NASA Contractor Report 202310

010276

Space-Based Information Infrastructure Architecture for Broadband Services

Kent M. Price, Tom Inukai, Rajendev Razdan, and Yvonne M. Lazeav
Space Systems/LORAL
Palo Alto, California

December 1996

Prepared for
Lewis Research Center
Under Contract NAS3-27559

National Aeronautics and
Space Administration

NASA Contract No. NAS3-27559
NASA Contractor Report 202310
Loral Technical Report No. SS/L TR01223

Technical Support for Defining
Advanced Satellite Systems Concepts

Final Report for Task Order 1

Space-Based Information Infrastructure Architecture for Broadband Services

Prepared by
Space Systems/Loral
3825 Fabian Way
Palo Alto, CA 94303-4606

Program Manager: Kent M. Price, Space Systems/Loral
Major Contributors: Tom Inukai, COMSAT Labs
Rajender Razdan, COMSAT Labs
Kent M. Price, Space Systems/Loral
Yvonne M. Lazear, Space Systems/Loral

Prepared for
NASA, Lewis Research Center
Cleveland, OH 44135

December 1996

CONTENTS

ILLUSTRATIONS

TABLES

SUMMARY

Satellite communications play an important role in providing universal service and accessibility at national and global levels. Satellite designs and architecture are highly need-driven and planned far in advance to maximize performance and fully utilize limited resources. The success of second and third generation satellite network systems will depend on their ability to provide interoperable emerging NII/GII applications. The spectrum and technology requirements for these systems need to be identified as soon as possible. This study addressed four tasks:

1. Identify satellite-addressable information infrastructure markets
2. Perform network analysis for space-based information infrastructure
3. Develop conceptual architectures
4. Economic assessment of architectures

Important conclusions drawn from this study are as follows:

a. Satellites will have a major role in the national and global information infrastructure. NII/GII applications are expected to create a tremendous demand for bandwidth-intensive communication links.

b. Seamless integration of terrestrial and satellite networks requires that quality of service not be degraded by a satellite link, and that cost be competitive.

c. The main distinguishing characteristics of satellite networks with respect to terrestrial networks are their delay, delay variations, poorer link quality and the need for beam/satellite handover. These characteristics vary widely for the proposed LEO, MEO and GEO satellite systems.

d. The barriers against seamless interoperability of satellite and terrestrial networks can be categorized as follows:

    1. An administrative hurdle is the lack of implementation by the vendor and terrestrial carrier community of the parameters and protocols necessary for efficient operation of satellite links.

    2. Important broadband standards, such as ATM and SONET, as well as data protocols (X.25/X.75) and transport protocols (such as TCP/IP) have, in the past, been designed assuming the use of terrestrial links. Similarly, congestion control schemes, too, have been designed with implicit assumption of terrestrial networks. These protocols behave poorly when the network involves large

propagation delays. Many of these standards can be greatly improved to include application to satellite communications by a few slight changes to the protocol.

3. Bandwidth efficiency, which is so crucial in satellite networks, has been given very low priority in standardizing the next-generation digital transport architecture (e.g., ATM/SDH). From very early on, a fundamental decision was made to base ATM and SDH on the characteristics of fiber optic media. Many characteristics of these networks result in very high percentage of overhead. This is in direct contrast to currently designed satellite networks, which attempt to maximize the use of the limited satellite bandwidth by applying advanced signal processing techniques to reduce the bandwidth requirement of each service.

e. While both technical and administrative barriers do exist towards a seamless integration of satellite and terrestrial networks, there are ways to overcome these barriers. For one, not all satellite networks have the same degree of interoperability problems with terrestrial networks. For example, LEO systems might often experience less delay than intercontinental fiber links, and their Quality of Service (QoS) may not differ too much from such terrestrial systems. Another welcome trend in reducing the interoperability barrier, is the fact that some of the problems experienced by satellite networks are also being experienced in other areas, such as inter-connection of high-speed LANs.

f. New standards are being formulated which attempt to address the aforementioned concerns. More participation by satellite service providers, equipment vendors, satellite manufacturers, and satellite service users is needed in the communications standards organizations.

g. There is growing awareness within the communications industry regarding the need and importance of satellite networks. The exponential rise of wireless network connections and the realization of the enormous costs involved in introducing fiber to the home, or even fiber to the curb, is renewing interest in global satellite networks.

h. GEO and LEO/MEO space-based architecture concepts for broadband services were designed to support future NII/GII applications and operate seamlessly with terrestrial networks as a part of the third-generation communication systems. The GEO architectures are preferred due to greater available bandwidth and less potential for interference with other systems.

i. Economic assessment of the concepts indicates that the wholesale circuit price for a switched T-1 (1.5 Mb/s) simplex circuit can be around ten cents per minute.

The report is organized in four sections and two appendices:

Section 1	Market assessment
Section 2	Network analysis
Section 3	Architecture concepts
Section 4	Economic assessment
Appendix A	Enabling technologies
Appendix B	NASA statement of work
Appendix C	Derivation of satellite earth coverage diameter

This page intentionally left blank.

SECTION 1 – SATELLITE ADDRESSABLE MARKETS

There has been considerable discussion in recent years about the need for a National as well as Global Information Infrastructure (NII/GII) which would act as a catalyst in providing a seamless integration of the telecommunications, computer and entertainment industries and would help facilitate universal access to vast amounts of information — be it voice, data or video — as well as present an unprecedented degree of connectivity to businesses and individuals. Some have extended this a step further to stress the need for even an Inter-planetary Information Infrastructure (III).[1] What has spurred the need for these information infrastructures has been the exponential growth in the demand for voice, data and video services; the proliferation of end-user devices (PCs, faxes, telephones, video display units, etc.); the explosive growth in Internet usage; and the increase in the mobility of people who desire to be connected.

In the NII, GII and III arenas, satellite systems will undoubtedly play a crucial role, and in some cases may even play a pivotal role. While it is generally useful to assess future markets while designing NII/GII terrestrial telecommunication services; it becomes crucial when it comes to designing satellite systems. Unlike ground-based systems where one may use optical fiber links having bandwidth capabilities far exceeding present and near-future needs, with satellites one has limited resources (in terms of transponder capacity). In addition, deployment of satellites is costly and design lead times are long compared to that of terrestrial systems. Finally, unlike terrestrial networks which can be easily upgraded without dismantling the existing system, with satellites there is very little leeway in terms of adding resources. Often the only way to upgrade is by launching new satellites at greater costs. For these reasons, it becomes imperative that satellites be designed only after assessing the market for their services well in advance.

It is the aim of this report to address the NII/GII markets and assess where it makes economic sense for satellites to be used and where it does not. In this report, current, planned and emerging NII/GII services and applications will be examined from the point of view of technical feasibility, Quality of Service (QoS) requirements, and the planned time-frame for deployment. The market trends in these services/applications will also be examined. Any viable satellite system design will have to fulfill not only technical feasibility criteria, but also economic feasibility (i.e., market) criteria as well as financial availability criteria (see Figure 1-1)[2].

---

[1] "An initial design asessment for a communications relay satellite to support the Interplanetary Information Infrastructure", Timothy G. Howard, AIAA-96-1057-CP, pg. 566.

[2] "Technologies to Enable Low Cost Satellite Communications", Kent M. Price & Yvonne Lazear, AIAA-96-1109-CP, pg. 1065.

