The exploratory system control model multi-loop network

by DANIEL J. PAULISH
Burroughs Corporation
Paoli, Pennsylvania

PURPOSE

The Exploratory Systems Control Model (ESM) Multiloop Network consists of the original three-loop ESM network delivered in 1977 and the fourth Exploratory Systems Control Model Development (ESMD) loop delivered in 1978. The ESM also includes a fifth loop supplied under the Modular System Control Development Model (MSCDM) project. The ESM provides a flexible tool for simulating and comparing a wide range of system control architectures and their related procedures and protocols. The ESM has been designed to model the class of the system control architectures that have the characteristics of decentralized operation, modularity, easy modification and upgrade capability, high reliability, high survivability and fail-soft operation.

BACKGROUND

The Defense Communications System (DCS) is a global, multiple-user system composed of leased and government-owned transmission media, relay stations and switching centers deployed in support of the National Command Authorities and the services, including command and control, intelligence and early-warning, as well as administrative and logistical communications. In order to increase the reliability and availability of these DCS services, it is essential to improve the responsiveness and robustness of the System Control (SYSCON) process as much as possible. This requirement demands a DCS SYSCON subsystem possessing such design features as modularity and "fail-soft" operation. Modularity implies a subsystem that is capable of being upgraded, modified and reconfigured easily, and "fail-soft" implies a subsystem that tolerates partial failures, yet is relatively immune to total collapse. To afford these capabilities, the future DCS SYSCON subsystem is expected to consist of many semi-autonomous, mutually-supportive, geographically-dispersed control centers.

During FY 75, Burroughs Corporation began development of an Exploratory System Control Model (ESM) which capitalized upon the inherent flexibility of multiple, interconnected data transmission rings and microprocessor-based host/ring interface nodes to provide an initial capability for experimental validation of various candidate SYSCON subsystem architectures characterized by distributed control and graceful degradation under stress. This capability to model apparently dissimilar architectures is a consequence of the universal physical connectivity provided by the ring structure coupled with flexible protocols that permit definition of different logical connectivities through selective routing of transmitted data.

In the broader context of the DCS SYSCON Program, the longer-term joint purpose of this effort and the separate-but-related "Modular System Control Architecture Study and Feasibility Development Model" is to provide DCEC with the necessary integrated means to evaluate through hybrid simulation a variety of candidate SYSCON subsystem the architecture(s) thereby identified as being suitable for implementation. The technical and performance information obtained from the unified hybrid simulation model will ultimately be used in the preparation of performance specifications for the future DCS SYSCON subsystem.

LOOP OR RING COMMUNICATIONS SYSTEMS

General operation

A communications loop is a closed, ring-connected set of nodes providing data flow unidirectionally from one node to the next. Each link between nodes is a single twisted pair of wires carrying a serial data stream in a self-clocking code. Full connectivity is achieved by associating a destination address with each packet of data carried on the loop. A node to whom a packet of data is not addressed acts simply as a "delayed repeater," having no effect on that data other than introducing some delay. The concept of a data exchange loop has been described extensively in the literature of computer communications by Reames,[1] Jafari[2] et al. Loops may be distributed such that each node contains its own power supply and cabinet and is located near the equipment it interfaces or locally such that all nodes are connected within a single cabinet with cable connections to interfaced equipment.

A functional block diagram of a communication loop node is given in Figure 1. The Loop Interface Unit (LIU) is responsible for reading data addressed to the node and writing data on the loop. The Control and Interface Processor (CIP) is a microcomputer that provides a data communications interface to the external device. The memory is used

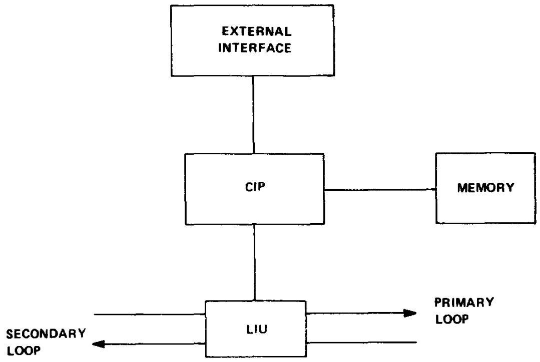


Figure 1—Communication loop node.

for program storage, routing tables and intransit queue storage. The external interface provides a hardware connection to the external equipment to be connected to the loop.

