obtained from the expression for a single-queue, single-server system with vacations:[12]

$$E(W_r) = \frac{\rho_r v_r^{(1)}}{2(1 - \rho_r)} + \frac{v_r^{(2)}}{2v_r^{(1)}} + \frac{\lambda_r h_r^{(2)}}{2(1 - \rho_r)} + h_r^{(1)} \qquad (5)$$

where $v_r^{(1)}$ and $v_r^{(2)}$ are the first and second moments of the server vacation distribution. This expression is appropriate for a system in which the server provides gated service to the queue and departs on vacation after serving customers in queue when the server arrived. As before, we can derive t_r, $h_r^{(1)}$, and $h_r^{(2)}$ from Figure 5. The moments of the vacation distribution, which are the moments of the polling periods whose initiation times are assumed renewal points, are derived in Appendix A. The results are

$$v_r^{(1)} = \frac{1}{S} \sum_{i=1}^{f_{\max}} p_0^{i-1}(1 - p_0) \times \left[(i - 1)(c_0 + f_{\text{neg}}) + \frac{\rho_e c_e^{(1)}}{(1 - \rho_e)(1 - p_0)} \right] + \frac{1}{S} p_0^{f_{\max}} f_{\max}(c_0 + f_{\text{neg}}) \qquad (6)$$

Table I. Baseline Configuration Parameters

Parameter	Value
Private line:	
Number of cluster controllers	5
Access line	4.8 kb/s
Propagation delay	10 ms*
Modem RTS/CTS delay	20 ms
Modem propagation delay	15 ms
NEGPOLP	0.1 s
SERVLIM	3
ACK/EOT/RVI	5 bytes
POLL/SEL	9 bytes
BSC overhead	7 bytes
ENQ message	100 bytes/block, 1 block/message
RESP message	1000 bytes/block, 1 block/message
Network:	
T1 trunks	3
T1 utilization	50%
Switch processing	5 ms/switch†
TA processing	2.5 ms/TA
LAPD frame	47 byte total (maximum)
LAPD overhead	7 bytes

* This represents a distance of approximately 1000 miles.
† This number is a conservatively high estimate for frame relay switching.

$$v_r^{(2)} = \frac{1}{S}\sum_{i=1}^{f_{\max}} p_0^{i-1}(1-p_0) \times \left[(i-1)(c_0 + f_{\text{neg}}) + \frac{\rho_e c_e^{(1)}}{(1-\rho_e)(1-p_0)} \right]^2 + \frac{1}{S} p_0^{f_{\max}} \left[f_{\max}(c_0 + f_{\text{neg}}) \right]^2 \quad (7)$$

This completes the analysis of the private-line BSC model.

The models of the VCDN virtual private-line replacements are similar to the private-line model outlined above. A simple network model consisting of M/M/1 queues for trunks and mean delays for TAs and switches is used. In remote polling, service and changeover times increase because of delays within the network and at its edge. In edge polling, the model is simply that of two independent, private access lines connected by a network transport portion. Again, the requisite moments are derived from timing diagrams similar to Figure 5, appropriate to the specific VCDN configuration and TA functionality.

Results

Each configuration will be discussed separately, starting with those configurations whose installation is transparent to the end systems and continuing with configurations that need modifications in the end system, such as changes in access line speeds and FEP SYSGEN parameters. All performance results presented below are compared to the performance of the typical private-line configuration in Figure 1. The configuration contains five cluster controllers multidropped onto a 4.8-kilobit-per-second (kb/s) analog line. The analog line is assumed to be full-duplex. Reference 13 provides tuning guidelines for the FEP system parameters from which the values for SERVLIM and NEGPOLP are taken. The remaining parameters—for example, modem timing elements, block sizes, and protocol overhead—are in Table I. When customer-premises-based TAs are used, we assume they also multiplex other traffic onto the network access line, resulting in a 50 percent load on this circuit. Both TAs and frame relay switches store and forward LAPD packets.

