NEAR EAST

1988

UNIVERSITY

Faculty of Engineering

Department of Computer Engineering

COM 400 Graduation Project

RADIO PACKET SWITCHING NETWORK

Submitted to : Prof. Dr. Fakhreddin Mamedov

Submit by : Gürcan Yılmaz 940454

Tülay Avcı 940867

CONTENTS

CHAPTER 1

CHAPTER VIII

CHAPTER I
RADIO LAN TECHNOLOGY

Within the past year or so a number of manufacturers have begun to offer local area networks based on very-low-power radio communication at speeds of 1 Mb s and aboved. Radio is one way to achieve "wireless" communication. The other common method uses infrared optical broadcast.

The task of a radio LAN is the same as that of any LAN - to provide peer-to-peer communication in a local area. Ideally, it should appear to the user to be exactly the same as a wired LAN in all respects (including performance). The radío medium is different in many ways to wired media and the differences give rise to unique problems and solutions. This section will concentrate on the aspecgts unique to the radio medium and will discuss only ìn passing aspects that are held in common with wired media.

1.1 Characteristics of the Indoor Radio Medium

1.1.1 Common Medium

There is only one broadcast space. That is in principle a radio signal transmitted any where in the world on a given frequency could be received anywhere else in the world (of course depending on propagation and signal strength). In practice, the strength of a radio signal decreases as the square of the distance from the transmitter (in some systems the decrease is with the fourth power of the distance!). It is this that enables us to re-use the same frequencies when transmitters are far enough apart. Contrast this with the wired environment where each pair of wires is a separate "signal space" with minimal interference.

Thus in a radio system every one shares essentially the same space and this brings about the biggest problem sharing the limited available bandwidth.

1.1.2 Multi-Path Effects

At the extremely high frequencies involved, radio waves reflect off solid objects and this means that there are many possible paths for a signal to take from transmitter to receiver. In this case both transmitter and receiver are in the same room. Part of the signal will take the obvious direct path but there are many other paths and someof the signal will follow each of these.(Reflection from the floor is especially significant.)

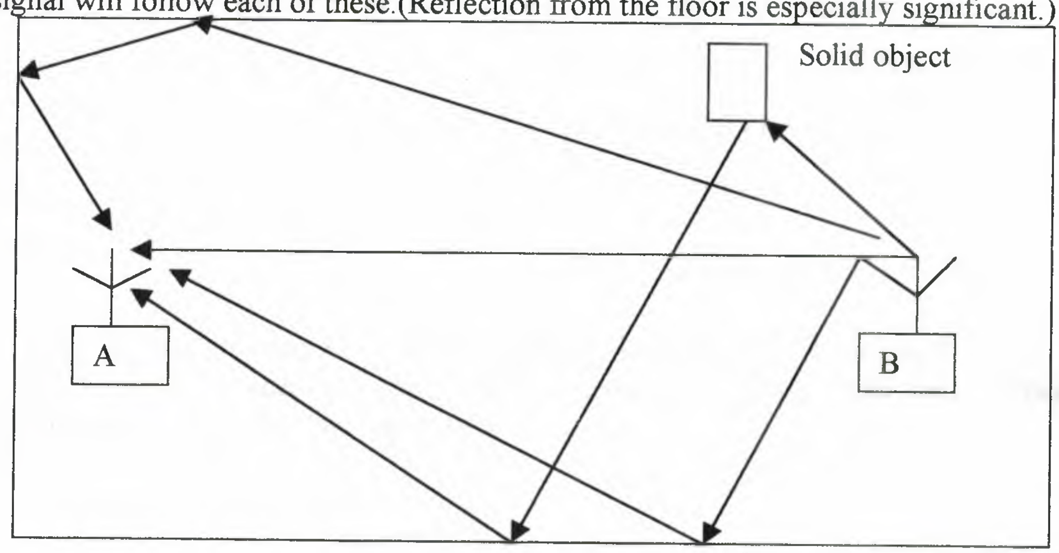


Figure 1.1.2 Multi-Path Effect. The signal travels from transmitter to receiver on multiple paths and is reflected from room walls and solid objects.

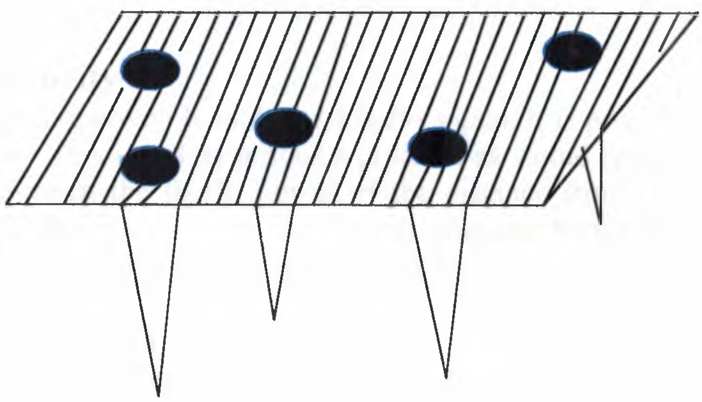
This has a number of consequence.

1. To some extent the signal will travel arround obstacles (and through soft ones). This is what gives radio its biggest advantage over infrared tranmission in the indoor environment.
2. As shown in figure a signal arriving on many paths will spread out in time (because some paths are shorter than others). More accurately, many copies of the signal will arrive at the receiver slightly shifted in time .

In the office and factory environments, studies have shown that the delay spread is typically from 30 ns to 250ns, of course depending on the geometry of the area in the question.(in the outdoor, suburban environment delay spread is typically between 5 microsec and 3 microsec). Delay spread has 2 quite different effects which must be countered.

Rayleigh Fading:

When two signal components arrive after travelling different distances they add together in the receiver. If the difference in the length of the paths they traveled is an odd multiple of half the wavelength of the carrier signal, then they will cancel one another out (if it is an even multiple they will strengthen one another). At 2.4 Gbps the wavelength is 12.5 cm.



DEEP RADIO FADES

The signal strength pattern in an indoor area can look like this.The strength can be relatively uniform except for small areas where the signal strength can fall to perhaps 30 dB below areas even one meter away.

In a room there can be dozens or even hundreds of possible paths and all the signals will add in quite complex ways.
The result is that in any room there will be places where little or no signal is detectable and other places, a few meters away, where the signal could be very strong. If the receiver is mobile, rapid variations in signal strength are usually observed.

Inter-Symbol Interference

When we are digitally modulating a carrier another important consideration is the length of the symbol (the transmission state representing a bit or group of bits). If we are sending one bit per symbol and the bit rate is 1 Mbps then the "length" of a bit will be a bit less than 300 meters. In time, at 1 Mbps a bit is 1 microsec long. (If the delay spread is 250 ns then each bit will be spread out to a length of 1.25 microsec and will overlap with the following bit by a quarter of its length.

This is called Inter-Symbol Interference (ISI) and has the effect of limiting the maximum data rate possible. ISI is present in most communications channels and there are good techniques for combating it (such as Adaptive Equalization). It is most severe in the radio environment.

Most people are familiar with this effect since it is the cause of "ghosts" in television reception - especially with indoor antennas.

3. When people move about the room the characteristics of the room (as far as propagation is concerned) change.

Overcoming multi-path effects is the most significant challenge in the design of in radio systems.

1.1.3 Intermittent Operation

In an office or factory environment people move about the area and oceasionally most large objects about. This can cause intermittent interruption to the signal, rapid and the like.

1.1.4 Security

Because there are no bounds for a radio signal, it is possible for unauthorized people receive it. This is not as serious a problem as would appear since the signal strength decreases with the fourth power of the distance from the transmitter (for systems which the antenna is close to the ground - such as indoor systems). Nevertheless it is a problem which must be addressed by any radio LAN proposal.

1.1.5 Bandwidth

Radio waves at frequencies above a few GHz do not bend much in the atmosphere(the travel in straight lines) and are reflected from most solid objects. Thus radio at this frequency will not normally penetrate a building even if it is present in the outdoor environment. Inside the building this means that there is a wide frequency space available,which could be used for local applications with very little restriction.

1.1.6 Direction

In general radio waves will radiate from a transmitting antenna in all directions. By smart antenna design it is possible to direct the signal into specific directions or even in to beams. In the indoor environment, however, this doesn't make a lot of difference because of the reflections at the wavelengths used.

1.1.7 Polarization

Radio signals are naturally polarized and in free space will maintain their polarization over long distances. However, polarization changes when a signal is reflected and effects that flow from this must be taken into consideration in the design of any indoor radio system.

1.1.8 Interference

Depending on which frequency band is in use there are many sources of possible interference with the signal. Some of these are from other transmitters in the same band (such as rather sets and microwave installations nearby). Electric motors, switches, and stray radiation from electronic devices are other sources of interference.

1.1.9 Characteristics of ISM Frequency Bands

The "ISM" (Industrial, Scientific and Medical) bands were allocated for indoor radio applications. Spread spectrum techniques must be used in these bands, but if transmitter power is very low (less than 1 watt), equipment does not need to be licensed in most countries. Note that there is some variation between countries on the boundaries of these bands.

Table 1.1.9. Indoor Radio Frequencv Band Characteristics

	915 MHz	2.4 GHz	5.8 GHz	18 GHz
Frequency	902-928 MHz	2.4-2.48 GHz	5.73-5.85GHz	18 GHz
Wavelength	32.8 em	12.5 cm	5.2 cm	1.6 cm
Width of Band	26 MHz	80MHz	120 MHz	
Usage	ISM-SS	ISM-SS	ISM-SS	Narrowband
Range	Greatest	%95	80%	
Status	Crowded	Low Use	V Low Use	V Low Use
Interference	High	Low	Low	Low

The 18 GHz band is for narrowband Microwave applications and is not wide enough for spread spectrum techniques. Nevertheless, one major radio LAN system on the market uses this band.

1.2 Sharing the Bandwidth

With many workstations in the sarne area wanting to communicate a method is needed to share the bandwidth. Different LAN designs use quite different methods of operation and of bandwidth sharing. However, most use a combination of the methods outlined below:

Frequency Division Multiplexing (FDM)
Time Division Multiplexing(TDM)

Polarization Division Multiplexing (PDM)

Provided that polarization can be maintained potentially we could use the direction of polarization as a multiplexing technique. In the presence of multiple reflections, however polarization changes and is essentially unpredictable. Thus polarization is not usable as a multiplexing technique in the indoor radio environment. (In the outdoor environment polarization is widely used as a way of doubling capacity on high speed digital microwave links).

Space Division Multiplexing(SDM)

Using directional antennas and reflectors we can (roughly) shape radio signals in to beams. Signals can be beamed from one location to another and the same frequency can be used for many beams between different locations. Typical outdoor microwave systems operate this way.

Channel seperation is far from perfect but a radio Lan system could be built with carefully selected frequences and directional antennae such that the same frequency is reused for many connections.

Structuring the network in a cellular fashion is also a form of SDM. This is described in "Cellular systems".

Code Division Multiplexing(CDMA)

In a spread spectrum system (with some techniques) it is possible to transmit multiple signals at the same frequency at the same time and still separate them in the receiver. This is called CDMA and is discussed later.

1.3 Conventional Narrowband Radio (Microwave)

It is perfectly sensible to build a radio LAN using conventional narrowband microwave radio. The Motorola 'Altair' product is an example of such system. There are a number of problem however.

1- The use of microwave radio even at low power usually requires licensing of the equipment to a specific user and allocation of a unique frequency.
2- The ISM bands(by regulation)can only be used by spread spectrum systems.

1.4 Spread Spectrum and Code Division Multiple Access (CDMA)

The concepts of spread spectrum and of CDMA seem to contradict normal intuition. In most communication systems we try to maximize the amount of useful signal we can fit into a minimal bandwidth. In spread spectrum we try to artificially spread a signal over a bandwidth much wider than necessary. In CDMA we transmit multiple signals over the same frequency band, using the same modulation techniques at the same time! There are of course very good reasons for doing this. In a spread spectrum system we use some artificial technique to broaden the amount of bandwidth used. This has the following effects:

Capacity Gain

Using the Shannon-Hartly law for the capacity of a bandlimited channel it is easy to see that for a given signal power the wider the bandwidth used, the greater the channel

capacity. So if we broaden the spectrum of a given signal we get an increase in channel capacity and/or an improvement in the signal-to- noise ratio.
This is true and easy to demonstrate for some systems but not for others. "Ordinary" frequency modulation (FM) systems spread the signal above the minimum theoretically needed and they get a demonstrable increase in capacity.Some techniques for spreading the spectrum achieve a significant capacity gain but others do not.

The Shannon-Hartly Law:

The Shannon-Hartly law gives the capacity of a bandlimited communications channel in the presence of "Gaussian" noise. (Every communications channel has Gaussian noise.)

$$\text{Capacity} = B \; \text{Iog2} (1 + Ps/2NoB)$$

Where P represents signal power, N noise power and B available bandwidth. It is easy to see that with P and N held constant, capacity inereases as bandwidth increases (though not quite as fast). So, for a given channel capacity,the required power decreases as utilized bandwidth increases. The wider the bandwidth the lower the power we need to use for a given capacity.

Security

Spread spectrum was invented by military communications people for the purpose of battlefield communications. Spread spectrum signals have an excellent rejection of intentional jamming (jammer power must be very great to be successful). In addition, the Direct Sequence (DS) technique results in a signal which is very hard to distinguish from background noise unless you know the peculiar random code sequence used to generate the signal. Thus, not only are DS signals hard to jam, they are extremely difficult to decode (unless you have the key) and quite hard to detect anyway even if all you want to know is which something is being transmitted.

Immunity to Multipath Distortion

Some spectrum spreading techniques have a significantly better performance the presence of multipath spreading than any available narrowband technique. This will be discussed later.

Interference Rejection

Spread spectrum signals can be received even in the presence of very strength narrowband interfering signals (up to perhaps 30 dB above the wanted signals)

Multiplexing Terhnique (CDMA)

Some techniques of frequency spreading enable the transmission of many completely separate and unrelated channels on the same frequency and at the same time as other, similar signals.
There are two major techniques for generating SS signals:

1. Direct Sequence (DS) - also called Pseudo Noise (PN)
2. Frequency Hopping (FH)

1.4.1 Direct Sequence Spread Spectrum (DSSS)

Also called "Pseudo Noise" (PN), DSSS is a popular teehnique for spreading the trum. Figure shows how the signal is generated.

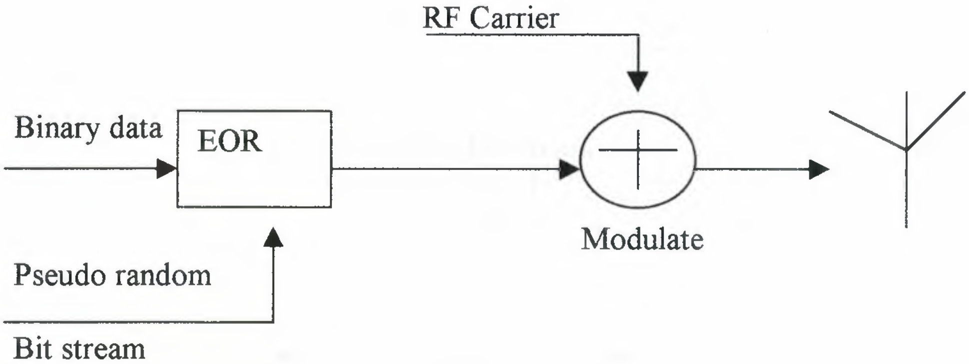


Fig.: direct sequence spread spectrum modulation transmitter

1. The binary data stream (user data) is used to "modulate" a pseudo-random stream. The rate of this pseudo-random bit stream is much faster (from time) than the user data rate. The bits of the pseudo-random stream are called. The ratio between the speed of the chip stream and the data stream is called spread ratio.
2. The form of "modulation" used is typically just an EOR operation performed between the two bit streams.
3. The output of the faster bit stream is used to modulate a radio frequency(RF) carrier.
4. Any suitable modulation technique can be used but in practical systems a very simple bi-polar phase shift keying (BPSK) approach is usually adopted.

Whenever a carrier is modulated the result is a spread signal with two "sidebands" above and below the carrier frequency. These sidebands are spread over a range plus or minus the modulating frequency. The sidebands carry the information and it is common to suppress the transmission of the carrier (and sometimes one of the sidebands). It can be easily seen that the width (spread) of each sideband has been multiplied by the spread ratio.

At first sight this can be quite difficult to understand. We have spread the spectrum but in order to do it we have increased the bit rate by exactly the signal spread ratio. Surely the benefits the spectrum (such as the capacity gain hypothesized above) are negated by the higher bit rate?

The secret of DSSS is in the way the signal is received. The receiver knows the pseudorandom bit stream (because it has the same random number generator).Incoming signals (after synchronization) are correlated with the known pseudo-random stream. Thus the chip stream performs the function of a known waveform against which we correlate the input. (There are many ways to do this but they are outside the scope of this discussion.)

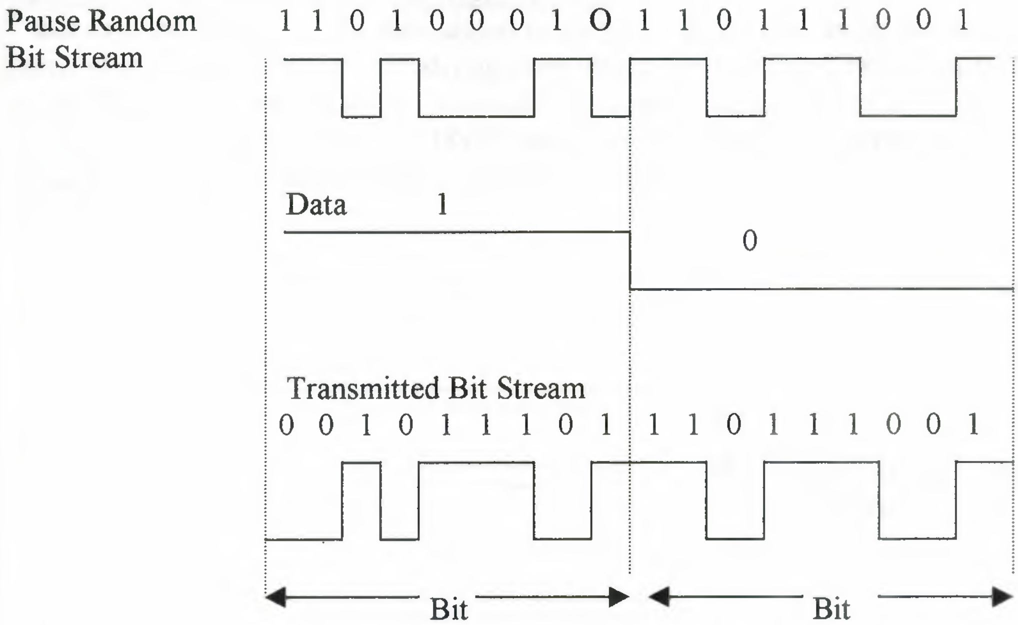


Figure : Direct Sequence Spread Spectrum Modulation.

A pseudo-random bit stream much faster (here 9 times the speed) than the data rate is EORed with the data.The resulting bit stream is then used to modulate a carrier signal. This results in a much broader signal.

DSSS has the following characteristics:

Capacity Gain

The capacity gain predicted by the Shannon-Hartly law is achieved. This means that for the same system characteristics, you can use a lower transmit power or a higher data rate (without increasing the transmitter power).

Improved Resistance to Multi-Path Effects

Above it was mentioned that the length of a data bit at 1 Mbps is about 300 meters. We can think of this as a notional "data wavelength". ISI is most difficult to suppress when the delay spread is less than this data wavelength.Because we have introduced "chipping" we can perform equalization at the chip wavelength. This chip wavelength is significantly less than the data wavelength (by the spread ratio).

It turns out that we can remove delayed signals (where the delay is longer than chip time) very effectively using adaptive equalization. This gives extremely good compensation for ISI.

Rayleigh fading is reduced with DSSS. The location of radio fades within an area is critically dependent on the wavelength. Since the wavelength at one side of the band is different (slightly) from the wavelength at the other side, tbe location of radio fades is also different. The wider the bandwidth used, the less the problem with fading. This mitigates the Rayleigh fading problem somewhat but does not entirely eliminate it.

Immunity to Narrowband Interference

Because the energy of the data signal is spread over a wide range, the presence of a narrowband signal (even a very strong one) within the wideband range has little effect on the DSSS receiver (all it sees is a small increase in the signal-to- noise ratio.

It is even possible to transmit a DSSS signal "over the top" of a group of narrowband signals (using the same frequency space). This is seen in Figure .

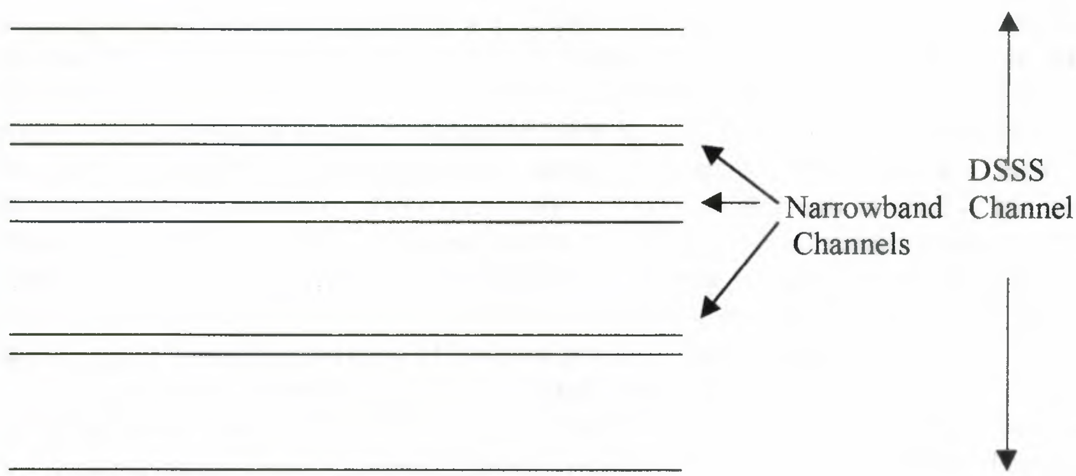


Figure: DSSS over Narrowband Channels

The narrowband channels see the DSSS signal as an increase in noise level (which, if kept within reason will have little effect). For metropolitan area cellular radio systems, DSSS has been seriously suggested for use "overlaying" existing analog FDM cellular radio channel space.

Security

Because the signal is generated by a pseudo-random sequence a receiver must know the sequence or it can't receive the data. Typically such sequences are generated with shift registers with some kind of feedback applied. Unless the receiver knows the key to the random number generator it can't receive the signal.

The biggest problem with DSSS is synchronizing the receiver to the transmitter pseudo-random sequence. Acquisition of synchronization can take quite a long time. Radio LAN systems are not as sensitive (from a security point of view) as a military communication system and it is feasible to use a short, predictable, bit sequence instead of a pseudo-random one. Security is not as good (to receive it you still need a DSSS receiver but you don't need the key anymore), but synhronization can be achieved very quickly and the correlator in the receiver doesn't have to be as smart.

Near-Far Problem

While DSSS is extremely resistant to narrowband interference it is not very resistant to the effects of being swamped by a nearby transmitter on the same band as itself (using the whole bandwidth). A signal from a far away transmitter can be blanketed out by a nearby transmitter if the difference in signal strength at the receiver is only about 20 dB.

1.4.2 Code Division Multiple Access (CDMA)

The DSSS technique gives rise to a novel way of sharing the bandwidth. Multiple trans mitters and receivers are able to use the same frequencies at the same time without interfering with each other! This is a by-product of the DSSS technique. The receiver correlates its received signal with a known (only to it) random sequence - all other signals are filtered out.

This is interesting because it is really the same process as FDM. When we receive an ordinary radio station (channels are separated by FDM), we tune to that station. The tuning process involves adjusting a resonant circuit to the frequency we want to receive.That circuit allows the selected frequency to pass and rejects all other frequencies. What we are actually doing is selecting a sinusoidal wave from among many other sinusoidal waves by selective filtering. If we consider a DSSS signal as a modulated waveform,when there are many overlapping DSSS signals then the filtering process needed to select one of them from among many is exactly the same thing as FDM frequency selection except that we have waveforms that are not sinusoidal in shape. However, the DSSS "chipping sequences" (pseudo-random number sequences) must be orthogoncil (unrelated). Fortunately there are several good simple ways of generating orthogonal pseudo-random sequences.

For this to work, a receiving filter is needed which can select a single DSSS signal from among all the intermixed ones. In principle, you need a filter that can correlate the complex signal with a known chipping sequence (and reject all others). There are several available filtering techniques which will do just this. The usual device used for this filtering process is called a Surface Acoustic Wave (SAW) filter.
CDMA has a number of very important characteristics:

<u>"Statistical" Allocation of Capacity</u>

Any particular DSSS receiver experiences other DSSS signals as noise.This means that you can continue adding channels until the signal-to-noise ratio gets too great and you start getting bit errors. The effect is like multiplexing packets on a link. You can have many active connections and so long as the total (data traffic) stays below the channel capacity all will work well.For example, in a voice system, only about 35% of the time on a channel actually has sound (the rest of the time is gaps and listening to speech in the other direction). If you have a few hundred channels of vnice over CDMA what happens is the average power is the channel limit - so you can handle many more voice connections than are possible by FDM or TDM methods.

This also applies to data traffic where the traffc is inherently bursty in nature. However, it has particular application in voice transmission because, when the.system is overcommitted there is no loss in service but only a degradation in voice quality. Degradation in quality (dropping a few bits) is a serious problem for data but not for voice.

<u>No Guard Time or Guard Bands</u>

In a TDM system when multiple users share the same channel there must be a way to ensure that they don't transmit at the same time and destroy each other's signal.Since there is no really accurate way of synchronizing clocks (in the light of propagation delay) a length of time must be allowed between the end of one user's transmission and the beginning of the next. This is called "guard time". At slow data rates it is not too important but as speed gets higher it comes to dominate the system throughput. CDMA of course does not require a guard time - stations simply transmit whenever they are ready.

In FDM systems, unused frequency space is allocated between bands because it is impossible to ensure precise control of frequency. These guard bands represent wasted frequency space. Again, in CDMA they are not needed at all.

Smooth Handoff

In the mobile environment perhaps the key problem is "handoff' where one user is passed from one cell to another. In an FDM system this is performed by switching frequency. In CDMA, all you do is pass the key (random sequence generator) to the next cell and you can get a very smooth handoff.

At this time existing radio LANs do not allow for fully mobile operation. When a station is moved it makes a new connection with a new base station. However, there are many applications in factories (large plant areas) and warehouses which need continuous connection to the system. Any system which aims to provide this will require a method for smooth handoff.

Requirement for Power Control

As mentioned earlier (the near-far problem), DSSS receivers can't distinguish a signal if its strength is more than about 20 dB below other similar signals.Thus if many transmitters are simultaneously active a transmitter close to the receiver (near) will blanket out a signal from a transmitter which is further away.

The answer to this is controlling the transmiz power of all the stations so that they have roughly equal signal strength at the receiver. It should be noted that this implies a "base station to user" topology, since in an any-to-any topology power control cannot solve the problem.

Easier System Management

With FDM and TDM sy`stems users must have frequencies and/or time slots assigned to them through some central administration process. All you need with CDMA is for communicating stations to have the same key.