Modularity-adaptability features

The basic Burroughs loop node is a module made up of the LIU, microprocessor CIP, memory and external interface. The nodes are identical except for the external interface and the external device interface software used to handle the protocol between the microprocessor and the external device. In the ESM external devices include processors (PDP11/40, PDP11/70, B776), terminals (TD802, TD832), gateways (between the multiple rings) and data communication interfaces (SDLC, AUTODIN II, TCCF). The nodal external device interface software provides code conversion, flow control, intransit queueing, logical attachment capability and emulation of various communications protocols for the devices. The interfacing capability of the nodal modules provides communication capability between devices in the heterogeneous system.

When a module fails, the loop will recognize the failure and cut the failed module out of the system by forced loopback from the module's nearest neighbors. The module may then be replaced and the loop will return to normal operation. In the meantime, the other modules will still be in operation so that degradation will be graceful in that only the operation of the failed module will have been lost.

Loop throughput capability

Loop throughput (total number of message units that can be sent over the loop per unit time without undue message delay to the receiving modules) is a function of line speed, loop discipline and the definition of "undue" message delay. Various loop disciplines have been developed and compared.

The Newhall loop which uses a special control packet called a Write Token can transmit only one message on the loop at a time and has the lowest throughput but has the advantage of simplicity and cannot be clogged by misdirected messages. It also shows some advantage in ease of detection and deletion of certain types of faults.

The Pierce loop which uses fixed size slots in which data packets can be placed can transmit multiple messages, but the small fixed packet size causes greater overhead than in the DLCN loop.[1] The DLCN loop uses queues within each LIU that can expand or contract to hold upstream messages in temporary storage. This allows packets of variable size to be transmitted and allows multiple transmissions. Loop clogging is possible in both cases, however, and special means must be employed to "declog" the loops under certain error conditions. Loop clogging is the deadlock situation when packets cannot be written onto the loop until previously written packets are removed.

The Jafari loop[2] is a double loop, one used for control and the other for data. The data loop is segmented such that a switched point-to-point circuit is set up when requests for communication are issued on the control loop.

The ESM uses a Newhall protocol at a loop frequency of one mega-baud. Simulation studies and queueing analyses[3] indicate that this loop can support a throughput in excess of 750K baud without undue delay. The Pierce, DLCN and Jafari throughputs can be even higher, due to simultaneous conversations.

Suppose the average node writes 15 packets of 2000 bits each per second for a total of 30,000 bits/second. At that rate a Newhall loop can support 25 nodes. The worst-case time that a node will have to wait for a write token is given by

$$T_{WT} = \frac{MP}{C_L} \qquad \text{(Eq. 1)}$$

where M is the number of nodes=25, P is the packet size—2000 bits, and C_L is the loop frequency of $1M$ bits/sec. Thus T_{WT} is 50 msec. The average wait for a write token is given by

