VSATs

very small aperture terminals

Edited by

JOHN EVERETT

Published by: Peter Peregrinus Ltd., London, United Kingdom

Peter Peregrinus Ltd.,
Michael Faraday House,
Six Hills Way, Stevenage,
Herts. SG1 2AY, United Kingdom

British Library Cataloguing in Publication Data

A CIP catalogue record for this book
is available from the British Library

ISBN 0 86341 200 9

Contents

Chapter 1

Introduction to VSATs

J.L. Everett
Government Communications Headquarters, Cheltenham,
Gloucestershire, UK

This chapter provides a broad overview of the subject of Very Small Aperture Terminals (VSATs) in terms of technology, systems and related issues: it also provides a context for the succeeding chapters.

1.1 Historical perspective on VSATs

The multiaccess and broadcast capabilities of satellites have been historically recognised but it has proved difficult to realise their full potential because of technology limitations. The advent of the VSAT type of system, whether one-way data broadcasting or two-way interactive, represents the congruence of recent advances in several technological areas including higher gain and higher power satellites, relatively inexpensive microwave and RF components, digital modems and protocol processing. This evolution in satellite and earth terminal technology will ensure a role for the VSAT in most telecommunication architectures, whether it is a business data network supporting a major company in the United States or as the backbone of a basic telecommunications service in a Third World country.

Most of the early developments in VSAT systems and service concepts evolved in the United States encouraged by a liberal regulatory environment and the availability of space segment at very competitive tariffs. These systems were developed, mainly, to support business requirements for data distribution and two-way interactive data communications where reconfigurability and rapid deployment are important. More recently, the potential for VSATs in support of communications in developing countries has been recognised.

1.2 What is a VSAT?

A more liberal interpretation of the term VSAT can include any form of small terminal system, irrespective of whether it is part of a dedicated business data network or a data network based on terrestrial television distribution or whether it is military or civil in application.

While most commercially available VSAT systems have been designed to carry data, the underlying traffic can be digitised voice, facsimile, reduced rate video or even narrow strands of data within the composite capacity of the system.

The term VSAT is generally assumed to refer to fixed installation systems; while mobile systems also have small aperture antennas and are based on related technology, many aspects are outside the scope of this book.

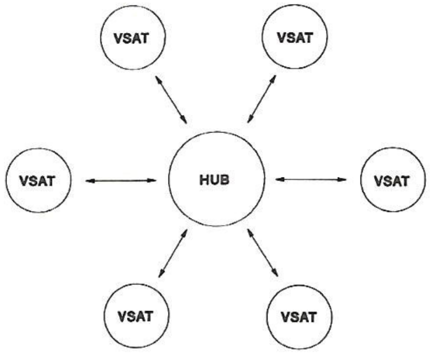

Fig. 1.1 *VSAT star configuration*

The VSAT, or microterminal as it is sometimes called in Europe, is generally assumed to be the remote terminal in a dedicated data network based on the star configuration depicted in Fig. 1.1. This configuration comprises a hub earth station, with a larger aperture antenna, controlling a cluster of VSATs, with small antennas, typically in the range 0·45 to 2 metres (m) diameter.

1.3 Satellite communication frequency bands

Frequency bands for all forms of radio spectrum usage are allocated by the International Telecommunications Union (ITU) with the first allocations for satellite communications made in 1959. These have been revised and extended at subsequent World Administrative Radio Conferences (WARCs) held in 1963, 1971, 1977, 1979 and 1987.

In general, fixed, commercial VSAT systems use satellite transponders operating at C-band (uplink 6 GHz, downlink 4 GHz) or Ku-band (uplink 14 GHz, downlink 11 or 12 GHz) within the Fixed Satellite Service (FSS). More recently, the Broadcast Satellite Service (BSS) in Ku-band has been used. There is also an allocation for fixed services at Ka-band (uplink 30·0 GHz, downlink 20·2 GHz). L-band (uplink 1·6 GHz, downlink 1·5 GHz) is used extensively by Inmarsat for the provision of mobile satellite services.

1.4 Space segment to support VSAT services

Satellites used for communications are almost exclusively in the geostationary orbit, located on an arc 36 000 km above the equator.

Space segment is available from organisations which have procured satellites, arranged launch and preliminary tests in-orbit and who then operate these satellites on a commercial basis. Initially, these organisations were international, e.g. Intelsat offer satellite capacity at C- and Ku-band, but as requirements for satellite communications have increased, regional and domestic systems have appeared on the scene. In addition to these international organisations, a number of companies, mainly in the USA, own or lease satellites which are used to carry their own or their customer's traffic. In most cases, the military use dedicated satellites operating in different bands from civil satellites.

Transponders currently operating in the FSS band typically extend from 36 to 72 MHz in bandwidth with EIRP levels from 30 to 52 dBW. Equivalent isotropically radiated power (EIRP) is the power transmitted from the satellite: it is the product of output power from the satellite's amplifier and antenna gain. In most VSAT systems power rather than bandwidth is the limiting resource in the satellite transponder.

Frequency re-use is usually achieved by using mutually orthogonal right hand circular (RHC) and left hand circular (LHC) polarisation beams in C-band, and by mutually orthogonal linear polarisation beams in Ku-band.

When selecting space segment it is important to ensure that the characteristics of the satellite transponders are suitable for operation with small terminals. The Intelnet [1] service offered by Intelsat and the Satellite Multi-Services (SMS) [2] from Eutelsat are particularly suitable for small terminal applications.

1.5 Network configurations

A star network comprising a hub operating in conjunction with a population of VSATs is the configuration upon which all one-way and most two-way systems are based. The use of the star configuration has been predicated by limitations in the performance of currently available components and systems: it requires the inclusion of a relatively expensive hub terminal.

The star configuration has the advantage that the hub can maintain effective control of the network and it is compatible with most business traffic requirements with the hub either colocated or directly connected to the company head office with individual VSATs serving field offices or retail outlets.

Satellites with higher gain together with improvements in low noise amplifier (LNA) and solid state power amplifier (SSPA) performance now offer the prospect of direct VSAT-to-VSAT communications within a mesh architecture [3] as shown in Fig. 1.2. This offers the feature of reduced propagation delay (typically 0·25 s compared with 0·5 s for the star configuration) which is especially advantageous when the link carries voice traffic.

It may be necessary to designate one VSAT as the 'master' terminal, possibly on a rotational basis, to ensure the exercise of proper control of the network. Ultimately, a mesh network without a master terminal should become a reality when suitable control protocol techniques have been developed.

1.6 A representative VSAT system

The configuration of a typical two-way VSAT terminal identifying the constituent sub-systems is shown in Fig. 1.3 and a representative commercially available product (developed by Multipoint Ltd.) in Fig. 1.4. The earth terminals and their sub-systems are described in Sections 1.7 and 1.8.

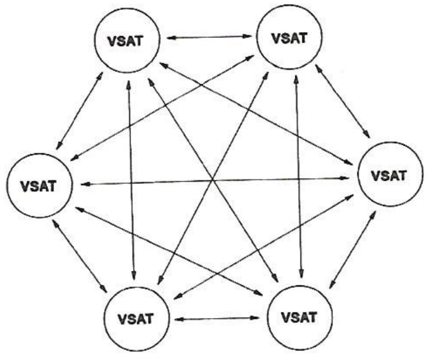

Fig. 1.2 *VSAT mesh configuration*

1.7 Earth terminals in a VSAT network

The principal elements in a VSAT network are the hub and VSAT earth terminals: they perform similar functions in terms of signal handling, the difference being in levels of power transmitted and the performance in the receive mode.

Antenna performance is specified by EIRP in the transmitting mode and the figure of merit G/T in the receiving mode. Both of these parameters depend on antenna size; in general terms, the larger the antenna the better the system performance. These terms and their meanings are defined in Chapter 2.

1.7.1 *Hub earth terminal*

The hub earth terminal can support either one or a number of VSAT networks

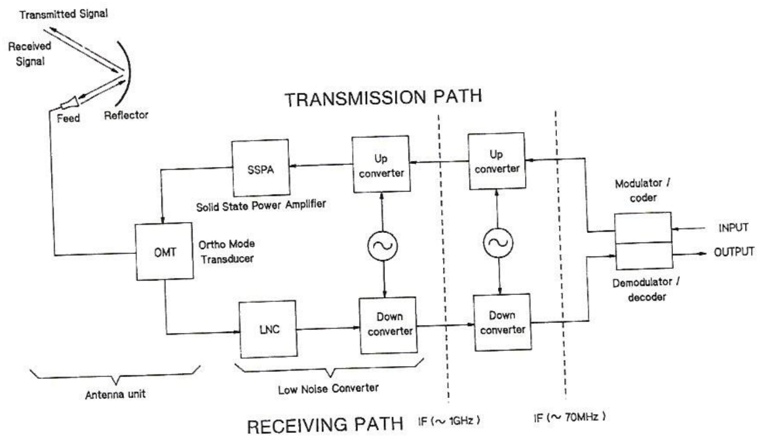

Fig. 1.3 *Configuration of a typical two-way VSAT terminal*

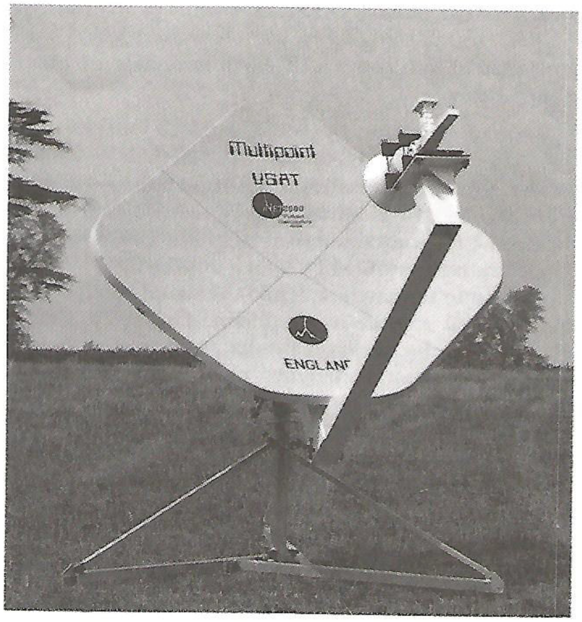

Fig. 1.4 *A representative commercially available VSAT system (courtesy Multipoint Ltd.)*

operating over a given satellite: the shared hub concept is attractive as it allows the high cost of a hub to be amortised over several networks. Earth terminals of this type are expensive, costing between $300 000 and $5 000 000 (at 1991 prices) depending on features and performance.