1.5 Frequency Hopping (FH)

In a Frequency Hopping spread spectrum system, the available bandwidth is divided up into a number of narrowband channels.The transmitter and the receiver "hop" from one channel to another using a predetermined (pseudo-random) hopping sequence. The time spent in each channel is called a "chip".The rate at which hopping is performed is called the "chipping rate".

Fast Frequency Hopping

A Fast Frequency Hopping system is one where frequency hopping takes place faster than the data (bit) rate. FFH demonstrates exactly the capacity gain sugested by the Shannon-Hartly law.

Unfortunately, while FFH systems work well at low data rates they are difficult and expensive to implement at data rates of 1 Mbps and above; thus, while they are theoretically important, there are no high -speed (user data rate above 1Mbps) FFH systems available.

Slow Frequency Hopping

Slow Frequeney Hopping is where hopping takes place at a lower rate the user data (bit) rate. To be considered an SFH s stem (from are regulatory point of view) hopping must take place at least once every 400 ms and it must statistically cover all of the available channels.

There are many advantages to SFH. However, the capacity gain achieved by other spectrum spreading methods is not demonstrated in SFH systems.

When encoding data for transmission over an SFH system the same requirements apply as for regular narrowband transmission. That is, the data stream must contain frequent transitions and should average the same amount of time each symbol state These characteristics are usually inherent in the data encoding scheme. If the encoded data is not in this form then it is necessary to descramble it on reception.

1.5.1 CDMA in FH Systems

Sharing the wideband channel between multiple FH systems is possible and can be considered a form of CDMA.With two or more systems hopping over the same bandwidth collisions do occur. When there is a collision, data is corrupted and lost.

In an FFH system (say 10-100 hops per bit) then corrupted chips will have little effect on the user data. However, in an SFH system user data will be lost and higher layer error recoveries will be needed. One way of avoiding the problem in the SFH environment is to arrange the hopping patterns so that each system uses a different set of channels so that collisions cannot occur.

1.6 DSSS and SFH Systems Compared

There is some discussion in the industry over which system of spread spectrum operation is the most cost effective. The technology is not at all mature yet and researchers are still trying to settle the matter but there are some early indications.

1. In a paper presented to the IEEE, Chen and Wu (1992) report a performanee comparison between the two systems.The study uses two kinds of mathematical channel models and studied two speeds of operation.Systems were compared without equalization or error correction techniques being applied.

Their conclusion was that at speeds of 1 Mbps the SFH system was superior to DSSS in almost every respect and significantly so in most. At speeds of 10 Mbps their conclusion is the opposite. That is, that DSSS is better, again under most simulated conditions.

As manufacturers bring more systems to market, experience will show the difference.

2. An assessment of manufacturing cost shows that SFH "should" cost less to manufacture and operate at lower power than DSSS. (This all depends on the system design.)

It should be noted that there are many ways to implement either system.For example, you can have a DSSS system which uses only a very short pseudo-random sequence. This saves significant cost in the adapter but limits the potential for CDMA operation and removes much of the security advantage.

3. An SFH system can easily avoid local sources of strong narrowband interterence. All it needs to do is to modify the hopping pattern to avoid the particular frequency band. The ISM bands have many uses and sources of narrowband interference are relatively common.While, in general, a narrowband interferer will not bother DSSS, a strong local interferer (such as a nearby microwave system) will. An SFH system can detect and avoid the frequency bands involved.

4. Laboratory tests of a DSSS-based LAN system collocated with an SFH-based system have shown that the SFH system is more robust. That is, the SFH system was not affected by the DSSS system but the DSSS system ceased to function! This was caused by the signal level of the SFH system being too high for the DSSS one (because the two systems were interspersed in the same room).

1.7 Building a Radio LAN System

There are many possible radio LAN topologies and three of them are illustrated in the following:

Direct User-to-User Transmission

This is where there is no base station and traffic is directly transmitted from user to user.

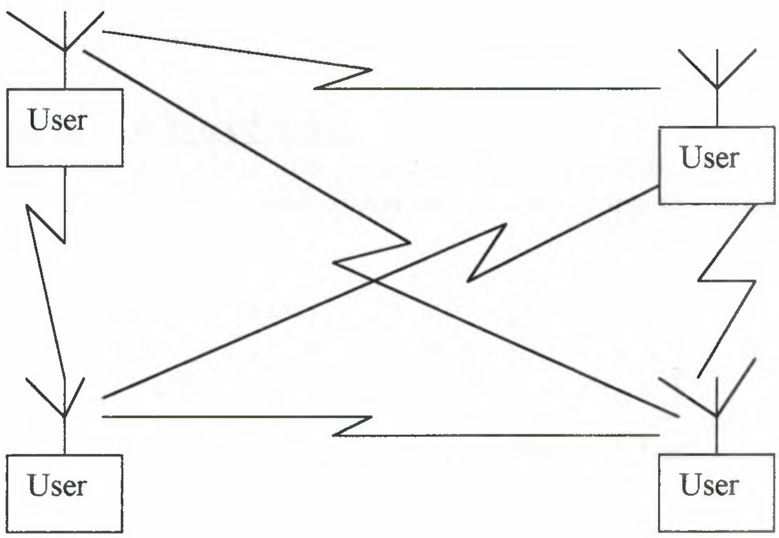


Figure : Radio LAN Topologies - An "Ad-Hoc" Structure

This mode of operation is considered essential because many users will want to set up ad-hoc connections between small numbers of stations on an as-needed basis. LANs such as these might be set up, used and dispersed again within the space of an hour. Such a system could use multiple channels (FDM or CDMA) between pairs of users or a single channel with TDM or CSMA operation to share the available capacity between users.

Use of a Base Station

When a base station is used, all transmissions from workstations are routed to or from the base station. The base station performs the function of a bridge (optionally) connecting the radio LAN segment to a wired LAN.

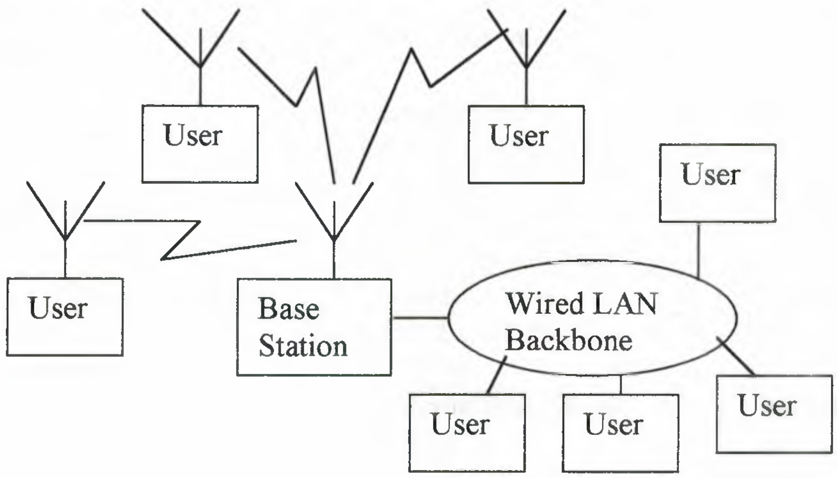


Figure : Radio LAN Topologies – Use of a Base Station

Connection to a Wired LAN

Most often a radio Lan will require a connection a wiredLAN. In this case the base station should perform the function of bridge connecting the radio LAN segment to the wired LAN segment.

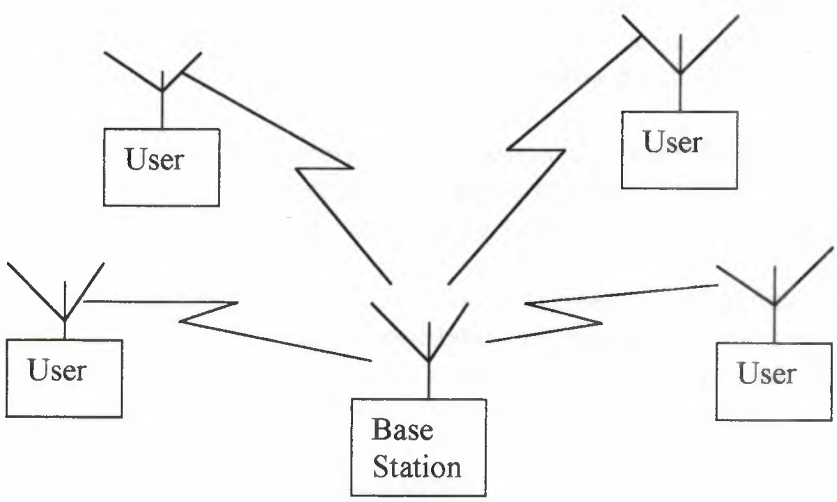


Figure : Radio LAN Topologies – Connection to a Wired LAN

In future LANs it will be possible to have multiple base stations connected together by a wired LAN with a mobile user moving around and being passed from one base station to another much like a cellular telephone user. There are many applications in large plant environments where this would be a very useful function. Currently there are no radio LAN products on the market which will do this.

1.7.1 Topologies Compared

Table 14-2. Comparison of Peer-to-Peer versus Base-to-Remote Operation

	Peer-to-Peer	Base-to-Remote
Coverage	Unpredictable (Hidden Terminals)	Predictable (Base to Remote)
Area Covered	Transmission Range = Network Diameter	Transmission Range = Network Radius
Access Points (to Network)	Multiple	1 per Cell
Security	Single Level (Netwurk O/S Only)	Multi-evelBase,MAC and Physical Control
Management	Unpredictable (Hidden Terminals)	Predictable(Mgmt function through base
Expansion	Limited - Difficult	Multi-Cell Design
Future Upgrades	Ivlanual Distribut.ion	Automated (through Base

In comparing a user-to-user (ad hoc) configuration to a base station confguration the following points should be considered:

1. In the user-to-user configuration the maximum size of the LAN is a circle of diameter equal to the maximum transmission range. In the base station approach the maximum size of the LAN is a circle of radius equal to the maximum range of the transmission. Thus the base station approach allows a single radio LAN to be geographically much larger than the user-to-user approach (all else being equal).

2. If the traffic pattern is genuinely peer-to-peer and evenly distributed the user-to-user approach offers much greater capacity and efficiency. (If you go through a base station the data must be transmitted over the air twice - halving the system capacity.)

However, in practical LANs this is almost never the case. Communication is usually from workstation to server or from workstation to gateway. In the radio case where there is a connection to a wired LAN, a significant proportion of the traffic will probably need to go between the radio users and wired LAN users.

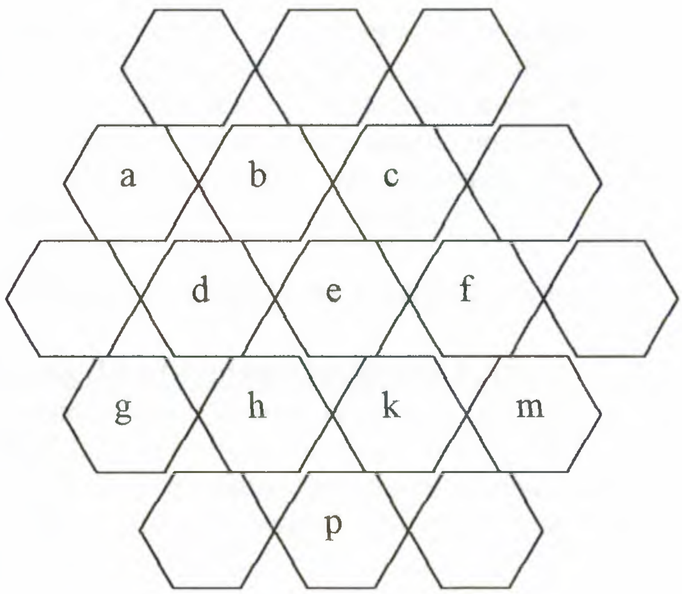
Thus systems with a base station will usually be a better approach. A good system might put the base station and the bridging function in the same machine as the most used server.

1.7.2 Cellular Systems

The big problem with radio systems is that the electromagnetic spectrum is shared by everyone and in consequence is a scarce resource. The cellular concept arose throught the need to get significant increases in capacity from that resource.

If a restricted amount of bandwidth is available for a particular function it doesn't matter much how you multiplex it into channels (FDM, TDM or CDMA); there will be a finite number of channels available for that function. If a large geographic area is covered by high-power equipment then the capacity of the total system is just the number of channels available.

The large area can be broken up into smaller areas using lower power (short range) transmitters. Many transmitters can then use the same frequency (channel), provided they are far enough apart so that they don't interfere with one another. Everyone is familiar with this since radio stations in different cities transmit on the same frequencies and rely on the distance between stations to prevent interference. This gives a significant capacity increase for the whole system. The penalty is that each user can only communicate directly with close by stations - if a user needs to communicate over a longer distance then a method of connecting hub (base) stations within each cell must be provided.



F·igure :Cell Structure.

Figure shows the notional construction of a cellutar radio system. The problem here is that the boundaries in reality are fuzzy following geography rather than lines on a diagram. The concept here is as follows:

**Users within a cell are allocated a channel or set of channels to use and operate at sufficient power only to reach other users (or a hase station) within the same cell

**Because transmissions within one cel) will still have significant strength in adjacent cells, the channels (frequencies) allocated to users in adjacent cells must be different.This is because of the uncertain nature of the boundary between cells. In addition, a user located near a boundary would experience equal strength interfering signals if the same frequency was re-used.

One of the most im ortant design problems is that a user in one cell can be physically close to a user in a neighboring cell.

** If there is a base station within a cell and users communicate only with the base station, then the transmit power used only has to be sufficient to get between any user and the base station.

If the sygtem calls for direct user-to-user communication then the transmit pover must be great enough to allow good signal strength for a distance of the diameter of a cell.

Thus, in a base station to user configuration we can re-use channels much sooner than we could in a user-to-user configuration. In the figure, in a base station configuration the same frequencies could be re-used in cells p, m, e, g, and a. If this were a user-to-user system frequency re-use would be limited to cells that were farther apart such as cells m and b or p and a.

** Cellular systems work because of the very rapid decrease in signal strength as distance from the transmitter is increased. In free space signal strength declines as the square of the distance from the transmitter. Where the antennas are "close to the ground" signal strength declines with the fourth power of distance!

In practical cellular radio (telephone) systems there is a capacity trade-off.The smaller the cells the greater the capacity of the total system but the greater the cost, since these systems have a base station in every cell.

In other kinds of systems there may be no wired backbone and the cellular structure is only used as a means of increasing capacity. (Of course in this case, communication limited to users within individual cells.)

1.7.3 Radio LAN Systems Considerations

Collocated but Unrelated Radio LANs

One important requirement for a radio LAN system is the need to allow multiple unrelated LANs (of the same type) to be located in the same area. This requires a system for channel sharing and LAN administration that allows each independent LAN to be set up and operated without the need to consider other (similar) LANs in the same space.

Countering Multi-Path Effects

As mentioned above, multi-path effects are the most serious problem for the indoor radio environment. The following are some of the general approaches used to counter them:

Antenna Diversity

Antenna diversity can mitigate both ISI and Rayleigh fading effects.There are many different wayds of approaching this:

1. Multiple Directional Antennas

An example of this is the Motorola "Altair" radio LAN system, which uses a six-segment antenna. Antenna segments are arranged in a circle at an angle of 60 degrees from each other. Each antenna segment is highly directional; thus, signals coming from different directions (reflections, etc.) are received on different antenna segments.

As signals are received the system selects the antenna segment receiving the strongest signal and uses that signal alone. This severely limits the number of possible paths. Moreover,surviving paths will have lengths that are not too different from one another.This provides an excellent solution to the ISI problem but does not do a ìot for fading. Notice however that the Altair system is a narrowband microwave system (1.6 cm wavelength) and at that wavelength Rayleigh fading is not as significant as it is at longer wavelengths.

2. Multiple Antennas

Other systems use multiple dipole antennas separated from one another by more than half a wavelength. Signals are added together before detection and this does provide a measure of protection against fading (but not against ISI).

3. Multiple Coordinated Receivers

Two antennas are situated exactly 1/4 of a wavelength apart. Each antenna services a different receiver circuit. The signals are then combined in such a way as to minimize multipath effects.

4. Polarization Diversity

Because the signal polarization changes as reflections occur, a good counter to fading is to provide both horizontal and vertically polarized antennas acting together. The signal is transmitted and received in both polarizations.

Data Rate

The ISI problem is most severe when the delay spread covers more than one data bit time. The easy way to avoid this is to limit the data rate to less than the inverse of the delay spread. However, if the objective is to operate at LAN speeds (above 1 Mbps) then this will not always be practical.

Spread Spectrum Techniques

As discussed above, spread spectrum techniques provide a good measure of protection against ISI and fading. Moreover, spread spectrum is mandatory in the ISM bands. Thus in the indoor radio situation spread spectrum is a preferred method of controlling the multipath effects.

Frequency Diversity

Fading in a narrowband system can be combatted by transmitting the signal on two different frequencies with sufficient separation for the channels to have different fading characteristics. When the signals are received, the station just picks the strongest.

You could call this "half-baked spread spectrum" since all it is doing is spreading the spectrum through an ad-hoc method.

Adaptive Eyualízation

Adaptive equalization is a very good way of countering the ISI form of multipath interference. It is, however, relatively expensive to implement at high speed.

There is some disagreement among specialists as to under which circumstances (if any) adaptive equalization is needed. However, to our knowledge, no current radio LAN system uses adaptive equalization.

1.7.4 Media Access (MAC) Protocols

At the present time all of the available radio LAN systems are proprietary. This is because there is no available standard as yet.The IEEE LAN standardization committee has commenced an effort to develop such a standard.This will be known as IEEE 802.11.As yet there is no draft standard but there are a number of technical proposals before the committee. The fnllowing description is of an IBM contribution (proposal) to the IEFE 802.11 committee (Document IEEE 802.1 Il/92-39).

The task of the MAC protocol is to control which station is allowed to use the medium (transmit on a particular channel) at a particular time. It does not define the overall operation of the LAN system. (For example, in this case the MAC must allow for mobile stations but cannot prescribe how stations are handed off when a cell boundary is crossed.)

The proposed MAC supports the following LAN functions:

1.Slow Frequency Hopped communications system - but it will also operate with a DSSS system
2. Transmission rate between 1 and 20 Mbps
3. Support for broadcast and multicast operation
4.Asynchronous frame delivery
5.Asynchronous (time bounded) frame delivery
6.Multiple, collocated LANs
7.Operation with a base station

8.Ad-hoc operation without a base station
9.Direct station-to-station transmission under control of the base station

Characteristic

The proposed MAC protocol has the following characteristics:

1. A single channel is used for each LAN segment. A channel may be either a unique frequency (narrowband system), a CDMA deived channel or an SFH hopping pattern. Multiple separate LANs or LAN segments may be collocated by using separate channels.
2. A base station (BS) schedules access. The system can operate without a base station in which case one of the mobiles performs the functions of the base station. All mobiles contain the basic base station functions.
3. A hybrid or reservation based protocol and random access protocols are used.

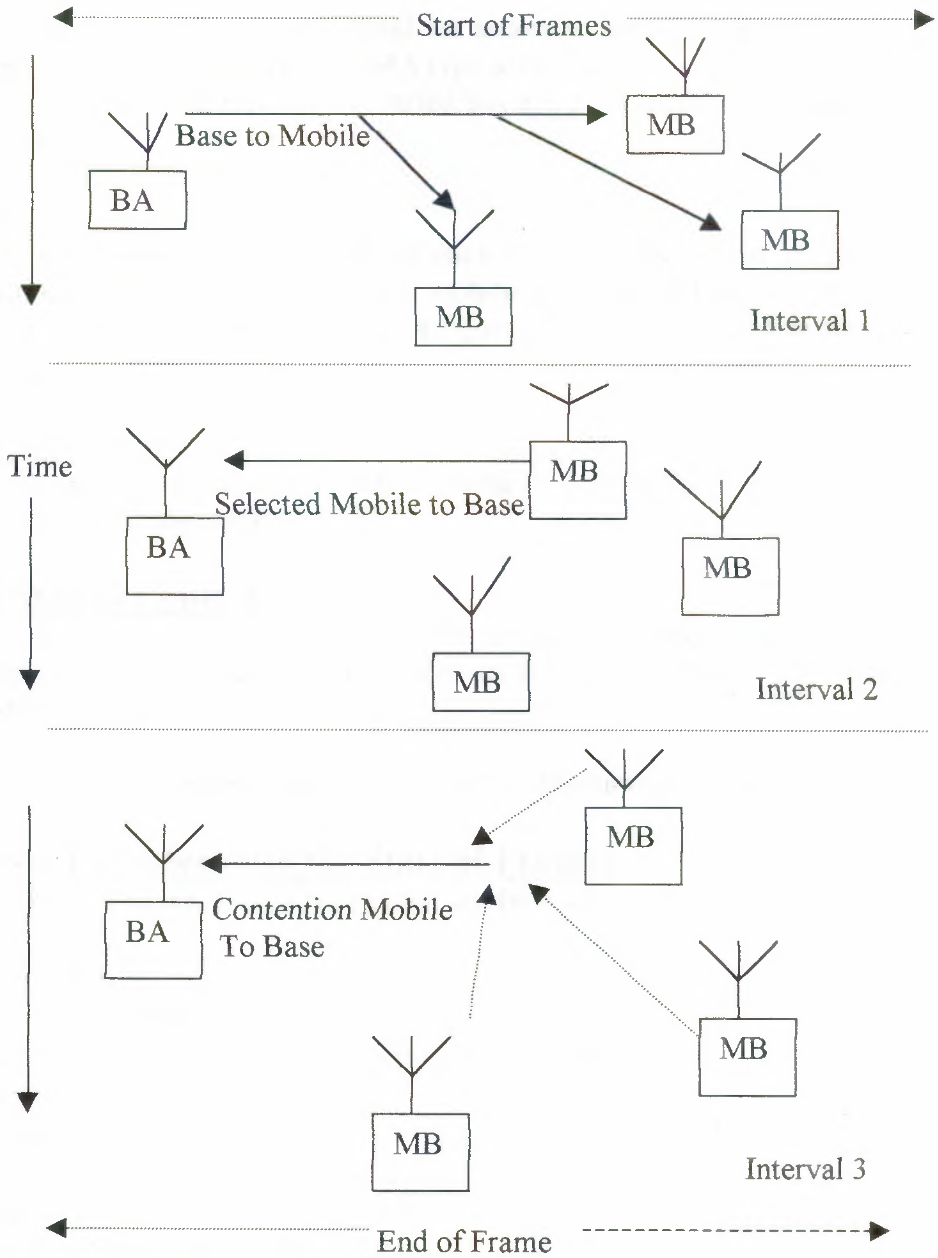


Figure : MAC Operation – Concept

Operation

An overview of the method of operation is shown in Figure 14-11.Operation proceeds as follows:

The Frame

Different from the usual conception of a TDM frame, MAC frames are a concept in time only. That is, a frame is not a synchronous stream of bits but rather an interval of time. The length of a frame is variable and controlled by the base station.

Frame Structure

The frame is struetured into three intervals:

1. In Interval 1 the base station sends data to mobiles (in MB).
2. In Interval 2 mobiles that have been scheduled for transmission by the base station may transmit.

3. In Interval 3 mobiles may contend for access to the air.This period can use either ordinary contention (Aloha) or CSMA type protocols.

The length of the frame can be varied but a typical length is thought to be around 50,000 bits.

Slots

For allocation purposes the length of each interval is expressed in numbers of slots. When time is allocated to a particular mobile it is done in numbers of slots. The size of a slot is yet to be determined but a figure of 500 bits is used in the performance studies.

Data Framing

Data blocks sent to air use HDLC framing to delimit their beginning and end and also to provide transparency.

Base Station Control

At the beginning of the frame and at the beginning of each interval the base station transmits a header to inform all stations of the characteristics of the next interval. Operation proceeds as follows:

Notice that there is a time delay whenever the direction of transmission changes.

Interval 1 Header (also the Start of Frame)

The frame header doubles as the header for Interval 1 and contains the following information:

- Network ID
- Access Point ID
- Frequency to be used for next hop (if SFH operation)
- Remaining length of this hop
- Length of interval 1
- Length of interval 2
- Length of interval 3
- Length of each interval header
- List of receiving stations

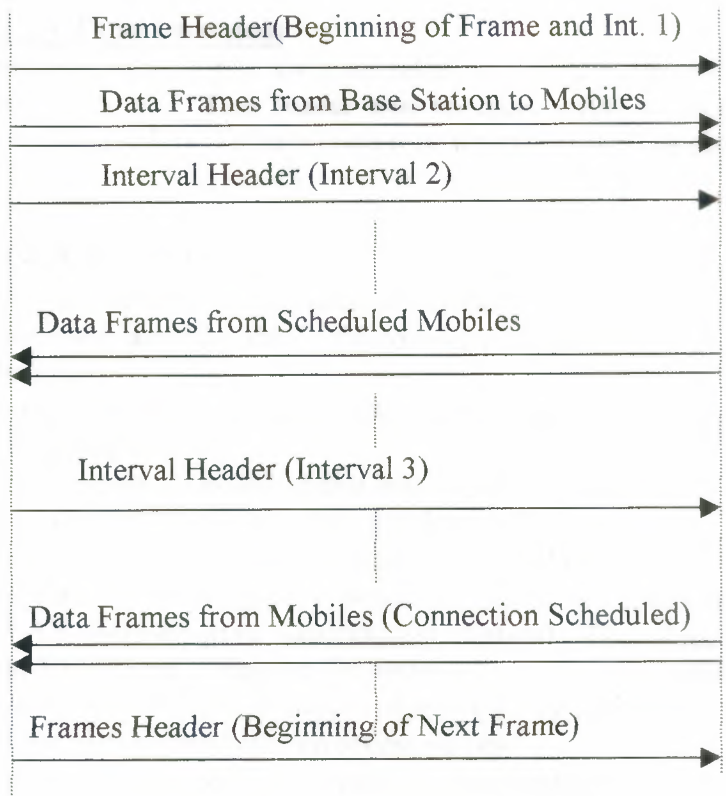


Figure : Radio Lan Operation

Interval 2 Header

This contains the following;

- Length of Interval 2
- Length of Interval 3
- A number representing the number of mobile stations that are allowed to transmit in this interval
- A list of user numbers paired wilh a number of slots

Each entry in the list represents a mobile station and the allocated number of slots it is allowed to transmit in this interval. Mobiles transmit in the same order as the list.

Interval 3 Header

This contains only the length of the interval. Mobiles use contention to decide which one is to send.

Registration with Base Station

When a mobile is switched on it makes a request for registration with the base *station during the next Interval 3.*

Reservation Reqaests

Mobiles send reservation requests during Interval 3. The request can be for some immediate capacity or for regular assignment.

Destination of Transmission

Frames transmitted hy a mobile are prepended by a header containing the origin and destination addresses, etc. This header also contains a flag to indicate whether the frame should be received by the base station or whether it is to be directly received by another mobile.

1.8 Radio LAN Systems

Described above have been two aspects of a radio LAN communication system, the physical transmission and the MAC. Of course, to build a usable system you need much more than this.

1. A management scheme is needed.to control which stations are allowed to become members of a particular LAN.
2. Network management is needed so that errors can be found and fixed quickly and so that time people spend in administrative tasks is minimized.
3. If the users are to be mobile, then you need to build a cellular structure. Within each cell there is a base station (access pointj and the hase stations are interconnected by a wired LAN infrastructure (distribution system). The objective is to allow continuous and transparent operation for users who are moving around. This means that a user must be able to continue a session (connection) to a server without interruption as the user moves between access points.

