IEEE Transactions on Nuclear Science, Vol. NS-28, No. 3, June 1981

FPSS - A FAST PACKET SWITCHING SYSTEM

H. Frese, G. Hochweller, W. Krechlok, R. Schmitz
DESY
Notkestr. 85
2000 Hamburg 52, West Germany

Summary

FPSS is a local packet switching network connecting up to 256 computers or intelligent controllers. Transmission hardware is realized as a star network operating at an internal speed of 160 Mbit/s. Packet handling overhead is less than 3 usec per packet. The network software employs the virtual channel concept to provide data transmission between various sub processes on a time shared basis. The network operating program has been implemented as an extension of the SINTRAN III operating system of the NORD-100 computers used at DESY. A version for the PADAC microcomputer (TMS 9900) is under development.

Introduction

During the setup phase of the PETRA control system[1], the different control tasks (rf, vacuum, magnets etc.) were assigned to individual computers, where each task could be executed independently. Computer-computer links were of very limited importance, since only a small number of process parameters had to be accessible from different control tasks. Data exchange was accomplished by bridges between the SEDAC systems[2] of the different tasks. Further system development is dictated by the integration of other accelerator control systems (DESY, DORIS, HERA). It is marked by a trend towards employing a large number of microcomputers for independent execution of sub tasks. Since these microcomputers are not equipped with standard peripherals and need access to each other in the course of executing their control tasks, the installation of a powerful communication system is inevitable.

Hardware

Design Objectives

- star network, all users clustered within ~100 m
- end user to end user connection without intermediate 'store-and-forward' elements
- packet switching with variable packet size
- user to user transmission at DMA speeds
- internal system speed high enough to allow several 'simultaneous' user - user connections at DMA speeds
- low packet switching overhead
- centralized request polling and bus arbitration
- simple hardware protocol
- position independent addressing scheme
- 'reasonable' number of participants (8-bit address)
- utilization of existing mechanical and electrical standards
- realization in standard LS-technology, RS-422 transmission lines

Realization

In order to achieve the desired internal speeds (roughly 10 times DMA speed or 160 Mbits/s) with standard LS technology, the internal data bus had to be kept short. This was accomplished by concentrating the user packet buffers at the network center. Over the longer distances between the participating computer and its buffer storage, data is transmitted at (rather uncritical) DMA speed, which in the 100 m range can be performed within normal RS-422 standards. The significantly higher data rates within the network occur only on the lines between source buffer and destination buffer within the network crate. PADAC[3] crates, already in use at DESY in large numbers, were chosen as mechanical standard without any extra wiring. The use of the standard PADAC bus backplane as the network's internal data bus permits the use of 32 parallel data lines plus some control signals. This results in a memory cycle time of 200 ns at the desired 160 Mbit/s data rate, well within the range of currently available memory products.

Network control and bus arbitration for individual packets (maximum length 4 kB) is handled by a micro-programmed controller in the FPSS main crate. Since every user is represented by a module plugged into a FPSS crate, a crate hierarchy system was implemented to provide the necessary number of slots within a master crate and a number of sub crates. Crate hierarchy is expandable to any depth and is transparent to the user; maximum inter-crate distance is limited to ~ 10 m.