We give results corresponding to configurations 0 through 5 in Figures 1 through 3. Configuration 0 shows the customer's present mode of operation: private-line network circuits. Configuration 1 shows the simplest VCDN virtual private-line configuration that requires no changes in the customer premises equipment. For this transparent VPL replacement the TAs may perform remote or edge-of-the-network polling.

Configurations 2 and 3 show the TA_F on premises and an increase in the FEP access speed, requiring some changes to the customer's equipment. In addition, configuration 3 shows that logically, M tail circuits can be multiplexed

Table II. Comparison of Simulation and Analytical Results for Private Line

Load (%)	Transactions per cluster controller per hour	Analysis	Simulation	Error (%)
20	75	3.18	3.16	0.6
40	150	3.97	3.94	0.8
60	225	5.29	5.48	3.5
80	300	8.26	11.33	27

onto a single FEP port. To the FEP, the M circuits appear as a single multipoint circuit with $5 \times M$ cluster controllers and require some modifications to the FEP's polling and selecting service tables.

Configuration 4 is similar to configuration 3 except that the TA_Cs are placed on premises. Here, as in configuration 3, the TAs may perform either remote or local polling; however, for consolidating tail circuits, only local polling gives good performance (see below).

Finally, configuration 5 shows switched access through the local exchange carrier into VCDN with the TA_F on premises. This configuration allows use of inexpensive voice-data modems, which cannot be utilized in other configurations because of their distance limitations. Here, the cluster controller access speed can be increased to 19.2 kb/s, the maximum transmission rate of the voice-data modems. As we will now discuss, the customer can improve performance by choosing configurations that are less transparent (for example, configuration 3 versus configuration 1).

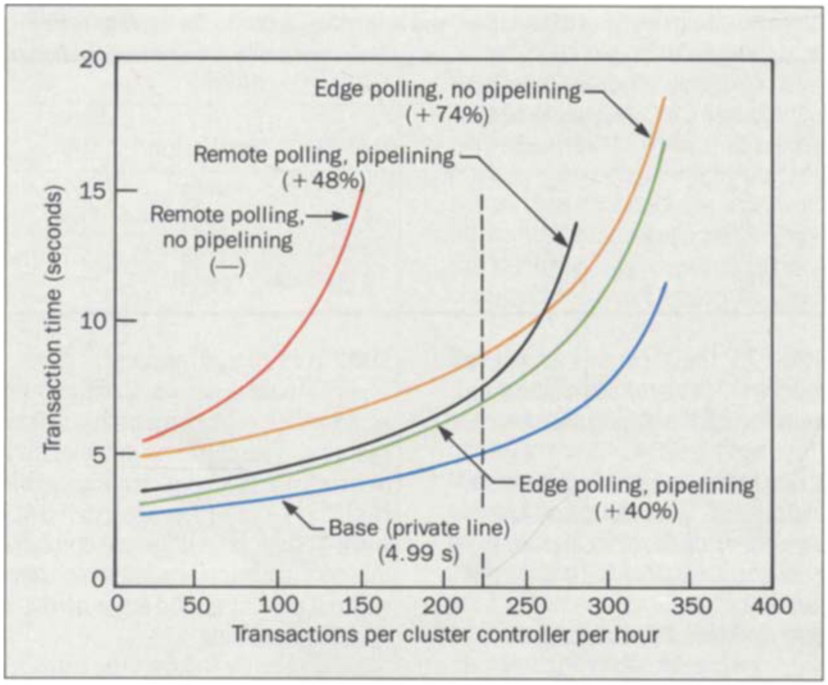
Private-Line Analysis and Simulation. A simulation of the private-line model was written in SLAM™ language (trademark of Pritsker and Associates, Inc.) to serve as a check on the analysis.[14] For loads of 60 percent or less, the analysis agrees well with the simulation results. However, at high loads (greater than 80 percent) delays can be 30 percent higher than the analysis predicts—because, we speculate, the independence assumptions in the analysis break down. However, for both the analytical model and simulation results, delay-load curves are very steep at this high loading, tending to exaggerate the differences when expressed in terms of percent relative error. Table II presents the results of simulation and analysis of the private-line configuration for system loads from 20 to 80 percent.