A typical C- or Ku-band earth terminal used as a hub may have an antenna diameter in the range 5·6 to 11 m.

1.7.2 *VSAT earth terminal*

A VSAT earth terminal is characterised by a much smaller antenna, typically less than 2 m in diameter. Consequently, the unit cost is appreciably lower than that of the hub, typically below $12 000. The VSAT has both lower antenna gain and lower transmit power than a hub, with the power normally generated by semiconductor devices of the type described in Chapter 3.

It is not normally convenient to integrate all the electronic equipment into a single package, for engineering reasons outlined in subsequent chapters describing individual systems. A compromise is reached where the sub-systems carrying microwave frequency signals comprise an outdoor unit while all others are accommodated in an indoor unit, i.e. the interconnections are at an intermediate frequency (IF).

1.8 Earth terminal sub-systems

The sub-systems used in hub and VSAT earth terminals are identified and their functions outlined below.

1.8.1 *Antennas*

The term antenna embraces the reflector, feed and a device to separate transmitted and received signals at the antenna. Any two-way earth terminal has two separate paths for the transmit and receive signals. These paths combine at the orthogonal mode transducer (OMT) which routes the transmit signal to the antenna for uplinking to the satellite, while the signal received from the satellite by the antenna is passed into the receive chain. The feed has to transmit signals at one frequency and receive them on another, usually slightly lower, frequency; the feed/reflector combination is designed as a composite unit to provide these required features.

A major objective in antenna design is to achieve high efficiency and high gain in the direction of the satellite commensurate with low radiated signal levels in other directions, i.e. the production of a narrow 'main' beam with low level 'sidelobes'. Offset feed configurations are generally used in small terminals to avoid beam blockage, giving high efficiency together with low sidelobe levels.

The antenna assembly must be of robust construction, adequate to survive under specified wind loading conditions.

Ku-band links use linear polarisation so the antenna assembly must allow for rotation to enable the feed to be aligned with the polarisation of signals transmitted to and received from the satellite. This feature is not required in the case of circular polarisation used in C-band.

Antennas used in VSAT systems are described further in Chapter 2.

1.8.2 *High power amplifiers (HPAs)*

Travelling wave tube amplifiers (TWTAs) are generally used in hub earth terminals as they can give power output levels up to several kilowatts and have the capability of being tuned across an individual satellite uplink band. Various types of C- and Ku-band TWTs and TWTAs are reviewed in Chapter 4.

1.8.3 *Solid state power amplifiers (SSPAs)*

Improvements in semiconductor technology have yielded SSPAs providing output powers of 50 W in C-band and 20 W in Ku-band: these power levels are adequate for final stage amplification in VSAT earth terminals.

1.8.4 *Low noise converters (LNCs)*

Low noise converters perform the dual function of amplifying and downconverting the received signal while minimising the noise added to the signal. Careful design is essential to ensure that no spurious signals are generated and phase noise in the oscillators is kept to acceptably low levels. Typical noise temperature values of 75 K (C-band) and 100 K (Ku-band) can be obtained, corresponding to noise figures of 1·0 and 1·3 dB, respectively. Especially low noise devices are used in the first stage of the converter as the noise performance of this stage determines the overall noise performance of the converter unit. Low noise converters (sometimes termed downconverters) are described in Chapter 5.

1.8.5 *Up- and downconverters*

Upconverters are used to translate the signal intended for transmission from an intermediate frequency (typically 70 MHz) to a microwave signal (6 GHz in C-band, 14 GHz in Ku-band) where it is amplified in an HPA (or SSPA) for transmission to the satellite. Downconverters translate the microwave signal received by the earth terminal (4 GHz in C-band, 11 or 12 GHz in Ku-band) to a similar intermediate frequency (IF).

1.8.6 *Modems and codecs*

As there is a requirement for a *mo*dulator and a *dem*odulator in each terminal they are usually incorporated in one unit for convenience, this unit being referred to as a *modem*. Similarly, the *co*der and *dec*oder functions are built into another unit called a *codec*.

A digital signal applied to the input of a modulator appears, typically, at the output as a phase shift keying signal (PSK) centred on an IF around 70 MHz. A demodulator in the receive path accepts a PSK signal at IF and converts it back to a baseband digital signal.

The functions of modems and codecs are outlined in Section 1.9 and described in Chapter 6.

1.8.7 *Network interface unit (NIU)*

This unit is needed to implement the user protocol interface and access to the satellite.

1.9 Modulation and coding schemes

In a typical VSAT network a modulator is used to convert digital data to analogue form for transmission over the satellite, with the demodulator at the receiving end of the link used to extract the information even when the signal has been distorted and corrupted by the addition of noise.

Phase modulation schemes are preferred for satellite communications applications since these require a constant power level irrespective of the data transmitted: this avoids the need for transponder load adjustment and smoothing which would be required for a non-constant envelope modulation scheme.

Since VSAT networks are predominantly designed to carry data, digital modulation schemes are invariably employed with phase modulation universally used, i.e. phase shift keying (PSK). This can take the form of binary phase PSK (BPSK) or quadriphase PSK (QPSK).

In power limited systems (always the case with the downlink to a VSAT), the under utilised bandwidth in the satellite transponder can be made available for digital encoding. In encoding for forward error correction (FEC), redundant bits are added to a bit stream so that errors can be detected and corrected at the receiving end of the link. Some modulators are used in conjunction with external coders while others have integral coders: thus, the boundary between digital

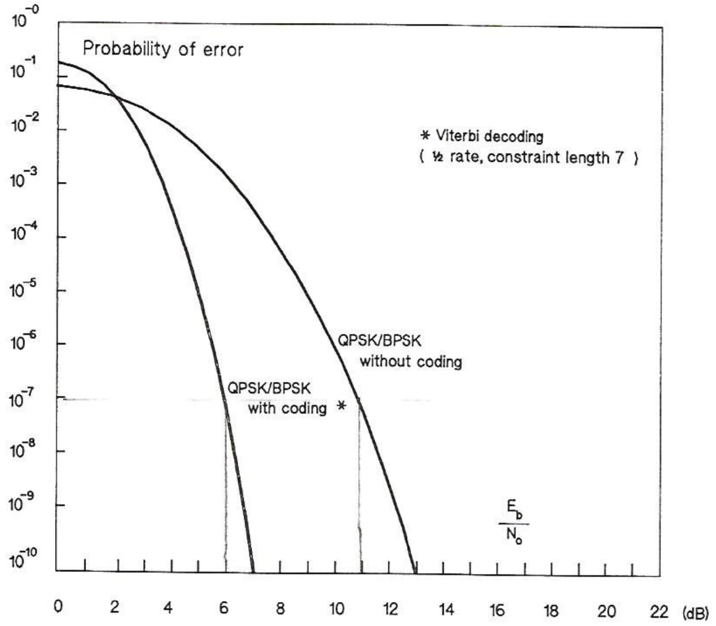

Fig. 1.5 *QPSK/BPSK with and without coding*

modulation and digital coding may not be clearly defined. Bit error rate (BER) values obtained for various E_b/N_0 ratios with and without FEC coding are shown in Fig. 1.5. This relationship indicates the performance of a modem or modem/codec combination.

1.10 The communication of data across the VSAT network

The objective of the VSAT is to provide an end-to-end communications link in a way which is as transparent as possible to the user. From a user perspective, external interfacing is required to allow the protocol on the user's terminal equipment to be understood and processed by the VSAT system. The network protocol employed by the VSAT facilitates the efficient transfer of user data over the satellite link, while the multiple access scheme allows many users to share the satellite transponder resource.

1.11 Multiple access

1.11.1 *Multiple access schemes*

As one transponder may be required to handle transmissions from a number of

---

different
access.
The tra
the use of

1.11.1.1
transponder
able to
to avoid
cause

1.11.1.2
the full
to another
is that the
running
TDMA
way
operating

different earth stations, it is necessary to use techniques which allow multiple access.

The transponder resource can be shared by frequency or time allocation or by the use of non-interfering codes.

1.11.1.1 *Frequency division multiple access (FDMA)* In FDMA users share the transponder by prior allocation of individual channels: this technique is applicable to analogue and digital services. Transponders are operated in a linear mode to avoid non-linearities which would generate intermodulation products and cause interference to all users.

1.11.1.2 *Time division multiple access (TDMA)* In TDMA each user is assigned the full bandwidth of the channel for a short period which is then made available to another user for the next period, and so on. The attraction of this technique is that the transponder is operated at high power levels (and high efficiency), even running close to saturation without interference being caused to other users. TDMA is intended primarily for very high bit rate digital modulation schemes.

Conventional TDMA has high capacity (typically 120 Mbit/s) which is a long way outside the capabilities of VSATs. However, partial transponder TDMA operating in the range 32–128 kbit/s is used in some VSAT systems.

1.11.1.3 *Code division multiple access (CDMA)* CDMA is based on the properties of direct sequence spread spectrum. Spread spectrum is a means of transmission in which the transmitted signal occupies a bandwidth in excess of that of the data signal. This 'spreading' is achieved by combining the data signal with a high bit rate code signal which is independent of the data. Reception at the distant end of the link is accomplished by mixing the incoming composite data/code signal with a locally generated and correctly synchronised replica of the code to obtain the data. The codes used for this purpose have characteristics allowing any individual code to be distinguished from others: it is this property which permits several composite signals to share a common frequency band, i.e. the code division multiple access feature.

There are practical restrictions in the use of spread spectrum which means that it is only employed for interference rejection or for security reasons in military systems.

1.11.2 *Selection of access scheme*

It has to be remembered that the multiple access scheme is simply providing a channel through the transponder for the traffic.

Selection of an access scheme has to take account of the requirements of (what may be) a changing population of VSATs to access a satellite in a way that optimises satellite capacity, satellite EIRP and spectrum utilisation in a flexible and cost effective manner. Clearly, since it is unlikely that all these factors can be optimised, some compromise is necessary.

Further details of multiple access schemes and the basis of their selection for specific applications will be found in later chapters describing two-way interactive VSAT systems.

1.12 Network or multiaccess protocols

Once the multiple access scheme has been selected there is a need to define network protocols which provide the means of interface between the VSAT and hub station. A proprietary protocol layer is introduced between the remote VSAT and the hub to support the end-to-end communications link.

It will be apparent from later chapters that there is considerable interaction at the design stage between the multiple access schemes and network protocols selected for use in a system.