For this to occur, the access points must communicate with each other and there must be some method to determine when handoff is to occur, to which cell the user is to be handed and to synchronize the necessary system changes in order to do it smoothly.

At the present time, while there are several radio LAN products on the market, we are notaware of any one that enables full transparent mobility as described above. Of course the theoretical problem has been extensively studied in relation to cellular telephone networks.

CHAPTER II

SATELLITE PACKET COMMUNICATIONS

The majority of satellite communication systems have been designed for voice and data traffic with fixed or demand assignment using a multiple access protocol such as FDMA or TDMA. Such systems thus work as circuit switching networks (a voice telephone network is an example of a circuit switching network) where a complete physical path is established from the sender to the receiver that remains in effect for the duration of the connection. The process of selecting a path or circuit establishment may take on the order of seconds for a complex network. Once the circuit is established, data transfer is continuous through the network, and no delays are added by the switches. End-to-end transmission time through the network is limited only by the propagation time of the circuit medium employed, which is dominated by the satellite propagation delay.

Circuit switching systems are efficient for voice calls or data with long messages compared to the time required to make new circuit allocations. Data traffic, however, has more diverse characteristics than voice traffic. In particular, data traffic generated in many data processing applications has great variability in its transmission requirements. The length of the message ranges from a single character to thousands of bytes. One such message is often made available instantly by a cuntrol signal and must be transported to the source within specific delay constraints. As such, data traffic in which a given data source duty cycle is low is often characterized as "bursty" (having a large peak-to-average ratio of the data rate). That is, if one were to observe the user's transmission for a period of time one would see that he requires the communication resource infrequently, but when he does, he requires a rapid response. If fixed-assignment capacity allocation of the communication resource is employed, then each user must be assigned enough capacity to meet his peak transmission rate, with the consequence that the resulting channel utilization is 1ow because of the large peak-to-average ratio of the data rate. To efficiently transmit bursty data traffitc where a fast response is required, the data is formatted into one or more fixed-length packets which are routed through a shared communication resource by a sequence of node switches. Packet switching makes no attempt to store packets for a prolonged period of time while attempting delivery. Rather, packets are discarded if difficulties are encountered in their delivery, in which case they must be retransmitted by senders. Packet switching systems are designed to rapidly forward packets to their destination with the only delay in the node because of a finite transmission capacity.

The use uf satellite packet switching for data traffic can have great econumic advantages over conventional satellite circuit switching, especially when there are a large number of geographically distributed users. A shared broadcast satellite channel uffers full connectivity between users within the satellite global beam, thus eliminating routing and node switches. Furthermore, each user can listen to her own transmission and thus receive autumatic acknowledgment. This allows the implementation of special multiple access protocols for dynamic allocation of satellite capacity to all users to achieve statistical averaging of traffic loads. The key performance of a multiple access protocol for satellite packet communications is that of the satellite channel throughput versus the average packet delay characteristic. The throughput of a satellite channel is defined as the rate at which packets are successfully transmitted.

2.1 PRELIMINARIES

To study packet communicatlons, the following traffic model is assumed:

1- Each user generates messages according to a Poisson process with an average arrival rate equal to (λ) messages per unit time.

2- The message consists of one or more packets and has an average length of 1/mu time units. Each packet carries a destination address so that. when it is transmitted over the satellite channel with no interference from other users, it will be received by the proper addressee.

It is known that telephone traffic can often be modeled as a Poisson process. This too has been verified for data traffic. A message input to system is characticterized by its average arrival rate (λ) and its average length or service time $1/\mu$, (μ being the average service rate). The average arrival rate multiplied by the $1/\mu$ service time is called the traffic intensity and represents the average load to the system:

$$P=(\lambda) / \mu$$

The probability of the arrival of exactly k messages during an interval length t is given by the Poisson distribution:

$$Pr[k]=((\lambda)*t^k / k!)*\exp(-\lambda*t)$$

Queueing, .systems are commonly used to model processes in which messages arrive, wait in a buffer for service, are serviced by servers, and leave. Examples of queues are theater ticket lines and supermarket checkout cashier lines. Queueing systems are characterized by the arrived process (interarrival time probability density function, message length probability density function), service discipline (priority scheme), number of servers (outgoing trunks), and buffer size (finite, infinite).In this chapter we will concentrate on an infinite buffer and a single server using a first-come first-served discipline. Queueing systems are usually symbolized by the notation A/B/C, where A is the interarrival time, B is the message length distribution or service time distribution, and C is the number of servers. The distributions A and B can be of the following three types:

1. "M" stands for "Markov" and is used for Poisson arrival or the equivalent exponential distribution. (A Markov process is a stochastic process whose past has no influence on the future if its present is specified.) Note that Poisson arrivals generate an exponential probability density.
2. "D" stands for "deterministic" and is used for a constant service time.
3. "G" stands for "general" and is used for arbitrary distributions.

Thus an M/M/1 queue has a Poisson arrival, an exponential service distribution, and one server. An M/G/1 queue has a Poisson arrival, a general service distribution, and one server.

2.2 MESSAGE TRANSMISSION BY FDMA: THE M/G/1 QUEUE

In this section we analyze the average message delay versus the satellite channel throughput performance using fixed-assignment FDMA as a multiple access protocol. The analysis is based on the M/G/1 queue as shown schematically in Fig. where messages arrive, according to a

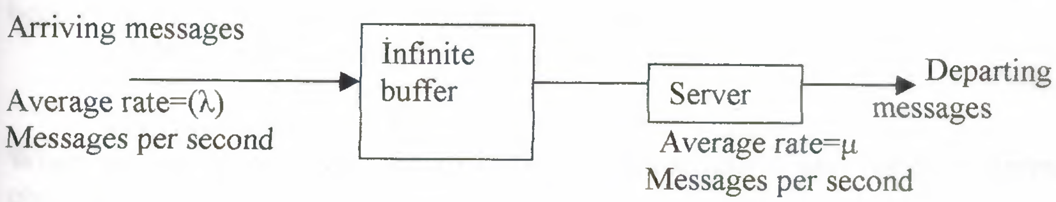


Figure A M/G/1 queue

erality we assume the message length or service time has a general distribution; that is, each is of a randomly varying length. The server works on one message at a time until completion, and then service begins on the next message (first-come first-served or first-in first-out basis).

Assume that the system is already in its steady state, that n messages exist in the buffer at the departure time t, and that one of n messages served at time t departs after an interval t. Also, let k be the number of messages arring during this interval t; then the number of messages existirlg in the buffer at the end of the interval t is

$$N' = \max(n-1.0)+k$$
$$= n-1+(\delta)+(\lambda) \qquad \text{eq.1}$$

where

$$(\text{sigma}) = \begin{cases} 1 & n=0 \\ 0 & n>0 \end{cases} \qquad \text{eq.2}$$

The exprcted (average) value of n' is expressed as

$$E\{rr'\} = E\{n\} + E\{\text{sigma}\} + E\{k\} - 1 \qquad \text{eq.3}$$

$$E(k/t) = \sum k\Pr(k) \qquad \text{eq.4}$$

$$= \sum k!(\lambda t/k!)\exp(-\lambda t)$$

$$= \lambda t$$

Hence

$$E(k) = \int E(k/t)\, g(t)\, dt$$

$$= \int \lambda t\, g(t)\, dt = \lambda / \mu = p \qquad \text{eq.5}$$

where g(t) = arbitrary probability density function of service time with mean (average) $1/\mu$, and variance σ.

Note that $p < 1$, otherwise the buffer will build up indefinitely and the system will become unstable. Using the steady-state condition, that is

$E\{n'\} = E\{n\}$, we have

$$E\{\delta\} = 1 - E\{k\} = 1 - P \qquad \text{eq.6}$$

When we square both sides of eq.1, take the expectation, and rearrange terms we obtain

$$E\{(k-1)^2\}+E\{\delta\}+2E\{n(k-1)\}+2E\{\delta(k-1)\}=0 \qquad \text{eq.7}$$

taking into account the steady-state condition, that is,

$$E(n'^2) = E\{n^2\} \qquad \text{eq.8}$$

and that

$\delta^2 = \delta$
$n\,\delta = 0$

Also, we note that the messages arrive with a Poisson distribution; that is k is independent of n or δ, hence

$$E\{n(k - I)\} = E\{n\}\, E\{k - 1\} \qquad \text{eq.9}$$
$$E\{\delta\,(k- I)\} = E\{\delta\}\, E\{k- I\} \qquad \text{eq.10}$$

Substituting eq.6, eq.9 and eq.10 into eq.eq.7 yields

$$E\{n\} = p+E\{k^2\}-p / 2(1-p) \qquad \text{eq.11}$$

To evaluate the mean square value of k, that is, $E\{k^2\}$, we note that

$$E\{k^2\} = \int E\{k^2 \mid t\}\, g(t)\, dt \qquad \text{eq.12}$$

And,

$$E\{k^2 \mid t\} = k^2\ (\lambda t)/k!\ \exp(-\lambda t)$$

$$=\lambda t+ (\lambda t)^2 \qquad \text{eq.13}$$

Therefore

$$E\{k^2\}= \lambda \int t g(t)\, dt + \lambda^2 \int t^2\ g(t)\, dt$$

$$= \lambda /\mu+ \lambda^2\, (1/\lambda^2+ \sigma)= p+p^2 +\sigma^2 \lambda^2 \qquad \text{eq.14}$$

hence

$$E\{n\} = p+ p^2+ \lambda^2\sigma^2/ 2(1-p) \qquad \text{eq.15}$$

The result in eq.15 is known as the Pullaczek-Khinchine equation and represents the average number of messages waiting in the buffer including the one being served; it is often called the average M/G/I queue length or the average buffer occupancy.
The average message delay (a message delay is defined as the time elapsing between the arrival of a message at the buffer and the departure of the complete message) can be found using Little's [1] result:

$$T= 1/\mu +\lambda\, (1/ \mu^2+ \sigma^2)/2\, (1-p) \qquad \text{eq.16}$$

The average time spent in a queue waiting to be served or the waiting time of messages is simply the average message delay less the average service time; that is,

$$W = T - 1/\mu = \lambda\,(1/\mu^2 + \sigma^2)/\,2\,(1-p)$$

When the service time is exponentially distributed that is when $\sigma^2 = 1/\mu^2$,

$$T = 1/\,\mu - \lambda \qquad \text{eq.17}$$

When the service time is constant that is when $\sigma^2 = 0$

$$T = 2-p/\,2\mu\,(1-p) \qquad \text{eq.18}$$

From (8.18) it is seen that the message delay increases quickly as p approaches 1.

Now consider a satellite channel of capacity R bits per second used in the FDMA mode by N users. Each user is assigned a channel of capacity R/N bits per second. Assume that the average message length is b bits per message; then the average service rate of the FDMA channel is

$$M = R/Nb$$

Let λ be the average message arrival rate for each user. Then the traffic intensity for each channel is $p = \lambda / \mu$

The message delay for the FDMA channel including the satellite roundtrip delay T may thus take on one of two models.

1. An exponentially distributed message length:

$$T = 1 / R/Nb - \lambda + Tr$$

2. A constant message length:

$$T = 2-\lambda / (R/Nb)/\,2\,(R/Nb-\lambda) + Tr$$

2.3 PURE ALOHA: SATELLITE PACKET SWITCHING:

The Aloha protocol is a random access scheme pioneered at the University of Hawaii for interconnecting terminals and romputers via radio and satellites. In the Aloha system, a satellite channel of capacity R bits per second is shared by a large population of ,M users whose traffic is very bursty; that is, it has a high peak-to-average ratio and a low delay constraint. Each user station transmits packets "randomly" at the channel bit rate R whenever its buffer contains one packet. Each packet contains parity bits for error detection. Assume that the satellite channel has a brondcast capability (i.e., all stations are within its downlink antenna beam); then a station can receive its own transmitted packet on the downlink after a satellite roundtrip delay. If the previously transmitted packet is received correctly, assuming that the satellite link has a low error rate, the transmit station can assume that the packet has been received correctly at the destination station and consider the transmission successful. In the situation where packets from different stations overlap at the satellite channel (called packet collision) the transmission error can be detected at the transmit stations on the downlink. The stations then retransmit the packets until they are free from overlap. If two packets from two transmit stations collide at the satellite, they will surely collide again if they are retransmitted after a fixed time-out. To avoid repeated collisions, the interval of packet retransmission is randomized for each station. The Aloha protocol is shown schematically in Fig. To analyze the average packet delay versus the satellite channel throughput in the Aloha system, we assume an infinitely large number of user stations that collectively generate packets according to a Poisson process with rate λ packets per second. Also, we let the packet length be τ seconds; then the average channel input rate or channel throughput is

$$S = \lambda \tau \quad \text{(packets per packet length)}$$

It is seen that $0 < S < 1$ because, if $S > 1$, the user population will be generating packets at a rate higher than that which the channel can handle and nearly every packet will collide. An infinite population assumption is necessary to ensure that S does not decrease as users wait to find out

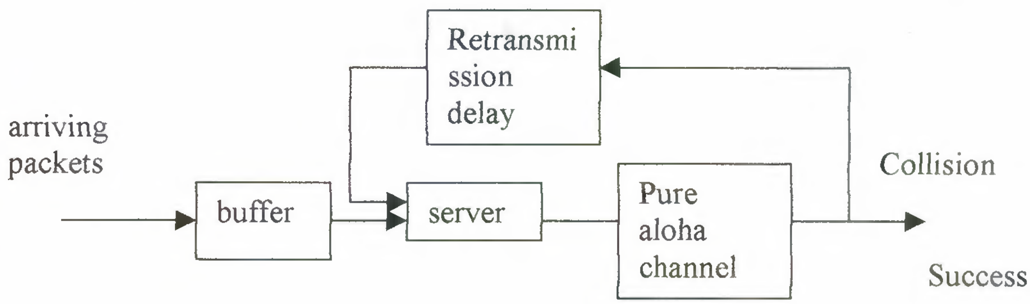


Figure Representation of an Aloha multiple access protocol.

whether their packets have been successfully transmitted. In addition to the newly arrived packets, the satellite channel also contains retransmitted packets.l.et G denote the average satellite channel traffic (newly arrived and retransmitted packets) in packets per packet length and assume that this traffic is also a Poissun prucess with mean G (this is true if the randumized retransmission delay is suffciently large).Then the probability that k packets will arrive at the sattellite channel during any interval of t packet lengths is

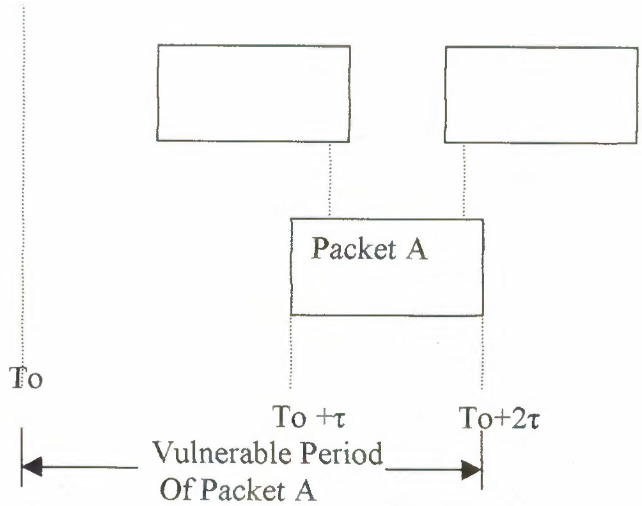
$$\Pr[k,t] = ((Gt)^k / k! \exp(-Gt)$$

Assume that even a partial overlap may cause a collision, as shown in Fig. then the probability that no collision will occur when a packet is transmitted is exactly the probability that no other packet will be generated during an interval of two packet lengths.

Pr[newly generated packet is successfully transmitted]

$$= \Pr[k = 0, t = 2] = \exp(-2G)$$

Since the channel throughput S is just the channel traffic G times the probability that a newly generated packet will not suf□er a collision, we



have

$$S = G \exp(-2G)$$

It is noted that the maximum throughput occurs at a channel traffic of G= 0.5:

$$S_{max} = 1/2e = 0.84$$

This shows that the channel throughput of the Aloha system is very poor, but this is expected since every user is allowed to transmit at will. But, as will be seen later, the Aloha protocol is more appropriate for serving a large number of earth stations whose traffic is very bursty and when satellite channel capacity is limited. In these situations, the average packet delay of the Aloha system is much better than that of a TDMA or FDMA system. A plot of the Aloha channel throughput versus the channel traffic is shown in Fig.

From it is seen that $G = f(S)$ is a double-valued function; for a given value of S, there are two values of G, namely, G and G' > 0.5 > G, such that $S = Ge^{-2G} = G'e^{-2G}$. This indicates that, as the channel traffic increases past 0.5, the throughput drops because the number of packets that suffer a collision increases (which means more packet retransmission). Hence there is a further increase in channel traffic and consequently a decrease in channel throughput, creating a runaway effect.This instability is an inherent characteristic of the Aloha channel and can be prevented only by operating it well below the maximum throughput with enough margin for peak traffic or with some sort of control. The latter method will be studied in the next section when we deal with the slotted Aloha channel.

The average packet delay in an Aloha channel consists of the service time τ (packet length), the average retransmission delay E{T}, and the satellite propagation delay TR:

$T_{Aloha} = TR + \tau + E\{ T \}$

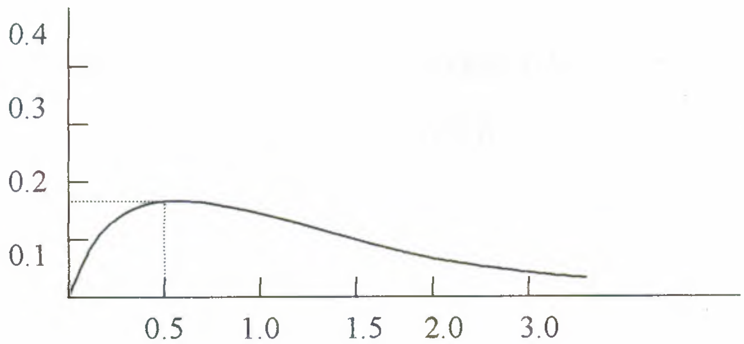


Figure : Throughput versus channel traffic for an Aloha channel.

Since τ and T,t are known, it remains to evaluate the average retransmission delay $E\{T\}$. As mentioned previously, if two packets collide at the satellite channel, each station involved must initiate a retransmission. If the timing of each retransmission is the same, then the collision will persist.Thus some strategy must be used to avoid persistent collisions. One strategy is to assign each station a fixed time-out delay. This approach has the advantage that it completely avoids persistent collisions. It has the disadvantage that some stations will experience a large delay. The second strategy is to use a randomized retransmission approach, where two interfering stations select a retransmission interval from a random sequence of retransmission delays. If each station has a ditferent sequence, then there will be a low probability of persistent collisions. This approach has an advantage over the fixed time-out approach in that it shortens the retransmission delay, but it has the disadvantage that there is a nonzero probability of r·epeated collisions.

Consider the randomized r·etl·ansmission strategy where the random time delay introduced is uniformly distributed over 1 to K intervals of τ seconds each (i.e., the number of intervals between the first and secondtransmissions may be $TR/\tau + I$,$TR/\tau + 2$, . , $TR/\tau + K$, each with probability I/K). The average delay before retransmission is $(K + I)\tau/2$, and
the retransmission delay of a packet after r retransmission is

$$T = r[TR + (K + I)\tau / 2]$$

Let Qr be the probability of a successful retransmission after r retries; then the average number of retransmissions of a packet is

$$E\{r\} = \sum_{r=1} r\, Q_r$$

and the average retransmission delay of a packet is

$$E\{T\} = E\{r.\}[Tr + (K + I)\tau / 2]$$

Now let q be the probability ofa successful transmission given that a new packet has been generated and q' be the probability of a successful transmission given that a retransmitted packet is tr-ansmitted. Then it is seen that, for $r \geq 1$,

Qr = (I -q)(I - q')'^r-1q'

E{r} = I – q / q'

If we assume that the probability of success is the same on any try, that is,

q = q', which is reasonably correct for sufficiently large K, then

E{r} ~ 1-q / q

From eq. we have, for K > > 1,

q~ exp(-2G)

Therefore

E{r} ~ exp(2G) - 1

Thus the average retransmission delay of a packet is

E{T} ~ [exp(2G) - 1l [T+ (K+ I) τ /2]

Putting all these results together we obtain the average packet delay as

T ~ Tr + τ + [exp (2G) –1][Tr+ (k+1)τ/2] K >> 1

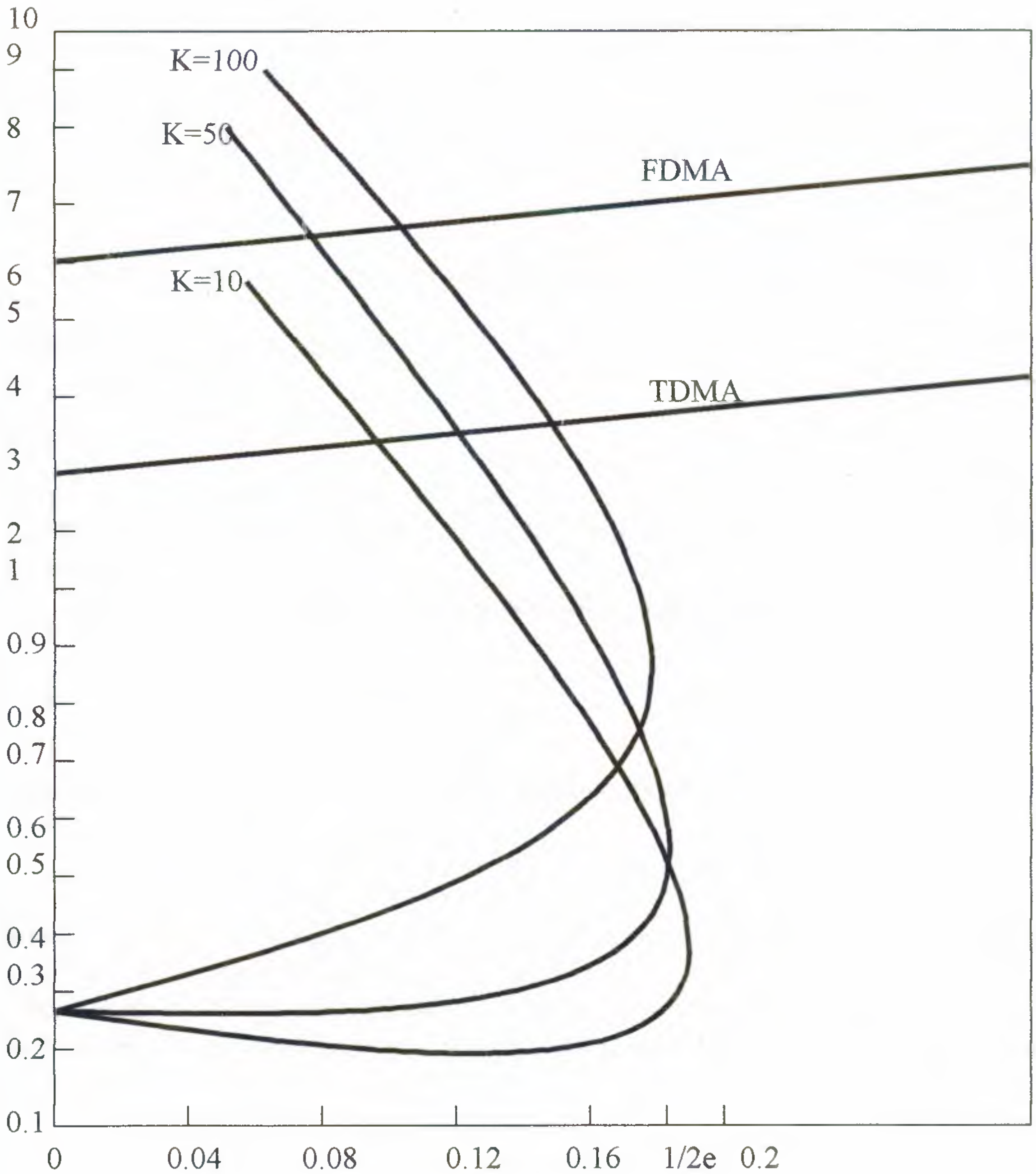
Note that, since the channel traffic G is a double-valued function of the channel throughput S, there are two average packet delays T(aloha) and T'(aloha) > T(Aloha) corresponding to channel traffic G < O.S and G' > 0.5.

Note that T(Aloha) can become arbitrarily large when the Aloha channel traffic exceeds 0.5, creating a runaway effect which further increases G if the peak load persists.

A plot of the average packet delay versus the channel throughput of an Aloha system is shown in Fig. with K as a parameter, for R = 250 kbps, b = 1000 bits, τ = b/R = 4 ms, Tr = 250 ms, and user population N =∞. Note that the average packet delay does not change substantially for values of K between 10 and 50, which means that the actual value of K selected is not critical. The average packet delays using FDMA and TDMA protocols are also plotted for comparison using N = 1500. Note the clear advantages of the Aloha channel in terms of the packet delay.

2.4 SLOTTED ALOHA:

The advantage of the pure Aloha protocol is its simplicity, which can result in low-cost user stations since no synchronization is required between stations in the system. Each station transmits a packet whenever its buffer has one. The disadvantage is somewhat ineflïcient channel utilization; that is, the maximum channel throughput is only 8.4% of the available capacity. One strategy for improving pure Aloha channel throughput is to coordinate transmissions between stations by synchronizing the



transmit and retransmit timing as in TDMA, in effect providing time slots of the same duration as a packet length. A station can transmit a packet only at the start of a slot. Thus when two packets collide at the satellite channel, they will overlap completely as shown in Fig., where the vulnerable period of a packet is one packet length τ. Partial overlaps like those occurring in pure Aloha channel will never occur. This modified. Aloha scheme is called slotted Aloha is a synchronized random access method.

If we assume that the user population N is infinitely large, as in the pure Aloha analysis it can he shown that the slotted Aloha throughput is

T $\quad$ $t + \tau$ $\quad$ $t+2\tau$ $\quad$ $t+3\tau$

Figure : Vulnerable period (one packet length) of a slotted Aloha protocol.

related to the channel traffic by

$$S = G \exp(-G)$$

This is because the channel throughput is just the channel traffic G times the probability that a newly generated packet will not suffer collision (is successfully transmitted)

Pr[newly generated packet is successfully transmitted)

$$= \Pr[k = 0, t = 1]$$

$$= \exp(-G)$$

The slotted Aloha maximum of throughput occurs at G = 1:

$$S_{max} = 1/e = 0.368$$

which is twice the maximum capacity of the pure Aloha channel. A plot the channel traffic G versus the throughput S is shown in Fig.