Configuration 1. Configuration 1 (see Figure 1) represents a VPL replacement that is transparent to the end systems. The cluster controller and FEP access lines to the network are balanced, having identical access line speeds. Both the TA_C and TA_F are located at service nodes at the network's edge. Four different combinations of TA functions for this configuration are discussed here: remote polling with and without pipelining and edge-of-the-network polling with and without pipelining.

Figure 6 shows the transaction times of these TA functionalities as a function of loading compared to the private-line transaction times. The vertical dashed line indicates 60 percent utilization of the private line. Indicated in the legend in parenthesis is the transaction time performance for each case at 60 percent utilization relative to the performance of the private line. For example, +10 percent indicates that the mean transaction time is 10 percent greater than that of a private line, and −10 percent indicates 10 percent less than a private line. The dash for the remote polling without pipelining indicates that this combination is unstable—that is, not capable of supporting the 60 percent utilization throughput.

Without pipelining, the end-to-end transport time of an ENQ or RESP message is essentially doubled because of the double insertion time: once at the transmitting device and once on the network edge at egress. Remote polling with pipelining gives transaction times of +48 percent relative to the private line, while edge-of-the-network polling without and with pipelining gives transaction times of +74 percent and +40 percent, respectively. With remote polling, pipelining is a necessity. With edge-of-the-network polling, pipelining yields benefits as well. The performance of remote

Figure 6. Transaction time results for configuration 1 in Figure 1.



polling and edge-of-the-network polling, both with pipelining, is not significantly different at and below 60 percent utilization, yielding times nearly 40 percent greater than the private line.

Configuration 2. Configuration 2 in Figure 2 features unbalanced access lines with the TA_F located on customer premises. The FEP access line to the TA_F is at 19.2 kb/s and the TA_F line into the VCDN network is a 56-kb/s digital data service (DDS) circuit. The cluster controller multidrop line, or tail circuit, remains as in the private-line case. The TA_F on customer premises can act as a concentrator for other traffic from the host data center. This traffic is multiplexed onto the 56-kb/s DDS circuit. Given that the TA_F and the FEP are collocated, it is a simple matter to increase the line speed from TA_F to FEP to 19.2 kb/s. This requires no hardware change to the FEP, such as may be required for a 56-kb/s FEP port.

Table III summarizes the performance results for configuration 2, which supports the four basic combinations of TA functions: remote and edge-of-the-network polling without and with pipelining. The performance (at 60 percent utilization) of the remote polling functionality relative to the private-line performance is +102 percent and +37 percent without and with pipelining, respectively. The transaction time performance for edge-of-the-network polling without and with pipelining is +2 percent and −6 percent, respectively. Thus, this configuration allows for some multiplexing on the 56-kb/s TA_F-to-VCDN line while maintaining good performance for edge-of-the-network functionality. In summary, remote polling is insensitive to the increased FEP speed, while local polling shows a marked improvement in performance.

Configuration 3. Configuration 3 in Figure 2 builds on the previous configuration by adding consolidation of multidrop lines onto a single FEP port. Tables III and IV summarize the performance of this configuration as a function of the number of multidrop lines consolidated onto the FEP port and only for edge-of-the-network polling with pipelining. Remote polling is incapable of maintaining satisfactory performance when consolidation of tail circuits is employed. The relative performance of this configuration is good for the case when two multidrop lines (10 cluster controllers) are consolidated (+7 percent), but degrades when three (+28

Table III. Comparative Performance of VCDN Virtual Private-Line Configurations at 60 Percent Load