Within a specified multiple access scheme consideration must be given to the proportion of capacity allocated to each earth terminal. When this is decided in advance of transmission it is fixed or preassigned access: when it is in response to changing traffic demands it is demand access. Preassigned access schemes are appropriate for systems containing a limited number of VSATs where traffic levels are relatively constant. In contrast, demand access is used in systems with large populations of VSATs where traffic is highly variable in terms of density, origins and destinations.

In those applications where traffic is bursty (i.e. short data bursts at random intervals) there are incentives for using channel sharing protocols which can support a large number of remote terminals on each inbound (i.e. VSAT to hub) channel. This is achieved by allowing an earth terminal to transmit when traffic is available and to wait for an acknowledgment: if no acknowledgment is received within a specified time period, due to interference or collisions with other data bursts, the traffic continues to be transmitted at intervals until an acknowledgment is received. This random access approach avoids the need for the overhead required to support a demand access scheme for the supply of dedicated channel resources.

Not surprisingly, there are trade-offs between channel utilisation efficiency and the complexity and cost of implementing protocol schemes at the VSAT stations. In satellite systems power and (to a lesser extent) bandwidth are constrained; in these power limited situations, a relatively inefficient protocol may be used without major adverse effect on system throughout. It should be noted that in those applications where the traffic is quasi-continuous (over 0·5 s duration) preassigned access schemes are more efficient and have the added advantage of simplified implementation and consequent lower cost.

The important aspects to be considered in selecting a channel sharing or network protocol for a VSAT system are:

(*a*) Data type, transactional or quasi-continuous
(*b*) Channel sharing efficiency (i.e. the percentage of time that data is carried on multiple access channels)
(*c*) Delay in transmitting data (average and peak) resulting from traffic dependent queueing delay
(*d*) Robustness when channel errors and equipment failures occur
(*e*) System operation—start up and recovery factors
(*f*) Implementation complexity and cost

Many of the data communications protocols used in VSAT systems were developed originally for terrestrial links and have had to be modified to operate with the propagation delay associated with geostationary satellites, i.e. 0·5 s for a system based on a star configuration.

As will be apparent from succeeding chapters, a number of user protocols can be supported by commercially available VSAT networks.

Network protocols (sometimes termed multiaccess protocols) and the protocol software used in VSAT networks are described in Chapters 7 and 8 respectively, followed by an overview of system design in Chapter 9.

1.13 Network management

Network management is an essential feature in any VSAT network in order to maintain system integrity. The management system must be integrated into the network at the design stage and be sufficiently versatile so that it can respond to changing requirements of users and be able to monitor the health and status of the system: the health monitoring function should be performed periodically in a manner transparent to the user. In general terms VSAT network management embraces administrative, operational and planning functions, several of these being identified and discussed in Chapter 10.

1.14 One- and two-way VSAT systems

In generic terms these systems can be considered in three forms:

(*a*) Data distribution: one-way
(*b*) Data gathering: one-way
(*c*) Interactive: two-way

In functional terms these three forms are represented in Figs. 1.6*a*, 1.6*b* and 1.6*c*.

1.14.1 *Dedicated one-way data systems*

The hub in a star configuration data distribution system broadcasts to a network of dispersed VSAT terminals with antenna diameters within the range 0·45 to 2·0 m. More sophisticated systems embody selective addressing to the sub-group or even individual terminal level. BPSK is normally used to support data rates up to 1 or 2 Mbit/s [VSATs designed to handle a T1 carrier (1·544 Mbit/s) are sometimes termed TSATs].

In applications where extremely small antenna size (typically below 0·6 m diameter) is critical, direct sequence spread spectrum [4] is used to enable the receive-only terminal to reject unwanted signals (representing interference) from adjacent satellites on the geostationary arc. The small terminal scenario is represented in Fig. 1.7.

The Equatorial Communications company made substantial progress in the 1980s in developing and installing data distribution receiving systems, based on spread spectrum, in business premises across the United States.

It should be noted that the unwanted signal problem with small terminals becomes less severe as the link frequencies are increased because the beams become narrower and, ultimately, the protection afforded by spread spectrum is not needed.

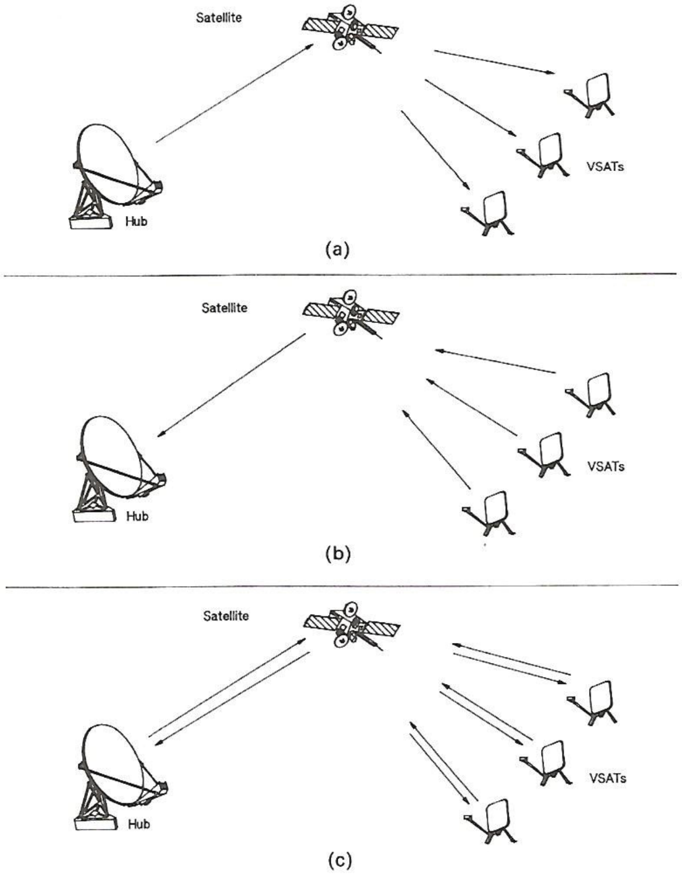

Fig. 1.6 *(a) Data distribution system. (b) Data gathering system. (c) Two-way interactive system*

A number of one-way systems have been developed and some are commercially available. Representative systems from Ferranti and Polycom are featured in Chapters 11 and 12, together with APOLLO, an advanced document distribution system, developed by the European Space Agency (ESA), in Chapter 13.

Data gathering requirements are satisfied by two-way systems. A scenario based on the use of dispersed VSAT units relaying data into a hub terminal without a return link is unacceptable to satellite operators as the hub would not be able to communicate to the remote VSATs nor exercise any control over them. This latter factor represents a serious flaw in the operating concept since a given VSAT could suffer a malfunction resulting in it operating on a frequency allocated to another user, causing incurable interference. For this reason, principally, two-

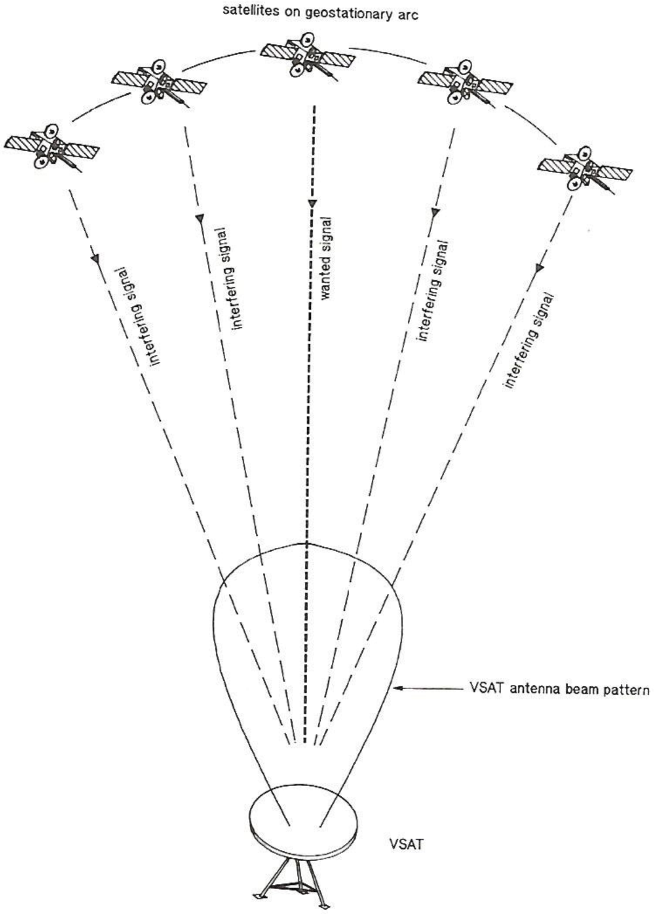

Fig. 1.7 *Small terminal scenario*

way interactive VSAT systems are used for data gathering applications even though this results in a more expensive implementation.

1.14.2 *Data distribution based on television broadcasting*

Data distribution based on television broadcasting can be considered in two classes: that based on conventional transmission schemes e.g. PAL (or NTSC, SECAM) and that based on Multiplexed Analogue Component (MAC) [5] packet systems. The former are relayed over satellites operating in the FSS band while MAC systems are intended for use on dedicated direct broadcast satellites (DBS) operating in the BSS band. However, a limited number of MAC signals are broadcast over Europe from satellites operating in the FSS bands.

Data can be broadcast in conjunction with conventional television schemes by using either Vertical Blanking Interval (VBI) or sub-carrier techniques. In the VBI scheme data are transmitted at a rate of 9·6 kbit/s per line with up to 5 lines normally being used; this technique is well established and supports teletext services on terrestrial television. In the sub-carrier case data are combined with the television video and audio components in an FM modulator. Data rates up to several hundred kbit/s can be transmitted but the data signal 'steals' power from the television signal and so data rates are usually restricted to around 100 kbit/s.

The MAC packet concept represents an entirely new television broadcast system providing a highly flexible medium for the transmission of video, sound and data in separate fields in the MAC packet format.

A proof-of-concept data distribution system based on the MAC scheme is described in Chapter 14.

1.14.3 *Two-way systems*

A range of two-way systems has been developed to satisfy different requirements and several are now available as products. A representative realisation is based on a hub in a star configuration transmitting a time division multiplex (TDM) stream to all VSATs in the network. The multiplex contains network control instructions and messages for some or all of the VSATs. A VSAT with a message for the hub (or for another VSAT via the hub) will transmit a short duration burst on a 'calling' channel requesting access to a channel to transmit its message: the hub acknowledges the request, assigns a channel and the VSAT changes frequency and transmits its message. In a busy network there will be collisions between some access request bursts and the VSAT may not get an acknowledgment from the hub. Under these circumstances the VSAT will retransmit its burst request after a pseudo randomly determined interval, and continue doing so until it receives an acknowledgment and is assigned a channel. At the end of the message the transmission is terminated and the message channel becomes available for reassignment.