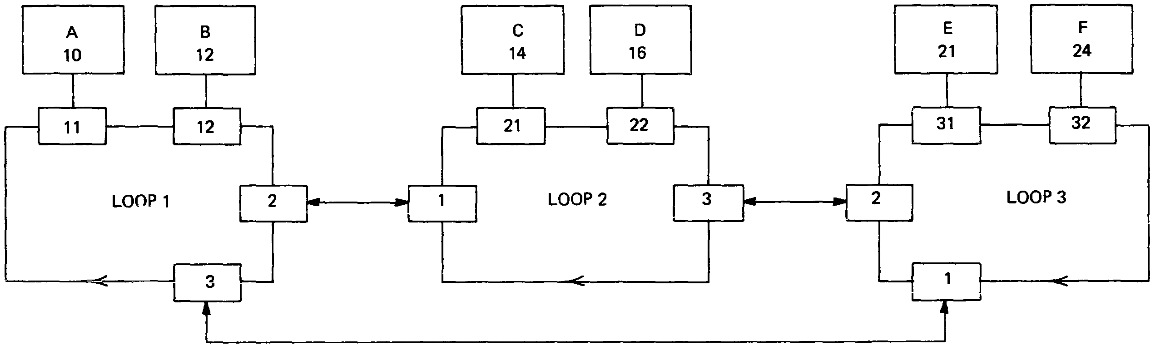
$$T = \frac{\rho}{2} T_{WT} \qquad \text{(Eq.2)}$$

where ρ is the loop utilization=0.75 for our example, thus T=19 msec.

Multiple loops—addressing schemes

The ESM system has proved the capability for providing multiple loop systems and has acted as a vehicle for testing multiple loop addressing schemes. Figure 2 exemplifies a multiple loop system. Three loops are shown connected via gateway nodes. Gateway 2 of Loop 1 connects to Gateway 1 of Loop 2 via a hard-wire connection independent of the loops. Similarly, Loop 1 connects to Loop 3 and Loop 3 connects to Loop 2 via gateways. Each loop is independent of the other loops.

The small boxes are nodes and the numbers within the boxes represent the "functional address" of the node. The functional address (FAD) is the local address unique within the loop. In addition, each node has a "logical identifier" (LID) unique within the system.



LID	FAD	ALT
10	11	–
12	12	–
14	2	3
16	2	3
21	3	2
24	3	2

LID	FAD	ALT
10	1	3
12	1	3
14	21	–
16	22	–
21	3	1
24	3	1

LID	FAD	ALT
10	1	2
12	1	2
14	2	1
16	2	1
21	31	–
24	32	–

3	M	21	10	→					→	31	M	21	10
11	A	10	21	←					←	1	A	10	21
2	AM	21	10	→	3	AM	21	10	→	31	AM	21	10
11	AA	10	21	←	1	AA	10	21	←	2	AA	10	21

Figure 2—Indirect method of addressing.

An example of how alternate routing is implemented with a multi-loop architecture using indirect addressing is given in Figure 2. Let us assume that Host Processor *A* on Loop 1 wishes to send a message to System Process 21. Host *A* sends a packet to its CIP with 21 as the destination LID and 10 as its source LID. The CIP formats a packet using an FAD or loop address equal to 3. The packet is sent out onto the loop, bypasses Nodes 12 and 2 and is read by Gateway Node 3. Gateway Node 3 sends the information part of the packet across the 1–3 link. Gateway 1 in Loop 3 formats a packet having Loop Address 31. The packet bypasses Node 2 and Node 31 reads the packet. An ACK packet is sent out on the loop using LID 10, and the packet is linked to the input queue for deliverance to Host *A*.

If Node 11 had not received an ACK message after a specified number of retransmissions, it would utilize alternate routing. It would do this by marking the packet indicating that alternate routing was used and changing the loop read address (FAD) from 3 to 2. Gateway Node 2 in Loop 1 would read the packet and send it across the 1-2 link. Gateway Node 1 in Loop 2 would use an FAD of 3. The packet would bypass Nodes 21 and 22 and be read by gateway Node 3. The packet would be sent across the 2-3 link and gateway Node 2 in Loop 3 would use an FAD of 31. The acknowledgment message would be sent via the alternate route.

Node 11 would also report to one or more network control processors who could remove the 1-3 link from service for repair. This would involve sending special broadcast control packets to Loops 1 and 3 so that Link 1-3 would not be used.