The slotted Aloha channel can also be analyzed for a finite earth station population N. Let Gi be the probability that a station will transmit a packet in a given slot. The average traffic due to the i th station is therefore

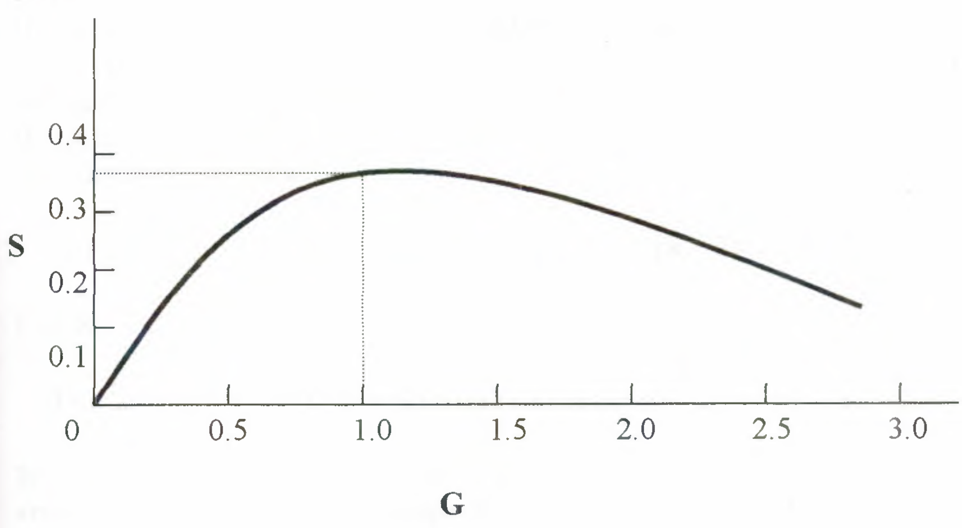


Figure : Throughput versus channel traffic for a slotted Aloha channel.

Gi Hence the average.channel traffic is

$$G = \sum Gi$$

Let Si be the probability that a newly generated packet will be successfully transmitted. Then the average throughput due to the i th station is Si, and the channel throughput is

$$S = \sum Si$$

The probability Si that a packet from station i will be successfully transmitted is simply

$$Si = Gi \prod (1-Gi)$$

When all the stations are statistically identical. that is, Si = S/N and Gi = G/ N, we have

$$S = \sum Si = \sum Gi \prod (1-Gi)$$

$$= G(1-G/N)^{n-1} > G \exp(-G)$$

Note that, as $N \longrightarrow x$ $(1- G/N)^{n-1} \longrightarrow \exp(-G)$ and $S = G \exp(-G)$, which is the case of an infinite population. To analyze the average packet delay of the slotted Aloha channel, we note that in the pure Aloha operation a packet arriving at an empty buffer is transmitted immediately; however, in the slotted channel, this situation occurs with a probability of zero; that is, a newly arrived packet must await the beginning of a slot to be transmitted. Given that a packet arrival occurring in a time slot is uniformly distributed over the interval (O,τ),where τ is the duration of the slot or packet length, then the additional average waiting time is simply $\tau /2$. This also applies to any retransmitted packet. Thus the average packet delay of the slotted Aloha channel for an infinitely large numbef of users can be expressed similarly to that for
the pure Aloha channel:

$$Ts\text{-}Aloha=TR + \tau + T / 2 + 1 - q / q' [TR + (K+ 2)\tau / 2 + \tau / 2]$$

For K >> 1,Ts-Aloha can be approximated as

$$Ts\text{-}Aloha = TR + 3 \tau / 2 + [\exp (G) - 1] [TR + (K+2) \tau / 2] \quad K>>1$$

It is noted that there are two possible delays for a given S and K. They correspond to a small delay Ts-Aloha and a large delay T's-Aloha· We will refer to the equilibrium point given by So, K, and Ts-Aloha as the cahannel operating point, since this is the desired channel performance given So and K. The existence of two channel equilibriun points suggests that the slotted Aloha channel (or pure Aloha channel) is inherently unstable. Starting with an initially empty system (no traffic) the channel moves toward equilibrium. It stays at the channel operating point for a finite period until stochastic fluctuations gives rise to high channel traffic and push the channel throughput toward its maximum value, consequently reducing it; this in turn increases the channel trafiïc, hence the packet delay. As the fluctuations continue, the channel becomes overloaded with retransmitted packets, the throughput vanishes to zero and channel saturation occurs. Thus the delay-versus-throughput performance at equilibrium as studied above is not sufficient to characterize the dynamic behavior of the slotted Aloha channel. An accurate measure of the channel's performance must include a stability study of the delay throughput trade-off.This is what we will consider in the following discussion.

To study the delay-throughput-.stabilty performance we will adopt the following model. Consider N earth stations, each of which can be in one of two states: thinking or blocked. In the thinking state, a station transmits newly generated packets according to a Poisson distrubution with a mean of $\alpha \ll 1$ packet per slot (one slot is a τ-second interval where τ is the packet length); that is, α can be thought of as the probability that a given station will transmit a newly generated packet in a given slot If the mean think time (interarrival time) of a packet is ta, then $\alpha = \tau / ta$. From the time the station transmits a newly generated packet to the time the packet is successfully received, which includes retransmission time if the packet suffers a collision, the station is blocked in the sense that it cannot transmit any newly generated packet. A packet which suffers a collision at the channel and is waiting for retransmission is said to be backlogged. It is this waiting time that determines the performance of the channel. As in the previous study of delay-throughput performance, the waiting time consists of K slots where the retransmission time is uniformly distributed. Including the satellite roundtrip delay of TR/τ slots the number of slots between the first and second retransmissions may be $TR/\tau + 1$, $TR/\tau + 2, \dots,$ $TR/\tau + K$, each with probability I/K. If K is small, the chance of recollision is large, but the average retransmission delay is small; for example, if two stations have collided packets and each must wait from 0 to 10 slots with equal probability, then the chance that the two packets will collide a second time is 0.1. On the other hand, if the retransmissions are spread out uniformly over the next 50 slots, the chance that the two packets will collide again is 0.02. Of course, the average retransmission delay in the latter case is larger than in the former. This is the delay-throughput trade-off of the slotted Aloha (or pure Aloha) system.

Unfortunately, the randomization retransmission strategy is difficult to analyze because of the inclusion of the satellite roundtrip delay TR/τ slots, which necessitates a model with memory consisting of a channel history of at least TR/τ consecutive slots. Instead, a probabilistic model called a Markov model is adopted, in which a backlogged packet is retransmitted with probability p in a given slot following the original transmission until it is successfully received.The average delay before retransmission is now assumed to be geometrically distributed with the probability of an n-slot delay given by $p(1 - p)^{n-1}$. The average delay before retransmission attempts is

$$\sum_{n=1} np(1-p)^{n-1} = 1/p$$

Note that the average retransmission delay (the time the station is blocked), which is denoted by E{T} and henceforth will be called the average backlog time, is different from 1/p since a packet can be retransmitted many times.

The assumption that the delay before retransmission attempts has amemoryless geometric distribution (TR = 0) permits an analytical study using a Markov model. Simulation has shown that the slotted Aloha channel performance in terms of average delay and throughput is dependent primarily on the average delay between retransmission attempts and is quite insensitive to the exaet probability distribution considered. In order to use the analytic result of the Markov model to analyze the delay-throughput performance with TR≠ 0 to approximate the slotted Aloha channel) with randomization retransmission uniformly distributed over K slots, it is necessary to equate the average delay 1 /p with the average randomization retransmission delay; that is, we must let (in the subsequent discussion we will use a slot instead of a second as the unit of time)

$$1 / p = TR / \tau + 1 / 2 + K + 1 / 2$$

$$= TR / \tau + K + 2 / 2$$

Thus, if the value of p is determined for a channel operating point, the corresponding value of K can be derived, and vice-versa.

The slotted Aloha channel state at any time t can be described by the total number of backlogged packets. In state n there are n packets backlogged with retransmission probability p in the current slot, yielding an average retransmission traf□c of np packets per slot. Besides the n blocked stations (each with one backlogged packet), there are N – n thinking stations which are busy transmitting newly generated packets at a collective rate of (N - n) α packets per slot. The quantity S = (N – n) α is called the channel input rate and must be equal to the average channel throughput (the fraction of slots that contain exactly one packet) at equilibrium.

Now consider the behavior of the Markov process. Let r be the number of backlogged packets retransmitted from i blocked stations in a current slot, with 0≤ r ≤ i. Also, let rn be the number of newly generated packets transmitted by N - i thinking stations during the same slot, with 0≤ m ≤ N - i. The transition probability pij that the channel in state i will move to state j in the next slot is given by

$$Pij = \begin{cases} 0 & j \le i - 2 \\ Pr[m=0]\ \ Pr[r=1] & j = i - 1 \\ Pr[m=0]\ \ Pr[r \neq 1] + Pr[m=1]\ \ Pr[r=0] & j = i \\ Pr[m=1]\ \ Pr[r \ge 1] & j = i + 1 \\ Pr[m=j-i] & j \ge i + 2 \end{cases}$$

where Pr[m = x] , x = 0, 1, ... , N - i, is the probability that x newly arrived packets will be transmitted in the current slot and Pr[r = 0], Pr[r = 1] , Pr[r ≥ 1], and Pr[r ≠ 1] represent the probabilities that none one, one or more than one, and none or more than one blocked stations will attempt a retransmission in the current slot.These probabilities are given below for the case of finite N and α and for the case of N→∞ and α→0 such that Nα = .S <∞. (In fact, Nα < 1 since the total rate of newly generated packets must always be less than one packet per slot.)

Finite cuse		Infinite case
$Pr[m=0]$	$= (I - \alpha)$^N-i	exp(-S)
$Pr[m=1]$	=(N-i) $\alpha(I-\alpha)$^N-i-1	S exp(-S)
$Pr[m=j-i]$	$= \begin{bmatrix} N-i \\ j-i \end{bmatrix}$ α^j-i $(1 - \alpha)$^n-i	S^j-i/(j-i)!*exp(-S)
$Pr[r=0]$	= (I - p)^i	(1 – p)^i
$Pr[r=1]$	= ip(1 - p)^i-1	ip(1 - p)^i-1
$Pr[r \geq 1]$	=1-(1-p)^i	1-(1-p)^i
$Pr[r \neq 1]$	=1-ip(1 -P)^i-1	1 -ip(1 -p)^i-1

Assume that N and α are both time-invariant; the Markov process that represents the number of blocked stations n can be described by a stationary probability distribution Pn, that is, the probability of finding the system in equilibrium state n (time-invariant). Pn can be found from a set of linear N + 1 dependent equations:

$$Pn = \sum_{i=0} Pi\ Pin \qquad n=0,1.......,N$$

subject to Ihe constraint

$$\sum_{i=0} Pi = 1$$

which means that the probability that the system will move from its current state to sume state is unity.

Given the equilibrium state probabilities, the average number of blocked stations (or the average backlogged packets) can be obtained from

$$E\{n\} = \sum nPn$$

The throughput of the channel in any state n, S(n), is the probability that a packet will be successfully sent, given that n stations are blocked and N - n stations are thinking. A successful transmission occurs when either exactly one thinking station transmits a packet and none of the blocked stations attempts a retransmission or exactly one blocked station attempts a retransmission and none of the thinking stations transmits a newly arrived packet. Therefore S(rr) can be expressed as

$$S(n) = Pr[r = 0]\ Pr[m = 1] + Pr[r = 1]\ Pr[m = 0]$$

For the finite case, $N < \infty$, $\alpha > 0$,

S(n) =(1 – P)^n (N - n) α (I - α)^N-n-1 + np (1 - P)^n-1 (1 - α)^N-n

For the infinite case $N \to \infty$, $\alpha \to 0$, $(N-n)\alpha \to N \alpha = S$,

$$S(n) = (1- p)^n S \exp(-S) + np(1-p)^{n-1} \exp(-S)$$

The steady-state channel throughput S is simply the sum of the throughput of each state weighted by the equiliblrium state probability; that is,

$$S = \sum PnS(n)$$

By using Little's result, the average backlog time (average retransmission time) in slots is given by

$$E\{n\} / S = \sum nPn / \sum PnS(n)$$

Using the previous notation, the average backlog time E{T} in seconds is

$$E\{T\} = \tau * E\{n\} / S = \tau * \sum nPn / \sum PnS(n)$$

Consequently, the average packet delay Ts-Aloha can be expressed as

$$\text{Ts-Aloha} = TR + 3\tau / 2 + E\{T\}$$

$$= TR + 3\tau / 2 + \tau * \sum nPn / \sum PnS(n)$$

Normally the computation of the equilibrium state probabilities Pn, n = 0,1, . . , N, is quite involved. Simulation has shown that S and E{n} are closely approximated by the throughput So and backlog no at the channel operating point for a stable system, which we will discuss next.

As shown above, the channel throughput S(n) at any state n, which represents the number of blocked stations, varies with the retransmission probability p. Assuming that N and a are time-invariant, a plot of S(n) in packets per slot versus the backlog n with K (hence p) as a parameter is for TR = 0.250s.τ = 4ms (TR/τ = 62.5 slots). N = 500, and $\alpha = 7.36 \times 10^{-4}$. Also plotted is the ,channel load line, representing the channel input rate versus the backlog,.S = (N - n) α.The channel load line intersects the n axis at N and has n slope of $-1 / \alpha$. At equilibrium the channel throughput must be equal to the channel input rate. Thus when the S axis and the S (n) axis have the same unit scale, the intersections of the curve S (n) versus n and the channel load line represent the equilibrium points. Four cases are considered in Fig. In Fig. the channel load line intersects the throughput-backlog curve at exactly one point. The channel is said to be stable, and the channel operating point is said to be a globally stable equilibrium point.The arrows on the channel load line point in the direction of decreasing backlog size, since the channel throughput is always greater than the channel input rate. If N is finite, a stable channel can always be achieved by using a sufficiently small p or, equivalently, a sufficiently large K. Of course, a large K implies that the equilibrium backlog no is large, and so is the average packet delay.The globally stable equilibrium point is obtained for

K = 100 at no≈37 So≈0.342

which yields the average backlog time (average retransmission time) in seconds as

$$E\{T\} = \tau \ no / So = 0.004 * 37 / 0.342 = 0.433 \text{ s}$$

hence the average paeket delay time is

$$Ts\text{-}Aloha = TR + 3\tau / 2 + E\{T\} = 0.250 + 0.006 + 0.433 \quad = 0.689 \text{ s}$$

A necessary condition for the stable channel in Fig. is that the max-

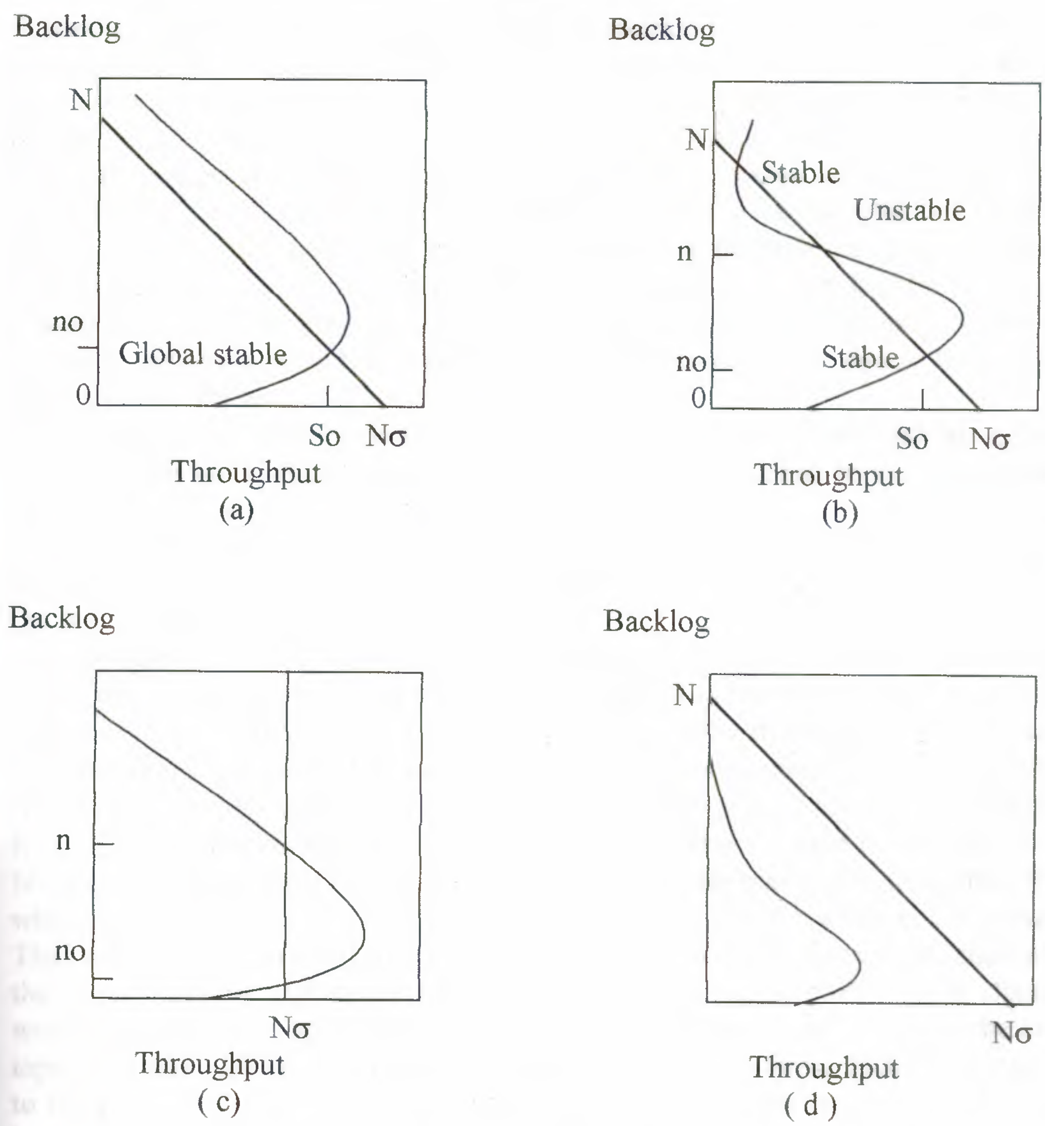


Figure Stability of a slotted Aloha channel. (a) Globally stable. (b) Bistable. (c) Unstable. (d) Overloaded.

imum channel input rate be below the maximum channel throughput, that is,

$$S_{max} = N\alpha \le (1 - 1/N)^{N-1} \approx 0.368$$

N

Thus for a given α (the probability that a newly generated packet will be transmitted in a time slot) the upper bound on the number of station is given by

$$N \leq Smax / \alpha \approx 0.368 / \alpha$$

For example, if a packet contains b = 1000 bits and the channel capacity is R = 250 kbps, then each time slot (one packet length) is 4 ms long.For an average thinking time (packet interarrival time) ta = 10s, $\alpha = \tau$ / ta = 4 x 10^-4, and N ≤ 0.368/4 x 10^-4 = 920.

In Fig. b, the channel load line intersects the throughput-backlog curve at three points which can be considered equilibrium points. But only two of them are stable equilibrium points which we refer to as local stable equilibrium points. One stable equilibrium point corresponds to a low backlog (hence a short delay) and is the desired channel operating point; the other corresponds to a huge backlog where the channel is almost saturated. The third equilibrium point is unstable in the sense that the backlog at this point will remain there for a finite but short period of time. If the rhannel backlog is close to n", the channel will remain at its operating point. Since n is a random process the channel cannot maintain the local stable equilibrium (So, no) for an infinite time. Fluctuations in the input rate will drive the channel toward the unstable equilibrium point. As soon as the backlog surpasses no the channel input rate will exceed the channel throughput, the backlog will keep building up, and the channel will settle at the high-backlog local stable equilibrium point for a finite but probably very long period of time. For this case the channel is said to be bistable, as shown in Fig. for K = 2,10,50.

The case in Fig. c corresponds to a channel with $N \rightarrow \infty$,$\alpha \rightarrow 0$, and $N\alpha \leq Smax$ = 0.368. This is an unstable channel, since it will certainly fail as soon as n surpasses n, and the backlog will grow without bound.

Figure d shows the case of an overload channel where the input rate is always greater than the throughput. This system certainly will fail as soon as it starts up, and the only way to correct it is to reduce N.

In designing a stable slotted Aloha channel for a given average retransmission delay, that is, for a fixed K or p, the designer has no choice but to limit the number of stations N such that the channel load line intersects the throughput-backlog curve at only one point, assuming that $N\alpha$ will never exceed its maximum value of 0.368. This results in inefficient utilization of the channel capacity per station. As an example, for K = 50 in Fig. the number of earth stations in the network must be reduced to about N= 425 to obtain a stable channel as compared to the maximum attainable N= 500 which can be achieved only if K is raised to about 100 (giving sume safety margin). The probelm can be solved by some sort of adaptive control. Since K or, equivalently, the retransmission probability p actually determines the stability of the channel, it would be ideal if K could be increased (or p could be deereased) to make the channel input rate less than the maximum throughput when the channel backlog is high so as to reduce it, and if K could be decreased (or p could be increased) when the channel backlog is low to achieve a shorter delay.

One way to control the channel is to estimate the channel traffic G. Recall that the channel throughput is maximum at G = 1. When G > 1, the channel throughput drops and can fall below the channel input rate, creating a higher backlog; thus K should be increased by the stations. When G < 1 the channel throughput exceeds the channel input rate and

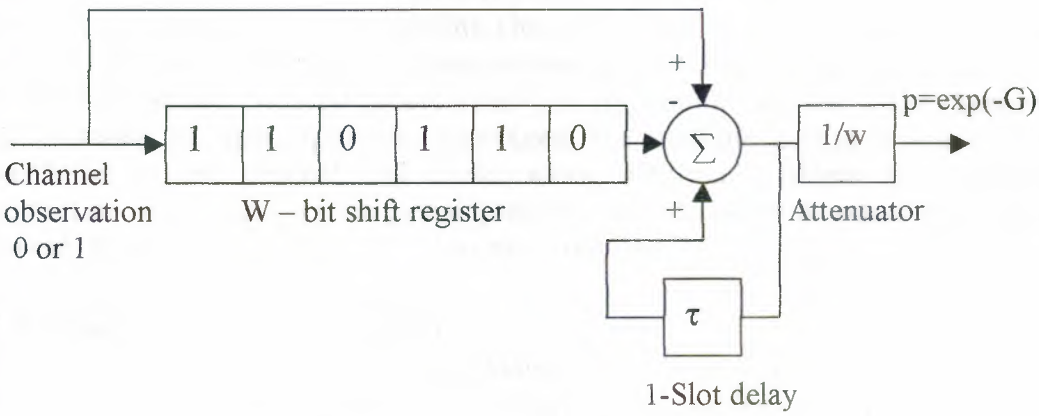


Figure :Controlled slotted Aloha protocol.

K should be decreased to reduce the delay. To estimatte G. note that the probability that a packet will be successfully transmitted is Po = exp (-G) or G = -1n Po. Po can be estimated by maintaining the channel history for W slots and recordin g only the fraction of empty slots (no collisions or transmissions).The size of W for each station is very important for successful estimation. If W is too large, the channel behavior may change when action is taken. If W is too small, large errors may occur in approximating the probability of zero channel traffic based on the fraction of empty slots. A good initial estimate is that W should be larger than the satellite roundtrip propagation delay TR/τ slots. Since it is only necessary to record the empty slots, a W-bit shift register as shown in Fig. would suffice. The stored bit string represents the channel history in a window of W slots. An empty channel slot is represented by 1, while a nonempty slot is represented by 0. The value of p can be selected as p = e^-G / (TR / τ + 1), which decreased (K increases) when G > 1 and increases (K decreases) when G < I. This gives K = 2 (e - 1) (TR/τ + 1).

2.5 PACKET RESERVATION

We have studied in detail two main types of satellite multiple access schemes for packet transmission: channel fixed assignments such as FDMA and TDMA which are appropriate for a small user population with high traffic density but perform poorly for a large population of bursty users, and random access systems such as pure Aloha and slotted Aloha which are independent of the user population but yield poor throughput.

To overcome the drawbacks of the above two schemes, a number of protocols called packet reservation have been proposed that work like the slotted Aloha when the channel throughput is low and move gradually to TDMA when the channel throughput exceeds 1/e. Packet reservation, like random access, is intended for use when a satellite channel is shared by a large number of earth stations. The satellite capacity is demand-as-signed to individual packets or messages from earth stations. The objective, of reservation is to avoid the collisions of packets that occur with random access, and the pr-ice paid is an increase in the packet or- message delay of at least twice the satellite roundtrip delay excluding the service time (one roundtrip delay due to making a reserva tion if it is srlccessful). To coordinate all the stations, a global queue is maintained for satellite channel access. When it has a packet in the buffer, each station generates a request to reserve a place in the queue. A fraction of the satellite capacity is used to accommodate reservation request traffic. Since the earth stations are geographically distributed, the multiple access problem in the reservation request channel still exists, along with the problem of maintaining information on the status of the global queue to make decisions on when to access the channel. Readers may refer to [2-5] for- mure information.

2.6 TREE ALGORITHM

Recall that pure Aloha and slotted Aloha are random multiple access protocols where the conflict (packet collision) is resolved by randomized retransmissions. In this section we present an alternative algorithm called a tree algorithm for resolving the conflict. As in slotted Aloha, the channel we consider is divided into slots of equal length. The length of each slot equals the packet length. Also, we consider the satellite to have a broadcast capability, hence each station can detect a packet collision in any slot. Also, we assume that the user station population is infinitely large and that packets arrive at the satellite channel according to a Poisson process with an average arrival rate of S packets per slot. Thus S is the channel input rate. Since the numher of users is infinite but the channel input rate is tinite, the probability that any station will have more than one packet arrival in a finite interval is zero. Also, if a second packet arrives at a station. it will not be trrrnsmitted rlntil the fir-st pricket has been successfully transmitted.

The tree algol-ithm for conflict resolution is based on the observation that the cvntention among active stations (stations with one packet to transmit) can be resolved if all the stations can be divided into groups such that each group contains only one active station. The tree algorithm can be best described by an example.