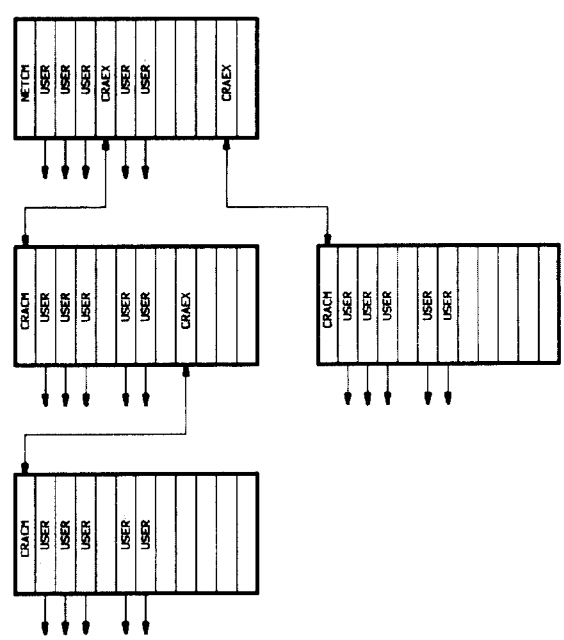

Fig. 1: FPSS Crate Configuration

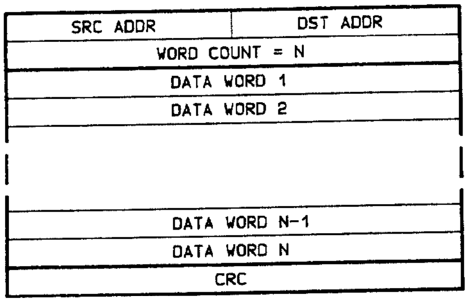

Fig. 2: Hardware Packet Format

The transmitting process has to supply destination address and word count, source address and CRC are added by hardware. This simple packet format has the advantage of packing the complete hardware routing information into less than 32 bits, which are presented to the network controller as the first 32 bit word of a packet buffer. This leads to a simple transmission protocol with a maximum packet switching overhead of 5 usec.

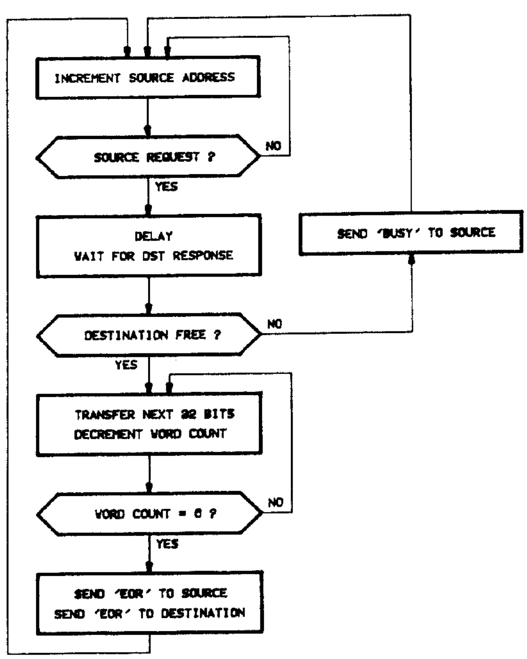

Fig. 3: FPSS Microprogram

Software

The software effort was two-fold: definition of a NETWORK OPERATION PROGRAM (NOP), which is host independent, and its implementation on NORD-100 computers under the operating system SINTRAN III. The purpose of the NOP is to allow communication among processes on different computers without undetected transmission errors. The basic transmission concept is a CONNECTION, which is a full duplex virtual packet transmission channel. A connection exists between two processes and is protected from other connections. Since none of the participating computers is distinguished as network master, the successful operation of the system depends on the proper execution of the NOP on all computers currently engaged in data transfer.

Using a Connection

Within a participating NOP, a connection is represented by a CONNECTION CONTROL BLOCK (CB). To establish a connection, it is necessary for the two communicating processes to obtain a CB and to contact each other in an asymmetric way: One process announces to its local NOP a contact name it will listen to, while the other process issues a REQUEST FOR CONNECTION (RFC) to the listener's NOP with the same contact name. Once a match is found, neccessary information is exchanged between the participating NOPs and the connection is declared OPEN. Now the contact name is no longer neccessary and can be reused in another LISTEN/RFC match. One process can open more than one connection.

A connection can be used for three levels of data transmission:

- byte transmission (high message rate)
- packet transmission
- buffer transmission (max 64 kB, highest data rate)

Packets can be flagged as CONTROL PACKETS to facilitate the implementation of user protocols.

Each NOP contains a watchdog process, which tests all open connections at regular intervals to detect inconsistencies. A connection is terminated if

- one of the processes has issued a CLOSE request
- a transmission error has occured
- one of the processes has been aborted
- the remote NOP doesn't reply within a certain time.

An attempt is made to inform both processes about connection termination.

A window mechanism has been utilized to implement flow control. The source window limits the number of filled but yet untransmitted packets on the transmitter side, while the destination window limits the number of packets received but not yet emptied on the receiver side. This scheme prevents one connection from consuming all packets and enables a slow receiving process to control the data rate on the transmitter side.

Software design permits packet transmission between users on different local FPSS networks through BRIDGES.

Software Implementation

A NOP for the NORD-10/NORD-100 has been written in MAC assembly language and NORD-PL. NOP functions are implemented as new monitor calls, while existing monitor calls have been upgraded to allow their use in communication services.

The resulting code (excluding buffer space) occupies 9 kB of real storage.

NOP implementation on the PADAC TMS-9900 microcomputers is under way.