PSK modulation is normally employed for the TDM broadcast and the VSAT transmissions, often in conjunction with FEC coding on the hub to VSAT link. A variety of multiple access techniques and multiaccess protocols are in use, most of them proprietary which means that systems from different vendors are not generally compatible with each other.

Representative two-way systems available from AT&T Tridom, Hughes Network Systems, NEC and Multipoint are described in Chapters 15 to 18; in contrast, a satellite based messaging system, where much longer transaction delays can be tolerated, is described in Chapter 19.

1.15 Ka-band VSAT systems

Increasing requirements for satellite communications will ultimately result in C- and Ku-band allocations becoming fully utilised. In anticipation of this eventuality, the next band allocated for satellite communications, Ka-band, is being investigated and prototype equipment is being developed and evaluated.

The allocations in Ka-band are 27·50 to 30·00 GHz for uplinks and 17·70 to

20·20 GHz for downlinks, i.e. a bandwidth of 2·50 GHz compared with only 0·5 GHz for uplinks in C- and Ku-bands.

The Advanced Communications Technology Satellite (ACTS) [6] being acquired by the National Aeronautics and Space Administration (NASA) in the United States will have a number of advanced features including a transponder with an on-board baseband processor to facilitate thin route experiments and service demonstrations in Ka-band.

A prototype VSAT system developed to support a range of experiments on the Ka-band transponder of the Olympus satellite, together with representative results, is described in Chapter 20. The conclusions from initial experiments are that availability levels are higher than predictions and Ka-band satellite communications, including those based on small terminal systems, will have significant impact in the future.

1.16 Military VSAT systems

Military satellite communication links are expected to operate under conditions significantly different from those encountered by commercial systems. Communications service may have to be established within a very short time of a terminal arriving at a location, often in inhospitable geographical areas with no existing terrestrial infrastructure and an intentionally hostile electromagnetic environment in the form of jamming.

In order to satisfy their communications requirements, military authorities in several countries have acquired dedicated space segment resources and specially developed small terminals which can be of various forms; e.g. shipborne, airborne, land mobile (vehicle) and even manpack units.

The general principles of military satellite communications systems have been outlined [7] and small aperture military ground terminals are described in Chapter 21.

1.17 Link budgets

Prior to the establishment of a satellite link an essential preliminary step is the calculation of signal levels through the system (from originating earth terminal, through the satellite and into the receiving earth terminal), to ensure that they are adequate to provide the quality of service required. Impairments to system performance arise from noise entering the link at various points; this noise is generally dominated by thermal and shot noise arising in the front end of receivers but the radiation temperature of the earth may also be significant. Unwanted signals from other earth terminals and satellites not associated with the link under consideration will degrade the link performance. A simplified block diagram of a satellite link upon which the link budget approach is based is shown in Fig. 1.8.

When computing link budgets, account must be taken of the form of modulation and coding (if used) and atmospheric/propagation losses which are dependent on the frequencies used for the uplink and downlink, as these determine service availability (i.e. the percentage of time that the link provides acceptable service).

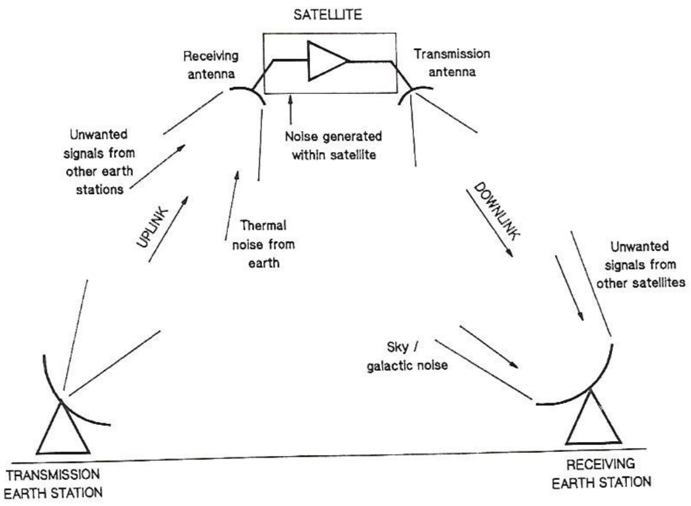

Fig. 1.8 *Simplified block diagram of satellite link for link budget purposes*

A derivation of a satellite link budget in tutorial form together with example VSAT scenarios is presented in Chapter 22.

1.18 VSAT applications and services

VSAT systems have developed rapidly, particularly in the United States, where they have supported the introduction of new communications services that have become indispensible to the efficient operation of many corporations and businesses. One of the consequences of this large market is that production costs of small terminals have fallen to relatively low levels. Nevertheless, it could be argued that VSATs still represent a niche market and have yet to achieve their full potential. In those countries where the regulatory environment is liberal, there is very substantial competition from other technologies, e.g. optical fibre, conventional cable, microwave links and wireless personal communication networks (PCNs).

VSATs also have a major role to play in those Third World countries with virtually no terrestrial telecommunications infrastructure: as an example, VSAT data services based on spread spectrum are now available from the telecommunications administration in India.

Studies have confirmed that in those developing countries where satellite communications have been introduced, there has followed substantial economic growth, a clear indication that low cost low complexity VSAT systems can be of potential benefit to the Third World where a communications infrastructure may be lacking. This theme is pursued in Chapter 23.

Two-way VSAT networks offering voice and data services or even two-way video conferencing, to satisfy domestic and international requirements, are now widely used. These developments are causing communications services to evolve around customer requirements, either within a country or, increasingly, on an international basis. Examples are news distribution services, voice and data links to isolated locations (mineral exploration sites and oil platforms) and non-standard service options not available from terrestrial communications networks. A review of applications and a brief survey of the VSAT market worldwide is presented in Chapter 24.

The liberal regulatory environment prevailing in the United States has permitted organisations to provide a wide range of communications services (for themselves and others on a commercial basis) by allowing acquisition of earth terminal equipment and acquisition or leasing of satellite transponder capacity from various satellites (to provide coverage of service areas with adequate power levels) and the ability to procure VSATs with the appropriate features (voice and data at different rates). Also, the resources to provide end-to-end service (including equipment installation, interfacing between VSATs and customer terminal equipment and maintenance support). These aspects are described in Chapter 25.

1.19 Economic considerations

The substantial VSAT base in the United States enables service offerings to be made available at acceptable tariffs: in other parts of the world, including Europe, this is not necessarily the case. Costs for a given terrestrial service vary significantly across Europe depending in which country the service is provided: these costs have been ascertained and compared with costs associated with a VSAT solution (based on two different access schemes, TDMA and CDMA) offering a similar service. Results presented in Chapter 26 indicate that the introduction of VSAT networks into Europe can be expected to offer communications service at costs below those currently charged for terrestrial communications service. However, it should be noted that tariffs in Europe have traditionally been based on what the market will bear rather than the true cost to provide a service.

An important economic factor is that interactive/transactional data services are often more viable than stream type services with high volume traffic per terminal. This follows from the ability of VSAT systems to multiplex signals from a large number of low volume traffic terminals cost effectively, a facility which can be difficult to achieve with terrestrial network technologies.

1.20 Regulatory considerations

Early satellite communication systems were based on very large earth terminals (with antennas of 30 m diameter or more) at both ends of the link and, as a consequence, they were owned and operated by national organisations, usually those responsible for Posts, Telegraph and Telephone (PTT) communications. International agreements regulated communications across national borders in a relatively harmonious manner. Improvements in technology have facilitated the realisation of earth terminals sufficiently small and economic that they can be located on customer premises. The control of networks based on these small terminals is a major issue for many governments and PTT organisations, as the networks can readily 'bypass' well established terrestrial networks with consequences for the diversion of revenue to a service provider other than a PTT.

Liberalisation has occurred in several parts of the world, principally in the United States, where an 'open skies' policy allows any competent operator to provide communication services, while in other countries control remains firmly vested in the PTTs. The effect of this imbalance in control, or regulation as it is normally termed, is that VSAT services have proliferated in countries with minimal regulation but with very modest penetration elsewhere. It is not surprising, therefore, that most of the advanced systems and innovative services described in later chapters have appeared in the United States. This situation is changing as several countries with insubstantial terrestrial telecommunications infrastructure are pursuing VSAT solutions to satisfy their needs.

It is instructive to examine the situation in Eastern Europe where there is now significant penetration of VSAT networks to establish a basic telecommunications service in the absence of any credible terrestrial infrastructure. The VSAT provides an interim, rapidly deployed network which will provide service until a permanent terrestrial system can be established.

The regulatory environments prevailing in the United States, the United Kingdom and continental Europe are reviewed in Chapter 27.

1.21 Future developments

Developments in VSAT and satellite technology in the future will be driven by the market requirements to reduce the system cost still further and to expand the range of services. Normal technical and manufacturing developments can be expected to reduce costs to around $1000 to $2000 per VSAT unit. The trend towards the ultra small aperture terminal (USAT) will provide extremely cost effective point-of-sale and home personal computer (PC) data services based on DBS terminals.

In the medium term (by the mid-1990s) these developments are expected to result in more integration of microwave functions based on monolithic microwave integrated circuits (MMICs) and more integration of digital functions, especially in respect of higher performance and lower cost voice and video codecs.

In the longer term (late 1990s) developments in space segment technology will lead to the use of high power spot beams (possibly scanning across dispersed geographical areas), on-board processing (OBP) and inter-satellite links (ISLs).

The ultimate development in small terminal technology may be represented by the IRIDIUM [8] concept proposed by the Motorola company to provide a world-wide network of hand-held terminals operating in conjunction with a constellation of 77 low earth orbiting satellites.

A European Space Agency perception of the future of VSATs is presented in Chapter 28.

1.22 Conclusions

The growth of VSATs has resulted from the complex mix between technology, regulation and market forces which, increasingly, allow the exploitation of the capabilities offered by these small and versatile systems.

The subsequent chapters cover in depth the many issues which have contributed to the considerable success of VSAT systems.