Configuration	Private-line time (s)	Percentage relative to private line: Remote polls, no pipelining	Remote polls, pipelining	Edge polls, no pipelining	Edge polls, pipelining
1	4.99	—	+48	+74	+40
2	4.99	+102	+37	+2	−6
3	4.99				−6
4	4.99				−20
5	4.99				−68

Table IV. Comparative Performance of VCDN VPL Configurations at 60 Percent Load with Consolidation and Local or Edge-of-the-Network Polling

Configuration	NEGPOLP value (s)	Private-line time (s)	Percentage relative to private line: 1 tail circuit	2 tail circuits	3 tail circuits	4 tail circuits	6 tail circuits
3	0.1	4.99	−6	+7	+28	+57	
3	0	4.25	0	+2	+6	+11	
4	0	4.25	−6	−4	−3	−1	
5	0	4.25	−62		−40		−3

percent) or four (+57 percent) multidrop tail circuits (15 and 20 cluster controllers) are consolidated. This is mainly due to the assumption that the FEP system parameters remain unchanged from the private-line values. This may not be a valid assumption, and therefore Table IV also shows the performance of the comparable configurations with the NEGPOLP parameter set to zero. The performance improvement over the case where the NEGPOLP value is 0.1 second is dramatic, the worst performance being +11 percent for four consolidated multidrop tail circuits (20 cluster controllers). Throughput limitations on the 19.2-kb/s FEP access line keep the maximum consolidation to no more than four multidrop tail circuits, approximately.

Configuration 4. Configuration 4 in Figure 3 has unbalanced access lines, TA_C and TA_F on customer premises, and tail circuit consolidation. In fact, several levels of network multiplexing exist in this configuration. Explicitly, multiple tail circuits are multiplexed onto a single FEP port. Implicitly, the TA_C on customer premises can multiplex traffic from other applications at the same location onto the 9.6-kb/s network access line.

Tables III and IV present results for various degrees of consolidation. Only results for local polling and pipelining are given. For all cases, the performance of the VCDN replacements is better than that of the private line. Placing the TA_C on premises greatly reduces the polling cycle time. This is because the following delays are eliminated:

- Modem propagation delays
- Modem RTS/CTS delays
- Physical propagation delays in polling between the TA_C and the cluster controllers.

Configuration 5. Configuration 5 in Figure 3 has unbalanced access lines; the cluster controller access line speed is 19.2 kb/s, and the FEP access line speed is 56 kb/s. Logically, the cluster controller multidrop line is broken up to the full fanout configuration, but physically, the lines are multiplexed at the serving central office through a hypothetical central-office-based local-area network (LAN). Because of the usually short distances to the central office, inexpensive voice-data modems operating at 19.2 kb/s can be utilized in such switched access architectures for cluster controller access. The TA_C is located at the service node. The TA_F is located on customer premises and connected by a T1 trunk to the service node. The circuit through the switched access service consists of two switches connected by a 56-kb/s DDS line and a T1 trunk to the TA_C at the service node.

Because the cluster controllers are consolidated onto a single FEP port, only edge-of-the-network polling, without and with pipelining, is supported. The performance results for this configuration are shown in Tables III and IV. The relative performance of the increasing levels of consolidation is -62, -40, and -3 percent respectively. The exceptional performance in relation to the private line is a result of the increased access line speeds on both the cluster controller and FEP sides of the network.

Conclusion

New models of 3270 BSC performance incorporate the important effects of FEP system parameters. The models are used to study VCDN virtual private-line replacements of customers' current private-line communications and demonstrate that pipelining and edge-of-the-network polling minimize transaction times for VPL replacements by introducing a high degree of parallelism into the transport network. Further improvements are obtained when terminal adapters are placed on customer premises. Edge polling also lends itself to many new networking capabilities such as network consolidation and multiplexing of tail circuits.