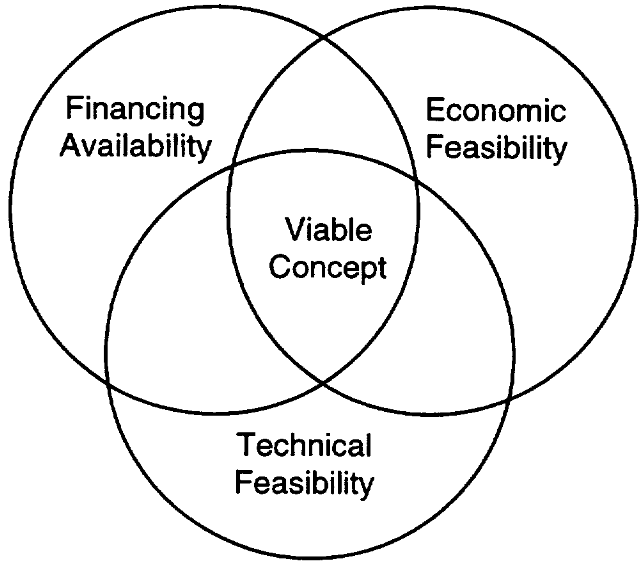


Figure 1-1. Requirements for a Viable Concept

We will touch upon these three aspects in this report, though the greater emphasis falls on technical and economic feasibilities.

1.1 NII/GII SERVICES/APPLICATIONS MODEL

In examining the NII/GII applications market, it is helpful to make use of the NII architectural framework model as defined by National Institute of Standards and Technology (NIST)[3]. Figure 1-2 depicts the three-layer NII/GII architectural model, with examples of each layer. Essentially, this model views NII as having three layers of functions: applications, services and bitways. Applications are defined as information technologies that can be used to accomplish tasks accessing a range of services. Services provide the building blocks for these applications as well as the interfaces for displays, sensors and other input/output devices. Finally, bitways are the "data pipes"; i.e., the cable, fiber optics, satellites and other means of transmission; plus the controlling software to transmit data from one place to another.

[3] "Framework for National Information Infrastructure Services", NISTIR-5478, July 1994.

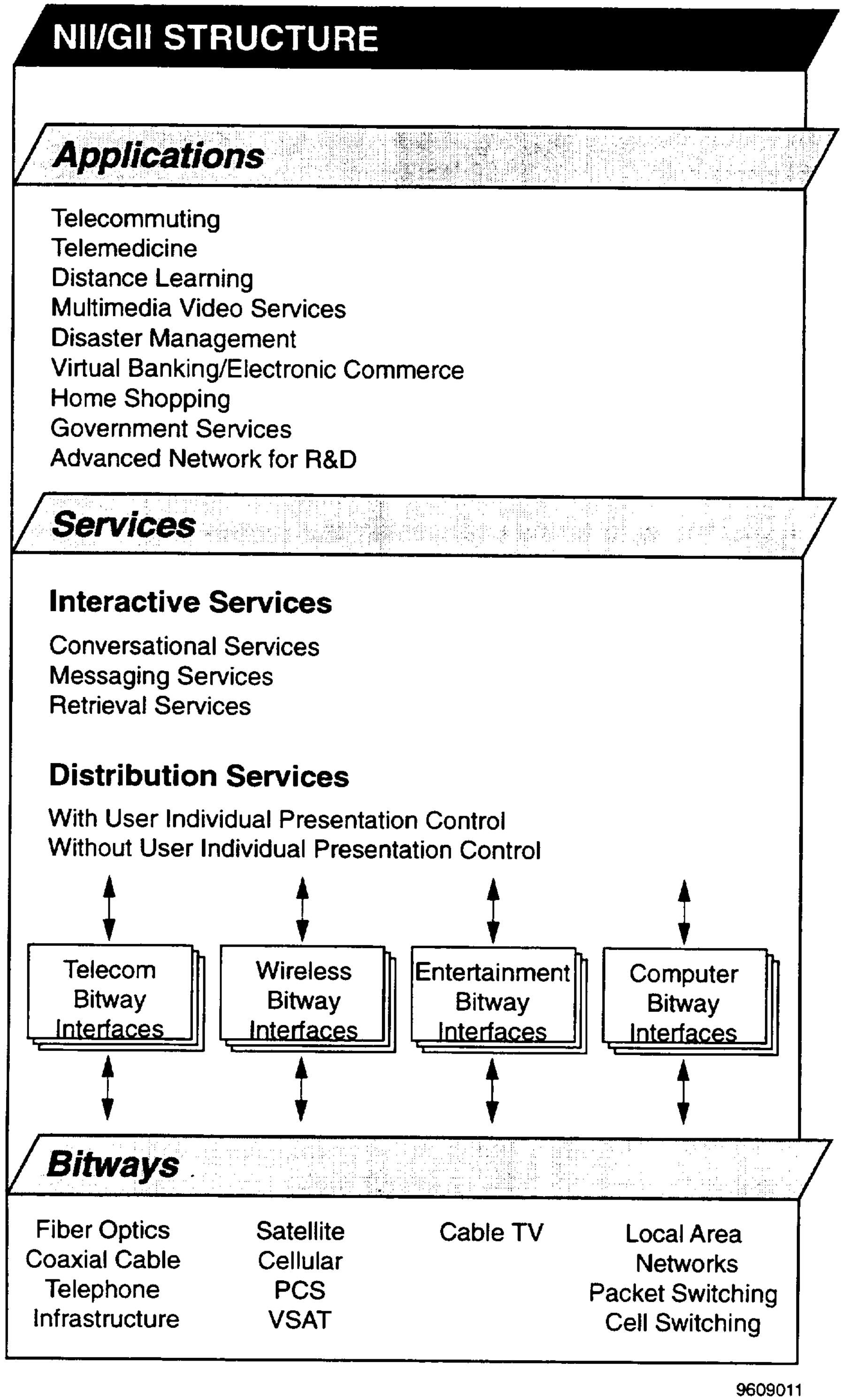


Figure 1-2. Functional Diagram of NII/GII Applications

In this hierarchy, application functions depend on various service functions which, in turn, can be carried over various bitways. Thus, a given application, such as Telemedicine, may require many different services, such as composite imaging (for transfer of high definition X-ray charts, for example), client-server database file access graphics (for transferring patient medical records), virtual reality visualization (for surgery practice) and video teleconferencing (for doctors to discuss a patient's case with each other). Some of these services may also be used in other applications, such as home-banking or distance learning. However, all the applications and services depend on the bitways as a physical medium for carrying information.