The above method of indirect addressing can be used for resource allocation such that processes could be moved around the network so that spare or less utilized processors can be utilized. For example, let us say that Host *E* is brought down for service and thus Process 21 is to be moved to another processor. Let us say that it is determined (possibly by some bid-quotation scheme) that Host *D* of Loop 2 is to handle Process 21. In order to move the process, control packets would be broadcast in each loop.

ESM/ESMD IMPLEMENTATION

System elements and connectivity

The ESM is a communications system used to interconnect devices (e.g., terminals, host processors, data communications lines) so that each device can interface with any other device for information transfer. To accomplish this, each ring is supplied with nodes that act as interfaces from ring to device and from ring to ring. The ring-to-ring nodes are called "gateway" nodes. Each node is the same physically as any other node except for a small amount of special separable hardware for each type of node. The major difference between nodes is in the software of the nodes. The nodes provide all the necessary communications functions of queueing, parity checking, ACKing, NAKing, retransmitting, alternate routing, etc. The hosts and terminals need only supply the data processing functions and need not be concerned with the communications functions.

The gateway node interchanges are via cables in the ESM configuration, but in principle can be via any communications medium such as telephone, microwave relay, optical transmission or satellite relay.

The terms "loop" and "ring" are interchangeable. Each loop is housed in a separate cabinet in this implementation, but this is not a necessity. A loop could, as easily, extend throughout a building or facility.

The ESM Multiloop Network is illustrated in Figure 3. Loops 1, 2 and 3 were delivered in 1977 as part of the ESM Contract.[4] Loop 4 was delivered as part of the ESMD Contract in 1978.[5] Loop 5 was delivered as part of the MSCDM Contract in 1979.

Features of the ESM

The ESM is designed to be transparent to the user. Regardless of the CRT used and the host on which a particular activity takes place, the activity will take place for the CRT that calls for it. When a message is transmitted from a CRT, suitable control bytes are added to the message by the CRT node and directed to the node of a nearby host. When the host receives the message, it will either handle the message completely if it can or it will pass it on to another host, via the ESM, for cooperative handling of the message. This is done under program control using the content of the CRT message and the added control bytes. The CRT will then receive a response from one of the hosts. A CRT can "ATTACH" itself to any node in the network via user command.[6]

Responses will generally be part of the "user language" which is designed to provide directions for further dialog as well as replies to previous messages. The language is designed to be modular so that it can be easily updated and