1.23 References

1 JAMSHIDI, S., and NGUYEN, L.N.: 'INTELNET services—A global data distribution and collection scheme', *Int. J. Satellite Communications*, 1986, **4**, pp. 83–87

2 PAPO, M.: 'The European satellite business service—EUTELSAT SMS, Satellite integrated communications networks'. Proceedings of the Third Tirrenia International Workshop, 14–16 September 1987 (North-Holland, Amsterdam, 1988) pp. 305–311

3 CROSS, M.A., and FLEETWOOD, T.: 'Hubless VSAT networks'. Proceedings of the European Satellite Communications '89 conference (Blenheim Online, London, 1989) pp. 93–106

4 TRI. T. HA.: 'Spread spectrum for low cost satellite services', *Int. J. Satellite Communications*, 1985, **3**, pp. 287–293

5 STALLARD, W.G.: 'An overview of the D-MAC/packet system', *IBA Tech. Rev.*, 1988, **24**, pp. 11–15

6 WRIGHT, D.L. *et al.*: 'Advanced communications technology satellite (ACTS) and potential systems applications', *Proc. IEEE*, 1990, **78**, pp. 1165–1175

7 HOWARTH, D.P.: 'Military satellite communications', *in* EVANS, B.G. (Ed): 'Satellite communication systems', IEE Telecommunications Series 24 (Peter Peregrinus Ltd, 1991 2nd Edn.) Chap. 15

8 PARKER, I.: 'Iridium turns to gold?', *Space*, 1991, **7** (Shephard Press Ltd, London) pp. 24–26

Chapter 7

Multiaccess protocols for VSAT networks

D. Raychaudhuri and K. Joseph
David Sarnoff Research Center, Princeton, New Jersey, USA

7.1 Introduction

In the past few years there has been increasing interest and commercial activity in the area of VSAT satellite data networks. VSAT networks employ inexpensive customer-premise earth stations with relatively small antennas (typically less than 2 m) to provide direct interactive data services to geographically dispersed remote locations. Such satellite based 'wide area networks' potentially offer several economic and technical advantages over conventional terrestrial technologies. This potential can now be realised because of a combination of recent advances in satellite and earth station technology. These include the deployment of high powered Ku-band satellites, the development of low cost solid state power amplifiers and earth station electronics, along with advances in shared channel access (multiple access) techniques.

VSAT data networks generally have a star configuration as shown in Fig. 7.1. In a typical interactive inquiry-response application, the remote stations communicate with a central host computer which is physically located at, or directly connected to, a large hub earth station. Transmissions from the hub to the terminals are sent out in a time division multiplexed (TDM) broadcast over one or more frequency multiplexed medium to high speed (56 kbit/s to 1·544 Mbit/s) satellite channel. In the terminal to hub direction, channel spacing considerations and transmit power limitations along with the delay performance requirements of the VSAT network constrain the typical transmission speed to the region of 32 kbit/s to 128 kbit/s. At these channel speeds, transmissions from remote stations supporting most interactive applications are very bursty, motivating the use of channel sharing protocols which support a large number of stations on each inbound (terminal to host) frequency. These multiaccess protocols are a critical element of VSAT systems, since their characteristics have a significant impact on network performance, space segment cost and VSAT equipment complexity.

The VSAT-based interactive data network scenario differs significantly from previous satellite systems employing multiaccess protocols. Specifically, in contrast to the earlier generation of satellite systems which used frequency division multiple access (FDMA) or time division multiple access (TDMA) to link large earth stations with high traffic cross-sections, multiaccess technology for VSAT networks is not primarily driven by the efficiency of bandwidth use. In the VSAT scenario, each station generates a relatively small amount of traffic, so that transmission cost per VSAT is typically overshadowed by the cost of the VSAT station and its operation. In addition, as will be seen in Chapter 9, power rather

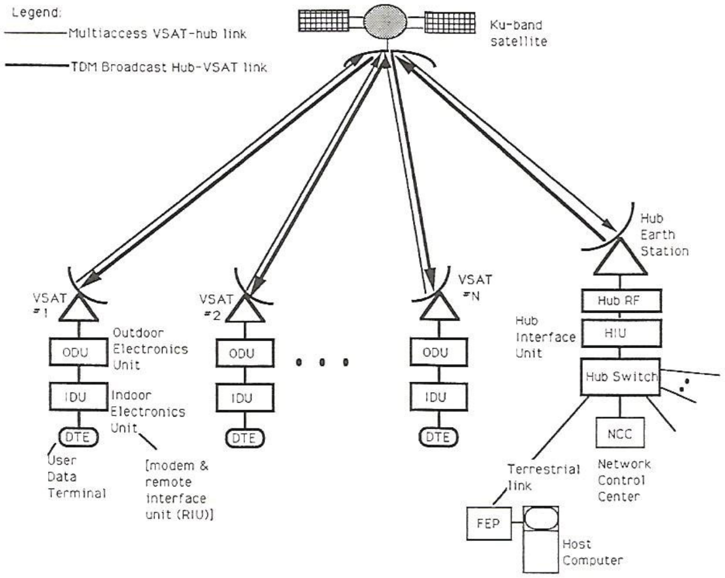

Fig. 7.1 *Typical star architecture VSAT network*

than bandwidth is often the limiting resource in VSAT networks, permitting the use of relatively inefficient protocols without significant impact on system capacity. Thus, for the interactive data VSAT network environment, low delay, simplicity of implementation and robust operation are generally of greater importance than the bandwidth efficiency achieved. Observe that some of these goals are analogous to those which drove the evolution of channel sharing protocols for local area networks (LANs). It may be expected that, as the industry matures, a limited number of 'standard' VSAT protocols (analogous to CSMA/CD [1] and Token Passing [2] for LANs) will emerge.

The problem of developing suitable multiple access techniques for ground radio, LAN and satellite wide-area network scenarios with bursty interactive traffic has attracted considerable research attention since the introduction of the ALOHA protocol (which offers low delay, minimal implementation complexity and robust operation at the expense of low throughput) in 1970 [3]. A detailed survey of multiaccess protocols that have been developed for all of the above broadcast channel scenarios has been given by Tobagi [4]. Another review paper by Lam [5], emphasises the satellite scenario, which is characterised by channel propagation delays that are much greater than the message transmission time. From References 4 and 5, it is clear that the desirable combination of high throughput, low delay and simple implementation is readily achieved in local area and ground radio networks with short channel propagation delays, but is much more difficult to achieve in the satellite environment. In Section 7.2, a survey is presented of many of the available alternatives for VSAT multiaccess, covering both established approaches discussed in References 4 and 5 as well as some more recent developments in the area. The review is followed by Section 7.3, which presents a fairly detailed performance comparison between a selected

---

number of
network sc

7.2 Revie
netwo

The discu
attributes
techniques. T
for the VSAT

(a) The
traffic
denote
to com
becau
(b) The
the sta
averag
transm
transm
(c) The
conges
(d) *Robus*
(e) *Operati*
types
(f) The
requir

The class
protocol is a
communica
protocols as
synchronisa
different di
message tra
types: *fixed*
between co
such protoco
discussion. T

*Note that al
7.1, the proto
ed) satellite ne
In the discuss
of a broadca

Section 7.2 has
data networks
Comms., 1987
Reproduced b

number of 'first generation' VSAT protocols applied to an interactive data network scenario.

7.2 Review of satellite multiaccess protocols for VSAT networks

The discussion is commenced by defining some of the important performance attributes which will serve as the basis for comparing dissimilar multiaccess techniques. The key issues to be considered in selecting a channel sharing protocol for the VSAT scenario are:

(a) The *efficiency* or *throughput* of channel sharing (i.e. the fraction of time useful traffic is carried by the multiaccess channel). The term *capacity* is used to denote the maximum operating throughput of the protocol. It is important to compare protocols on the basis of *net*, rather than nominal capacity, because of wide variations in transmission overheads.
(b) The *access delay* (i.e. time between the arrival of a message at a VSAT and the start of its successful transmission on the channel), both in terms of average and distribution characteristics. Note that the *total* delay in transmitting the message is the sum of the access delay and the message transmission time.
(c) The *stability* properties, relating to the possibility of undesirable long-term congestion modes.
(d) *Robustness* in presence of channel errors and equipment failures.
(e) *Operational properties* relating to start-up, addition of new stations or traffic types etc.
(f) The *implementation cost/complexity*, related to the VSAT hardware/software required.

The classification of access protocols is determined as follows. A multiple access protocol is a set of rules by which a number of distributed remote stations communicate reliably over a shared channel*. It is convenient to classify such protocols as 'slotted' or 'unslotted' depending on whether or not continual time synchronisation between stations is required for the specified operation. A different dimension for classification is based on the qualitative nature of the message transmission discipline used. Typically, message access can be of three types: *fixed assigned*, *contention* (*random access*) or *reservation* (*controlled access*). Hybrids between contention and reservation are frequently proposed, and for simplicity, such protocols will be placed in the reservation category for the purposes of this discussion. Table 7.1 provides a listing of many conventional and new satellite

---

*Note that although VSAT networks are most often based on the star architecture of Fig. 7.1, the protocols discussed in this Chapter are equally applicable to mesh (fully connected) satellite networks which may become more viable as satellites become more powerful. In the discussions that follow, the usual convention of describing protocols in the context of a broadcast (i.e., mesh) channel will be observed.

Section 7.2 has been adapted from: RAYCHAUDHURI, D., and JOSEPH, K.: 'Ku-band satellite data networks using very small aperture terminals — Part 1: Multiaccess protocols', Int. J. of Satellite Comms., *1987, pp. 195–212*

Table 7.1 *Summary of satellite multiaccess protocols*

Message access type	Channel synchronisation	
	Unslotted (1)	Slotted (2)
Fixed assignment	SCPC–FDMA CDMA	TDMA
Contention (random access)	Pure ALOHA SREJ–ALOHA Time-of-Arrival CRA* Unslotted RA–CDMA SREJ–ALOHA/FCFS*	Slotted ALOHA Tree CRA ARRA Slotted RA–CDMA
Reservation (controlled access)	Locally synchronous* reservation with: Time-of-arrival access SREJ/ALOHA access ALOHA access	DAMA with: TDMA access Slotted ALOHA access Tree CRA access Hybrid reservation/ random-access

* Uses channel event based local timing

access protocols* within the framework of the above classification. Also, to summarise the detailed discussions that follow, Fig. 7.2 presents a side-by-side performance comparison of 'typical' throughput (S) versus average delay characteristics for each of the protocols discussed in the next section. Note that the curves shown are intended only for qualitative assessment of relative performance characteristics, and should not be used as an absolute reference point†. In addition to the throughput-delay summary in Fig. 7.2, for convenient reference, a qualitative side-by-side comparison of the protocols considered (in terms of the key attributes listed earlier) is given in Table 7.2.

