

1 Transputers and Routers: Components for Concurrent Machines

1.1 Introduction

This chapter describes an architecture for concurrent machines constructed from two types of component: ‘transputers’ and ‘routers’. In subsequent chapters we consider the details of these two components, and show the architecture can be adapted to include other types of component.

A transputer is a complete microcomputer integrated in a single VLSI chip. Each transputer has a number of communication links, allowing transputers to be interconnected to form concurrent processing systems. The transputer instruction set contains instructions to send and receive messages through these links, minimizing delays in inter-transputer communication. Transputers can be directly connected to form specialised networks, or can be interconnected via routing chips. Routing chips are VLSI building blocks for interconnection networks: they can support system-wide message routing at high throughput and low delay.

2 The T9000 Communications Architecture

2.1 Introduction

This chapter describes the communications capabilities implemented in the IMS T9000 transputer, and supported by the IMS C104 packet router, which is discussed in chapter 14. The T9000 retains the point-to-point synchronised message passing model implemented in first generation of transputers but extends it in two significant ways. The most important innovation of the T9000 is the virtualization of external communication. This allows any number of *virtual* links to be established over a single hardware link between two directly connected T9000s, and for virtual links to be established between T9000s connected by a routing network constructed from C104 routers. A second important innovation is the introduction of a many-one communication mechanism, the resource. This provides, amongst other things, an efficient distributed implementation of servers.

3 DS-Links and C104 Routers

3.1 Introduction

Millions of serial communication links have been shipped as an integral part of the transputer family of microprocessor devices. This 'OS-Link', as it is known, provides a physical point-to-point connection between two processes running in separate processors. It is full-duplex, and has an exceptionally low implementation cost and an excellent record for reliability. Indeed, the OS-Link has been used in almost all sectors of the computer, telecommunications and electronics markets. Many of these links have been used without transputers, or with a transputer simply serving as an intelligent DMA controller. However, they are now a mature technology, and by today's standards their speed of 20 Mbits/s is relatively low.

Since the introduction of the OS-Link, a new type of serial interconnect has evolved, known as the DS-Link. A major feature of the DS-Link is that it provides a physical connection over which any number of software (or 'virtual') channels may be multiplexed; these can either be between two directly connected devices, or can be between any number of different devices, if the links are connected via (packet) routing switches. Other features include detection and location of the most likely errors, and a transmission speed of 100 Mbits/s, with 200 Mbits/s planned and further enhancement possible.

Although DS-Links have been designed for processor to processor communication, they are equally appropriate for processor to memory communication and specialized applications such as disk drives, disk arrays, or communication systems.

4 Connecting DS-Links

4.1 Introduction

Digital design engineers are accustomed to signals that behave as ones and zeros, although they have to be careful about dissipation and ground inductance, which become increasingly important as speeds increase. Communications engineers, on the other hand, are accustomed to disappearing signals. They design modems that send 19200 bits per second down telephone wires that were designed 90 years ago to carry 3.4KHz voice signals. Their signals go thousands of kilometers. They are used to multiplexing lots of slow signals down a single fast channel. They use repeaters, powered by the signal wires.

Digital designers do not need all these communications techniques yet. But sending 100Mbits/s or more down a cable much longer than a meter has implications that are more analog than digital, which must be taken care of just like the dissipation and ground inductance problems, to ensure that signals still behave as ones and zeros.

Actually, it is easy to overestimate the problems of these signal speeds. Engineers designing with ECL, even fifteen years ago, had to deal with some of the problems of transmitting such signals reliably, at least through printed circuit boards (PCBs), backplanes, and short cables. One of the best books on the subject is the Motorola 'MECL System Design Handbook' [1] by William R Blood, Jr., which explains about transmission lines in PCBs and cables. This shows waveforms of a 50MHz signal at the end of 50ft (15m) of twisted pair, and of a 350MHz signal at the end of 10ft (3m) of twisted pair, both with respectable signals.

This chapter first discusses the signal properties of DS-Links. PCB and cable connections are then described, followed by a section on error rates: errors are much less frequent on transputer links than is normal in communications. A longer section introduces some of the characteristics of optical connections including optical fibre, which should be suitable for link connections up to 500m, using an interface chip to convert between the link and the fibre. A pointer is given towards possible standards for link connections. Appendix A describes a connector that will assist standardization of transputer link connections. Appendix B shows waveforms of signals transmitted through cable and fibre. Appendix C gives detailed electrical parameters of DS-Links, and appendix D gives an equivalent circuit for the DS-Link output pads.