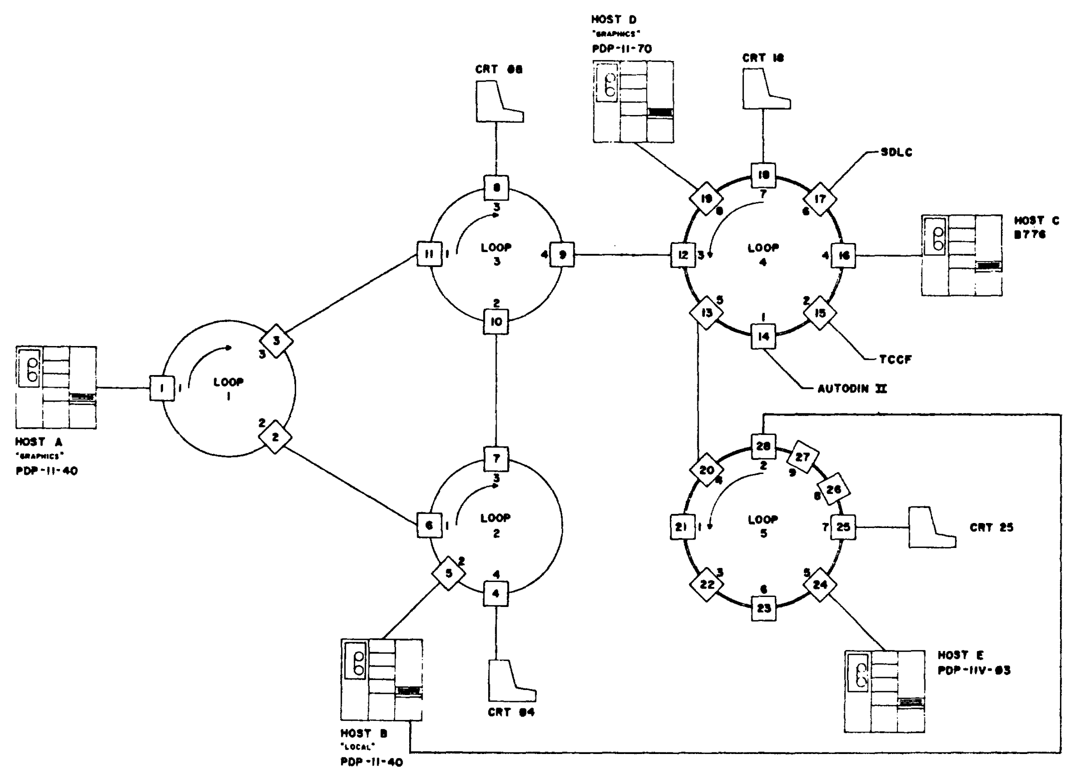


Figure 3—ESM multi-loop network.

enhanced. In addition CRTs can communicate directly with the operating system of the particular processor as if it were a local terminal.

Messages are sent in the form of packets of length not greater than 256 bytes. As each packet is sent, the sending node holds it for acknowledgment (ACK) from the receiving node. When an ACK is received by the sending node, it frees the packet space.

If a non-acknowledgment (NAK) is received, the message is resent or sent by an alternate route. Absence of an ACK or NAK after a timeout period is considered to be a NAK. After a suitable number of resends without an ACK, the message may be reported "not sent." Nodes automatically provide input and output queueing for the external device. Sufficient extra memory space is provided in each node to permit receipt of system control commands from the loop and to act on these commands. This is done to prevent a deadly embrace condition within the node. If the input queue (from the loop to the external device) is full, new input messages are rejected. Room always exists for the receipt of ACKs and NAKs and other control messages. These are acted upon with dispatch so that they do not reside in data memory for a long period. If the output queue is full, the external device is prevented from sending to the node.

The loop protocols are designed to be non-blocking and self-polling. Each node in the loop has its turn to write onto the loop and if any noise exists on the loop from prior transmissions, it is overwritten by the new transmission. Nodes share the polling activity and any loss of polling is restarted automatically.

ACKs and NAKs are generated by end-user nodes when they receive packets. Each packet is tested against cyclic redundancy bytes in the packet. A good check results in an ACK and a bad check results in a NAK.

Examples of use

The ESM Multiloop Network is part of DCA's Hybrid Simulation Facility (HSF). The ESM provides DCA with a System Control Simulation Facility. The uses of the system are to be outlined. The various applications are in different stages of development; some of the system uses are a direct result of the original implementation, others require additional modeling application and demonstration software.

User language

The User Language provides the human interface to the system and demonstrates many of its modeling capabilities. The User Language is an application program running on the various Host processors in the network. All loop connected terminals can communicate with the User Language on any processor.

The User Language consists of four major modes of operation. The first mode, CRT-to-CRT, provides users with the capability to send messages to each other. This simulates communication between System Controllers at different sites who must talk to each other in order to isolate certain faults. Mode 2, System Inquiry, allows the user to examine the nodal configuration tables; the tables are monitored on disk on a host computer. Mode 3, System Control, allows the user to modify the nodal configuration tables; this feature provides the capability to model different network architectures. Mode 4 implements a distributed data base on the PDP 11/40s. The TOTAL Data Base Management System is used to distribute records of files on the two processors. The data base appears to the user to reside completely on one machine.

File transfer

A file transfer utility has been written to transfer files between host processors. The program allows peripheral sharing by providing a capability to send files to another machine's disk, printer, or tape. Files can be obtained from another machine's disk. In addition, terminals can be ATTACHed to the various host processors in the system.