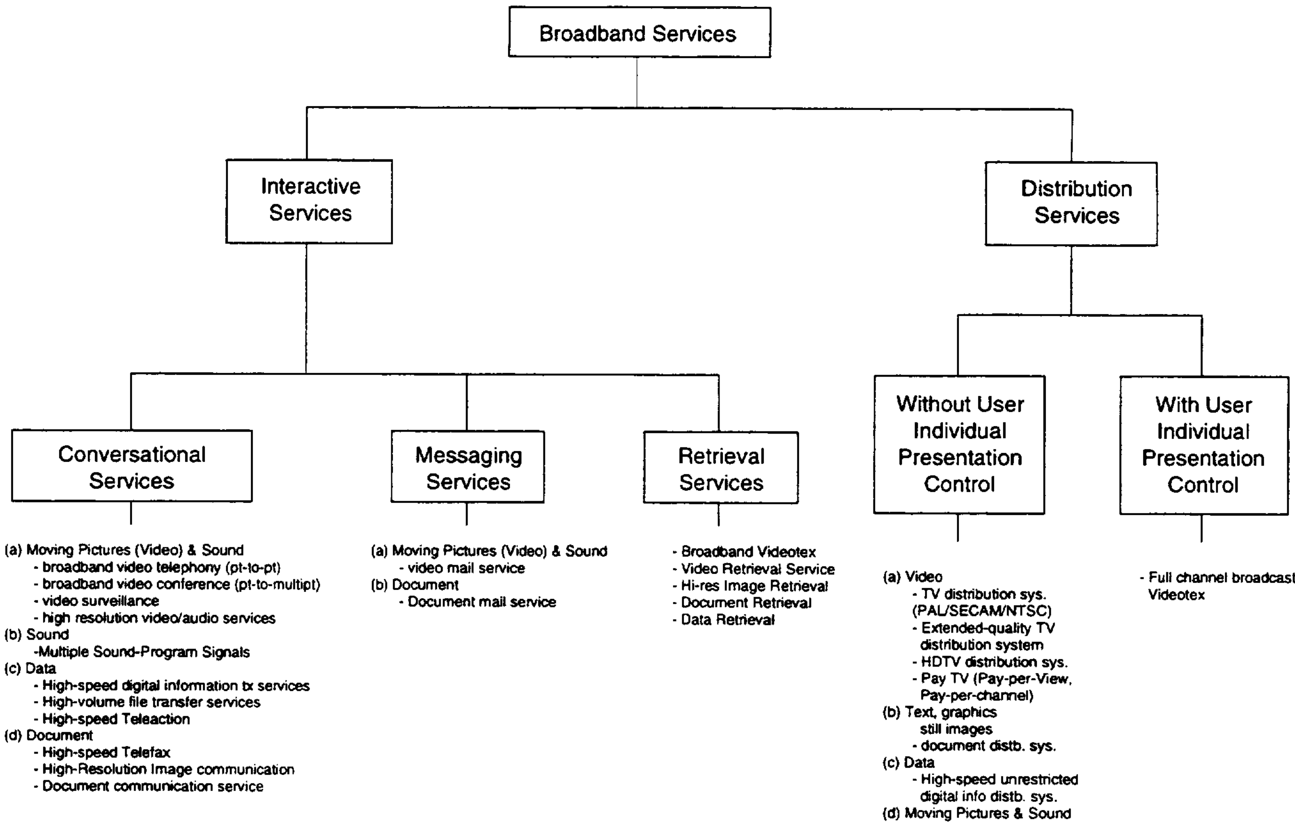
The NII/GII architectural model provides us with a smooth framework within which we can make assessments of the future markets for satellite systems. We will first examine the various NII/GII applications in terms of their service requirements and needs as well as determine a time frame for their deployment. We will also examine the underlying factors facilitating the growth in the telecommunications market. A survey of growth in the various bitways will also be analyzed so that comparison can be made between the market available for satellite systems as compared to the market available for other media. Finally, we will look at the overall telecommunications market and projected growth in the USA as well as other regions of the world.

1.2 BROADBAND SERVICES FOR NII/GII APPLICATIONS

Broadband communications demand higher bandwidths and information bit rates than does a regular telephone. The various services that make up broadband communications can be classified according to function, feature and characteristic (i.e., peak bit rate, QoS parameters, etc.).

1.2.1 Classification of Broadband Services

As per CCITT Recommendation 1.211, broadband services can be classified into either interactive-type or distribution-type services. Interactive services are subdivided into conversational, messaging and retrieval services. Distribution services are sub-categorized into those with and without User Individual Presentation. These various services, as well as their sub-categories are shown in Figure 1-3.



9608351

Figure 1-3. Classification of Various Broadband Services Used in NII/GII Applications

Interactive Services

Interactive services involve a two-way exchange of information (this does not include signaling information).

Conversational services, a sub-set of interactive services, represents a real-time end-to-end transfer of information. Examples of conversational services include: broadband video telephony and conference (point-to-point and point-to-multipoint), video surveillance, high quality video transmission (e.g., TV signal transfer, video/audio dialog), multiple sound-program signals (e.g., multilingual commentary channels), high-speed unrestricted digital information services (e.g. LAN/ MAN interconnections, still-image transfer, multi-site CAD/CAM, distributed processing), high volume file transfer, high speed teleaction (e.g. real-time control, telemetry, alarms), high-speed telefax and high-resolution image communication (e.g., medical images, remote games) and document communication services.

Messaging services, another interactive service, involves sending information to a database for subsequent retrieval. Examples include electronic-mail, voice mail, and image mail.

Retrieval services, is the complement of messaging services, and involves extraction of information from a database. Examples of retrieval services include: broadband videotex, video retrieval, high-resolution image retrieval, document retrieval and data retrieval. Bit-Error Rate (BER) requirements are generally lower for retrieval services compared to messaging services the user can recover from some errors by retransmission).

Distribution Services

Distribution services are one-way information transfer from service provider to user. Distribution services without user individual presentation control include broadcast services (such as TV and radio program distribution) and provide a continuous flow of information distributed from a central source to an unlimited number of authorized receivers connected to the network. Distribution services with user individual presentation control include full channel broadcast videotex which are repeated cyclically and which the user can control the start and order of the presentation.

The characteristics of these services (in terms of data rate, information density, etc.) will be described in subsection 2.1.

1.3 NII/GII APPLICATIONS

Some key NII/GII applications that have been identified include:

a. Telecommuting
b. Telemedicine
c. Distance learning
d. Multimedia video services
e. Disaster management
f. Home banking/electronic commerce
g. Government services
h. Advanced networks for research and development

Each of these application areas are examined in detail in the following subsections.

1.3.1 Telecommuting

Telecommuting refers to the partial or total substitution of telecommunications technology for the trip to and from the primary workplace. Computers, cellular phones, fax and advanced communication links have removed the physical barriers that once required workers to work from their offices.

There are essentially four types of telecommuters:

1. Those that work at home (this is the most popular form)
2. Those using satellite offices (i.e., remote office locations normally placed within a large concentration of employee residences)
3. Those using neighborhood work centers (i.e., a common work space for employees from different companies)
4. Those that are mobile (i.e., workers who are constantly on the road using technology as the main link to centrally located resources).

The number of telecommuters of all types in the U.S. is expected to grow from 5.5 million in 1991[4] to 13 million by 1998[5] . The Department of Transportation predicts a full 15 percent of the U.S. work force will be telecommuting by 2002. These trends are depicted in Figure 1-4 which indicates that the number of corporate employees participating in telecommuting programs will grow by almost 50 percent over the next five years[6]. The federal government hopes to have 60,000 telecommuting employees (3% of the federal work force) by end of fiscal year 1997, saving up to $150 million annually, according to the Emerging Technologies Research Group.

The forces driving the telecommuting trend are: the federal incentive to reduce air pollution, corporate efforts to cut costs, and the need to provide flexible working conditions for the highly-skilled work force. Management consultant, Jack Nilles[7] estimates that, on average, a management-level employee who telecommutes twice a week saves the company anywhere between $6,000 to $12,000 a year. The most common reasons for implementing a telecommuting program are:

a. Increased productivity
b. Reduced absenteeism
c. Improved employee retention
d. Savings on office space costs
e. Reduced employee stress

---

[4] Estimate based on study by Link Resources a NY based company - Forbes, 9 Oct, 1995, pg. 133.