7.2.1 *Fixed assigned multiaccess*

7.2.1.1 *SCPC/FDMA* The simplest form of satellite multiaccess is SCPC/FDMA [6], in which the available frequency of a transponder is partitioned into multiple low to medium speed SCPC (single channel per carrier) channels. Although the nominal capacity of SCPC can approach the ideal 1·0 (along with zero access delay), its use for bursty applications is generally inefficient because of a fundamental mismatch between the required burst speed (based on transmission delay considerations) and the average data rate of a terminal.

---

*Owing to space limitations, the list of protocols discussed here is intended to be representative, rather than exhaustive. For more complete coverage, the reader is urged to supplement the discussion with References 4 and 5.

†Note that the curves in Fig. 7.2 are based on the simple interactive-station/fixed-message-length traffic model, commonly used in the multiaccess literature.

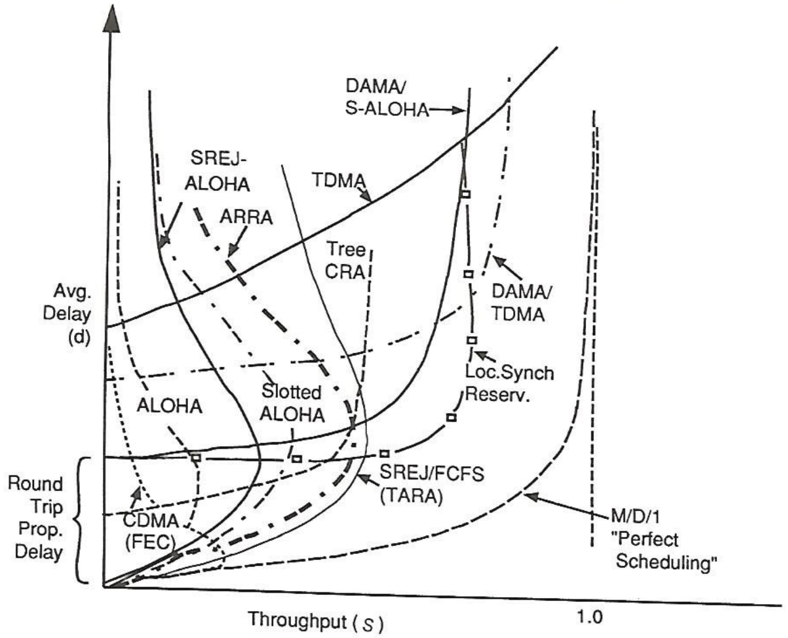

Fig. 7.2 *Summary of delay-throughput characteristics for typical VSAT multiaccess protocols*

7.2.1.2 *CDMA* It is also possible to subdivide the resources of a satellite transponder using code division multiple access (CDMA), in which the information signal is spread into a wide bandwidth using spread spectrum techniques, e.g. References 7 and 8. Fixed assigned CDMA is generally characterised by low bandwidth efficiency and is used in situations where the other properties of spread spectrum (such as low power spectral density, rejection of adjacent satellite and/or terrestrial interference and anti-jam capability) are of importance. Thus, CDMA may be an appropriate approach for power density and interference limited scenarios associated with very small remote station antennas (< 1 m). Typical capacities (maximum throughput normalised to bandwidth occupancy) achieved by fixed assigned CDMA without forward error correction (FEC) are in the region of 0·05–0·1, increasing to 0·2–0·3 when powerful FEC is used. Although access delay for fixed assigned CDMA is zero, the transmission delay may be substantial in systems based on relatively low effective station transmission rates (typically 1·2–9·6 kbit/s).

7.2.1.3 *TDMA* TDMA (time division multiple access) [9, 10] is based upon synchronisation of channel time into periodic frames, containing a time slot reserved for access by each station sharing the channel. Since fixed guard times are required between the TDMA slots, increasing the number of stations *N* will tend to increase the latency delay for access and/or decrease the proportion of

Table 7.2 *Summary comparison of satellite multiaccess protocols*

Multiaccess protocol	Capacity	Delay	Robustness	VSAT cost/ complexity	Remarks
Slotted ALOHA	0·37	low	good	med.	Simplest slotted technique, suitable for fixed length messages
Tree CRA	0·43–0·49	med.	poor	med.–high	Stable operation, high capacity for fixed length messages; interleaving delay and deadlock problems
ARRA	0·5–0·6	low	moderate	med.–high	Side information used to increase slotted ALOHA capacity without degrading delay
Packet CDMA	0·1–0·4	v. low	good	med.	Offers low delay with spread spectrum advantages; Complex FEC required for high end of capacity range
DAMA	0·6–0·8	high	poor	high	High throughput by reservation at the expense of latency delay and complexity; for long messages
Hybrid reservation/ random access	0·6–0·8	variable	poor	v. high	Low delay for short messages, high reservation delay for long messages. More complex than DAMA

SLOTTED SYSTEMS

ALOHA	0·13–0·18	low	good	v. low	Simplest possible VSAT implementation at the expense of throughput; suitable for short variable length messages
SREJ–ALOHA	0·2–0·3	low	good	low	Significant capacity increase over ALOHA with minimal added complexity; Suitable for short to medium length variable length messages
Time-of-arrival CRA	0·41–0·51	low	moderate	med.	ALOHA-like operation with higher capacity; for fixed length messages; Exploits CSMA or CSMA/CD type receiver capability
Asynchronous packet CDMA	0·1–0·3	v. low	good	low–med.	Low delay and spread spectrum advantages without slotting. Only slightly lower capacity than slotted CDMA. Complex FEC required for high end of capacity
SREJ–ALOHA/FCFS	0·45–0·6	low	moderate	med.	Semi-compatible upgrade of SREJ–ALOHA; Suitable for bimodal message length distribution
Locally synchronous reservation	0·6–0·9	high	moderate	med.	DAMA-type features without explicit time slotting; for long messages.
UNSLOTTED SYSTEMS					

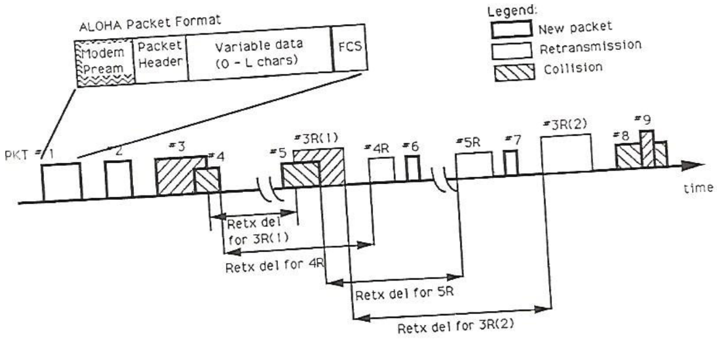

Fig. 7.3 *Illustration of channel events for the ALOHA protocol*

useful traffic carried. In addition, as the channel is partitioned into smaller fixed assigned segments, the effective channel speed seen by the stations will decrease, adversely affecting the delay performance. Thus, TDMA is suitable only for networks with a small number of medium to high volume VSATs per channel, typically achieving capacities in the region of 0·6–0·8 for such applications.

7.2.2 *Contention/random access protocols*

Random access (or contention) protocols, in which data messages transmitted on the channel may 'collide' with each other, have attracted considerable attention since the introduction of ALOHA in 1970 [3]. Contention protocols are motivated by the inapplicability of fixed assigned protocols such as TDMA to an environment in which the channel must be shared by a large population ($N \gg 1$) of bursty stations with low average data rate. It is clear that, for such traffic, some form of dynamic sharing of the channel resource, which does not incur overhead or delay penalties which increase with N, must be employed. One approach to dynamic channel sharing is to permit all stations random (i.e. nominally unrestricted) access to the channel, while maintaining channel loading at a level which is low enough to ensure a relatively low probability of contention. For the messages which do encounter collisions, a *collision resolution algorithm* (CRA) is executed, ultimately resulting in the successful transmission of all messages which were initially involved in the collision. Generally speaking, the specific nature of the CRA employed (which may vary from simple to fairly complex strategies) distinguishes one random access protocol from another.

7.2.2.1 *Unslotted ALOHA* Pure (asynchronous) ALOHA [3] is the first and simplest possible random access technique, requiring absolutely no timing or logic co-ordination among stations on the channel. In ALOHA, stations transmit new messages on the channel as they are generated. Collision resolution is achieved simply by retransmitting colliding packets with a random delay, as shown in Fig. 7.3. ALOHA channels are characterised by the familiar throughput (S) versus channel traffic (G) curve in which S first increases, reaches a peak and then decreases with increasing G. This type of S/G curve opens the possibility of a long term operating mode with low throughput and high delay, a condition loosely

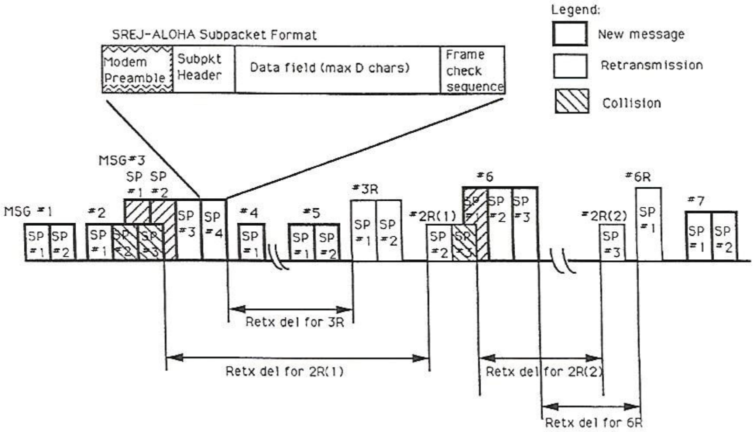

Fig. 7.4 *Illustration of channel events for the SREJ–ALOHA protocol*

referred to as instability. However, with proper design of fixed or adaptive [11, 12] retransmission procedures, ALOHA channels shared by a finite number of stations can be operated with satisfactory stability properties.

As one might expect, the simple ALOHA protocol provides relatively low efficiency of channel use, with maximum throughput typically varying between 0·13 and 0·18 (depending on the message length distribution [13]). The low throughput is offset by low access delay, the ability to handle variable length message traffic, robust operation and minimal equipment complexity, offering an overall combination of attributes that is quite attractive for many VSAT applications. Note that for very short messages (as in point-of-sale networks), even the net throughput supported may be quite competitive with other protocols because of the greater significance of transmission overheads, which are minimal in ALOHA.