Fail-soft operation

The system is designed to tolerate partial failures. LIU failures result in automatic loop-back performed to remove the failure from the network. The failure is detected by the nodes and reported to a System Control Monitor node. Alternate routing is automatically performed when a node in another loop fails to ACK a packet after a specified number of retries. Failure to ACK a packet is reported to the Monitor node. Queue overflows are also reported to the Monitor; a queue overflow results from the failure of an external device to respond to the node.

Security

A demonstration of the use of a Security Monitor node is performed using Loop 4. A CRT-B776 conversation is monitored by a PDP 11/70 to detect an invalid password. The node connected to the PDP 11/70 is commanded such that it does a non-destructive read on packets addressed to the B776. Thus the data sent by the CRT is read by both the B776 and the PDP 11/70. The PDP 11/70 monitors the CRT-B776 conversation. When a bad password is detected, the Security node sends control packets to the nodes directly upstream and downstream from the CRT node to perform loop-around, resulting in the CRT node being removed from the network.

Network architecture

The ESM can be used to study System Control network architectures. Since the logical connectivity of the network is maintained by the modifiable nodal tables, the system can be used to model other network architectures. Network control problems such as automatic channel reconfiguration can be studied.

Response time

Since each loop frequency is independently modifiable, response-time studies can be done on the system. The loop rates are modifiable via switches on the clock generator cards. In addition for Loops 4 and 5, an external clock generator can be connected to drive the loop.

Software development

The ESM can be used as a general software development facility. Since each loop-connected terminal can ATTACH to any processor, software can be written on a variety of machines with different operating systems and different language compilers. In addition, since files may be transferred between machines via the network, duplicate copies of files can be kept on different machines. Processors without certain resources (e.g. line printers) can utilize the resources of other machines via the network.

APPLICATION TO SYSTEM CONTROL

The ESM Multiloop Network will be used as a DCS System Control Simulation Facility. DCS System Control must accomplish the following functions[7]:

- *Network control*—Transmission and switched network configuration control, which includes network and extension supervision, reconstitution, restoral and satellite configuration control.
- *Traffic control*—Control of traffic routing and traffic flow.
- *Performance assessment* of the DCS and status monitoring of the DCS resources.
- *Technical control*—Includes quality assurance and monitoring, patching, testing, coordinating, restoring and reporting functions necessary for effective technical supervision and control over trunks and circuits traversing or terminating in a facility.

ACKNOWLEDGMENT

I wish to acknowledge that this work was supported in part by the Defense Communications Agency under Contracts DCA100-75-C-0054 and DCA100-76-0081. I also wish to acknowledge the contributions of personnel associated with the project at the System Control Group at the Defense Communications Engineering Center and the Advanced Development Organization of the Burroughs Corporation—in particular, the guidance given by D. Schutzer, O. Halpeny, J. Lynch and A. Meyerhoff.

REFERENCES

1. Reames, C. C., "System Design for the Distributed Loop Computer Network," Ph.D. Dissertation, The Ohio State Univ., Columbus, Ohio, March 1976.
2. Jafari, H., "A New Loop Structure for Distributed Microcomputing Systems," Ph.D. Dissertation, Oregon State Univ., December 1977.
3. Meyerhoff, A., D. Paulish, C. Tomlinson and T. Woodward, "Loop Communications Systems," *Burroughs Corp. Tech. Rept. TR77-3,* Paoli, PA, November 1977.
4. Burroughs Corp., "ESM User Manual," Doc. 66143-1, Paoli, PA, March 1977.
5. Burroughs Corp., "ESMD User Manual," Doc. 66146, Paoli, PA, March 1978.
6. Burroughs Corp., "Final Report for the ESMD," Doc. 64297, Paoli, PA, January 1978.
7. Fluk, M., "Technical Overview of the System Control Improvement Program," *DCEC Tech. Rept. TR4-78,* Reston, VA., May 1978.