With development of today's high-speed frame-relay networks, the functionality at the network edge affects user-perceived performance to a far greater extent than further decreases in network switching times. This is dramatically illustrated when the end systems use binary synchronous communications as the link layer protocol. Because BSC has little inherent parallelism, the network can achieve a high degree of parallelism only through pipelining and edge-of-the-network polling.

Acknowledgments

I would like to acknowledge the many people who contributed to this work. I wish to thank G. Garner, T. Hou, H. Josephson, A. Kumar, T. Schonfeld, and D. Sheng for their insightful comments and discussions, and, in particular, D. Schmuckler for his guidance throughout the course of this work.

References

1. "General Information—Binary Synchronous Communications," *IBM Systems Reference Library*, October 1970.
2. J. E. McNamara, *Technical Aspects of Data Communications*, Digital Press, Bedford, Massachusetts, 1977.
3. D. D. Sheng, "Virtual Private-Line Performance and Customer Cost Impacts," *AT&T Technical Journal*, Vol. 67, No. 6, November/December 1988 (to be published).
4. J. C. Kaufeld, K. A. Boakye, and J. W. Palmer, "Data Networking Architecture," *AT&T Technical Journal*, Vol. 67, No. 6, November/December 1988 (to be published).
5. G. Deaton and D. Franse, "Analyzing IBM 3270 Performance Over Satellite Links," *Data Communications*, October 1980, pp. 117-132.
6. T. L. Lo and J. Peng, "Performance Analysis of Line Capacity in a Multipoint Network," *Proceedings of Computer Measurement Group International Conference*, December 1982, pp. 86-95.
7. W. Chou, "Performance Evaluation of Polled Multipoint Teleprocessing Networks," *Proceedings of Computer Networking Symposium*, December 1981, pp. 1-16.
8. W. Chou, "Analysis of Data/Computer Networks," Chapter 10, *Computer Communications*, W. Chou, editor, Prentice-Hall, Englewood Cliffs, New Jersey, 1983.
9. A. Kumar, "Performance of SNA™/SDLC over Virtual Circuits in a Data Network" (trademark of IBM), *AT&T Technical Journal*, Vol. 67, No. 5, September/October 1988, pp. 27-40.
10. O. Hashida, "Analysis of Multiqueue," *Review of the Electrical Communication Laboratories*, Vol. 20, Nos. 3-4, March-April 1972, pp. 189-199.
11. A. Kumar, private communication.
12. S. W. Fuhrmann, "Symmetric Queues Served in Cyclic Order," *Operations Research Letters*, Vol. 4, No. 3, 1985, p. 139.
13. "Tuning and Problem Analysis for NCP BSC/Start-Stop Devices," IBM Document GG24-3093-00, October 1986.
14. A. A. B. Pritsker, *Introduction to Simulation and SLAM*, Halsted Press, New York, 1986.

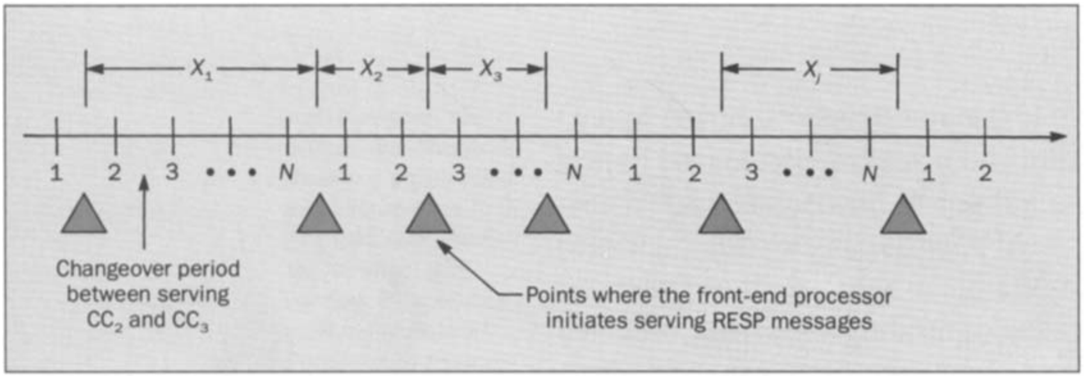


Figure A.1 Time line depicting vacation epochs as observed in the polling system.