First consider a binary tree where each node consists of only two branche, and all the stations in the network are arranged like the leaves of the tree, as shown in Fig. fol- eight stations (Xo, X1 X2, . . , X7). The

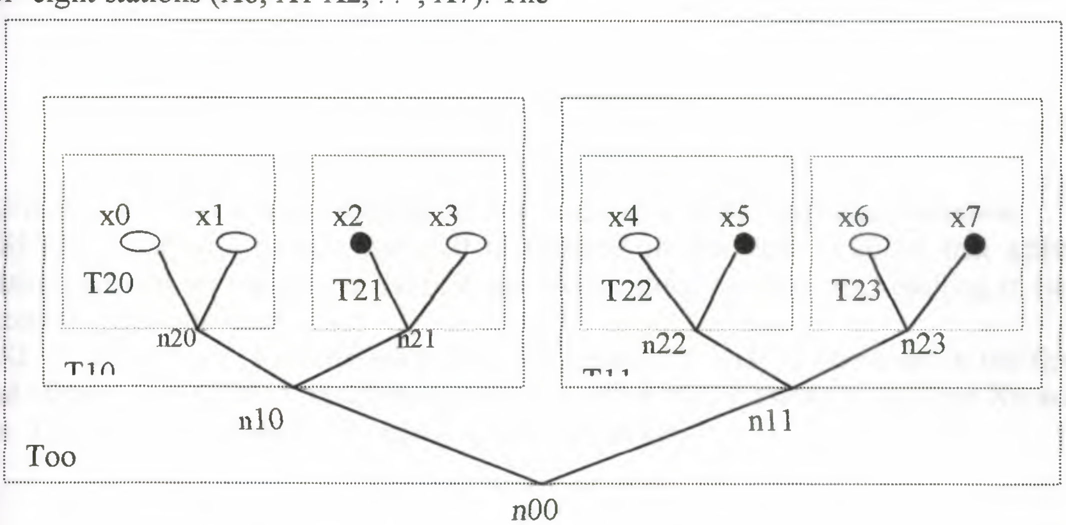
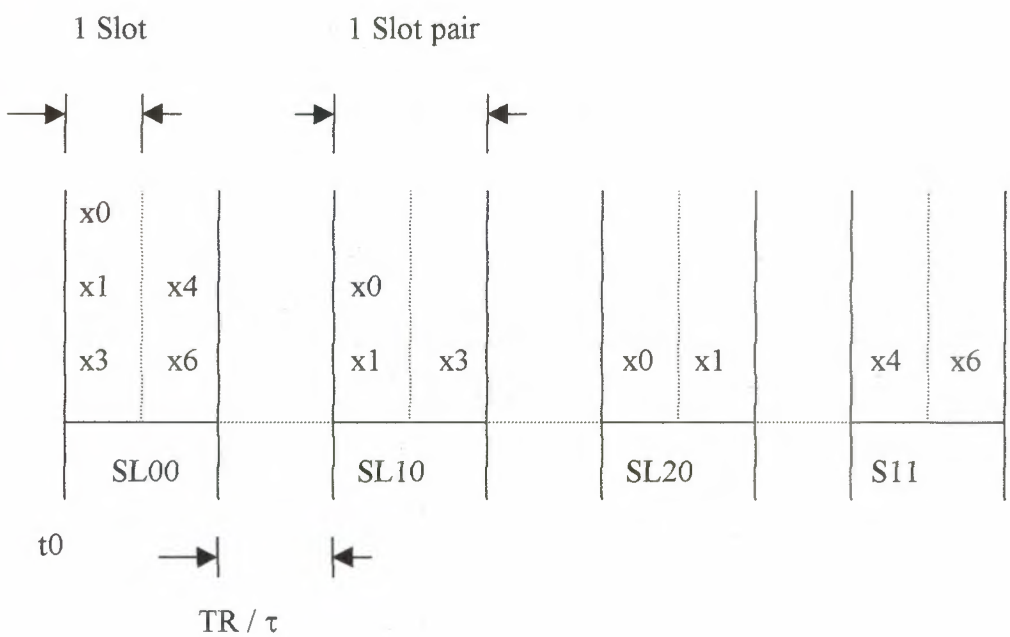


Figure: A binary tree

satellite channel time is slotted as shown in Fig. and the slots an paired. Each slot pair SLij is separated by TR / τ slots, where TR is the satellite roundtrip delay and τ is the slot length and both are in seconds. Note that $\tau = b/R$, where R is the channel capacity in bits per second and b is the packet length in bits. In the binary tree algorithm, the stations in the tree Tij transmit in the slot pair SLij. Now assume Ihat no collision exist until time to, the start of slot pair SLoo where stations X0, X1, X3, X4, and X6 become active. Then the binary tree algorithm takes the following steps in the indicated slot pairs:

SL00: All the active stations in tree Too (i.e., X0, X1, X3, X4, and X6) transmit is SLoo. Active stations in tree T10 (i.e., X0, X1 and X3) transmit in the first slot, and active stations in tree T11 (i.e.,X4 and X6) transmit in the second slot. This results in two collisions, one between X0, X1 and X3 and the other between X4 and X6. Since there is at least one collision in SLoo, new packets generated at these active stations will not be transmitted until the conflict is resolved.

SL10: Since there is a collision in tree T10, T10 transmits in the slot pair SL10. Active stations in tree T20 (i.e., X0 and X1) transmit in the first slot, causing a collision, while the only active station

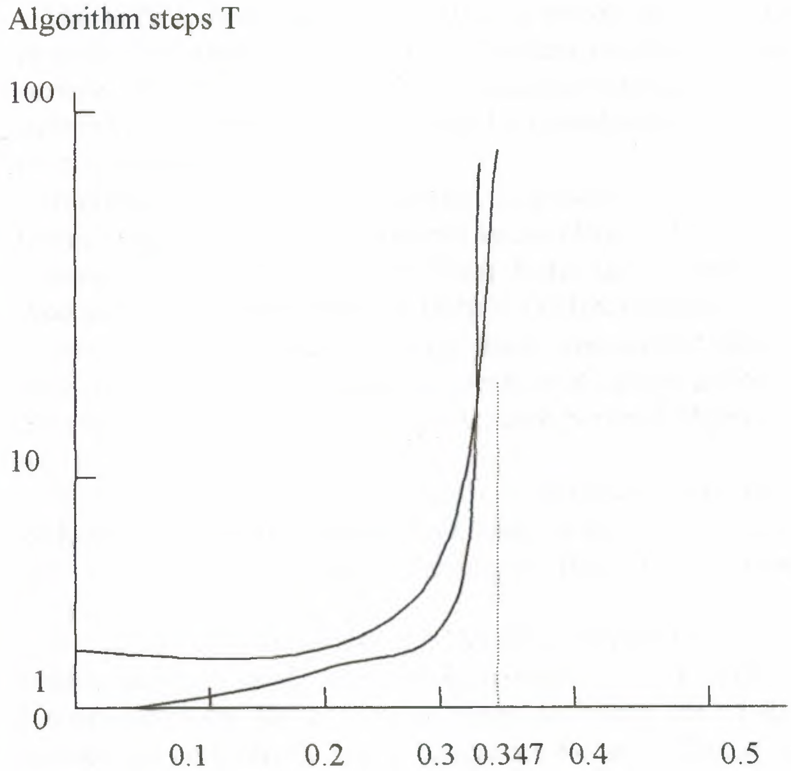


X3 in tree T21, transmits in the second slot, resulting in a successful transmission.

SL20: A collision causes tree T20 to transmit in slot pair SL20 so that active stations X0 and X1 transmit in the first and second slots, respectively, resulting in two successful transmissions.

SL11: Since there is a collision in tree T11, trees T22 and T23 transmit in the first and second slots of SL11, resulting in two successful transmissions by stations X4 and Xs. The process is repeated for newly generated packets.

For the binary tree algorithm it has been shown that the average throughput (the fraction of slots that contain exactly one packet each when the channel is at equilibrium) S' of the satellite channel can achieve a maximum of 0.347 packet per slot [6]. Note that at equilibrium the average throughput is equal to the average channel input rate. The average packet delay 7 versus the average channel input rate S for the tree algorithm is norm□tlly expressed in terms of the number of algorithm steps (an algorithm step is equal to a satellite roundtrip delay plus two slots). 'There is no closed-form expression for 7, however, 7 is characterized by its upper and lower bounds as given in Fig. 8.16.



Figure

CHAPTER III

ARCHITECTURE AND FEATURES OF A PACKET SWITCHING ISPBX

3.1 SYSTEM ORGANIZATION AND FEATURES

The ISPBX described is based on a set of modules- Interfaces and Servers, that provide the capabilities required for access to basic and supplementary ISDN bearer services and teleservices. ISPBX modules communicate through a packet switching subnetwork, that provides an highly modular system architecture, along with an efficicnt switching support.

Interfaces are modules oriented to provide external connectivity, both to local terminal equipment, or to a remote system like an ISDN exchange or another ISPBX. Examplcs of Interfaces are the Basic Rate Access Interface (BRI), the Primaty Rate Access Interface (PRI) and the ISPBX-ISPBX Interface.

Servets prowide Interfaces with shand specialized skills, required to offer service capabilities to communications.Examplcs of Server an the Directory Server, the X.25 Server, the Three Party Voice Conference Server (3PCS) and the Broadcast Server.

The ISPBX supports communications between local terminal equipment with an S Reference Point compatible interface, and offcis access to ISDN through a T Reference Point interface, accocding to the ISDN user-netcrork interface reference configuration.

For local communications the ISPBX provides ISDN-like baste and supplcmentary bearer services and teleservices,according to CCIIT Blue Book I Series of Recomendations. On communications accessing thc ISDN, the ISPBX provides a connection between S and T Reference Points, offering transparent acccss to ISDN based services and telesetvices to local terminal equipment.The call handling procedure appropriate to each situation is derived from the implicit or explicit dutination address.

The ISPBX supports on demand and permanent packet and virtual circuit switched calls, over standard ISDN Basic and Primary Rate interfaces. Although only 64 kbit/s B channel and 16/64 kbit/s D channel based communications are considered, the system architecture also supports H channel communications, provided that appropriate Interface modules are available.

All call setup procedures are based on logical to physical addresses convertion provided by the Directory Server. Incoming calls from the ISDN not providing identification of the destination terminal, must be routed to an operator compatible with the service capabilities required by the call. Voice calls, for example, not using Direct-Dialling-In (DDI) supplementary service should be routed to an human operator for approptiate handling.

X.25 calls an supported according to X.31 case scenarios . To accomplish to them, a X.25 packet handling function is available at the X.25 Server, proriding X.lS Data L.ink and Packet layers protocols support to B or D channel communications, between local, or local and remote, X.25 DTEs.

For local communications according to X.31 cases A and B, the X25 Server acts as a DTE-DCE-DTE relay, being active during all the call to support the non-simetrical procedure of X25 Data Link and Packet layers protocols.

On X.31 case B scenario or communications between Iocal and remote X.75 terminals wer B or D channels, the X.25 Server only acts to etablish and clear the virtual connection between S and T reference points, providing transparency for bidirectional information transfer between these two points.

The ISPBX may also providc a 64 Kbit/s Three Party Voice Conference Server, for support to voice conference among local or local and remote partners. This service *may only be requested and cleared by a local terminal participating in an already* established call.

The ISPBX may optionally include a Broadcast Server to provide information distribution services. A typical example is a 64 kbit/s digital music broadcast service, that may be used on voice calls to fulfill voice information gaps, originated by temporary call suspending procedures.

Concerning external connectivity, an ISPBX-ISPBX module is considercd for standard interfaring with other ISPBX,providing means for system upgrading and private networking. 'Iltis intreface is supposed to offer transparent service migration between ISPBXs, but because its features an still under study by Standardiration bodiu [1]. no further details on features or atchiterture are provided in this paper.

Along with the described services, the ISPBX may also provide teleservices based on dedicated optional servers. An example of such servers is a Message Handling Server. This server should provide an X.400 environment (or interpersonal message handling between local users, acting like a private Message Transfer Agent (MTA) for local communications. This server should require the availability of the X.75 Server for X.400 access based on X.25 network service.

All the ISPBX intermodule signalling is performed according to the Internal Signalling Protocol (ISP). ISP provides appropriate procedures and messages for resource allocation, as required to support service capabilities. Furthermore, ISP also provides a transport service to information packets concerning D channel packet communications.

The ISPBX's switching subnetwork is an extended version of a packet switching system based on a token passing parallel bus topology, already described as a switching approach to a LAN [2].

The shared commmunication medium is a passive parallel bus, comprising independent control and data lines. Bus mastering is allocated to nodes according to a token holding strategie. A node holding thc token, beyond being allowed to access the bus for packet transmition, is responsible for signal generation concerning both the packet transfer and the token passing procedure.

Packet transfer rate between nodes is a function of the number of lines of the shared data bus, and of the data clock used by each node a very conservative environment aed on an 8 bit bus and an 8Mhz clock, for example, provides a 64 Mbit/s instantaneous data transfer rats.

3.2 SYSTEM MODULES

The modules of the ISPBX are organized according to a general model comprising the following functional entities:

Communicating Agent (CA): internal or external entity acting as endpoint origin or destination of information; external CAs are not considered part of a module.
Channel Access Unit (CAU): providce the protocols required to support a CA.
Signalling Unit (SU): executes the intermodule signalling procedures, appropriate to the coordinate use and management of the physical and logical resources required to provide service capabilitis to a communication.
Switching Access Unit (SAU): provides a packet oriented access interfare (SAI) to the switching subnetwork.
Data Unit (DU): mailboxes used for information transfer between entities of a module.
Switching Access Interface Multi plexer(SAIM): provides shared acccss to the SAI.

Each module of the ISPBX offers communication support to a act of internal or external CAs, that an endpoints of generation and /or comsumption of information. CAs communicate through signalling and information channels. The CAU provides the protocols required to accomplish to each CA particular channel type.

According to this model, information transfer between functional entities is peformed through DUs. An unique address,in the address universe of the switching system, is allocated to each DU used, by an entity, as mailbox to receive information from the switching subnetwork, thus corresponding to a network (or physical) address of that entitiy.

The SAIM provides shared access to the SAI when a node includes more than one transmit or rereive channels. SAIM provides an access interface for packet transfer similar to the SAI. Whenever a node holds a token, the node's SAIM executes all pending packet transfer requests, on a first-come-first-served basis.

On broadcast transmitions the SAU always signals the SAIM of broadrast requests, also providing the relating addresses.

The SAIM then determines if any information or signalling channel is expeting information from the broadcast source identified by the address. This feature is based on a broadrast table offered by the SAIM, and maintained by the SUs of each module through an appropriate SAIM program procedure. When broadcast parket transfer occurs the SAIM provides its transmition to the appropriate DUs. The broadcast table at the SAIM may also provide permanent broadcast accpetance for packcts originated at special nodes.

According to the address structure of the switching subnetwork, each address holds node and channel identification. Node addess is an hardwired parameter of a SAU. The channel identification is normally an implementation dependent parameter, related to the address decoding mechanism used by the SAIM to address a particular DU.

DUs provide temporary storage capabilities for information transfer purposes between module's entities. Information store and retrieve is provided by appropriated access interfaces, chararteristics of each DU, resulting from the type of information conveyed.

DUs supporting B channels an organized on a FIFO basis, providing adequated packetizing and depacketizing support for information flow.The packetizing procedure includes packet ready to transmit detection based on FIFO occupation. Feature that is used to generate SAI's packet transmition requests.

DUs used for packet transmition to the switching subnetwork also include a Transmition Address Register (TAR), that holds the address corespondent to the next packet transmition. Access to the TAR is also provided over the DU input/output interfaces.

DU,s storage capacity is a function of packet Iength, channel rate and expected switching subnetwork access delay and transmition time. On 64 kbit/s circuit switched synchronous channels using the fixed delay procedure, a two packet storage capacity is required from DUs, both on transmition and reception paths.

3.3 EXAMPLES OF MODULES' ORGANISATION

The ISPBX modules Described in this section exemplifies now modules may be organized, according to the described general model, to provide the required system features. The modules presented are the Basir Rate Access Interface, and the X.25 Server.

Basic Rate Access Interface

The Basic Rate Access Interface (BRI) provides 2B+D basic access interfaces, for connection of terminal equipment at Reference Point S according with the appropriate CCITT Recomendations . The same general organization may be used on modules providing 2B+D ISDN interfacing at Reference Point T.

According to available technology, the organization of this module comprises one CAU by Basic Rate Interface.These CAUs, holding multiplaxing capabilities, provide the protocol handling for layers 1 and 2, required to support the two B information channels, and the D information and signalling channel.

Layer 3 protocol handling is provided to old Basic Access Interface by a SU, that uses the ISP service for intermodule signalling and D channel packet communications handling.

CONCLUSION

This topic presents the architecture of an highly modular ISPBX, based on a packet switching subnetwork. This topic includes a description of the system organization and features, and presents examples of organization of some relevant modules. Present work includes detail of functional provisions required to support basic and supplementary ISDN services and teleservices, namely a formal specification of ISP.

CHAPTER IV

Evolution of the SL-10 Packet Switching System

The evolution of the Northern Telecom SL-10 packet switching system closely parallels the development of public and corporate packet switching networks over past ten years. It is also closely linkedn to the successful introduction of truly universal standardsfor data comminications based on the CCITT x.25 recommendation and later CCITT and ISO standards. This paper outlines the steps in the evolution of the SL-10 which have made it a basis for the deployment and growth of many very successful packet switching networks in a large number of coutries. It shows how these steps have built the basis for networking products to meet the requirements of the 1990s.

4.1 INTRODUCTION

Ten years ago at this conference, the SL-10 packet switching system was demonstrated in Toronto. The system had been developed for the Canadian Datapac public packet switching network, the world's first network to be designed from the start as an x.25 virtual circuit network. The 1976 demonstration included switching nodes in Toronto and Ottawa, with an IBM host computer in Ottawa connected to the network via x.25. An interactive terminal PAD interface, the precursor to the then not yet standardized x.28, was also demonstated.

This paper descibes the evolution of the SL-10 packet switching system since that significant milestone in data networking. It follows the steps which were followed to make SL-10 a success in the Canadian Datapac public network. Datapac now stands as one of the largest and fastest growing packet switching networks in the world, a position which has been achieved while maintaining exeptionally good levels of reliability, availability,low delay , and high throughput.

The paper also covers the process which adapted the SL-10 to meet the requirements of international markets so that today, a single unified product line is serving as the national public packet switching network of five European countries, including DATAEX-P in Germany,and is installed and ready to provide service in Bell Operating Company and Interexchange Carrier networks throughout the United States. The capabilities which satisfy these public network requirements have also been deployed in large and important private networks, particularly in the banking and cooperate sectors, as the value of packet switching for these applications has become well-established.

4.2 ORIGINS OF THE SL-10 SYSTEM

The potential for packet switching was originally established in the ARPA network in the early 1970's. Bell-Northern Research, Northern Telecom's Research and Development company, explored this technology and the new HDLC link protocol with an in-house packet switching network. The network had minicomputer-based nodes in BNR laboratories in Toronto,Montrael, and Ottawa. It was described at the IFIP conference held in conjunction with ICCC 1974.

When Telecom Cananda decided in 1974 to build a nation-wide packet switching network, the designers of the NBR network were selected to build Datapac. The technology selectedfor the packet switch was based on the then new Northern Telecom SL-1 digital private branch telephone exchange. The proccessor and software technology were augmented by a uniquely new distributedmultiproccessor architecture, designed to meet the proccessing and memory intensive needs of packet switching.

The initial Datapac plan was to introduce a datagram service similar to that of the ARPA network. This would have required host computers to implement virtual circuit protocols end-to-end in order to ensure reliable data transmission. The network developed initially by BNR which was tested by Telecom Canada in a 2-city trial in 1975 was designed to supprt this type of service.

By the end of 1975 BNRand Telecom Canada planners had determined that a virtual circuit service within the network should be offered instead. Working with other data networking experts particularly from Britian, France and the United States they defined a new CCITTstandard for packet mode communications based on virtual circuit service. This was the original CCITT X.25 1976 recommendation . It became Data's Standart Network Access Protocol and BNR had to implement both X.25 and the virtual circuit capability in only 7 month. This culminated in the ICCC Toronto demonstration of Datapac and X.25 in July 1976.

After the ICCC demonstration in 1976, SL-10 switching nodes were installed in two additional Canadian cities for a field trial in late 1976. This trial included paticipation by a number of computer vendors who were developing X.25 network interface capabilities. It also included installation of the Network Control Centre node for real-time control, billing and statistics data collection, and software and configuration data downloading.

After the field trial, a few necessary enhancements were identified before commercial service could proceed.These included the async interface for which the CCITT X.3/28/29 recommendation was then being prepared. There were also some statistics, billing, and security enhancements needed before full public operation. These were accomplished early in 1977 and Datapac entered tariffed service in June 1977.

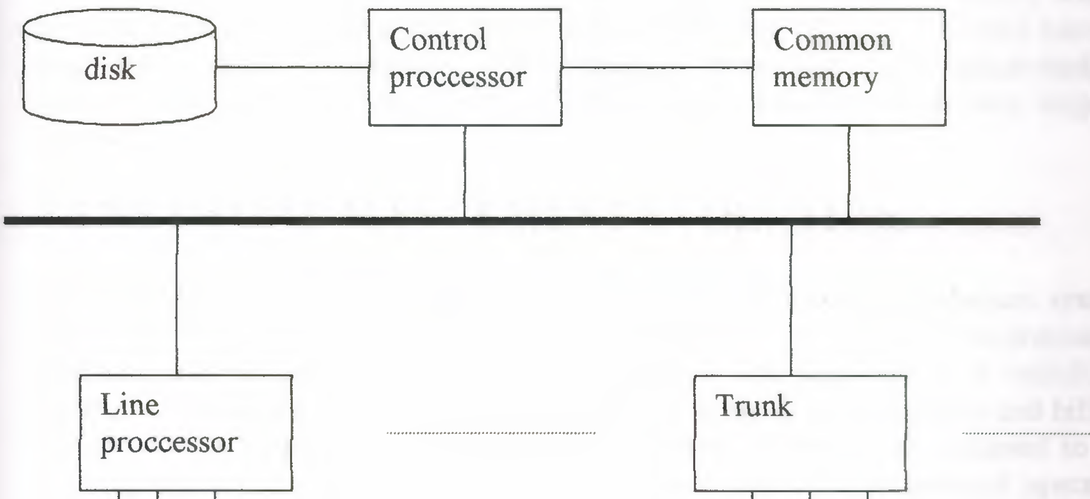
4.3 VIRTUAL CIRCUIT ARCHITECTURE

The SL-10multiprocessor architecture was originally designed to provide for a pool of processors to implement the datagram protocol and packet transfer functions. Other processors on the same common bus would provide large numbers of line interfaces and would relay user frames directly to and from a common message memory with minimal processing. The introduction of virtual circuits created a problem by requiring that common memory also be dedicated to holding the status information for each virtual circuit and that user packets be retained in common memory until the far end of the virtualcircuit acknowledged the transmission. This created a switch bottleneck which could have limited the planned Datapac network growth.

Therefore, the SL-10 architecture was changed in two stages to accommodate many virtual circuits and much higher throughputs. This work also included an improved operating system and a softeware architecture to support 24-hour 7-day service without interruptions when subscriber interfaces were installed or subscribed facilities were changed. This architecture was described at ICCC in 1978. It was also designed of many new interface services in the succeeding years, starting in 1978.

The key to the new architecture was that virtual circuits were moved to the Line Proccessors, each with its own memory address space, not constained by the capacity of the common memory module. The multiple proccessor which had originally performed the datagram and later the virtual circuit, protocol functions became Control Proccessors. Their function was to performed control functions including automatic topology determination, routing table maitenance, incomming call packet routing, subscription data management, and interface to the network control and management system. However, the Control Proccessor were no longer on the path of data packets.

Virtual circuit were implemented only in the endpoint Line Proccessor modules. The common equipment of the endpoint nodes and the transit nodes did not retain any status information for specific virtual circuits. An end-to-end internal protocol was designed which would request retransmission within the network if necessary and would ensure the non-lossy, non-bublicating,ordered properties of the virtual circuit service. The ability of the network to instantly alter a route at any switch or transit node without virtual circuit resets or any need to update path information provided a very highly reliable service. The line proccessors and the proccessors which served high-speed internodal trunks could route packets to each other directly without any common proccessing bottleneck. This ensured very high throughput, low delay and almost no impact from transit traffic on intermediate nodes as the network grew.



The internal protocol avoided the problems of pure datagram implementations by using abbreviated addresses in packet headers once the call was established and by never routing successive packets via different trunks unless a topology-affecting failure had occurred. This eliminated packet proccessing delays because no reordering at the destination would be needed. The architecture developed at BNR was published at a number of important conferences in 1979 and in a special IEEE issue of Transaction on Commin\u0131cationsin 1981. Its appropriatness for packet switching was shown by the use of very similar end-to-end virtual circuit protocols in a number of systems devloped later by other vendors and by the rework of many other packet switching systems based on single proccessor architecture, to a similar multiproccessor virtual circuit architecture.

4.4 PROCCESSOR PERFORMANCE IMPROVEMENTS

The next step was to redsign the proccessor module to povide much greather troughput and memory capacity. Packet switching and particularly virtual circuit operation is both demanding of memoy and of proccessor capacity for both call setup and data transfer. This differs from PBX teleohony requirements were these resources are used for call proccessing but not for the actual user information transfer. With the virtual circuit software moved to the same proccessor as the physical subscriber interfaces, it was necessary to greatly increase the proccessor speed and the memory addressability and capacity.
The new proccessor was designed using bit-slice technology with a much larger address space. It programming language as used in the SL-1 proccessor. This meant that all existing software could be readily moved to the new proccessors. The new proccessor were installed incrementally in to SL-10 switches and because they could co-exist with the old proccessor, the migration to higher switch capacity was never disruptive. The new SL-10 proccessors high performance design was subsequently also incorporated into Northern Telecom's SL-1 and DMS-10 products to increase capacity for voice call proccessing.
The larger address space was utilized by the introduction of denser memory boards, providing instages first a 2-fold, then a 8-fold, and most recently a 32-fold increase over the original components. These changes have also been accommodated incrementally with free intermixing of old and new board types at each stage of growth.

4.5 NETWORK MANGEMENT EVOLUTION

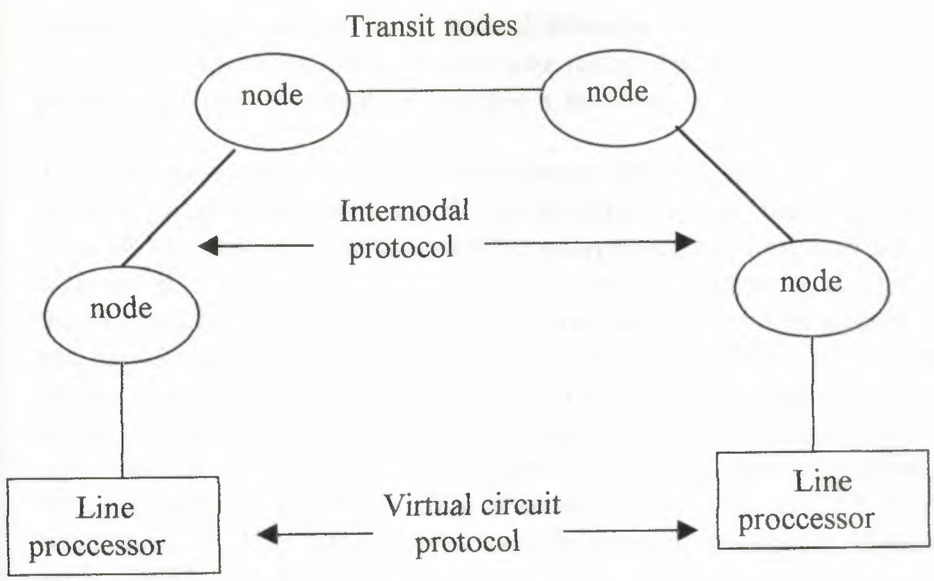
The initial SL-10 system included a complete network contol and mangement system. The real-time contol and surveillance element used the same multiproccessor hardware and software as the switching nodes. It was supprted by a mainframe computer running database applications to support network administration and billing. The network administration and maintenance system (NAMS) was designed to use network facilities as much as possible to achieve a true network oriented operation with both centralizesd and distributed functions. Operator terminalsall connected via the standard Async (x.28) interface augmented by a man-machine interface implemented in both the switching node and Network Control Centre software. These terminals emulated full screen displays high data rates (9600 baud) to give timely over views of network component status, traffic and alarms. Communication between the control centre and complementary software function in the switches were by standard virtual circuits, internal to the switches. These virtual circuits were used for both the collection of billing data, alarms and statsics and for the distribution of new generic softwre and subscription data. The only hardware added for network control purposes was a disk unit to temporarily spool the collected until it could be moved by an automatic file transfer procedure to the Administration computer via and x.25 circuit.
The operations, administration and maintenance capabilities of SL-10 have been enhanced frequently since the original deployment of the system. When SL-10

networks grew very large, it was necessary to divide the automatic network control functions such as billing data collection among multiple processing sites. This was achieved first by configuring more than one Network Control Centre in the network to divide the workload. Second, the data spooling function which collects billing and statistical data was decentralized to multiple switch sites in the network, each equipped with disk to support this function. Finally, the Network Control System was redesigned to support both fully distributed and centralized organizations. It now has the capabilitiy to support complex organizations with functions shared or partitioned among central, regional or local sites.