5 Using Links for System Control

5.1 Introduction

The T9000 family of devices includes processors and routers which have subsystems and interfaces which are highly flexible to match the requirements of a wide range of applications. In addition to the static configuration requirements of subsystems such as the memory interface of the T9000, the more dynamic aspects of a network of devices must be configured before application software is loaded. These more dynamic items include:

- cache organization;
- data link bit-rates;
- virtual link control blocks;

If T9000 processors are configured as stand-alone devices, the configurable subsystems will be initialized by instructions contained in a local ROM. When the devices are integrated as part of a network with a static configuration every processor in the network could also initialize these subsystems independently by executing code contained in a local ROM. Typically, however, networks of T9000 family devices contain routers as well as processors and executing code from a ROM is not an option for a routing device. As a consequence, routing devices must be configured under external control. During system development or for systems which are used for multiple applications a flexible configuration mechanism for processors is also required.

Debugging of software and hardware on networks consisting of many devices is not a simple problem. The major difficulty is in monitoring the behavior of the system as an integrated whole rather than observing the individual behavior of the separate components. A flexible mechanism which allows monitoring tools to observe and manage every device in a network in a simple manner is essential in designing a system-wide debugging environment.

6 Models of DS–Link Performance

This chapter contains analytic studies of the performance of DS-Links, the IMS T9000 virtual channel processor and the IMS C104 packet routing switch.

The first section considers the overheads imposed by the various layers of the DS–Link protocol on the raw bit–rate. Results are presented for the limiting bandwidth as a function of message size, which show that the overheads are very moderate for all but the smallest messages (for which the cost of initiating and receiving a message will dominate in any case).

The next section analyses the diminution of bandwidth caused by latency at both the token flow–control and packet–acknowledge layers of the protocol. The losses due to stalls at the packet level of the protocol when only a single virtual channel is active are plotted in the latter part of the section.

The final section considers the performance of the C104 routing switch under heavy load, both in the average and the worst case.

7 Performance of C104 Networks

The use of VLSI technology for specialised routing chips makes the construction of high-bandwidth, low-latency networks possible. One such chip is the IMS C104 packet routing chip, described in chapter 14. This can be used to build a variety of communication networks.

In this chapter, interconnection networks are characterized by their throughput and delay. Three families of topology are investigated, and the throughput and delay are examined as the size of the network varies. Using deterministic routing (in which the same route is always used between source and destination), random traffic patterns and systematic traffic patterns are investigated on each of the networks. The results show that on each of the families examined, there is a systematic traffic pattern which severely affects the throughput of the network, and that this degradation is more severe for the larger networks. The use of universal routing, where an amount of random behavior is introduced, overcomes this problem and provides the scalability inherent to the network structure. This is also shown to be an efficient use of the available network links.

An important factor in network performance is the predictability of the time it will take a packet to reach its destination. Deterministic routing is shown to give widely varying packet completion times with variation of the traffic pattern in the network. Universal routing is shown to remove this effect, with the time taken for a packet to reach its destination being stabilized.

In the following investigation, we have separated issues of protocol overhead, such as flow control, from issues of network performance.

8 General Purpose Parallel Computers

8.1 Introduction

Over the last decade, many different parallel computers have been developed, which have been used in a wide range of applications. Increasing levels of component integration, coupled with difficulties in further increasing clock speed of sequential machines, make parallel processing technically attractive. By the late 1990s, chips with 10^8 transistors will be in use, but design and production will continue to be most effective when applied to volume manufacture. A “universal” parallel architecture would allow cheap, standard multiprocessors to become pervasive, in much the same way that the von Neumann architecture has allowed standard uniprocessors to take over from specialised electronics in many application areas.

Scalable performance

One of the major challenges for universal parallel architecture is to allow performance to scale with the number of processors. There are obvious limits to scalability:

- For a given problem size, there will be a limit to the number of processors which can be used efficiently. However, we would expect it to be easy to increase the problem size to exploit more processors.
- There will in practice be technological limits to the number of processors used. These will include physical size, power consumption, thermal density and reliability. However, as we expect performance/chip to achieve 100-1000 Mflops during the 1990s, the most significant markets will be served by machines with up to 100 processors.

Software portability

Another major challenge for a universal parallel architecture is to eliminate the need to design algorithms to match the details of specific machines. Algorithms must be based on features common to a large number of machines, and which can be expected to remain common to many machines as technology evolves. Both programmer and computer designer have much to gain from identifying the essential features of a universal parallel architecture:

- the programmer because his programs will work on a variety of machines - and will continue to work on future machines.
- the computer designer because he will be able to introduce new designs which make best use of technology to increase performance of the software already in use.