[5] Business Week, 26 June, 1995 — source used in article: Link Resources.

[6] Interoperability, Nov. 1995, pg. 54, source: Link resources & Find/SVP.

[7] http://www.telecommute.org/jack.html; private communication.

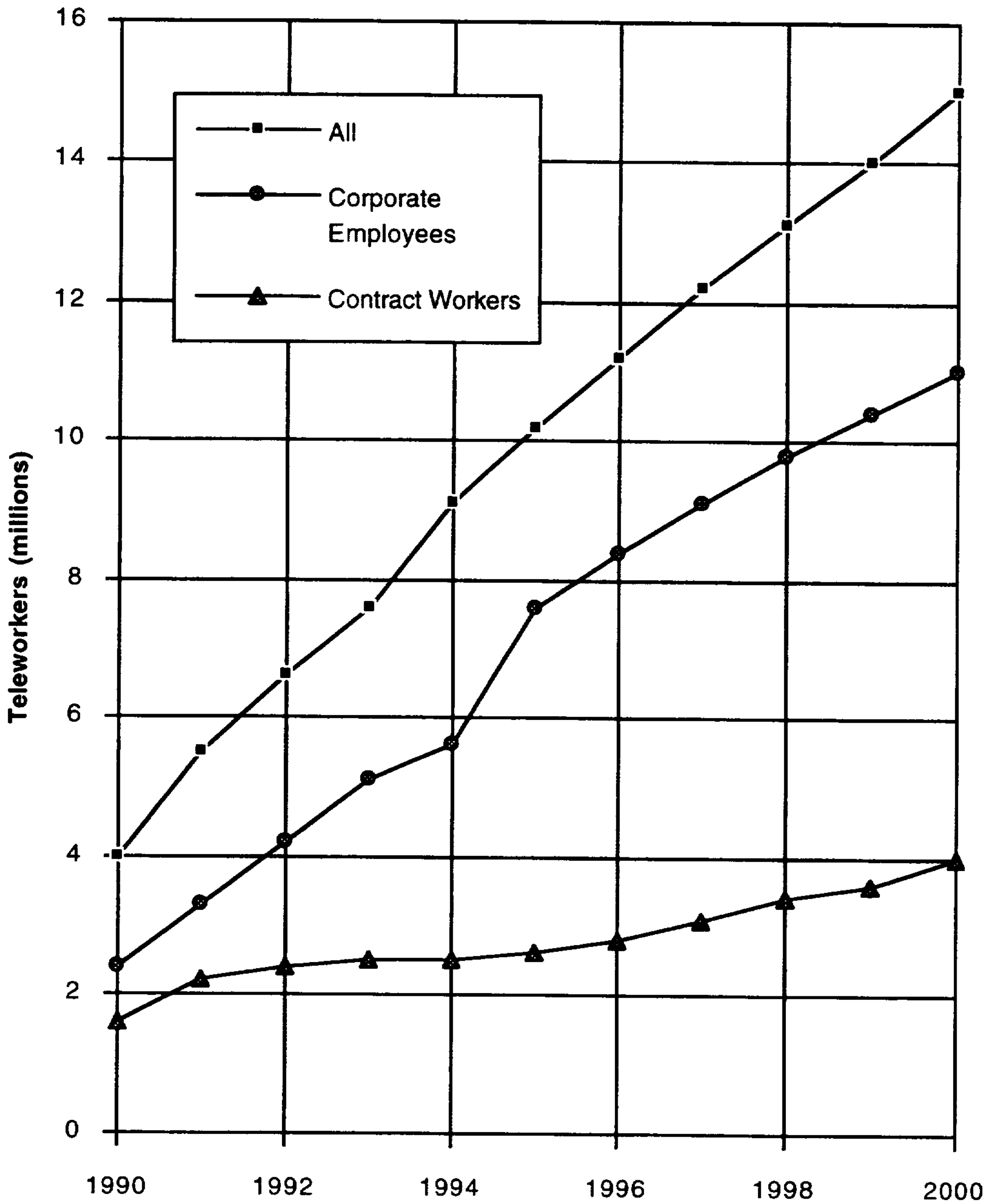


Figure 1-4. Telecommuting in the United States

f. Improved employee morale

g. Helps in recruiting outside geographical confines of office

h. Reduces pollution

For telecommuting to be effective, networks must connect the telecommuter's workstation to those of other telecommuters as well as to the corporate office. This need will accelerate demand for portable computers, groupware, digital phone lines and cellular systems. The services required by telecommuters will vary greatly depending upon the nature of the worker's job. For example, a telecommuting real-estate agent may need to access a high quality video database; whereas a telecommuting sales representative may require only a low-speed e-mail and voice connection to the corporate office. In general, telecommuting services will require the use of a low speed path from home or satellite office (64 kb/s or lower) to the corporate office and a bursty but fairly large channel from the corporate office (and other places) into the home (384 kb/s up to 5 Mb/s for some services).

1.3.2 Telemedicine

Telemedicine refers to the provision of healthcare services via electronic transmission of medical information among different sites. The healthcare industry, by any indicator, is one of the largest industries in the United States. Since 1980, our nation's healthcare costs have quadrupled. In 1994, the healthcare industry has accounted for nearly 14 percent of the nation's GNP over $1 trillion. This figure is expected to grow to $1.6 trillion by year 2000[8]. With a spending rate of more than $2 billion every day on health care, the United States grapples with a crisis[9]. Twenty-five cents of every dollar on a hospital bill goes towards administrative costs and does not buy any patient care. At the same time, over 36 million Americans are faced with restricted access to health care because they lack adequate medical insurance coverage.

These problems will not be solved without a comprehensive healthcare reform. However, better use of information technology and development of healthcare applications for the NII can make an important contribution to reform. The goal of telemedicine is to simultaneously improve the quality and reduce the costs of healthcare by facilitating communication of medical information amongst healthcare professionals, reduce costs, expand access, and enhance quality of services. The market for healthcare computing and information technology exceeded $25 billion in 1994 and is predicted to rise above $50 billion annually by year 2000. It has been estimated that 25 percent of the data traffic on the NII will be related to healthcare[8]. Experts estimate that telecommunications applications

---

[8] See MMS Business Overview report, pg. 14.

[9] See Arthur Little, "Health Care Cost Reductions: The InterLATA Component".

could reduce health care costs by $36 to $100 billion each year, while improving quality and increasing access.

Telemedicine represents a perfect market for satellite systems, since it allows physicians access to patients (and vice versa) in rural areas where poor terrestrial telecommunications infrastructure and even poorer healthcare facilities prevailed. With telemedicine, physicians can make use of the types of advanced imaging, diagnostic equipment and treatment procedures that were only available to well equipped and costly medical centers in urban populated areas. Thus, by using telemedicine, doctors and other healthcare physicians can consult with specialists thousands of miles away can continually upgrade their education and skills and can share medical records and x-rays. The patient, too, can make use of his/her PC and use the NII to promote self care and prevention by having health care information available 24 hours a day.

Present applications of telemedicine include:

a. Storage and dissemination of patients' records for diagnostic purposes

b. Image compression for storage and retrieval of image data

c. Image processing for diagnostic purpose

d. Digital transmission of large 2-D and 3-D medical images

e. Computerized control of medical equipment via the NII

f. High-speed networking for video-conferencing, large file transfers, and real-time computer visualization of data.