The growth of public packet switching networks over the last ten years has been dramatic. SL-10 and its management system have been enhanced continuouslyfacilitate this growth and at the same time provide steadily improving quality of service and performance. As an example, Datapac, which was originally planned as a 14-node network, has now grown to over 55 nodes in 17 cities. It serves the Canadian public through over 90 exchanges spanning over 5100km. Each of the nodes now has over five times the capacity of the original SL-10 equipment, so the actual increase is more than 17 times larger than the 14 nodes originally planned as Datapac's maximum size. There are now over 20.000 physical circuits connected to Datapac, carrying over 1.5 billion subsciber data packets on 10 million virtual calls every month. More than 40 percent of the circuits are now synchronous, the great majority being x.25 and this proportion is growing as the benefits of packet interfaces are realized and as the availability of x.25 from computer and PAD vendors becomes universal.

4.6 TOPOLOGY AND TRUNK CAPACITY EVOLUTION

To reach such capacity levels, more than just proccessor upgrades were needed. The design of high performance networks with minmal delay was undertaken using high-speed trunks throughout,a mesh topology,and strict minimization of end to end trunk hops by avoiding the use of dedicated tandem nodes. Initially 56 kbit/s trunk were used for all subscriber internodal traffic and the number of nodes increased, it was necessary to increase the number of high speed trunk proccessors accessing the SL-10 common bus in order to permit network design with these constraints. Maintaining this maximum hop objective as networks grew has ensured excellent end-to-end delay performance.



End-to-end Virtual Circuit

The traffic per trunk also increased dramatically. SL-10 was designed from the start with an internodal multilink capability in which parallel trunks could serve all packets on the same route. However, the preffered solution to add capacity on a route was higher speeds. Telecom Canada introduced 112 kbits/s digital trunks in 1985 and other countries, such as Germany with 128 kbit facilities, were expected to do the same. Therefore a higher performance trunk processor was developed in 1984. It supported not only the 112 kbit/s and 128 kbit/s intercity speeds but also even higher speeds between colocated switches. Its link level protocol was designed to achieve *high throughputs on long-delay satellite facilities. This required augmenting HDLC* with larger sequence numbers and a multiple selective repeat capability to avoid delaying frames due to full windows or to preceding retransmissions.

The SL-10 network routing system is fully automatic and adapts to changing topology as well as trunk or switching failures. It always finds the lowest delay primary and alternate route between any node pairs. For greater effectiveness in the deployment and use of network trunks SL-10 now allows the network planner to direct the routing system to favour specific routes of equivalent delay. Interactive tools were developed to support planning and analysis of complex network topologies. These made it possible to define and simulate future topologies for the largest SL-10 networks. The traffic flows in these future views were projected and the impact of topology changes and trunk or switch failures on traffic and on routing update messages were studied. Such planning makes it possible for SL-10 networks to continue to grow far beyond their present size, while at the same time handling increased traffic from existing subscribers.

4.7 EVOLUTION FOR MULTIPLE NETWORKS

The early success of the Datapac network and the SL-10 system on which it was based created opportunities outside Canada. In 1978 SL-10 was selected by the Deutsche Bundespost for a packet switching network trial in Berlin. After this trial plans went ahead for a German national packet switching network and SL-10 was selected for this network. At the same time in 1979 SL-10 was installed for a Belgian bank and in the following year was selected for the backbone packet switching network of the United States federal Reserve System. These networks are now all large networks in the 20 to 50 node size range. For example Datex-p in Germany has experienced growth equal to Datapac's but with a higher proportion of high-traffic X.25 lines.

These network had a common requirement for high speed and high throughput X.25 service. To meet this need the SL-10 architecture was augmented with an additional layer of processing. A dedicated 8-bit microprocessor line card was implemented for high speed X.25 or X.75 lines with direct memory access to the local memory bus of the SL-10 Line Processor. This device was used for 56 kbit/s lines in North America and for 48 kbit/s in Europe but was also tested at speeds of 72 kbit/s and greater.

Differences in the requirements of each of these networks required a number of software changes. For example, each network chose a quite different numbering plan with various address lengths, different area code formats, different international prefixes and different area code structures. Even the CCITT X.121 International Numbering standart allowed two country/network designator formats and both were used with SL-10 networks.

It soon became apparent that supporting different software streams for each network was becoming costly and time-consuming. It also delayed the universal availability of new features on SL-10. Therefore in 1980, a single streamgeneric approach was adopted in which all SL-10 networks would be supplied with the same generic software. National and network requirements would be met by a wide range of configuration options. Most of these could be selected and changed through the normal mechanism used for subscription data management. A Network generation tool was developed to allow the more basic configuration choices to be made by each administration as well.

The single stream software generic and the interworking capabilities of SL-10 facilitated the establishment of additional national networks in 1980 and succeeding years. There are now five such networks in Europe including Austria, Switzerland, Ireland and Portugal. In the corporate networking sector, SL-10 has now also been supplied to five major banks, with installationsin Britain, Belgium, the United States and Hong Kong. In adition, many corporations in other sectors, such as manufacturing, utilities, transportation, and communications have committed their data communications to dedicated SL-10 networks or to the public networks which use SL-10. The network customization and configuration options and tools allow the same single stream generic to meet the needs of these networks as well.

4.8 NETWORK INTERCONNECTION

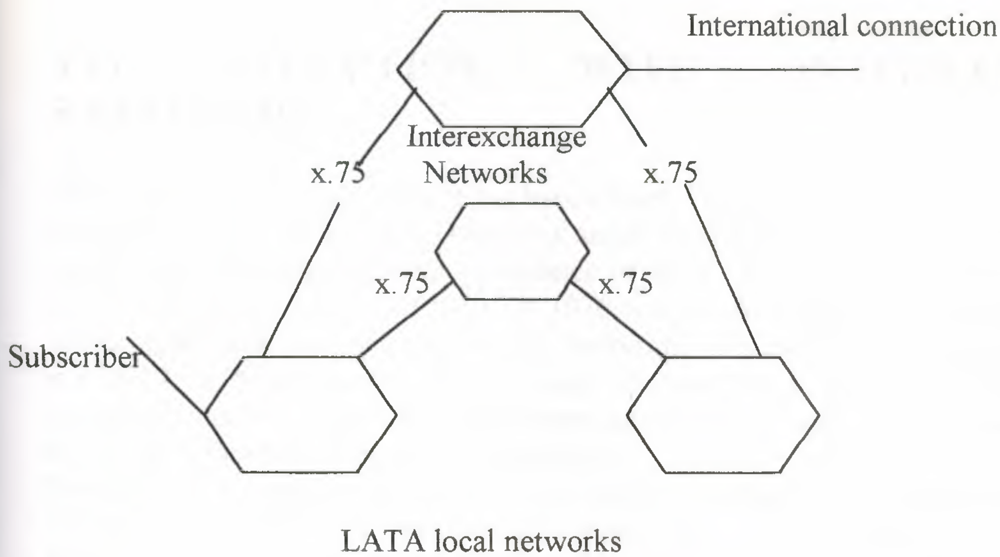
SL-10 was used to pioneer the interconnection of packet switching networks in different countries. At the IFIP Conference in Toronto in 1977 an experimental transatlantic connection of Datapac with RCP, the French predecssor network to Transpac, was demonstrated. Datapac co-operated with Telnet and Tymet to establish cross-border x.25 gateways in 1978 and these were described at ICCC. This experience led to the definition in 1978 of CCITT recommendation x.25 to meetthe additional interworking requirementsof public networks. This service was implemented on SL-10 in 1979 and used initially to establish the world's first x.25 link, between Datapac and Telnet, also described at ICCC. X.25 was implemented as a line proccessor service on SL-10 and therfore a full range of speedswas available using normal or high speed line cards. The service required enhancements to the routing system so that x.75 gateways to other networks could be installed on one or more nodes with the best route chosen based on distance, traffic, and failures.

Some countries established separate gateway network for their overseas connections. For connections outside North America, Canada's international carrier, Teleglobe selected SL-10 in 1980 to be a gateway switch carrying transit traffic. SL-10 supported this configuration as well as direct international connections to the national public network. Combining international and national routing in a single network presented some interesting challenges. For example different routing tables were used for international Data Network dentification Codes and for on-network areas codes. However the same underlying subnetwork routing mechanisms were used to ensured that x.75 calls had the same reliability and throughput on the SL-10 network portion as did normal calls. The gateway point however was a bottleneck and a single failure point and therefore load sharing and alternate routing mechanisms were provided at call setup time to take advantage of multiple equivalent routes. The later addition of multilink to X.75 provided SL-10 with a way to increase the reliability of internetwork calls, after call establishment.

4.9 CHALLENGES IN THE UNITED STATES

Since 1983, a new set of challenges for the SL-10 product have emerged in the United States. The divesture of the Bell System in to AT&T and separate regionally-based Bell Operating Companies has resulted in a unique new environment. Equal access legislation results in a situation where the 22 Bell Companies (BOCs), grouped into seven regions were required to establish logically separate networks in over 160 geographic local areas (LATAs). Customers are allowed to select the separate interexchange carrier of their choice, by subscription or per call, whenever a call crosses LATA boundaries. Neither the X.75 and X.121 standarts, nor the existing packet switching products had anticipated such a requirement.

SL-10 was selected initially to provide BOC public packet switching service in six networks in four of the seven regional Bell Operating Companies. Most of these are multi-LATA networks.To accommodate the equal access requirements SL-10 now support use by subscribers of the RPOA transit network selection facility as a per call or subscription option.A BOC can offer network services across LATA boundaries by of such facilities as "international" closed user group, reverse charging, fast select,and permanent virtual circuit facilities, provided via an interexchange carrier.Integration of network billing using the telephone network's Automatic Message Accounting formats has also been added to SL-10. These and other BOC-specific capabilities were described at the ICC-85conference.



LATA Network Internetworking

AT the same time as its deployment in local Bell networks, SL-10 has been supplied to a US interexchange carrier, MCI, to connect these networks. The fragmentation of the BOC system into many local networks has created special problems for numbering plan, routing, and X.75 circuit provisisioning in these networks. SL-10 addressed these problemswith a number of new features which made possible the first demonstrations of multi –LATA interworking in the US. These were organized by Pacific Bell at the TCA Conference in San Diego in the fall of 1985 and by Bell South at the Interface Conference in Atlanta in the spring of 1986. Using SL-10 equipment in multiple LATAs and in MCI's interexchange network, it was demonstrated that internetworking in the new BOC multi-LATA environment was feasible. X.75 made possible links to other vendor system and to international networks.

4.10 COST EFFECTİVE INTEGRATED ACCESS

The advent of powerful microprocessors and the high capacity of modern memory devices have made attractive the dispersion of processing of network access protocol and multiplexing of customer traffic. Significant transmission cost economies can be achieved with remote PAD and concentrator devices. The real challenge for large packet switching networks, however, is to keep the network manageable as such devicesprolifreate and to ensure that all usres obtain the same high quality of service, consistency of access procedures and availability of network features.

The first multiple micrproccessor access concentrator for SL-10 was developed for the US Federal Reserve System in 1980. It provided the basis for designing a network integrated concentrator which was delivered in 1984 to provide both cost-effective remote access and a wide range of protocol support both for local and remote access. This access technology is described in another paper at ICCC-86. The software architecture has been improved but the basic concept of multiple proccessors exchanging messages via common memory has been maintained. The software investment in SL-10 services has been preserved by developing compilers for the original SL-1 high-level block-stuctured programming language. Transporting the SL-10 service software to the new device ensures that the user sees a consistent interface and identical services.

4.11 EVOLUTION WITH INTERNATIONAL STANDARDS

The 10-year evolution of SL-10 has been closely linked to the international standards proccess. In 1975 BNR and Telecom Canada staff made major contributions to the development and international acceptance of the CCITT x.25 recommendation. The SL-10 x.25 demonstration at ICCC in 1976 was the first carrier. Since then BNR has made significant contributions to the definition of new recommendations such as x.3/28/29, x.75, and x.121, the evolution of these and x.25 in the 1980 and 1984 revisions, and most recently the development of the x.32 dial-access standard and the ISDN packet interworking recommendation.

Many of the features of x.25 1984 were initially proposed and implemented by BNR for SL-10 to meet the commercial requirements in countries operating SL-10 networks, prior to their standardisation. Examples of these include Call Redirection, Hunt Groups, Charging Information and Network User Identification, which were added to X.25 in 1984. When implementing such additional features it is essential to continue to support existing X.25 devices and not to make incompatible changes to the existing services. At the same time it is desirable to make new capabilities available to as wide a subscriber base as possible by not requiring new signalling by DTEs in order to use new features. In SL-10 this is supported by a very rich set of subscription options. They allow each port to be configured so that new facilities can be blocked or made transparent to old DTEs. On the other hand, new facilities may be accessed on a subscription basis if the DTE is unable to perform the signalling associated with their use on a per call basis.

SL-10 has implemented computer industry standards, as well as the CCITT defined services, to allow terminals supporting other protocols to use packet transport and to access to X.25 hosts. Some of these, such as the IBM 3270 display system and 3780 contention-mode bisynchronous protocols have been developed in to effective international standards by cooperation among such public networks as Datapac, Telnet and Tymnet. The multi-service architecture of SL-10 has provided these standard services for end-to-end transparent communication (using both a terminal and a host PAD) and for access to x.25 hosts which support the higher- level protocol. Additional services have been implemented to provide asynchronous polled support for point of sale terminals, and multileaving support for remote job entry systems. Finally, the IBM definition of an end-to-end signalling protocol QLLC for efficient support of SDLC peripherals over packet switching networks has been implemented on SL-10 to provide cost-effective transport functions to IBM SNA networks.

4.12 CONCLUSION AND VİEW TO THE FEATURE

The SL-10 packet switching system pioneered public packet switching and it has grown with the first and largest of the world's networks through the first ten years of X.25 packet switching. The technology of SL-10 has changed to keep pace with advances in processors, memories and transmission facilities and this change has always been available to upgrade in-place network switches in a non-distruptive way. However, the original highly innovative multiprocessor architecture, the end-to-end virtual circuit implementation and the dynamic adaptive routing system have proved very succesful and have adapted well to all new demands.

In September 1985, Northern Telecom announced the DPN Networking System. The architecture and characteristics of the DPN are described in other papers at ICCC-86. The principal characteristics of DPN are very high network availability, based on highly redundant architecture and high reliability component designs. A broad range of switch and network sizes are supported cost effectively with full integration of all network elemnts for control, administration and subscriber service. This new product line builds on the success of SL-10and provides a compatible evolution path for packet networks that can grow in to the 1990s and beyond. As with the changes that have been made to SL-10 over the years, the new products maintain software, network and service compatibility with existing SL-10networks.

DPN systems were first deployed in the existing Southern Bell SL-10 Pulselink network and then into two new BOC networks. The DPN Data Networking System will continue the evolution of high-performance and reliable packet switching technology begun ten years ago with SL-10.

CHAPTER V

MULTICAST COMMUNICATION FACILITIES IN A HIGH SPEED PACKET SWITCHING NETWORK

Abstract:

This paper discusses the multicast communication facilities in a High Speed Packet switched Network (HSPN) that can accommodate various kinds of terminals such as telephone sets, high speed data terminals and video terminals.

First, the architecture of the HSPN is discussed. Since packet size in the HSPN is determined for each type of information media and allowable network delays in each media are different, packet ordering control with priorities is required. Some packet multiplexing, which allow high-priority packets to interrupt long, low-priority packets, are proposed.

Next, multicast communication facilities in the HSPN are discussed. Our proposal is to provide those facilities as functions of the datalink and the network layers, instead of the session layer. Each packet has a special address part in which multiple addresses or grouped addresses are assigned. Packet switching nodes recognise such addresses and duplicate packets for following nodes. With these facilities, real-time broadcasting of information can easily be provided. The control mechanism for the multicast packet communications in the HSPN nodes is also described.

Finally, an information service using the HSPN's multicast communication facilities is presented. High speed packet transmission and an intelligent selection mechanism enable users to obtain abundant information at low cost and with short access time.

5.1 Introduction:

In recent years, the demand for diversified telecommunications services has been growing rapidly. With regard to communication schemes, not only point-to-point but also point-to-multipoint and multipoint-to-multipoint communications are coming more into widespread use. At the same time, many different kinds of information media, such as text, voice, picture and video are being transmitted on telecommunication networks.

In order to support to the above-mentioned communication demands, a High Speed Packet switched Network (HSPN) is being developed in our laboratory. The HSPN can exchange and transmit high speed packets at more than 10 Mbit/sec, so that various types of information including video can be accommodated in the network. Each node in the HSPN has an ingenious mechanism for exchanging packets with unfixed lengths and priorities.

Multicast communication (i.e., point-to-multipoint or multipoint-to-multipoint) facilities are other important elements of the network layer in the network, and network delay is not dependent on the number of destinations. Through them, various kinds of information providing services such as high-grade videotex can be provides at low cost and with quick response.

In this paper, we discuss the architecture of the HSPN and propose multicast communication protocols for it. The principal items described in this paper are :

1)Packet multiplexing methods with priorities
2)High speed packet processing mechanism
3)Multicast communication protocol and its control mechanism
4) Realisation method of multicast mode communication in arbitrary mesh network
5)Some applications on multicast communication

5.2 ARCHITECTURE OF HSPN

5.2.1 Why The HSPN?

In many countries, packet switched networks are being developed as the infrastructure of business communication systems. Particularly, considerable computer communication trafficis borne by the packet switched networks. Then, can packet switched networks satisfy all kinds of communication demands? Under the present conditions, the answer is negative. Major problems involve network delay and access line speed. The typical network delay in packet switched networks is several hundred milliseconds. This is too long to support conversational telephone service. The maximum access line speed of almost all existing packet switched networks is at most 64 kbit/sec. Consequently, terminals of more than 64 kbit/sec can not enjoy packet switching services.

In spite of those conditions, the current trend in telecommunications suggests that,in the near future, demands for mixed-media communications will increase and networks will be required to accommodate various kinds of terminals. Generally, for the purpose of transmitting various speeds of information on a transmission line, the statistical multiplexing method is more efficient than time division multiplexing. Therefore, if packet switched networks can accommodate wide-band terminals, such as video terminals, and network transmission delays can be shortened to less than one hundred milliseconds, almost all demands can be satisfied by packet switched networks more economically than by circuit switch based ISDNs.

Transmission delay in the packet switched networks depends on packet length and transmission speed. Shortening packet length can reduce transmission delay, but at the same time,it increases switching cost. On the other hand, if transmission speed is increased, shorter transmission delay is possible at the same cost. Furthermore, high speed transmission can improve transmission efficiency, since the capacity of transmission lines increases in proportion to their speed.

5.2.2 Design Principles Of The HSPN

The HSPN is an advanced packet switched network that will provide wide-band, shortly delay packet communication services. Design objectives are up to about 10 Mbit/sec for subscriber access rate and are less than 100 msec of network delay. This enables the HSBN to accommodate various kinds of terminals including telephones, precise image terminals and teleconference terminals. In order to realise such as short network delay and high tranmission efficiency packet transmission speed on trunk lines must be about 100 Mbit/sec. At present a laboratory model of the HSPN is under testing to verify the basic technology.
The design principles of the HSPN are as follows:
New multiplexing methods for mixed-media packet transmission and switching :

- they include the distance-indexed frame multiplexing method and preemptive priority packet multiplexing method.
- Simplyfied protocol for short delay transmission: no retransmission error control is applied to voice and image packets.

Multicast communication facilities in the network layer: multicost communications are provided without long delays or extra storage.
Under these consideration, we can construct the HSPN from two main elements. First is "core network" that handles high speed simplified packet transfer. Second is "edge interfaces" that convert to various user interface protocols with simplified in-core protocols.
The basic architecture of HSPN with these two elements. The core network is made from high-speed packet switching units with hardwired control and link control sub networks. The edge controllers are made from user interface controllers and protocol converters.
The core network is simplified switch network without error checking or retransmission control. On the other hand edge interfaces communicate each other through core network to sattisfy the user on-demand quality requirements with inter-edge protocol.

5.2.3 Packet Multiplexing Methods And Protcol

The current packet communication protocol, X.25, is standardized as the protocol for data communications and offers high transmission quality and efficiency. However, for some kindsof communication such as telephone and large file transfers, the X.25 protocol is too complicated and inefficient. Its more than 500 msec network delay, for example, is not tolerable for telephone conversation. As a protocol for the HSPN, therefore, a more simplified protocol is required to emulate quasi-transparent, real-time transmission. In addition to this point, packet flow control with priorities is essential to mixed-media communications, since a FIFO policy may cause extremely long delays when large-sized packets occupy the packet buffers andtransmission links. Two kinds of packet multiplexing methods (protocols) for the HSPN were proposed and studied. Both are being examined in the experimental system.

5.2.4 Distance-Indexed Frame Multiplexing Methods

The distance-indexed frame multiplexing method is a fixed-length frame multiplexing method similar to TDMA (time division multiple access).Information strings are divided into frames, each of which consists of a frame header and an information field and are multiplexed on the basis of the frame. Each frame, derived from one information string, is linked to another by a distance index, which indicates the number of frames between two linked frames, each frame's location can be reallocated in a frame stream so that urgent frames can interrupt the frame stream and be multiplexing within one frame period. (When the frame length is 128 bytes and the frame stream bit rate is 100Mbit/sec, one frame period corresponds to 10 microseconds.) Compared with other methods such as the direct addressing methods and the labelling method, the distance-indexed multiplexing method has the following advantages:

1. The distance index is independent of the number of subscribers. A large number of subscribers can be accommodated, but the number of connections at any time is small.
2. Unified label assignment to all multiplexed frames is unnecessary.

Very long but not urgent data strings and short but urgent strings can be transmitted on the same link, for instance, a mixture of facsimile and conversational voice.

5.2.5 Preemtive Priority Packet Multiplexing Method

Another possibilty is the preemtive priority packet multipexing method. The protocol of this method is based on the X.25 protocol, but it is considerably modifed for the HSPN.

Characteristic items of the modified protocol are:

1. Packet size is not fixed and it extends up to over 8 Kbytes so as to optimize packet size according to the characteristics of the media.
2. Preeptive packet ordering can be applied. High priority packets can be sent prior to low proirty packets, even if low priority packets are being transmitted on the links. A preemted packet is divided into two packets and the latter packet is sent out after the high priority packets.
3. The error control examination bounds can be limited to headers, so that those packets that require a short delay but are not affected by a few bit errors, for example voice data packet, can be transmitted efficiently.

The packet format of the preemptive priority packet multiplexing method. Text for low priority packets can be divided into several blocks. A division control field indicates whether this packet has been divided by high priority packets or not and error check control field indicates error check boundary.

5.3 Multicast Communication Facilities

5.3.1 Needs on Multicast Communication

Up to now, almost all communication on telecommunications networks have been carried out on a point-to-point basis. This fact, however, does not imply that there are few needs for multicast communications. A typical point-to-multipoint communication is broadcasting. In addition to mass communications like broadcasting, personal or business needs for multicast communications are rapidly increasing as social activities extend over a broad area. Typical examples are multicast delivery of messages/facsimiles and and a teleconference. Nevertheless, current telecommunication network do not have adequate functions for efficient multicast communications. Therefore, it is expected that telecommunication network will be drastically improved or even reconstructed.

5.3.2 Multicast Communication in Packet Switched Networks

In general, point-to-multipoint communications are one-way and can tolerate much more than the delay in ordinary packet switched networks. Therefore, a packet switched network, which has a store-and –forward mechanism and can easily duplicate packets, is more suitable for point-to-multipoint communications than a circuit switched network. On the other hand, multipoint-to-multipoint communications like teleconferences are often two-way and require a delay quality equal to telephone circuits. Existing packets switched networks do not satisfy such requirements, but it is not difficult for the HSPN to meet them. Since the HSPN is suited for mixed-media communications, teleconferences using various media such as voice, picture and video equipment can easily be performed over the HSPN.

5.3.3 Multicast Communication Facilities in The Network Layer

At present, multicast communication through public networks are in most cases, caried out by using the extra storage and higher layer protocols of the OSI model. On the contrary, the HSPN has multicast communication facilities in the network layer and provides real-time multicast communication facilities in the network layer and provides real-time multicast communication services. The advantages of the method in the HSPN are :

1. Multicast routing can be performed at the same speed as normal (point-to-point)routing.
2. Multiple addresses are directed in the same manner as network layer addressing.
3. No extra facilities for multicast communications are required.

5.3.4 Routing Table Alteration for Multicasting

Up to present time, multicast mode communications have been used in the networks using *multicast transmission media including* Ethernet type *common* bus system or wireless broadcasting. These networks have some advantages in network utilization efficiency and data transmission delay because of multicast characteristic of tranmission media itself.

According to the increase of communication speed, other type of network such as loop, may be available. In the case that we can use the HSPN mentioned in above section, to adopt cascade transfer of multicast data packets on node is not impractical in transmission delay any longer.

To release the restriction of multicast network topology such as common bus or loop, we suggest the method that make arbitrary mesh networks to those multicast networks equivalently. We use routing table to decide a route from a node A to other node B on mesh networks. Here we propose the strategy to make this routing table efficiently to perform multicast network topologies. One of these approach is to use the fixed routing table. However, this method is not so flexible to network traffic fluctuations. To improve this point, we consider the dynamic alternation of this routing table using specific rules.

In this paper, we suggest two generation rules for small delay multicasting (SDM) and for small traffic load multicasting (STM). The generation rules for the SDM and the STM are logically correspond to common bus network topology and loop network topology repectively.

The generation rules for small delay multicasting (SDM scheme) is as follows :

1- A multicast information source node (M-Node) sends R-packets (Routing table creation packets) to adjoining nodes simultaneously.
2- Receiver-Nodes (Nodes received R-Packet) sends same R-Packets simultaneously to all adjoining nodes except for sender-node (node that sending the R-Packet).
3- Receiver-Nodes send ACK reply back to the sender-node when no adjoining nodes were found or all adjoining nodes send back some reply.
4- Receiver-Nodes send NAK reply back to the sender–node when the same R-Packet has been received from other adjoining node already.
5- M-Node knows the completion of routing table creation when all adjoining nodes send back some reply. And each node registers adjoining nodes, which sent back ACK reply, to its routing table.

Using this routing table, we realize multibranch topology network with shortest transmission delay under that traffic conditions.
The generation rules for small traffic load multicasting (STM scheme) is similar to above mentioned rules except for next one point. Each receiver-node sends R-Packets for adjoining node not simultaneously but sequencialy. That is, each receiver-node sellects one adjoining node and sends R-Packet and waits for reply. After receiving some reply, the node repeats the similar sequences repeated to the other adjoining node, and so on.

Using this routing table, we realize thread topology network. That is, the network adopt this routing table is a graph drawn with a single stroke of the brush.
With these two routing tables, we can control multicast data transmission delay and traffic load.