7.2.2.2 *Selective reject (SREJ) ALOHA* SREJ–ALOHA [14, 15] employs subpacketisation of messages in conjunction with a selective reject ARQ retransmission strategy, to achieve significantly higher asynchronous random access throughput than pure ALOHA. As shown in Fig. 7.4, in SREJ–ALOHA, messages are formatted into a contiguous sequence of independently detectable fixed length subpackets (the last subpacket in a message may be shorter than this fixed length), each with its own header and acquisition preamble. The protocol exploits the fact that most collisions in an asynchronous channel result in partial overlaps, so that only the subpackets actually encountering conflict are retransmitted (analogous to the well-known SREJ mode in ARQ), as illustrated in Fig. 7.4. This seemingly simple modification of the CRA used in conventional ALOHA provides surprising performance advantages in terms of increased net throughput, better delay distribution properties and improved stability.

The maximum throughput (not accounting for overheads) of SREJ–ALOHA with arbitrarily small subpackets approaches the theoretical asynchronous capacity limit of 0·368 [16]. In practice, the need for acquisition preamble and

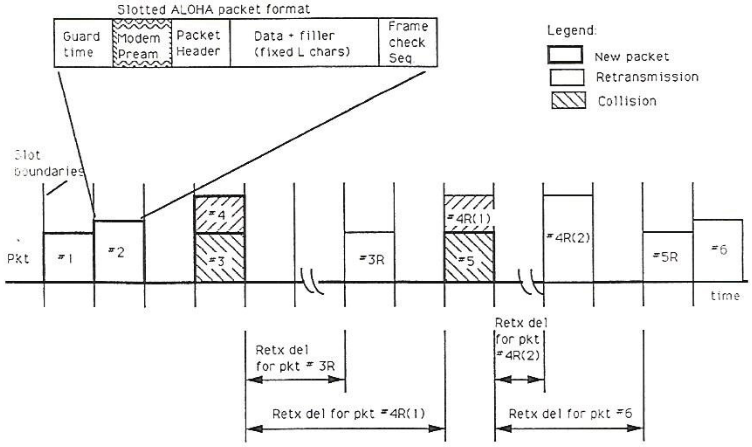

Fig. 7.5 *Illustration of channel events for the slotted ALOHA protocol*

header in each subpacket limits the maximum net throughput to the range of 0·2-0·3, typically a 2 : 1 improvement over pure ALOHA. In this context, it is important to note that burst modems with short acquisition preamble are a prerequisite for effective SREJ–ALOHA operation. This protocol retains the robustness and low implementation complexity of pure ALOHA.

7.2.2.3 *Slotted ALOHA* Slotted ALOHA [17, 18] is the most widely known and studied random access protocol, and is based on the principle of reducing the 'vulnerable period' of a packet in ALOHA by constraining station transmissions to be fixed length packets which begin and end at TDMA-like slot boundaries, as shown in Fig. 7.5. This has the well known effect of increasing maximum throughput to 0·368 with delay characteristics qualitatively similar to pure ALOHA. Thus, slotted ALOHA is a simple and effective random access technique for a fixed length traffic model, although, relative to ALOHA and SREJ–ALOHA, timing synchronisation between stations could add significantly to the implementation complexity. Note that for the important variable message length scenario, use of slotted ALOHA requires either a large packet size which accommodates most messages or the partitioning of messages into smaller subpackets, as in SREJ–ALOHA. With the former option, large overheads tend to lower the net capacity, while with the latter option, the advantage of slotted operation over the simpler unslotted mode tends to decrease [14], and may even be offset by the guard time overheads not required in asynchronous systems. Thus, while slotted ALOHA has a higher nominal capacity of 0·368, the *net* data throughput for typical variable length traffic scenarios may often be lower than that of SREJ–ALOHA.

7.2.2.4 *Tree CRA* For the slotted fixed message length scenario, a second major class of random access is based on the so called 'tree algorithm' type of CRA, first introduced by Capetanakis [19], and later refined by Massey [20] and Gallager [21]. In tree CRA based random access, messages involved in a collision par-

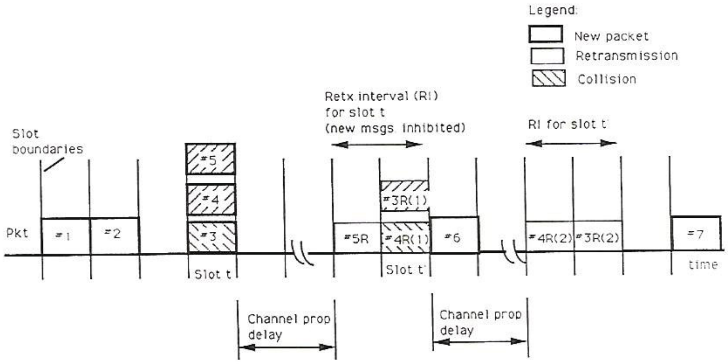

Fig. 7.6 *Illustration of channel events for slotted tree CRA based random access*

ticipate in a systematic partitioning procedure for collision resolution (usually based on a tree-structured decision process), during which new messages are not allowed to access the channel. An example of the operation of a typical tree CRA is shown in Fig. 7.6. In contrast to ALOHA protocols, the CRA procedure is guaranteed to converge, so that the channel is provably stable for both finite and infinite population models.

Tree CRA based random access systems typically have a capacity in the range of 0·43 to 0·49, depending on the specifics of the CRA used. In principle, they have capacity and stability advantages over slotted ALOHA, and appear to be attractive, at least for fixed message length traffic. However, while tree algorithm protocols are stable, they are subject to deadlocks in the presence of incorrect channel observation, due mainly to the highly structured nature of the collision resolution process [22]. In contrast, ALOHA protocols, which can also be operated with guaranteed stability for finite station populations, are free from such deadlocks because of the inherent simplicity of the CRA. A second difficulty with tree algorithms is the need for interleaving in satellite channels [19], which usually results in latency delay problems. The latency delay could potentially be eliminated, but usually at the expense of a reduction in maximum throughput [23]. Overall, the deadlock and interleaving problems as well as the higher implementation complexity tend to offset the potential capacity benefit of tree algorithms in the VSAT scenario. Finally, as for slotted ALOHA, tree CRA based protocols suffer from inherent inefficiency of employing a fixed packet size format for the variable length message traffic commonly encountered in VSAT applications.

7.2.2.5 *ARRA* Announced retransmission random access and its variants [24, 25] offer higher capacity than slotted ALOHA by adding a small amount of 'minislotted' side information known as an 'announcement' to each message transmission. ARRA makes use of the fact that retrials are actually pseudo-random (and hence predictable) so that an announcement of the future slot in which retransmission will be attempted (in the event of a collision) can be used to prevent new transmissions from interfering with retransmissions. In addition,

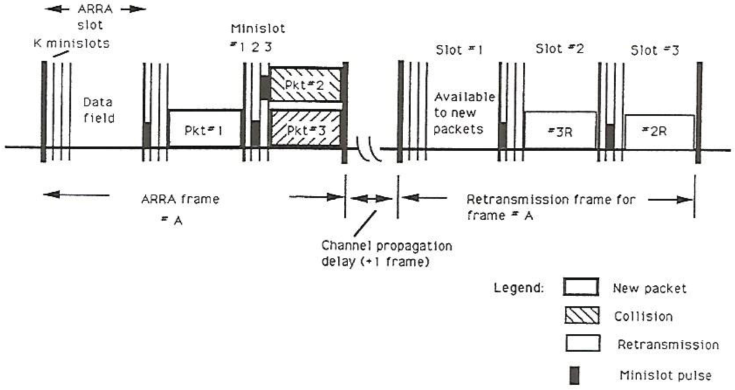

Fig. 7.7 *Illustration of basic ARRA operation* $(K = 3)$

by predicting future collisions between announced retransmissions, all unsuccessful retransmissions can be aborted and the unused slots made available to new messages. An illustration of channel events for the basic ARRA scheme is shown in Fig. 7.7. The capacity achieved is typically in the region of 0·53 to 0·6, but the usefulness for VSATs is limited by the implementation difficulties associated with minislotting, along with the characteristic problems of fixed packet size slotted formats.

7.2.2.6 *Time-of-arrival based random access (TARA)* Unslotted (locally synchronous) time-of-arrival random access protocols [26] are based on CRAs which exploit the packet arrival sequence feedback inherently embedded in unslotted ALOHA-type channels to expedite the collision resolution process. In particular, for fixed length packets, collision burst analysis with continuous signal detection (SD) only can be used to determine the 'first' and 'last' packet in a collision, thus permitting conflict-free scheduled transmission of two packets per collision. If stations also have collision detection (CD) capability, at least three packets per collision burst ('first', 'second' and 'last') can be resolved. Packets thus resolved by the CRA are retransmitted in a locally synchronous scheduled mode, made possible by using channel events as timing markers. To minimise unnecessary conflict, stations avoid transmitting new messages during the mutually known periods of scheduled retransmission. Fig. 7.8 illustrates the collision burst analysis and locally synchronous retransmission process for the case with both signal and collision detection (SD/CD).

The CRAs described result in random access systems with delay, throughput and stability properties superior to ALOHA-type protocols, both slotted and unslotted. Typical capacities are 0·41 with SD only and 0·47–0·51 with SD/CD, and are thus competitive with relatively complex slotted tree CRAs. Of course, as for slotted ALOHA and tree CRAs, the fixed packet size format implies lower net throughput for variable length traffic. TARA protocols may be attractive for second generation semi-compatible upgrades of VSAT channel protocol units originally intended for unslotted ALOHA. Since time-of-arrival CRAs use

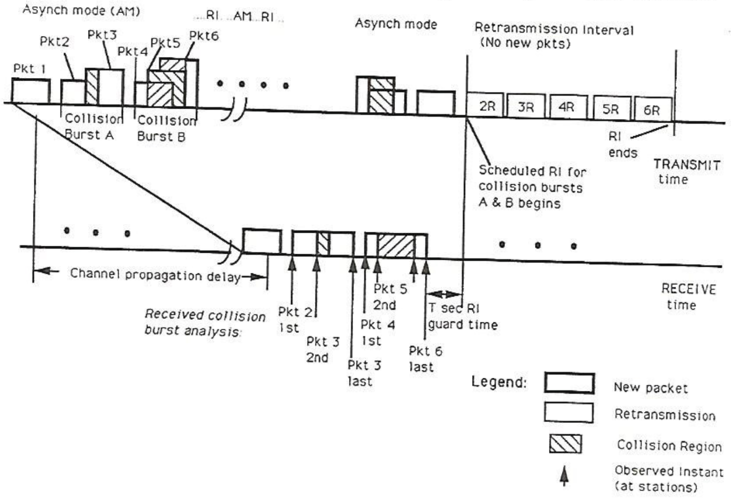

Fig. 7.8 *Illustration of channel events in time-of-arrival based random access (with SD/CD and RI initiation parameter $K = 2$)*

channel event based local synchronisation, it is expected that some of the operational and robustness advantages of unslotted random access would be retained.