Other applications to telemedicine will develop as this field matures. Some evolving and future applications include:

a. Real-time transmission of video images for physician-physician and physician-patient consultations

b. Networked super-computing resources for medical and educational applications

c. Direct transmission of medical data to hospitals from medical devices attached to patients at home

d. "Data mining" of large databases of patient records for use in medical education, diagnostics, and cost/benefit analysis

e. Wireless multimedia communications from portable equipment used by emergency personnel

f. Dynamic control of medical hardware by use of virtual reality tools and the NII. Virtual environments and related technologies add value to health care in the areas of cost savings, improved services, and savings in material resources. Table 1-1 summarizes the value added in these three areas.

Table 1-1. Value Added to Telemedicine by Virtual Environment Systems[10]

Value Added	Examples
Cost Savings	Trauma units in emergency rooms could improve operating efficiency and reduce costs by using telepresence. Doing so would conserve resources by limiting the need for part-time specialists to be physically present in the trauma units.
Improved Services	Simulations allow surgeons to develop new techniques, to practice unfamiliar techniques, and to predict results of particular surgical procedures. The success of joint replacement depends on the proper placement and fit of implants within bony structures. Surgical robots and preoperative planners using computer simulations can improve surgical techniques and accuracy. Advantages offered by telepresence systems include enhancing risk performance in remote manipulation; allowing controlled application of extremely large or small forces; improving operator perception of the task; and facilitating manipulation in hazardous environments.
Savings in material resources	The use of simulators saves precious resources such as cadavers and animals.

These various services differ widely in terms of their telecommunication requirements. For example, blood pressure monitoring requires data rates of 0.3 b/s whereas high-resolution image transfers (such as x-rays) require data rates in excess of 300 Mb/s. Table 1-2 lists some of these applications, and the necessary requirements in terms of data bandwidth and QoS.

These services will be offered by both wired and wireless communication systems, and will involve Local, Medium and Wide Area Networks (LAN/MAN/WAN). The communications satellites that provide mobile telemedicine services will require high channel capacity and high transmission reliability in real-time operation.

1.3.3 Distance Learning

Distance learning, which has started to become a core educational strategy in the 1990s, refers to the visual extension of the classroom to include video-conferencing, where multiple classrooms conduct interactive sessions; *broadcasting*, where one site communicates to multiple classes; and *personal conferencing*, where individuals can communicate with one other visually on their computers.

---

[10] "Virtual Environments in Health Care. A White Paper for the Advanced Technology Program", National Institute of Standards and Technology.

Table 1-2. Data Requirements for Telemedicine Equipment Operation[11]

Application/ Service	Requirements	Impact
Medical Examinations	Two-way video and electronic stethoscope, otoscope, opthalmoscope	TBD
Radiology	X-rays, CT scans, MRI, mammogram	File sizes range from 1 to 20 Mbytes. T1-T3 rates needed
Pathology	High resolution, color pathological imagery	Large files (10s of Mbytes) and High data rates
Medical Evaluation	One-way or two-way interactive audio and video	Point-multipoint high quality videoconferencing
Virtual Reality	Detailed, digitized, 3-D active models of human body for surgical and other training	Large data sets need to be manipulated in Real-Time, Possibility of OC-3 Rates
Smart Card Medical Data	Basic data transmission to multimedia data transmission	1.6 kb/s to 1.5 Mb/s and higher
Computer Aided Diagnostics	TBD	TBD

Distance learning is being applied in diverse applications such as:

a. Statewide cooperative educational programs involving many institutions offering curricula for grade levels from kindergarten to graduate school

b. Multi-campus administration of distance learning designed to share resources among participating institutions

c. Faculty workshops providing high-quality, cost-effective professional training that promotes improved communications amongst geographically dispersed instructors

d. College-level degree programs for non-traditional students who are employed full-time in the business community

e. Specialized job-skills training taught by highly qualified faculty members in disciplines such as engineering; medicine and business

f. Innovative curricula like global field trips where students can be taken almost anywhere through the magic of video.

---

[11] "Communications Satellites in the National and Global Health Care Information Infrastructure: Their Role, Impact, and Issues", John E. Zuzek and Kul B. Bhasin, AIAA-96-0993-CP.

Distance learning provides educational institutions with at least three important benefits:

1. Distance learning can extend the reach of an educational program far beyond the geographical constraints of a classroom- ensuring that students on remote campuses or attending rural schools have access to the same classes and teachers as those on the main campus or attending better funded schools. Additionally, distance education can bring more consistency to the curriculum, by making certain that all students — regardless of campus — take core courses from the University's two or three best instructors in each subject area.

2. Distance learning can deliver substantial economic benefits by reducing the time and expenses required to shuttle instructors/students from campus to campus around the state or a large school district; and by generating much-needed added revenues through increased student enrollments. It is estimated that the cost of delivering instructions via distance learning is about one-tenth the cost of moving people around[12].

3. Distance learning can provide a university with a strategic advantage in penetrating new market segments; such as corporate education, continuing adult education, and job training. This has important implications for institutions seeking to make up for lost revenues due to declining student populations and ever-increasing competition.

There are two primary approaches to distance education; i.e., two-way video-conferencing networks that provide face-to-face interaction for teachers and students, and one-way networks such as with Very Small Aperture Terminals (VSATs) that broadcast video instruction or training to a large number of locations, handling questions or feedback from students through audio hookups or fax machines.

A common feature for both distance education approaches is that they make use of digital video, which offers distinct advantages over analog-based networks of the past:

a. Digital video allows using digital compression techniques that permit transmission of digital channels in a fraction of the transponder bandwidth required for analog channels. This translates into lower costs/channel and allows an institution to offer more classes at a wider range of times to suit students

b. Digital video provides a clearer, crisper video and audio- with none of the ghosting, snow or other interference associated with analog broadcasts.

[12] "Distance Learning Spurs Growth in the VSAT Market", Satellite Communications, Sept. 1995, pg. 88.

The range of data rates for digital video can be anywhere from 300 kb/s (for video-conferencing type applications) to 800 kb/s (for high-quality video).

1.3.4 Multimedia Video Services

Multimedia video services refer to a multitude of services such as Video on Demand (VOD), home shopping, CD-ROM, games etc. Of these, VOD has received the greatest attention. VOD will be one of several types of multimedia services which can be distributed over broadband telecommunication networks. Those who are connected to such a digital network and who subscribe to a multimedia service incorporating VOD program services will be able to call up films and television programs on to their television screen from a "video library" (or video server) at times of their own choosing, and make use of such functions as pause, rewind and fast forward.

There are many ways of perceiving VOD. The most simplistic version of VOD allows the viewing of many channels of TV programming with an opportunity to choose among top movies. This view of VOD is actually implemented today; for example, GM-Hughe's DirecTV has over 100 channels of regular programming.

A more challenging view of VOD is one which allows the movies to be shown in real-time. There are a number of tough issues here. First, video data is delay-sensitive. The recording and playback of video are continuous operations, and a management system must provide continuity of the video streams. Second, even compressed video consumes a large amount of system resources such as disk and memory storage space, IO bandwidth and network bandwidth.