5.3.5 Multicast Communications Protocols

The major technical items for implmenting multicast communication functions are the addressing methods are the routing algorithm. For discussing those items, it is helpful to classify multicast communications into two categories. One is destination-designated multicast communication suc as multicast delivery of facsimiles and the other is receiver-selective multicast communication such a teletext.

In the former case, a grouped addressing method is effective to designate a large number of destinations. A call set-up packet, including a group label and its corresponding destination addresses, establishes multicast packet links in the network layer and following information packets, which include the group label in the layer 3 headers, are transmitted on the links. The packets are duplicated at the nodes where the links branch off.
On the other hand, a call set-up packet for a receiver-selective multicastcommunication includes a label and the contents of the following information packets and is distributed to all receivers. Receivers who need the information note down the label and extract packets by collating their labels with the noted label.

5.3.6 Multicast Communications Facilities in HSPN

The multicast communication methods mentioned in Section (Multicast Communications Protocols) are implemented in the HSPN. In the caes of the distance-indexed multiplexing scheme, each frame header includes a bit that indicates whether the frame is a multicast frame or not. When multicast frames arrive, all nodes except for an originating node receive it and at the same time, duplicate and send it to the next nodes. Extra delay is not added for those processing.

When the preemptive priority packet multiplexing scheme is adopted, each packet is examined for whetherit is a multicast packet or not and,besides if multicast, whether it is a destination-designated type or a receiver-selective type. For this purpose, all packets include the specified bits in their layer 3 headers. Multicast packets are routed by network layer control facilities in the switching nodes according to the contents of label-destination tables.

Since the multicats packets are duplicated when they are distributed simultaneously to multiple packet buffers corresponding to outgoing links, delay in the switching nodes does not rely on the number of reproduced packets.

5.4 SOME APPLICATIONS IN MULTICAST COMMUNICATIONS

Utilizing the multicast communication facilities of the HSPN, various advanced services can be provided. Two examples are presented in this chapter.

5.4.1-Multicast Information Providing Service

As a broadcasting information providing services, the teletext service has already been provided in some counties. A multicast information providing service , which is provided on the basis of the receiver-selective multicast communication protocol, is similar to the teletext but some differences exist between those two services. Characteristics of the multicast information providing services are:

1) A large amount of information can be provided at a low cost, since the band width of the HSPN is far wider than television channels.
2) Charging according to the value of information is possible, since selection of desired information from multiplexed packet flow is performed in the switching nodes.
3) An intelligent selection mechanism using AI technology is practical for selecting information.

5.4.2-Packetized Teleconference

Another application on the HSPN is a packetized teleconference. All information organized by teleconference terminals is packetized and distributed to virtually connected terminals by utilizing the multicast communication facilities in the HSPN. To compress the band width of video information, differential encoding of video frames is applied prior to packetization.

5.4.3 CONCLUSION

Up to now,it has been considered that packet switched networks must be special purpose networks for data communications only, while circuit switched networks are the fundamental facilities for telecommunications. However, the HSPN is not only for this special purpose, but is for future integrated services including telephone and video communication. Almost all services existing on the current circuit switched networks will be provided by the HSPN at less cost. Accordingly, intensive study and standardization of high speed packet communication protocols and multicast communication protocols must now be conducted by the CCITT.

CHAPTER VI

PERFORMANCE MODELLING OF A HIGHLY MODULARIZED PACKET SWITCHING NODE

A highly modularized, high speed network node for wide area packet networks is considered. The control is subdivided into four levels: Termination Unit (TU) containing Line Terminator Unit (LTU) and Terminator Group Controller (TGC) and the Switching/Routing Unit (SU) containing Switching Processor Controller (SPC) and Switching Processor Unit (SPU). The Tus and Sus are interconnected through a pipelined ring controlled by a special access protocol using a ring empty indicator. The exchange of packets between the LTU, TGC, SPC and SPU modules is controlled by internal protocols which provide for buffer reservation, acknowledging and proper retransmission in case of packet loss due to buffer overflow. The modular concept allows a gradual extension up to 255 TU/SU system units. The network node has been modelled by an extensive queueing network reflecting all important resources, protocol features and statistical aspects. tHe model has been analyzed by means of simulation as well as analytically. For the simulation of such a complex problem an aggregation technique has been developed and implemented where some modules and the central ring are represented in full detail, whereas the residual modules are replaced by simpler models, however with the identical load generation. The analytical performance evaluation is based on the decomposition of the total model into submodels of the type of a multiple-level processor with feedback and priorities; these models are anlyzed in isolation and the global results are found from the various submodel results. The performance evaluation of the network node yields results on the maximum nodal throughput and the cross-switch packet delay under different load conditions and allows the proper sizing of components and the evaluation of cross-network delays

6.1-PACKET SWITCHING NODE

6.1.1 Switch Structure

The increasing need for packet communications leads to a second generation of packet switching nodes to be introduced in wide area packet networks. This second generation of nodes is characterized by a mutiplicity of modules which are interconnected by high speed internal bussess or rings. Figure 1 shows the basic structure of such a node consisting of 4 global fuctional blocks:

Line Termination Unit	LTU	Termination Unit TU
Terminator Group Controller	TGC	
Ring Unit	RU	
Switching Unit	SU	

The LTUs interface the switching nodes to the subscriber lines and internodal trunks; they perform basically the functions of levels 1 and 2 of the ISO-basic refence model. The traffic of several LTUs is multiplexed on to one TGC. The TGCs communicate with the more sentralized SU's through the central RU. The SU consist of two parts, the Swiching Processor Unit SPU and the Switching Processor Controller SPC. The SPCs interface the more centralized SPUs with the RU. The SPUs perform the basic packet switching functions whereas the SPCs offload the SPUs from time critical transmission procedures.

The whole packet switching node consist of many SU's, TGCs and LTUs. Up to 16 LTUs mat be connected to one TGC. Each SU handles the traffic of several LTUs which are logically assigned to that SU. The logical assignment may be changed in case breakdown of a SU.
Virtual connection through the switch are established by a setup procedure involving two (originating/destinating) LTUs and their correspnding SU's. once a virtual connection is established, succeding data packets are routed through the switch from the incoming LTU to the outgoing LTU via one of the SU's being assigned to the connection setup phase.

6.1.2 Switch Operation

The basic internal routing scenarios for a virtual connection setup and a subsequent data transfer.
The virtual connection set up starts with a Call Request (CR-) Packet (being carried within a CR-Message) from the originating LTU 1 which is forwarded to the corresponding SPU 1 through TGC1 and SPC 1. The SPU 1 performs the routing function and forwards the CR-Packet to the destinating SPU 2 through SPC 2 within a correspnding CR-Message. SPU 2 sends the CR-Message via SPC 2 and TGC 2to LTU 2. The Call Confirmation (CC-) Packet runs as a CC-Message from LTU 2 the opposite way back across the system.
Within the data transfer phase a data packet is transferred from LTU 1 to SPU 1 as a Data Block (DB-) Message through TGC 1 and SPC 1. The SPU 1 buffers the data packet in the Main Memory (MM) and sends an Output Request (OR-) Message to LTU 2 which is acknowledged by LTU 2 with a Block Request (BR-) Message. Upon reception of the BR-Message the SPC 1 transfers out the data packet from the Main Memory of the SPU within a DB-Message. Finally, there are two different options for the internal acknowledgement of the data packets transferred out: In mode a, (LT-buffered Mode) the data packet is cleared within the SPU 1 directly after successful sending of the DB-Message across the RU, whereas in mode b, (CP-buffered Mode) the data packetis cleared after reception of an End of Transmission (EOT-) Message from LTU 2.

6.1.3 Objectives

The operation of the highly modularized packet switching node is characterized by parallel packet processing within the various units and a pipelined packet transmission over the central RU. The performance of the switch, i.e. the throughput and delay characteristics, depends largely on the individual processing and transmission times as well as on the various queuing delays caused by the resource sharing principle. The estimation of the throughput capabilities and cross-switch delays has therefore been subjected to an extensive performance evaluation study through modelling, simulation and analytical calculations. The results of the performance evaluation may be used for a paper

- configuration management
- sizing of buffer and processor capacities
- implementation of load-adaptive schedules
- throughput estimation
- cross-switch and cross-network delay estimation

6.2 MODELLING

6.2.1 Derivation Of the Switch Model

In this section a model of the central part, i.e. Switching Processor Unit (SPU), Ring Unit (RU) and Terminator Group Controller (TGC) of the high-speed packet node will be derived. The different units are modelled by their processing and scheduling phases. The data and control packets may wait for processing in the queues corresponding to these phases. All queues will be served according to the FIFO service discipline. Furthermore, an arbitrary scheduling cycle is defined for each particular unit to serve the waiting packet requests by taking the various load situations into account.

In the following, the various component submodels are shortly motivated. The global switch model is then composed from these component submodels.

6.2.2 Terminator Group Controller

The TGC is essentially made up from one microprocessor single server with 4 processing and 1scheduling phase. The phases correspond to the following functions:

Phase 1 : transfer from TGC-ring receive queue to LTU-send queue (DMA)
Phase 2 : transfer from LTU-send queue to LTU (DMA)
Phase 3 : transfer from TGC-transmit queue to TGC-ring send queue (DMA)
Phase 4 : transfer from LTU-receive queue to TGC-transmit queue
Phase 5 : scheduling

To account for a proper emptying of the high-speed ring receive and send queues, phase1 and 3 have nonpreemptive priority, whereas phases2 and 4 are scheduled for service cyclically. Phase 5 represents the operating system overhead. The scheduling strategy can be given arbitrarily. In case of nonempty queues, the cycle mat be
1 3 5 1 3   2   1 3 5 1 3 4 .....

The phase service time distributions can be arbitrarily fixed e.g , according to measured data of the implemented switch. Typically, the service time is composed of a fixed part (corresponding to DMA initialisation) and a variable part corresponding to the actual packet length.

6.2.3 Switching Processor Controller

The SPC consist of two microprocessors corresponding to the receive and transmit functions with respect to the high-speed ring. The Receive-SPC (RPC) submodel is a single server model with 6 service phases corresponding to

Phase 1 : EOT-transfer from RSPC-EOT queue to SPU-EOT queue
Phase 2 : BR-transfer from RSPC-BR queue to TSPC-BR queue
Phase 3 : EOT-transfer from RSPC-DB queue to SPU-queue and main memory
Phase 4 : DB-transfer from RSPC-ring receive queue to the queue belonging to the Message
Phase 5 : transfer from RSPC-ring receive queue to the queue belonging to the Message
Phase 6 : scheduling

Similarly to the TGC-Submodel, phase 5 is granted nonpreemptive priority. Scheduling and service time distributions may be fixed analogously as described in Section **(B)**

The Transmit-SPC (TSPC) submodel is also a microprocessor single server with 5 phases:

Phase 1 : transfer from TSPC-transmit queue to TSPC-ring send queue (DMA)
Phase 2 : OR-transfer from TSPC-OR queue to TSPC-transmit queue
Phase 3 : BR-transfer from TSPC-queue to TSPC-transmit-queue
Phase 4 : BR-transfer from TSPC-BR queue to TSPC-queue
Phase 5 : scheduling.

Phase 1 has nonpreemtive priority. All other details are similar as described above.

6.2.4 Switching Processor Unit

The SPU performs the packet switching functions and is modelled by a microprocessor single server system with 2 processing phases. The Phases correspond to

Phase 1 : Routing of Messages
Switching of DB-Message
OR-transfer to the TSPC-OR queue
Phase 2 : Processing of EOT-Messages

In addition to the phase-specific input queues a third buffer models the intermediate storage of packets within the Main Memory. Each packet buffered in the main Memory will be released upon processing of the corresponding EOT-Message. The scheduling is organised according to a clocked mode by which phase 2 is initiated perodically after a specified time.

6.2.5 Ring Unit

The RU operates according to a pipelined ring system. As such, a modified single server model is used where the scheme for the initiation of a message transmission is modelled by a polling mechanism and where the pipelined transmission process is modelled by scheduled arrival times at the various receive queues. This submodel is schematically depicted within the global switch model

6.2.6 Global Switch Model

The global of the switch representing explicitly only one TGC-Submodel, one SPU-Submodel and one pair of RSPC- and TSPC-Submodels. In the real model, an arbitrary number of these representatives can be given, e.g., 50 SU's , 50 RSPCs, 50 TSPCs and 200 TGCs.

The individual packet flow through the global switch model follows principally according to the scenario i.e., there are two types of messages corresponding to connection setup and data transfer. The resulting total packet traffic flow follows from a given origination – destination traffic matrix and can be arbitrarily unbalanced.

The packet traffic is generated by external traffic generators which are simply modelled as Poisson processes as a good approximation for the superimposed traffics of the various LTUs (these are not included in the central switch model).

6.3 PERFORMANCE EVALUATION TECHNIQUES

6.3.1 Analytical Performance Evaluation

The analytical performance evaluation is based on decomposition techniques where the Submodels are analysed in isolation under certain "environment" conditions, i.e., under simplified interface traffic assumptions. For the typical assumption of Poisson arrivals pure priority models and pure polling models can be analysed exactly or approximately with respect to mean delay based on renewal theory and Little's theorem. The models are somewhat more complicated; they can approximately be analysed by an aggregation technique for all equally treated processing phases. These methods are currently under study and will not be described here explicitly.

6.3.2 Simulative Performance Evaluation

Opposite to the analytical approach, simulation is an experimental performance evaluation technique. A computer simulation program implement the simulations model where the System State is represented by a complex data structure. Messages are represented by requests (job, customers) which require transmission or processing from service units (servers). They are generated by some traffic generators with individual interarrival time distributions. The generation of new requests and the service terminations are represented by discrete –time system events. The simulation is organised according to a time sequenced event list which is processed sequentially by the simulation program. Within a simulation run, usually hundreds of thousands of events are executed and measurements are taken simultaneously. By averaging over the various sample measurements representative results on the throughput delay or occupation of servers and queues are found.

Simulation, however has significant drawbacks when it is applied to models as the examples of fig.6. The simulation execution time depends linearly on the number of executed events. In order to obtain reliable results for a highly modular system as that of many millions of events have to be executed resulting in excessive CPU-times. Therefore various simplified simulation techniques which are less time-consuming have been applied:

1. Subsystem Simulation:

In this method only part of the global model is simulated as, e.g., RU-Submodel , TGC-Submodel, SPC-Submodel, or SPU-Submodel. This method allows a correct representation of the Subsystem-scheduling and service mechanisms, but suffers from scenarios. Finally a combined simulation applying aggregation techniques and Subsystem consideration provides performance measures and may be a basis for the decision process in the development and engineering of such complex packet switching nodes. Currently, the simulation of the LTU is being implemented where some aspects of the data link protocol are also taken into account as well as an analytical performance evaluation of the whole packet switch.

CHAPTER VII

THE DISTRIBUTED IPSS ARCHITECTURE:A HIGH RELIABLE SWİTCH FOR HIGH-PERFORMANCE PACKET SWITCHING NETWORK

This paper gives a general survey of the No. 1 Packet Switching System – 1PSS – and points out how the system meets high capacity and high reliability requirements for medium to large public packet switching networks. The system design philosophy is based on the principle strictly to implemented a X.25 / X.75 type packet switching network. Additional communication protocols are supported by means of adequate protocol conversion. The 1PSS system architecture uses a hierarchical approach with distributed processing capability. The system reliability is achieved through combined efforts in hardware and software architecture.

7.1-GENERAL

A rapidly increasing activity in the area of data communication has been observed during the last few years. This has taken place in manifold ways. First there is a higher demand for communication in general with a strong emphasis on data and text communication in particular. This has resulted in a higher amount of dialogue as well as bulk data transmission. Second the communication is not restricted to inhouse or local applications but includes a high portion of nation wide and even international traffic relations. Third communication needs no longer are restricted to data transport within a closed group of users but as well exist for a more open environment.
Those requirements led to the necessity to provide economic and standardised transport capabilities for data communication. Together with the demand for communication the market for data terminal equipment as well as for network equipment has grown by an average rate of 20 to 30 percent during the last years and is expected to continue to grow. This would not have been possible without introducing reasonable communication standards and agreeing on them. Important milestones on that way have been the standardisation in the field of packetized communication like X.25 / X.75 and the OSI standards. At the moment we all are preparing the next step into that direction-ISDN. There is no doubt that ISDN will evolve from the voice communication area but there is also no doubt that it heavily benefits from the experiences in standardisation of data communication. Moreover its

future success depends on the integration of already existing standards for data communications.

Bearing in mind the above mentioned facts and needs AT&T has put a lot of efforts and experiences into the data communication market. Naturally their first steps aimed at the US market. Being able to rely on a powerful communication computer at first network products have been introduced for wide area communication between high capacity data sources like hosts or local area networks. In a second step the product line has been extended by concentrator equipment to economically provide access and communication facilities for regionally spread data terminals of low up to medium data rates. Both together now have been integrated in to a unique product, known as No.1 Packet Switching System –1PSS.

A next step – planned within the joint venture of AT&T and Philip's aiming on the market for public communication products – will be the introduction of 1PSS on to the European market. For the European market the flexible and powerful ring based version of 1PSS will be used. How this version of 1PSS has evolved from its predecessors, its overall system architecture, modifications provided for the 1PSS application in Europe, the 1PSS capacity and performance and its ability to be an integrated part in a future ISDN environment will be outlined in the following chapters of this paper.

7.2-SYSTEM DESIGN PHİLOSOPHY

The system design philosophy has been influenced by several goals:

1- Providing a product for medium to large packet switching networks.

2- Using experience and approved components from development of public telephone switching.

3- Offering a product compatible to present CCITT standards for packetized data communication

4- Providing a technique which supports high performance and which in addition is price competitive both in initial costs as well as in operation and maintenance costs.

This has led to an approach with an elaborate mixture of centralised and distributed architecture. From the first release of 1PSS deployed in the early 1980's up to its today's ring based version this has been a successfully performed evolution strategy. During that period essential backbones of the productline and of its development environment have been :

1- AT&T's very reliable, duplicated switching processor 3B20D as central computer.

2- The 3B20D's operating system UNIX RTR (Real Time and Reliability) – an extension of the common UNIX operating system for realtime applications.

3- The kernel of the peripheral processors with its microprocessor based CPU and its realtime operating system RTX – also a UNIX – derivate.

Powerful software development tools embedded in a UNIX environment.

Functions of a packet switching network can generally be divided into user data transport for the network subscriber and into OA&M functions for the network provider. Transport functions involve interface and protocol handling, data switching, call control and call routing. OA&M comprises billing, service and network provisioning, network control and network management. This given sequence of tasks in a certain way corresponds to a grade of a feasible decentralisation. So in the 1PSS version presented dedicated peripheral processors only perform interface and level 2 protocol handling. They are star like connected to the 3B20D. thus the 3B20D is in charge for all other transport and for the OA&M related functions. With this comprehensive utilisation of the central processor, the 1PSS capacity was well suited performed on a virtual circuit basis. It is static under normal operation but is adaptive under failure conditions.

Flow control and congestion control strategies work together to provide an effective overload control strategy for the 1PSS system. Flow control is performed by the access protocols and by the IP at level 2 and 3. The coupling occurs on level 3 between the protocols. For satellite links extended flow control parameters are implemented. Congestion control, necessary to recover from overload conditions, is closely bound to buffer management procedures in the 1PSS system. Two design principles have been chosen : First buffers are dedicated on a per line basis and second terminating traffic is given preference to input traffic. Simulation as well as field experience have shown the validity of the chosen combination.

7.3.3 Gateway Functions

Internetwork connections between 1PSS and PSPDNs (packet switched public data networks) are based on 1984 CCITT X.75 and handled via STE (signalling terminal) functions. When PSPDNs are connected to the 1PSS network via one or several gateways they all may also work as subnets and from one common PSPDN. This is supported by the 1PSS packet switches through interpretation of the calling pattern.

7.3.4 Operation , Administration and Maintenance

Centralised management functions like trunk line provisioning or traffic data collection are established by one or more NCCSs. A hierarchical configuration of several NCCSs is supported. Functional distribution is given preference wherever reasonable. Thus subscriber line provisioning including the necessary databases are supported by the different PSs, but centralised access via an NCCS is provided.

7.4-Components of the 1PSS Packet Switching System

1PSS uses three basic building blocks out of which the components of a packet switching network can be configured in a flexible way:

- The Packet Administrative Module (PAM): an AT&T Technology 3B20D computer
- The Packet switching Module (PSM): a fast duplex token ring with Ring Peripheral Controllers (RPC) as interfaces to the 3B20D and with Packet Switching Units (PSU) as dedicated front-ends;
- The remote Packet Module (RPM): a multi-microprocessor system for protocol conversion and concentration.

PAM and PSM are the basic blocks out of which all packet switches are build up. To extend the subscriber line access capability of a PS, RPMs are used remotely or locally. . Additional features, e.g. post processing of billing data or enhanced OA&M functions at a NCCS location, can be achieved by adding 3B20S and 3B20D computers.

7.4.1 Packet Administrative Module

The PAM is a 3B20D computer, running the UNIX RTR operating System. The 3B20D is a fast 32 bit super minicomputer especially designed to achieve high reliability and fault tolerance. It uses redundant modules for the control unit, Input Output Processor Unit and the Disk File System Unit. Essential parts within a single unit are duplicated. Matching, self-checking and parity techniques are used. The mate processor works in a standby mode with real time memory update. In emergency situations an additional microprocessor based control unit- the Emergency Action Interface-provides access to the 3B20D. The PAM is left only with operations, administration and maintenance functions at the PS and at the NCCS.

7.4.2 Packet Switching Module

The PSM consist of a fast duplex ring with two types of ring processors-the RPCs and the PSUs. Token passing performs ring access. The ring transmission rate of the existing system version is 4Mbytes/s. An upgrade to 8 Mbytes/s will allow to handled more than 30000 user data packets. Thereby a user data length of 106 bytes is assumed. Including protocol header and acknowledgment packets this results in a ring utilisation of 70 %. The RPC is a microprocessor system providing an interface between the 3B20D computer and the ring. The number of RPCs is in the range of 2 to 8. It depends on the amount of OA&M traffic.

The PSUs are software configurable multiprocessor system. Using different application software packages they do protocol handling and switching. In a large size switch some hundred PSUs are devoted to either access line handling or trunk line handling. The packet throughput capacity of a PSU-L exceeds 150 data packets per second; that one of the PSU-T because of the reduced protocol overhead even 250 dpps. They support interfaces for transmission rates between 1200 bit/s and Mbit/s. Such a switch needs several PSUs performing routing (PSU-R), each of which has a capacity of handling about 50 call setups per second.

The pair of rings interconnecting the PSUs and RPCs operates in an active/standby mode. Only one ring is involved in data traffic, while the other one is used for maintenance purposes in a counter-rotating manner. On occurrence of a faulty situation the ring operates in simplex mode: the failed node is isolated and a redirection of traffic is done.

-Remote Packet Module

The RPM is a hierarchical multi-microprocessor system. It performs protocol conversion and concentration for asynchronous interfaces (X.3, X.28, X.29) and for binary synchronous interfaces like 3270 BSC or 2780/3780 BSC. Moreover it serves as a concentrator for low to medium speed X.25 access lines. A support of customised protocols like IBM's SNA can also be provided.

The packet throughput of an RPM will be more than 100 dpps assuming an average user data length of 64 bytes. Redundant mode of operation will be possible for the main CPU, the memory and for X.25 interface to the packet switch.

-Performance of the Packet Switching System

For a medium user data length of 64 bytes computations and measurements point out, that a packet switch configured with 500 PSUs and 200 RPMs reaches the following performance values: up to 10.000 subscriber can be connected, 600 to 800 call setups/s are supported and more than 40.000 data packets per second can be transmitted.

7.5-Software Architecture

Today's 1PSS software system is based on well defined software engineering and quality assurance guidelines, which are founded on ANSI/IEEE standards. A set of tools, incorporated into a UNIX system and running on the 3B20S builds a unique program environment for multiple processor types, supporting the whole product life cycle. Almost all software is written in the C programming language and the operating systems of the 1PSS basic components-PAM, PSU and RPM-are derivates of the UNIX system. System generation and management for the PAM, the PSU and the RPM is done by the agency of four major databases located at the PAM. These are the System Generation Database, the Equipment Configuration Database, the 1PSS and the AFP Databases. System software is down-and uploaded between PAM and PSU and between PAM and RPM.

7.5.1-Classification of Functions and Layers

The 1PSS software system of the NCCs, PCs and RPMs consist of four major functional subsystems and three layers.

- 1PSS Environment
- 1PSS Maintenance
- 1PSS Transport
- 1PSS Operation and Administration

These functions are mapped on different layers for the basic processors of the 1PSS packet switching system. Each of these layers has access to the services on the same level or any lower. The environment software consist of three layers providing a general set of operating system functions to the 1PSS maintenance, transport and operation and administration software. These layers are UNIX RTR, IMS and 1PSS Environment concerning to the PAM.

The 1PSS Maintenance Software is responsible for maintaining 1PSS software and hardware stability. Most of these functions are controlled by the PAM. The 1PSS Transport Software is concentrated in the Pus providing facility interface software and for receiving and transmitting packets. The PSU transport software is responsible for processing level 3 of X.25 with level 1 and 2 handled by the PSU hardware. Switching of packets from PSU to PSU is handled by PSU software. The 1PSS Operations and Administration software is concentrated in the PAM and consist of craft interface, measurement, billing, NCCS and RPM support.

7.5.2-Description of Essential Subsystems

Interprocess Message Switch IMS:

The IMS is a high performance fault – tolerant packet transport fabric and its main function is the Interprocess/interprocessor communication within a packet switch. It forms the second layer in the PAM as well as in the PSU.

Local Transport Service Subsystem LTSS:

The LTSS is the 1PSS Transport Software in the PSU. It uses the services of layer 1, the RTX(Real Time Executive), and of layer 2, the IMS. It performs the facility interface management, level 3 of X.25, X.75 and IP, packet transport, switching and routing functions.

Administrative Feature Package AFP:

The AFP is a part of 1PSS Operation and Administration Software in the PAM, and provides billing, traffic measurement collection and assists in overload control of the RPMs.

7.6-FAULT-TOLERANT PROFICIENCIES WITHIN THE 1PSS SYSTEM

During the design of the ESS systems AT&T Bell Labs has achieved a wide range of experiences in fault – tolerance systems development over the last two decades, which has strongly influenced the 1PSS design. Following we summarise main aspects by which the high reliability of a 1PSS based network is achieved.

7.6.1-Network Reliability

Errors are mainly caused by physical disturbance or software faults. 1PSS uses its high available components, the advantages of its Integral Protocol, failure adaptive routing, its measurements Subsystem as well as its maintenance and diagnostic subsystems to assure a high reliability.

The IP is based on the virtual circuit model. Its link level implements error detection and correction according to LAPB. The packet level includes a reconnect procedure that routes packets on an alternate path in case of a virtual circuit failure. It also separates the packet acknowledgment function into a flow control authorisation and a delivery assurance acknowledgment to achieve a high efficient buffer utilisation. A careful buffer management at the line buffers is also closely linked to the IP, because 1PSS follows the philosophy that the most effective method of controlling overload is throttling traffic at the endpoint PSUs.

As mentioned above 1PSS uses a failure adaptive alternative routing supported by tandem PSs. The tandem capability allows to setup virtual calls and permanent virtual circuits on two or more hop paths, through intermediate switches.