7.2.2.7 *SREJ–ALOHA/FCFS* For the variable length message environment that is encountered in most VSAT applications, time-of-arrival based collision resolution can be achieved within the framework of SREJ–ALOHA described earlier, without requiring additional hardware for signal or collision detection [27]. Specifically, with SREJ–ALOHA transmission, collision burst analysis at the subpacket level can generally provide sufficient resolution to develop an implicit FCFS retransmission ordering among many of the stations involved in channel collisions, based on the order in which successful subpackets from partially successful messages are received. The operation of SREJ–ALOHA with locally synchronous first-come-first-served (FCFS) retransmission scheduling is illustrated in Fig. 7.9. Note that use of this CRA results in conflict-free transmission at the second attempt of all messages with at least one successful subpacket. This implies that congestion due to colliding long messages (which is a problem in asynchronous ALOHA channels) is greatly reduced, since the probability of at least one subpacket getting through is quite high.

The subpacket feedback based CRA described above provides significant improvements in throughput, stability and delay relative to asynchronous SREJ–ALOHA. The achieved net capacity ranges from 0·4 to 0·5 depending on the message length distribution of the traffic carried. Thus, for the important variable message length traffic scenario, SREJ–ALOHA/FCFS generally outperforms the best slotted random access protocols (such as the tree CRA), approaching throughput levels typical of demand assigned protocols that are described later. SREJ–ALOHA/FCFS is thus an attractive semi-compatible upgrade option for ALOHA or SREJ–ALOHA equipment, providing improved throughput, stabil-

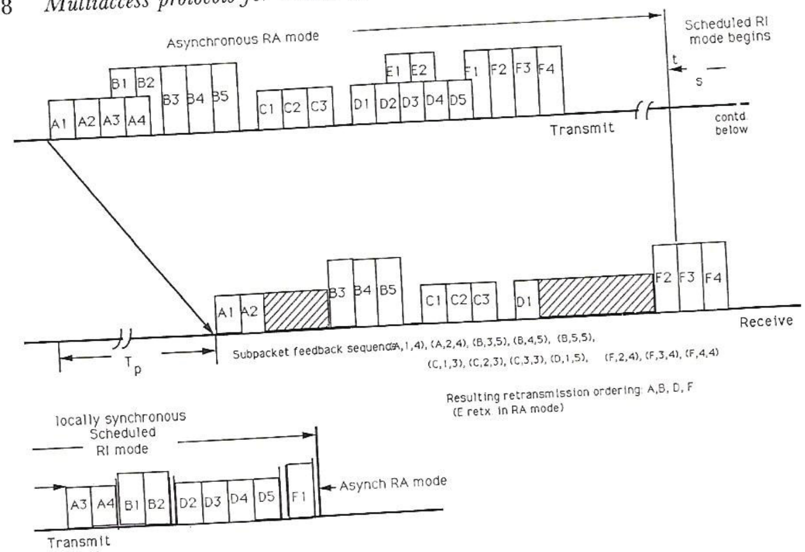

Fig. 7.9 *Illustration of channel events in SREJ–ALOHA/FCFS channel*

ity and delay (both average and distribution) properties over a wider range of traffic profiles.

7.2.2.8 *RA–CDMA* In addition to providing the ability to operate in power spectral density and interference limited environments, spread spectrum transmission can be used to modify (and possibly improve) random access performance. Specifically, spread spectrum coding permits 'colliding' ALOHA packets to encounter multiple interferences without being destroyed, potentially resulting in better random access throughput-delay and stability characteristics.

As illustrated in Fig. 7.10 for an asynchronous, unslotted, RA–CDMA channel, changes in the number of interfering transmissions results in a varying bit error probability over the duration of a transmitted packet. Packets that are received

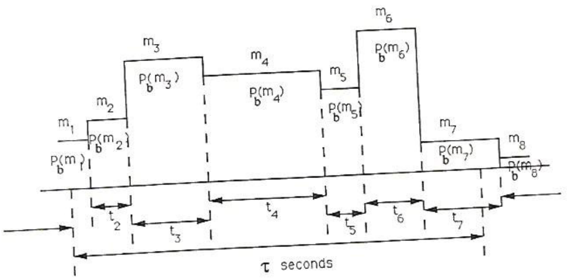

Fig. 7.10 *Example of interference pattern experienced by a τ second message packet on an asynchronous RA–CDMA channel*

with errors are retransmitted with random delays, similar to the strategy illustrated in Fig. 7.3 for pure ALOHA. The ability to support multiple simultaneous transmissions results in lower access delay than non-spread ALOHA, although the normalised capacity (i.e. maximum throughput divided by the bandwidth spreading factor) of uncoded random access CDMA is typically in the same region as pure ALOHA (i.e. ~0·1). However, with powerful FEC, it is conceivable that the capacity can be improved to the region of 0·2–0·3, as suggested by preliminary results [28, 29]. Of course, the associated spread spectrum/FEC hardware adds to VSAT complexity, relative to a system based on non-spread ALOHA.

As in non-spread systems, slotting can be used to reduce the collision probability of packets in random access CDMA systems. Slotted RA-CDMA has been considered in References 30 and 31. The qualitative performance is similar to unslotted CDMA discussed earlier, with achievable (bandwidth normalised) capacity being somewhat higher. It is interesting to note that, as the SSMA spreading factor increases, the advantage due to slotting becomes substantially less than the 2 : 1 advantage of slotted ALOHA over pure ALOHA. Particularly for variable length message traffic, there does not appear to be a compelling reason to add the complexity of slotting to a random access CDMA system.

7.2.3 *Reservation/controlled access*

Demand assigned multiple access (DAMA) is a widely advocated solution to bursty satellite data communications. In DAMA systems, there are two layers of channel access: the first level of access is for short reservation packets containing information regarding the station's demand, while the second level of access is for the actual data messages (which are typically longer than the request packets). Access at the reservation level may be provided using any of the fixed assigned or contention access methods described in Sections 7.2.1 and 7.2.2. Once a reservation is successful, data messages can be scheduled in a conflict-free manner by processing the reservation information in a distributed or centralized global queue. Thus, if the data messages are long relative to request packets, it is possible to achieve reasonably high overall net channel throughput. Appropriately designed DAMA systems are well suited to variable length message traffic, and can handle mixed interactive/file-transfer traffic. However, note that, as illustrated in Fig. 7.2, the higher throughput is accompanied by a relatively large (> 0.5 s) minimum latency delay associated with the reservation mechanism, although this is partially offset by low access delay variance.

The early applications of satellite DAMA were in the context of circuit switching, e.g. the SPADE system [32], in which FDMA/SCPC channels are assigned based on reservations made on a single TDMA channel. For the VSAT scenario, a typical DAMA system is based upon a TDMA-like channel format, with periodic frames containing a number of small slots for reservation and large slots for data messages. The relative proportion of data and reservation slots used depends upon a number of factors including the anticipated traffic profile and the specific fixed assigned/contention mechanism used for reservation access. A few popular DAMA approaches are discussed below.

7.2.3.1 *DAMA with TDMA reservations* In this technique, channel frames are formatted into request and message transmission intervals, as shown in Fig. 7.11,

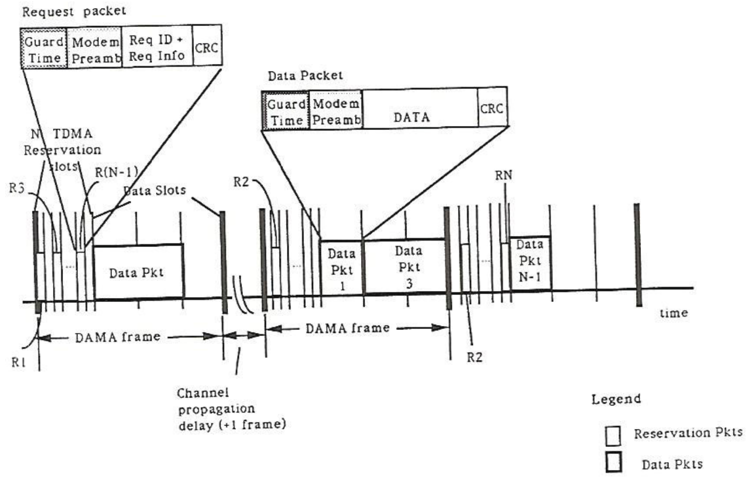

Fig. 7.11 *Illustration of channel events for DAMA with TDMA access*

with a short request slot allocated to each VSAT in every frame. Once in each frame, a VSAT may place a request for one or more data slots. In star VSAT networks, the request packet will be received by the central hub station, and an allocation (based on the status of a global queue) will be sent back to the station over the TDM broadcast channel. Note that, owing to satellite propagation delays and the TDMA frame structure, such protocols are generally characterised by a relatively high latency delay ($\sim 0{\cdot}6$–$0{\cdot}7$ s), and are limited in terms of the maximum number of stations that can be supported. Of course, since the actual transmission of data messages is conflict-free, the data portion of the frame can be utilised with high efficiency.

DAMA with TDMA access is thus appropriate for VSAT applications with a relatively high traffic volume per site, particularly when messages are long and variable in length. In spite of the lack of contention inefficiency, the need for sufficient guard time and modem acquisition preamble for reservation packets makes it difficult to achieve a maximum throughput much higher than 0·5–0·6. An example of such a DAMA system is the priority-oriented demand assignment (PODA) system [33], running in the FPODA or fixed assignment mode. This protocol has the additional feature that the relative capacity assignments for the reservation and data packets are changed in response to long term variations in input load. This type of adaptation in operating parameters is generally necessary for practical operation of a DAMA system of this class. As might be expected, the synchronisation, global queue maintenance and parameter adaptation requirements lead to significantly higher complexity and poorer robustness than for the random access protocols discussed earlier.

7.2.3.2 *DAMA with slotted ALOHA reservations* DAMA can be used to support large station populations when contention access is used instead of fixed assigned TDMA for the reservation messages. A typical implementation of DAMA for this