According to a study by WinterGreen Research, interactive video services (2-way communications via a TV set) will rapidly accelerate to a total of nearly $25 billion by the end of this decade (see Figure 1-5)[13]. The early consumer market drivers are primarily video games and VOD. Services such as distance learning, medical services, financial transactions, and videoconferencing are expected to follow. It is interesting to compare these figures with those in Europe (see Figure 1-6) which are much lower[14]. Unlike in the U.S. market which shows a bullish trend in the multimedia market right up to 2005, in Europe there will be an initial strong surge in multimedia market revenues followed by slow growth rates from year 2000 onwards. This is partly due to disparities in Gross Domestic Product (GDP), differences in regulatory regimes and cultural environments (Europeans tend to watch less TV than Americans).

---

[13] Interactive Video market, pg. 18 Telecommunications.

[14] "Jeux Sans Frontieres - Toward 2000", pg. 35, Communications International, October 1995.

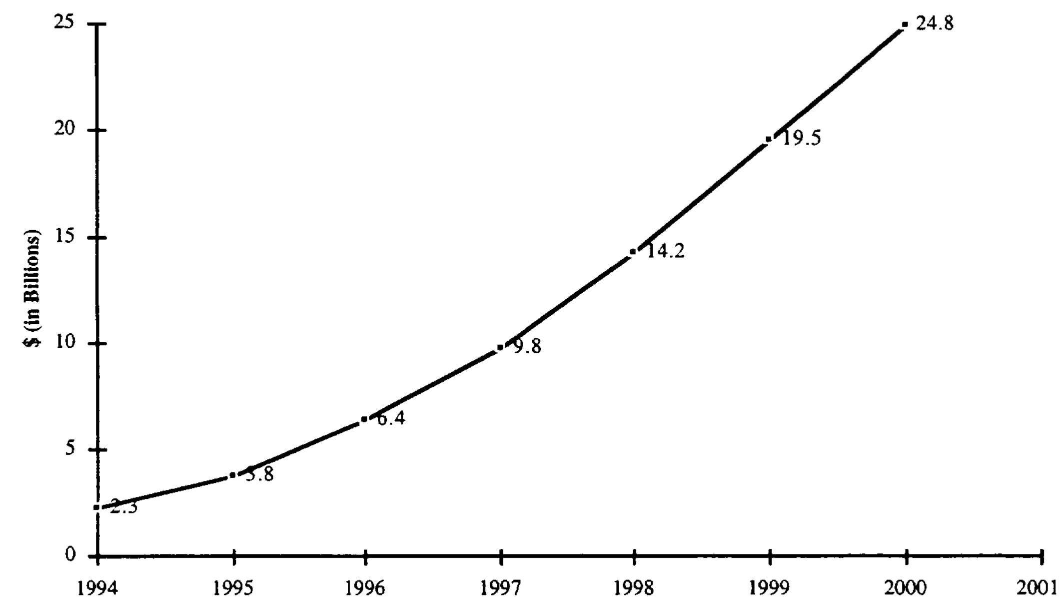
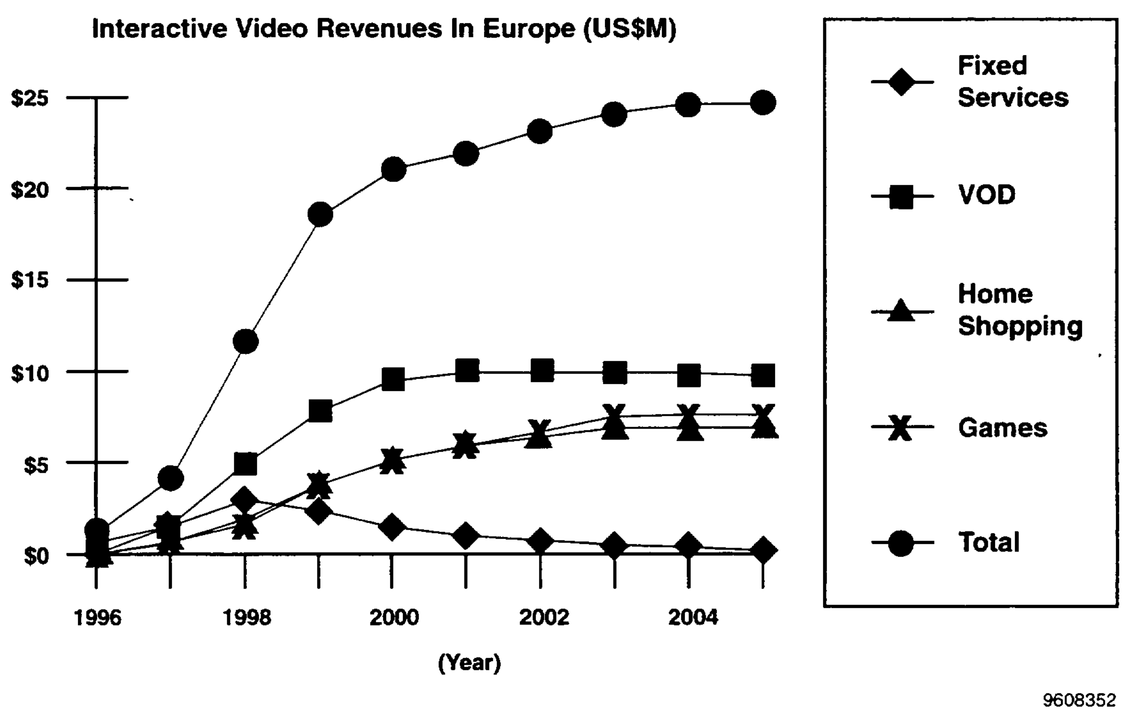


Figure 1-5. Interactive Video Total Market Forecast in the United States



Source: Datamonitor

Figure 1-6. Interactive Video Revenues in Europe

One of the key drivers to the multimedia market has been the sale of PCs with CD-ROM drives (see Table 1-3). It is expected that by the end of 1996, nearly 36 percent of all PCs will be equipped with CD-ROM drives. This has helped the market for CD-ROM software to grow to $1 billion a year from only a fifth of that two years ago. The total number of CD-ROM software titles sold in 1994 was 22.8 million. The total multimedia hardware and software are expected to rise to over $24 billion in sales by 1998, with a rate of growth of 25 percent per year estimated[15].

Table 1-3. Growth of PCs Equipped with Multimedia CD-ROM Drives[16]

Year	Number of PCs in the U.S. Equipped to Run Multimedia CD-ROMs
1994	20.9 million
1995	35.3 million
1996	48.5 million

Home shopping, like VOD, is spurred by growth on the Internet. There has been considerable activity in this field in the last few years with home shopping already a $1.2 billion industry in 1993. Most home shopper services are of the following types:

a. Users have a menu of on-line vendor brochures or catalogs which they can query for services or goods, much like using electronic "yellow pages"

b. Users can subscribe to multicast informational and product releases. When new catalog entries or changes to prices occur, this information arrives at the users' information port. This is analogous to a multimedia version of today's Internet news groups

c. The user can submit a search request which traverses a mesh of attached communication nodes seeking information that matches the instructions from the shopper. These searches will report relevant information to the buyer and seek confirmation of the purchase.

The data rates required for these services should be handled by a relatively low-speed link from the home (<64 kb/s) and a bursty but fairly large channel in return (384 kb/s to 5 Mb/s for some video clips and image services).

---

[15] "Multimedia market to grow to $25 billion by 1998", from Silicon Graphics World, Software Database, 03-01-1993, pp p5(1).