7.6.2-Availability of a Packet Switch:

A packet switch is not deactivated for maintenance purposes or growth. These functions are performed during the switch is online. Following we give the values for a packet switch configured with two RPCs and 200 PSUs. Using MTBF numbers and an MTTR of 2 hours, the number of minutes per year that components are unavailable are : PAM 2 minutes. Ring 0.35 minutes and two RPCs 0.002 minutes. This results for a packet switch supporting 200 PSUs in an availability of 99.99955%.

7.6.3-Recovery Strategy:

Early detection of hardware – faults is achieved by extensive use of self – checking circuitry, sanity timers and system integrity checks. Error counts are kept for this type of faults. If such a count exceeds its predefined threshold or if a serious fault is detected, the equipment unit involved is automatically removed from service and diagnostics are invoked. Throughout the recovery procedure, messages are printed out and logged on disk files so that subsequent analysis can be done easily.

Software errors are detected by the use of defensive checks and audits, running periodically or on demand.

Corresponding to the degree of the failure graduated recovery actions are necessary. As an example UNIX RTR recovery is given below. It uses a dedicated, progressive recovery strategy structured in five phases. It is implemented as a set of ordered pairs (x,y) whereby x is a UNIX RTR phase and y is the application level within that phase. Each application level takes its own special action environment. Below only the phases are described. They are started automatically or manually.

- PHASE 0: All recovery actions are application dependent .
- PHASE 1: All kernel processes are started, but no bootstrap.

- PHASE 2: Bootstrap, but the Equipment Configuration Database and the Protected Application Segment are preserved
- PHASE 3: Bootstrap, but the Protected Application Segment is preserved.
- PHASE 4: Bootstrap, this phase can only be requested manually.

7.7 SUMMARY

It has been shown that 1PSS – No.1 Packet Switching System – has the capabilities to fulfill today's needs for medium and large packet switching networks. It marked by the support of the 1984 CCITT recommendations, its distributed architecture and last not least by its high reliability and availability. 1PSS will also be able to meet the challenges of the ISDN era. Last not least it should be mentioned that the evolutionary development steps driving 1PSS from a medium to a supper high capacity system have been invoked by the needs indicated in a request for proposal of a European PTT. Thus some of the above indicated features –like upgrade to an 8 MHz ring, high capacity PSUs and RPMs, NCCS network structure – will not be available as standard features but only as special add on developments if a successful market introduction can definitely be foreseen.

CHAPTER VIII

RANDOM-ACCESS TECHNIQUES

In the introduction to this chapter, we noted that the two basic ways of providing access to a common medium are controlled-access using polling and random access. (We leave out deterministic access, such,as time division multiplexing, as noted in the introduction.) 'There are in turn a number of different types of random access strategies. We focus in this section on the two simplest types,pure Aloha and slotted Aloha. We provide some references to the literature for a number of variations on these simple schemes, techniques designed to improve the performance obtainable through their use. We do provide a brief introduction in this chapter to the CSMA/CD (carrier-sense multiaccess with collision detection) technique, adopted for use in random-access local area networks.The discussion continues in the next chapter, in which we describe CSMA/CD- type l.ANs.

Random-access techniques, as the name implies, are completely decentralized. A user will essentially transmit at will, with possibly a few constraints depending on the particular access method adopted. These techniques range from the pure Aloha technique, in which a user (a station) transmits whenever it has a message (packet) to deliver to some destination, to techniques where the user is constrained to transmit in certain time intervals only, to more sophisticated techniques where a user "listens" before it transmits and then does so only if it senses an idle medium. Other variations include reservation techniques where a user may, through random access, request permission to transmit a full message at some reserved time [SCHW 1977]. Many other schemes have been suggested as well.

Because the essential idea of random access is a very simple one-transmit at will,with possibly a number of constraints-the access algorithm is generally easily implemented and is relatively inexpensive in practice. Randorn access has thus received widespread interest and, in a more sophisticated form such as CSMA/CD, has been implemented widely. It is particularly, practical and useful at Iow levels of traffrc.

No rrratter what the technique used, because of the randorn times at which users may decide to transmit, there is always the chance that two or more users will decide to transmit at overlapping times. This results in a "collision," which must first be recognized as such and then resolved. As the traffic intensity increases, the probability of collisions increases as well, leading to possible instabilities in the operation of these mechanisms. Throughput is limited as a result to some maximum value less than the channel capacity, the particular value depending on the original access mechanism and collision-resolution algorithm proposed.For these reasons, random-access techniques have been primarily suggested for applications that involve many bursty, interactive users, with each user attempting to transmit only infrequently or, at the other extreme, for applications that involve relatively few host computers communicating with one another. In this latter case, because there are very few users, the chance of a collision is reduced as well. In applications such as manufacturing, assembly plants, and other factory operations that require tight control of access delay, the use of controlled access has been preferred.

8.1 Pure Aloha

We begin a more quantitative discussion of random-access techniques by starting with the simplest scheme of all, pure Aloha. As noted earlier, this scherne was first adopted as a common channel-access strategy by workers at the University of Hawaii in the early 1970s. It is the forerunner of many random-access strategies that have been proposed and/or adopted since. In this scheme a user wishing to transmit does so at will. As a result two or more rnessages may over-lap in tirne, causing a collision. There must be sonre way of recognizing the collision and so signaling to the users involved. This could be done by a central station designated for this purpose (in the original Aloha system all messages were beamed via radio to a central receiving site) or by the use of a positive acknowledgement with timeout method. In the case of a bus-type local area network this could be accomplished by having each station listen as it transmits and it.self detect a collision. In any case, on detecting a collision, colliding stations attempt to retransmit the message in question, but they must stagger their atternpts randomly, following some collision resolution algorithm, to avoid colliding again.

It turns out that this pure-Aloha access strategy, although very simple is quite wasteful of bandwidth, attaining at most 1 /2e = 0.18 of the capacity of the channel. To demonstrate this limit on throughput, we first require some definitions. Let there again be N stations contending for use of the channel. Each station transmits, on the average, packets/sec. Now take the specific case, for simplycity, where all messages transmitted are of the same fixed length. m, in units of time. (These messages again normally contain data plus necessary over head.) In keeping with the notation commonly used in discussing Aloha-type systems-see (ABRA) and (SCHW 1977) we now let the traffic intensity p, the fractional utilization of the channel by newly arriving packets, be written as a parameter S.

$$S=p=N(\lambda)m$$

(Recall that 1 /m, or mu, as used in earlier chapters, represents the channel capacity in units of packets/sec transmitted. N(λ)/mu = N(λ)mum is thus the relative utilization of the channel, or throughput normalized to mu =1 /m.) It is this parameter that we shall show is limited at most to 1 /2e = 0. I 8.

We now assume that the arrival of packets at each station obeys a Poisson process. The total arrival rate is then Poisson, with parameter (λ)N. As indicated earlier, the total traffic on the channel will consist of newly transmitted messages plus those that have been retransrnitted. We now make the additional assumption that the retransmitted messages are Poisson distributed as well. This is obviously not true, since they do depend on collisions having taken place. Simulation studies indicate that the assumption is valid if the random retransmission delay time is relatively long (LAM 1974) .The total rate of packets attempting transmission over the channel, newly generated plus retransmitted ones, is then some number (λ')>(λ). The actual traffic intensity or utilization of the channel is then a parameter G given by

$$G = N(\lambda')m$$

Consider a typical message m-sec long, as shown in Fig. It will suffer a collision with another message if the two overlap at any point. It is easy to see, by shifting the dashed (colliding) message in time, that collisions could come from an interval 2m-sec long. The probability of no collision in an interval 2m-sec long is just the probability that no Poisson messages are generated during that time. From Poisson statistics (Eq. 2-1), this probability is just equation

The ratio S/G represents the fraction of messages transrnitted over the channel that get through successfully. This must equal the probability of success,i.e., the probability of no collision. We thus have, very simply, the pure,Aloha throughput equation:

$$S = Ge2\lambda$$

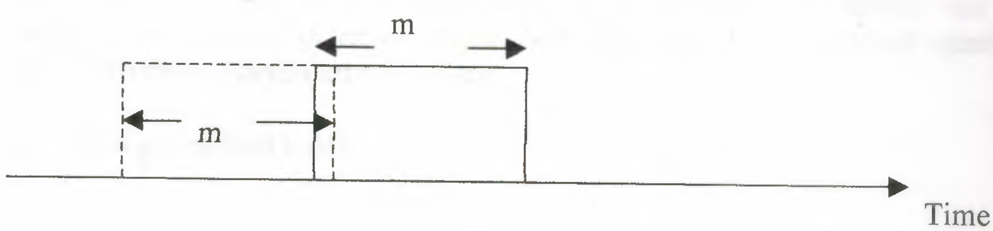


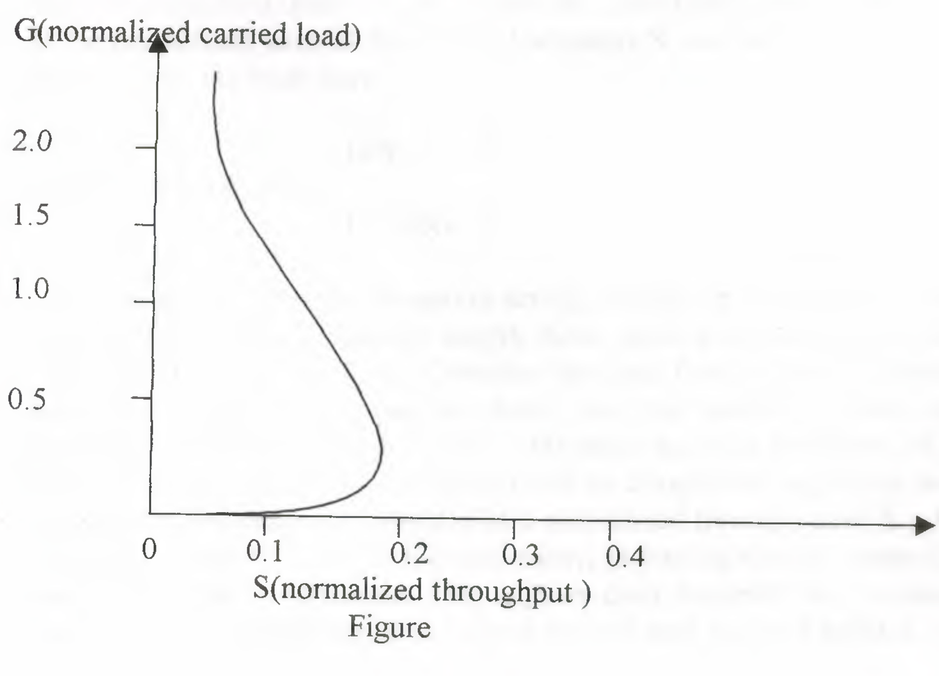
Figure : Collision between two mesages

S is the normalized throughput (average packet arrival rate divided by the maximum throughput 1/m) and G is the norrnalized carried load. S is thus the independent variable and G the dependent one. G plotted as a function of.S gives rise to the two-valued curve of Fig. Note that S is maximized at a value S = 0.5 1/e = 0.18 at G = 0.5. 'This is also shown by finding dS/dG from equation and setting it equal to zero.

Note the interpretation from either equation or F ig. For small offered load S there are few collisions and G = S. As S begins to increase; however, approaching its maximum value of 0.18, the number of collision increases rapidly, increasing the nurnber of r-etransmissions, which in turn increases the chance of a collision. The system becomes unstable, S drops, and G increases to a Iarge value.

The fact that the maximum throughput of the Aloha system is limited to rnost 18 percent of the line capacity may seem disconcerting at first. Yet in many applications this maximum transmission capability may be quite sufficient for practical use. Consider an example where interactive terminals are connected via a multipoint line to a central controller, as suggested by the various topolgies. Say that the line capacity is 4800 bps. Messages in bound to the controller contend for the common medium; messages sent in reply by the controller are obviously controlled and can occupy almost the full bandwidth of a separate return channel. Say that a human user at a tërminal inputs a 60-character message and receives a 400-character message in reply. Since the person at the terminal must take time to compose and type in his or her message and then wait to receive and read the reply, it is clear that there are human limits on the rate at which messages can be input. A typical figure might be to input one such message every 2 minutes. The average input rate per terminal is then 60char/ 120 sec or 1 char/2 sec. If these are asynchronous terrninals, with 10-bit characters [SCHW 1980a], the average input rate per terminal is 5 bits/sec! If 10 percent of the line capacity is made available for random access with the Aloha technique to ensure lying well below the 18 percent limit 96 interactive users can be accommodated on this channel. If the line sped used were 2400 bps instead, this would mean that 48 such users could be accommodated. If user statistic indicate that a message is inputted every 3 minutes on the average about 140 users could be accommodated on the 4800-bps line and 70 users on the 2400-bps channel. If ınput message lengths were reduced to 30 charactersthe number of users could be doubled.

The point to be made here is that in many applications involving highly bursty interactive traffic a simple scheme like pure aloha could be used quite successfully and very simple. The original experiment at the university of Hawaii was designed precisely for this type of environment. Note also that we have picked relatively low bandwidth those associated with voice



Figure

grade telephone lines. Much higher bandwidths-available for example on a CATV system or packet radio-would allow correspondingly more interactive users to access the common rnedium (SAADT). An example appropriate to metropolitan-area networks using CATV will be presented later when we compare random access with polling techniques.

What are the time delays involved in the pure Aloha scheme? To determine these we need first to define a retransmission strategy to use when collisions occur. As already noted, retransrnission should be randomized to stagger the retransrnissions and reduce the chance of a second collision or even more. One suggestion is to choose an arbitrary time interval and select a uniformly distributed random retransmission time within that interval. More precisely, let the time interval cover K message-unit times, of m units of time each. Retransmission then takes place within 1 to K such m-sec intervals after it is Iearned that a collision has taken place. Let the round-trip delay plus processing time required to obtain this knowledge be R m-sec intervals. (As noted earlier, a simple procedure would be one in which the destination system positively acknowledges each transmission. The source station would then timeout after R message-length intervals if a positive ack had not been received by that time. R must obviously be chosen, to account for round-trip delay plus processing time.) The average time required to successfully transmit a message is then

$$D=m[\ l+R+E(R+K+1/2)]$$

Here E represents the average number of retransmission attempts per message transmitted.

The average number of retransmission attempts per message transrnitted should depend on the retransmission interval K. In fact, one would expect that for small K there are more collisions and hence more retransmissions; for large K there are fewer collisions. For large enough K, however, the dependence on K disappears and E is given by a quite simple expression. Specifically, we defined the parameter G as the normalized sum of the original attempts S and the retransmissions. It is then apparent that we must have

$$G/S = 1 + E$$

and, from Eq. (8 -11),

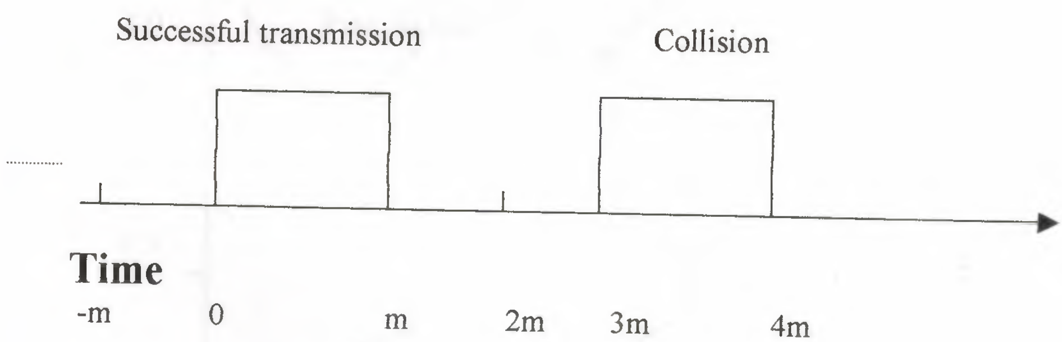
$$E = é2G - 1$$

As an exarnple, take the N-station string, which we discussed in connection with polling. We chose the message length there, quite arbitrarily, to be 1200 bits. For a 4800-bps line, m = 0.25 sec. Consider the case first of the 2000-mile system. The worst-case round-trip propagtion delay, for the station furthest away from the controller, is 80 msec using 2 msec/ 100 miles again as the speed of propagation of energy. This is one-third of m and so will be considered negligible here. Let K = 5, and say that the system is operated at a normalized throughput of S = 0.08. Then G = 0.1, and E = (G/S -1 = 0.25 the total delay, including time to transmit a message, is then D = 1.75m = 440 msec. This appears quite tolerable and is comparable to the results obtained in the previous section for hub and roll-call pollirig. (Recall that the entries there were for access delay.The average message length and propagation delay must be added to rnake those calculations comparable to the one here.) As the traffic load increases, approaching a maximum value of 0.18, the behavior of the pure Aloha scheme becomes worse.A detailed comparison of Aloha access with polling will be presented after we discuss an improved version of pure Aloha, the slotted Aloha scheme.

8.2 Slotted Aloha

The limitation on maximum possible throughput of the pure Aloha scheme can be doubled by the simple expedient of slotting the time scale into units of time m (a message width) wide and allowing users to attempt transmission at the beginning of each slot time only. This scheme requires, of course, that all users in a system be synchronized in time. A simple example of the operation of this technique appears in Fig. One rnessage is shown being transmitted successfully; another suffers a collision.

Since messages can only be transmitted in the slot intervals as shown, collisions only occur when two or more users attempt transmission in the same time slot. The probability of a successful transmission, again assuming that retransmitted messages obey Poisson arrival statistics, is then given by é-G (compare



with the same quantity in the pure Aloha case), and the throughput characteristic for slotted Aloha becomes

S=Gé -G

The normalized throughput S is easily shown to reach its peak value of 1 /e = 0.368 at G = 1. The carried-load-throughput characteristic for stotted Aloha is shown plotted in Fig. 8-12 and is compared there with the corresponding characteristics for pure Aloha. It is apparent from this characteristic with two values for G again possible for a given throughput S, that this access technique is subject to instability as well.

Studies of-the instability problem in Aloha systems have appeared in [LAM1974], [KLEI 1975b], [LAM 1975], [CARLA], and [FAYO]. A summary of the analysis appears in [HAYE,]. Control mechanisms designed to avoid this problem are suggested and analyzed in these papers as well. One sirnple control procedure is to increase the retransmission interval after each detected collision.This spreads user retransmission out and reduces the chance of a collision. It also increases the delay in the system and is the reason why very large retransmission intervals (very large k) are not chosen after the first collision. Interestingly, an algorithm similar to this is used with the CSMA/CD protocol designed for use on local area networks. The CSMA/CD collision-resolution algorithm is described in the last section of this chapter and in the next chapter.

The time-delay analysis of slotted Aloha is very sirnilar to that of pure Aloha. The retransmission strategy is the same as that suggested for pure Aloha: On learning that a collision has taken place R slots after the attempted transmission, a station retransmits with uniform probability anywhere in an interval K slots long. This procedure is diagrammed in Fig. Based on this retransmission procedure, it. is apparent that the time required to successfully complete a transmission in slotted Aloha is given by

$$D/m=1.5+R+E[\ R+0.5+K+1/2]$$

This equation is written in normalized form, in units of slots or message length

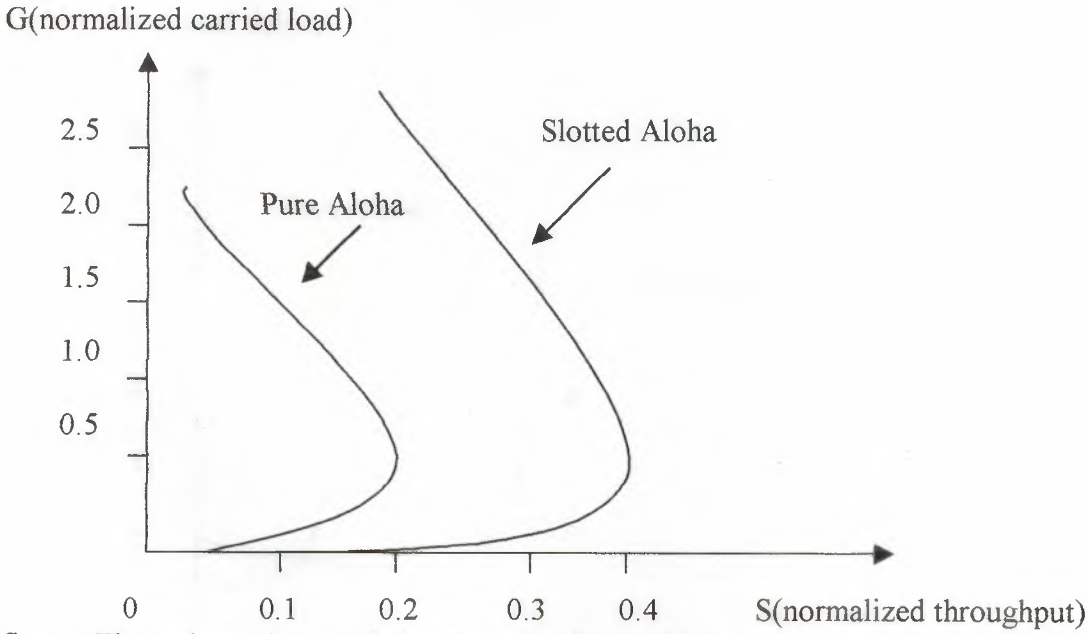


figure Throughput characteristics slotted and pure Aloha

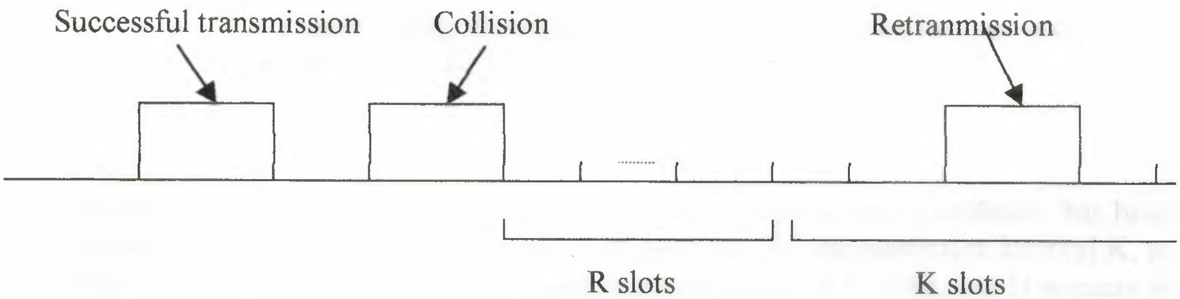


Figure: Collision resolution slotted Aloha m to enable it to be plotted more readily. The only difference between this expression for pure Aloha is the extra factor of 0.5 units of time that appears in two places. This factor is required to account for messages that arrive after slot interval has begun. Queueing (wait) time at a station has been neglected both here and in the prior expression for pure Aloha, since with maximum throughput just a fraction of a line capacity in either case, the probability that a message has to wait for transmission would be expected to be small.

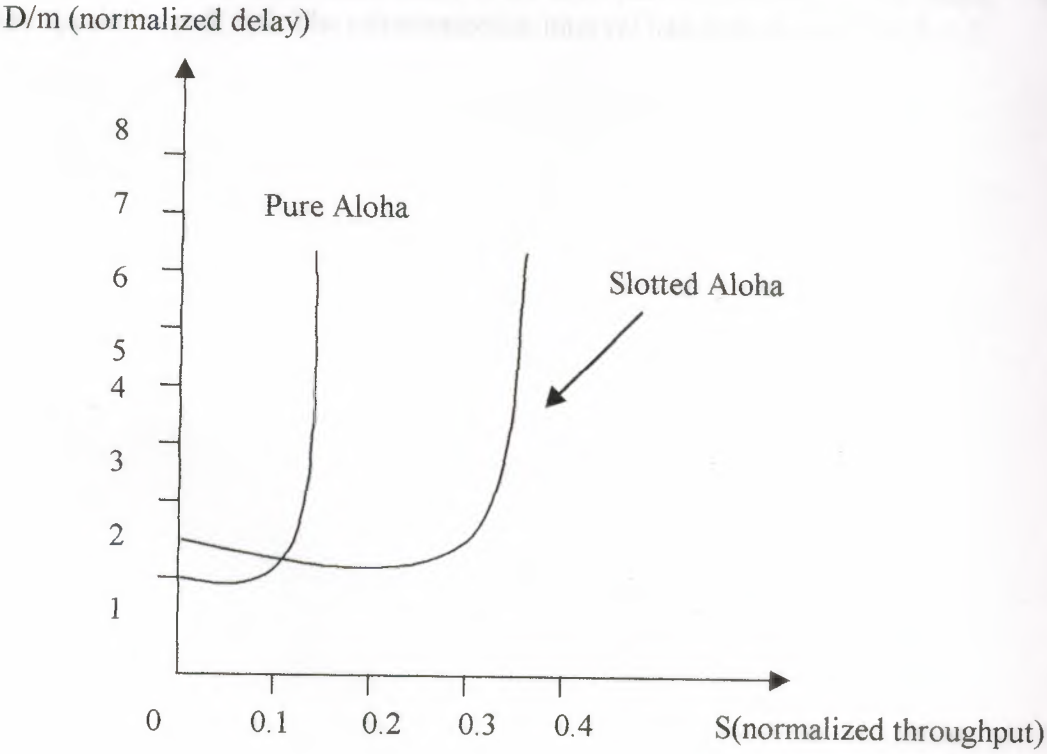


Figure :Delay throughput characteristic Aloha techniques zero Propargation delay K=5

The calculation of E, the average number of retransmissions once a collision has been detected, is again rather complex, since E depends on the retransmission interval K, as does S/G. A detailed analysis of the interrelations among K,E, S/G, and D appears in (SCHW I 977), and in [LAM 1974) and (KLEI 1975b), from which this analysis is taken. It turns out, as noted earlier, that for very large K, one obtains the throughput characteristic, independent of K. Since E = G/S - 1, E is then written simply as

$$E = G / S - 1 = \acute{e}G-1$$

for slotted Aloha. Both Eq. turn out to be good approximations even for K as low as 5 ([LAM 1974,]or [SCHW 1977]). We can thus safely use Eq. in Eq.to calculate the slotted-Aloha delay even for relatively small values of K.

An optimum value for the normalized retransmission interval K is found to exist just as in the case of pure Aloha. For, as noted earlier, with small K one reduces the time to schedule a retransmission but at the cost of more collisions,while with large K collisions are resolved more readily, but the scheduling time is increased [LAM 1974], [SCHW 1977,]. The optimum K, minimizing the time delay D, depends on the normalized throughput S as well as on the normalized round-trip propagation delay R. The variation of delay with K is quite insensitive to K, however, so long as one stays away from values of S close to its maximum value of 1/e = 0.368. A value of K = 5 appears to be a good compromise choice and will be used in the discussion to follow.

Figure plots the normalized delay D for both pure Aloha and slotted Aloha, for the special case R = 0. The retransmission interval has been selected as K = 5.

REFERENCE BOOK:

1- Authors: E.Mair , H. Hausman

EWSP – A High – Performance Packet Switching System

2- Authors: J.F. Huber , A.V. Kienlin

Functional And Performance Requirements In Future Packet Switching Networks

3- Author: N. Roacha , J. Alves, P. Oliviera

A Fast And Flexible Data Communication Switch

4- Author: F. Ramos

Analysis Of A Token Parallel Bus For Voice Transmission

5- Author: S.S. Lam

Packet Switching In A Multi – Access Broadcast Channel