Appendix A. Polling System Modeling

Consider the time line in the accompanying figure, where the tick marks denote service points of queues (cluster controllers), the solid triangles denote vacation points (where the FEP terminates polling and initiates selecting) and the x_is denote the distance (measured in terms of the number of cluster controllers served) between vacations. Vacations are taken when either (1) a polled cluster controller sends ENQs to the FEP [with probability $(1-p_0)$] or (2) f_{max} consecutive cluster controllers are polled without interruption. The parameter β is simply $E^{-1}(x_i)$. If we assume that the points of initiation of polling periods are renewal points, then the x_is are independent and identically distributed (iid) random variables between 1 and f_{max}. Thus, the probabilities for the values of the x_is are

$$\Pr[x^i = j] = \frac{1}{S} p_0^{j-1}(1-p_0) \qquad j = 1, 2, 3, \ldots, f_{max} - 1$$

$$\Pr[x_i = f_{max}] = \frac{1}{S} p_0^{f_{max}-1}$$

where the normalization constant is

$$S = (1 - p_0) \sum_{j=0}^{f_{max}-1} p_0^j + p_0^{f_{max}} = 1 \qquad \text{(A-1)}$$

and p_0 is the probability that a queue is empty upon the arrival of the server, given in equation (3) of the main text. The mean of x is therefore

$$E(x) = \frac{1}{S}\left[(1 - p_0) \sum_{j=0}^{f_{max}-1} (j + 1) p_0^j + f_{max} p_0^{f_{max}} \right]$$

and

$$\beta = \frac{1}{(1 - p_0) \sum_{j=0}^{f_{max}-1} (j + 1) p_0^j + f_{max} p_0^{f_{max}}} \qquad \text{(A-2)}$$

which is the expression for β in equation (4) of the main text.

We can now write expressions, as we did for β, for $v_r^{(1)}$ and $v_r^{(2)}$. The moments of the vacation distribution for the selecting system are just the moments of the sojourn time of the FEP in a polling period. For $v_r^{(1)}$ and $v_r^{(2)}$, ignoring the correlations between polling and selecting periods, we can immediately write

$$v_r^{(1)} = \frac{1}{S} \sum_{j=1}^{f_{max}} p_0^{j-1}(1 - p_0) \times \left[(j - 1)(c_0 + f_{neg}) + E(S_e \mid \text{occupied}) \right] + \frac{1}{S} p_0^{f_{max}} f_{max}(c_0 + f_{neg}) \qquad \text{(A-3)}$$

and

$$v_r^{(1)} = \frac{1}{S} \sum_{j=1}^{f_{\max}} p_0^{j-1} (1 - p_0) \times \left[(j-1)(c_0 + f_{\text{neg}}) + E(S_e \mid \text{occupied}) \right]^2 + \frac{1}{S} p_0^{f_{\max}} \left[f_{\max}(c_0 + f_{neg}) \right]^2 \qquad \text{(A-4)}$$

where the terms in brackets represent the sojourn time for the respective x_is $= 1, 2, \ldots, f_{\max}$ and $E(S_e \mid \text{occupied})$ is the server sojourn time for serving a single queue (cluster controller), given that the queue is occupied upon the arrival of the server. This is simply[12]

$$E(S_e \mid \text{occupied}) = \frac{E(S_e)}{1 - p_0} = \frac{\rho_e c_e^{(1)}}{(1 - \rho_e)(1 - p_0)} \qquad \text{(A-5)}$$

Equations (3.3) to (3.5) complete the expressions for $v_r^{(1)}$ and $v_r^{(2)}$ in equations (6) and (7) of the main text.

(Manuscript received August 26, 1988)