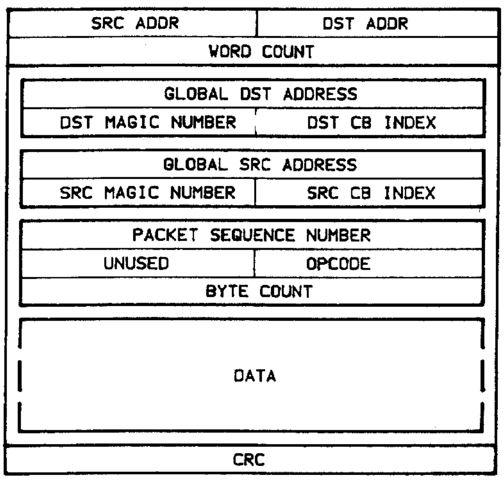

Fig. 4: Software Packet Format

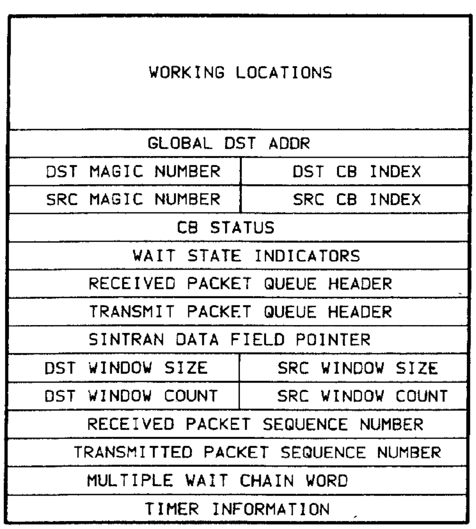

Fig. 5: Connection Block Format

Performanence Measurements

Measured packet and data rates strongly reflect operating system and buffer copy overhead times. Throughput dependance on packet size and buffer length has been measured using a connection between a NORD-10 and a NORD-100. Maximum rates are:
- 160 kB/sec process to process throughput using 4 kB packets.
- 190 control messages / second

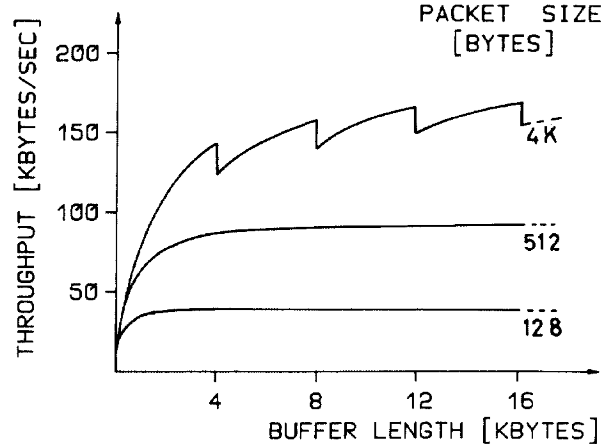

Fig. 6: Throughput as a Function of Buffer Length and Packet Size

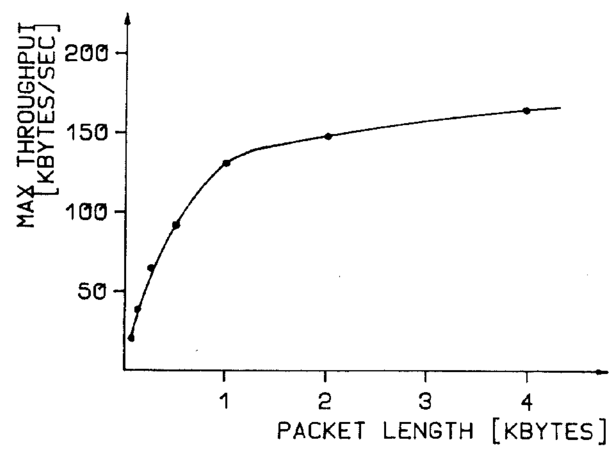

Fig. 7: Max. Throughput as a Function of Packet Size

References

1) H. Frese et al., First Experience with the PETRA Control System, IEEE Nucl. Sci. 26 (1979)
2) H. Frese, G. Hochweller; The Serial Data Acquisition System at PETRA, IEEE Nucl. Sci. 26 (1979), p. 3384
3) G. Hochweller, H. Frese; Tools for Man Machine Communication in the PETRA Control System, IEEE Nucl. Sci. 26 (1979), p. 3382