[16] Dataquest, 1994.

1.3.5 Disaster Management

The concept of disaster management is only recently getting recognized as an important NII/GII application market. While disaster management requires many implementation tools, one of the most critical is the use of information technology to support the efforts of emergency managers. The improved use of information technology tools, through the implementation of the NII, will greatly improve the capability of emergency managers to coordinate and disseminate vital emergency communications.

The number and impact of natural disasters around the world are increasing at a dramatic rate. Between 1963 and 1967, the world experienced 16 disasters that took the lives of 100 or more people and 89 disasters that caused a damage of 1 percent or more of national GNP of the countries affected. Twenty-five years later, between 1988 and 1992, the world experienced 66 disasters that killed 100 or more people and 205 that cost 1 percent or more of national GNP. Over three million people have been killed by disasters in the past decades[17].

Although 90 percent of the people affected (and 95% of all people killed) by natural disasters live in the developing world, the more developed countries are not immune from this deadly trend. Table 1-4 yields some disconcerting results as it presents the statistics of U.S. disasters over the last six years. For one, there seems to have been an increase in the number of federally declared disasters over the last few years. An even more alarming figure is the cost of these disasters as shown on Table 1-5. It becomes quite obvious that the federal government spends a great deal on disasters. For example, in the U.S., insurance payouts from natural disasters since 1990 has more than quadrupled payouts for all of the 1980s. And yet, the federal government's expenditures are often not even half the total costs associated with the disaster.

Table 1-4. Summary of Presidentially Declared Disasters or Emergencies

Fiscal Year	Drought	Earthquake	Fire	Flood	Human Caused	Hurricane/ typhoon	Severe Storms	Snow/ Ice	Tornados	Toxic Substances	Volcanos	Total
1989	0	0	1	14	0	5	2	1	7	0	0	30
1990	0	1	5	14	0	2	0	3	13	0	1	39
1991	0	0	2	17	0	10	1	5	6	0	0	41
1992	1	2	11	15	0	9	6	2	8	1	0	55
1993	1	1	7	20	1	2	3	18	12	0	0	65
1994*	0	2	2	2	0	0	14	6	2	0	0	28

* Results available only up to 22 June 1994.

17 "The Media and Disaster Reductionb: Roundtable on the Media, Scientific Information and Disasters at the United Nations World Conference on Disaster Reduction", Fred H. Cate... from WWW address http://www.annenberg.nwu.edu/pubs/disas/disas71.GIF/.

Table 1-5. Cost of Presidentially Declared Disasters

Fiscal Year	Number of Presidential Declarations	Federal Expenses (in Billions of Dollars)
1988	32	189.6
1989	30	138.5
1990	39	2026.2
1991	41	391.5
1992	55	1725.5
1993	65	2467.9
1994*	26	3081.4
Totals	288	10020.6
Average	40	1428.8

*Results available only up to 7 June 1994.

In most of these disasters, along with the toll on human life is the damage to power systems and standard telecommunication systems. However, it is absolutely vital that during all phases of the disaster (before, during, after), emergency telecommunications be operational. This will ensure quick safety procedures and evacuations of people and property, as well as provide assessment of damage and the tools for a quick recovery. The lack of a good emergency telecommunication system has even more devastating effects, in terms of life and property, in many other less affluent countries around the world. It is here that a good GII and NII could prove invaluable.

In the face of the extraordinary and increasing costs (both in human and economic terms) of natural disasters, the United Nations designated the 1990s as the International Decade for Natural Disaster Reduction (IDNDR). Effective, reliable communications are vital for disaster reduction and an important focus of IDNDR.

There are many different forms of emergency communication methods that can be deployed during a disaster as shown in Table 1-6[18]. Satellites can provide very robust and secure emergency communications; both as part of a NII as well as part of the GII structure.

[18] "Global Disaster Mitigation: An IRIDIUM Strength — Matrix of the main elements of Emergency Communications and Information", AIAA-96-1173-Cp, pg 1040 (originally from DHA News-September/December 1994).

Table 1-6. Main Elements of Emergency Communications and Information

	Mitigation (Pre-Disaster)	Information Management (Permanent)	Response (Post - Disaster)
Public Networks	• Information Exchange • Survival and recovery capability of networks	• Routine telecommunications (Phone, fax, telex, data) • Access to data networks	• International links • Local/ regional links (if available) • Distribution of reports
Satellite Telecommunications	• Local/ regional data transmission for remote sensing	• UN backbone network • UN thin-route network • Permanent digital links	• International links from disaster site • Backup for regional and international links • Transition from initial response to thin route network links
HF and VHF Radio Communications	• Decentralized means of telecommunication with high survival capability. • Frequency allocations for humanitarian needs.	• Interface between emergency telecommunications and routine systems	• Regional and local emergency telecommunications links • Backup for international links.
Regulatory and Political Issues	• Facilitation of the use of decentralized means of telecommunications and of data collection links • Collaboration with IDNDR projects containing telecommunications elements	• Licensing of land stations for UN backbone network and connections to public networks.	• Transborder use of emergency telecommunications equipment • Facilitation of emergency telecommunications for NGOs including their participation in UN networks
Satellite Observation	• Monitoring • Early warning • Scientific applications	• Access to and maintenance of databases/images	• Post-disaster data (high resolution images) • Future real-time information
Data Processing	• Maintenance of databases • Linking of databases	• Access to databases • Humanitarian information management systems and networks	• Data collection in the field • Access to HQS and outside databases
Public Information and Media	• Creation and upkeep of public awareness • Education • IDNDR promotional activities	• PR work of UN and other humanitarian institutions • Publications	• Information from media (Monitoring) • Information to media (on field level)

The main emergency communication requirements are:

a. Early warning signals to be passed to affected areas through news bulletins and emergency broadcasts based on sensor readout data and monitoring equipment (typically using satellite observation monitoring)

b. Real-time brief operational information among national and international teams and emergency-managers at the site of the event, typically through mobile or portable equipment

c. Fast exchange of consolidated information between the coordinating organizations on field level, such as an On-Site Operations Coordinator, the National Disaster Management team in the capital of the affected country, and their counterparts worldwide, typically through land-mobile satellite terminals and long-range HF links. Since the normal, wired, public switched telecommunications network is often the first "casualty" in a disaster, untethered satellite systems provide a natural backup function. Aside from low-speed sensor data and voice communications will probably be the need for transfer of video images. This will allow disaster management teams to quickly evaluate the extent of damage thereby helping them formulate a quick line of relief action. It is expected that the data rates required for these types of services will be fairly low to medium-speed, and the communication network will be bursty in nature

1.3.6 Home Banking/ Electronic Commerce

Electronic commerce is the ability to perform transactions involving the exchange of goods or services between two or more parties using electronic tools and techniques. Electronic commerce has substantial advantages over traditional, face-to-face, paper-based commerce:

a. It gives the customer more choices and customization options

b. It decreases the time and cost of search and discovery (of customers and "banks")

c. It expands the marketplace from local and regional markets to national and international markets with minimal capital outlay, equipment, space or staff

d. It reduces the time between the outlay of capital and the receipt of products and services

e. It facilitates just-in-time production and payments, reducing overhead and inventory through increased automation and shortened processing times

The driving force for home or *virtual* banking has been the growth on the Internet. Commerce is now firmly in the driving seat of future Internet developments. An important step has been to provide secure processing in these cyber-transactions. Providers of financial services and enabling software, electronic encryption and security services have been quick to follow up on these opportunities. Online purchases today account for only four percent of total global sales. However, electronic commerce is predicted to grow dramatically over this decade. Various predictions have been made about this growth. For example, within six years, global shoppers could use the NII to purchase as much as $500 billion of goods and services — or about eight percent worldwide. Similarly, it is now projected that by year 2005, the number of NII based transactions could rise to anywhere between $17 billion[19] to U.S. $24 billion[20] .