9 The Implementation of Large Parallel Database Machines on T9000 and C104 Networks

The design of large database machines requires the resulting implementation be scalable and cheap. This means that use has to be made of commodity items whenever possible. The design also has to ensure that scalability is incorporated into the machine from its inception rather than as an after-thought. Scalability manifests itself in two different ways. First, the initial size of a system when it is installed should be determined by the performance and size requirements of the desired application at that time. Secondly, the system should be scalable as processing requirements change during the life-time of the system. The T9000 and C104 provide a means of designing a large parallel database machine which can be constructed from commodity components in a manner that permits easy scalability.

10 A Generic Architecture for ATM Systems

10.1 Introduction

The rapid growth in the use of personal computers and high-performance workstations over the last ten years has fueled an enormous expansion in the data communications market. The desire to connect computers together to share information, common databases and applications led to the development of Local Area Networks and the emergence of *distributed computing*. At the same time, the geographical limitations of LANs and the desire to provide corporate-wide networks stimulated the development towards faster, more reliable telecommunications networks for LAN interconnection, with the need to support data as well as traditional voice traffic. The resulting increase in the use of digital technology and complex protocols has resulted in the need for enormous computing capability within the telecommunications network itself, with the consequent emergence of the concept of the *Intelligent Network*. With new, higher bandwidth applications such as video and multimedia on the horizon and user pressure for better, more seamless connection between computer networks, this convergence of computing and communications systems looks set to accelerate during the nineties.

A key step in this convergence is the development by the CCITT of standards for the *Broadband Integrated Services Digital Network (B-ISDN)*. B-ISDN seeks to provide a common infrastructure on which a wide variety of voice, data and video services can be provided, thereby eliminating (hopefully) the final barriers between the world of computer networks and the world of telecommunications. The technological basis for B-ISDN chosen by the CCITT is the *Asynchronous Transfer Mode (ATM)*, a fast-packet switching technique using small, self-routing packets called *cells*.

The single most important element which has driven the development of both distributed computing and the intelligent network is the microprocessor. Indeed, as systems such as telecommunications networks have come to look more like distributed computers, so microprocessor architectures which support distributed multi-processing have come to look like communications networks. A message-passing computer architecture, such as that of the transputer, shares much in common with a packet switching system and thus provides a natural architecture from which to build communication systems. The communications architecture of the latest generation transputer, the T9000, shares much in common with ATM and is thus a natural choice for the implementation of ATM systems.

In this Chapter we describe the application of the transputer, in particular the serial links and packet routing capabilities of the communications architecture, to the design of ATM switching systems. We discuss their use in public switching systems and present a generic architecture for the implementation of private ATM switches and internetworking applications. We look at terminal adaption requirements and develop some ideas for interfacing transputers, routers and serial links to ATM networks. Finally, we consider various aspects of the performance of this architecture.

11 An Enabling Infrastructure for a Distributed Multimedia Industry

11.1 Introduction

Advances in technology for telecommunication and new methods for handling media such as voice and video have made possible the creation of a new type of information system. Information systems have become an essential part of the modern world and they need to be made accessible to a very high proportion of the working population. It is therefore important to exploit all the means available for making the transfer of information effective and accurate. In fields such as computer assisted training, multimedia presentation is already well established as a tool for conveying complex ideas. So far, however, the application of multimedia solutions to information retrieval has been limited to single isolated systems, because the bulk of the information required has needed specialized storage techniques and has exceeded the capacity of present day network infrastructure. There do exist special purpose multimedia communication systems, such as those used for video-conferencing, but their cost and complexity separates them from the common mass of computing support.

If, however, *distributed* multimedia systems can be realized, many possibilities for enhanced communication and more effective access to information exist. The key to this new generation of information systems is *integration*, bringing the power of multimedia display to the users in their normal working environment and effectively breaking down many of the barriers implicit in geographical distribution. Now that significant computing power is available on the desktop, integration of voice and video is the next major step forward.

These integrated systems represent a very large market for components and for integrating expertise. It will probably be the largest single growth area for new IT applications over the next ten years. A coordinated set of components, conforming to a common architectural model with agreed interface standards, is required to allow the research and development of prototypes for new applications and to progress smoothly to the delivery of complete multimedia distributed systems. T9000 transputers, DS-Links and C104 routers provide a cost-effective platform on which this infrastructure can be built.