1.3.7 Government Services

The U.S. government is one of the largest markets for information technology. As shown in Figure 1-7[21], the U.S. government currently spends (as of 1995) about $17 billion on civilian-related applications of information technology and about $10.1 billion on defense-related applications. This is expected to rise to about $20 billion and $12 billion respectively by the end of this century.

The use of NII and GII should provide considerable savings in cost and increase efficiencies by reducing person-to-person transactions. Despite the fact that about 83 percent of federal and 65 percent of state government workers use e-mail, and about 57 percent are connected to the Internet, only about 6 percent of the communications between government and private sector is currently conducted via e-mail. This figure will improve considerably as the NII/GII picks up momentum. Just this year, for example, the IRS has included a home page on the Internet to help people with queries about their income taxes and to allow easy to transfer of government documents and forms. This will considerably reduce the tie-ups on the telephone lines that have been so commonplace when dealing with queries to the government. This is going to create greater efficiencies by allowing the NII to take care of routine jobs and allowing government personnel to handle more involved, and thought-provoking tasks.

Similar benefits can be obtained by the ability of law enforcement agencies to have access to central databases on crime statistics, training, etc. as well as the ability to transfer photographs, fingerprints, and other time critical information between sites.

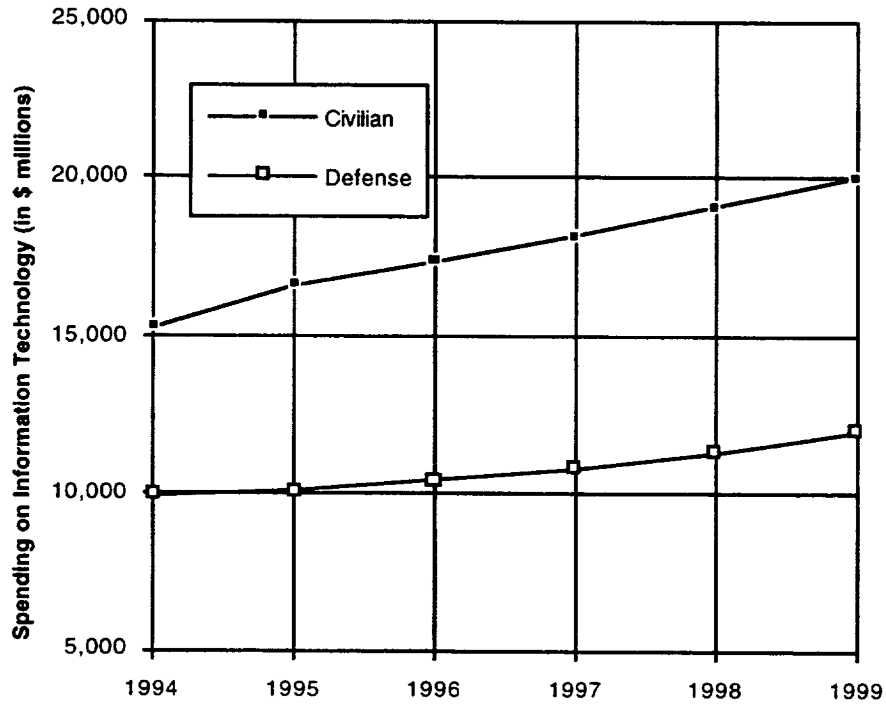
The data rates required for these services, as in the previous case, should be handled by a relatively low-speed link from the home (<64 kb/s) and a bursty but fairly large channel in return (384 kb/s to 5 Mb/s for some video clips and image services).

---

[19]Electronic Commerce in the NII from http://www.cnri.reston.va.us:3000/XIWT/documents/documents.html.

[20] Net Venture: Special Feature: Virtual Banking from http://www.web.co.za/cytec/future.html.

[21] "The Federal Market," Washington Technology Almanac, 1995, pg. A-13.



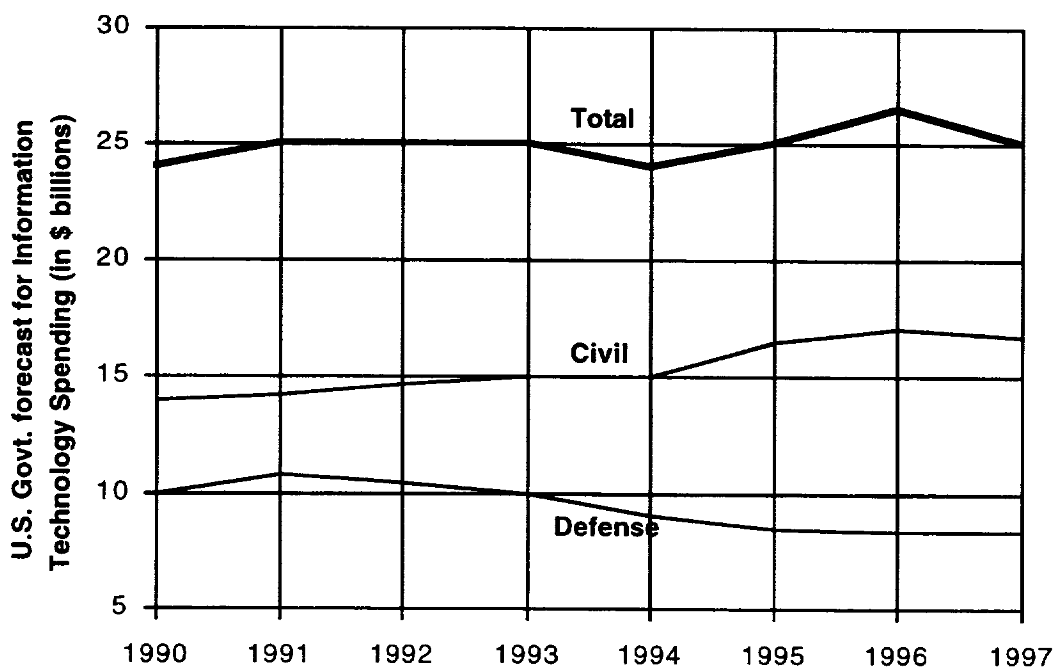


Figure 1-7. Civilian vs. Defense Spending on Information Technology

1.3.8 Advanced Networks for Research and Development

One of the very first applications for the NII/GII, in fact for the Internet, was creating a supercomputing environment for advanced research and development activities. The federal High Performance Computing and Communications (HPCC) program was developed to act as a driving force to advance computing, communications and information technology and their applications to challenges in science and engineering. Some of the key strategic focus areas of HPCC are:[22]

a. Global-scale information infrastructure technologies that build advanced application building blocks and widely-accessible information services

b. High performance/scalable systems to support high-performance and low-end applications in a seamless fashion

c. High confidence systems that will provide the availability, reliability, integrity, confidentiality, and privacy needed by the Nation's emerging NII

d. Virtual environments and simulations that will continue to transform scientific experimentation and industrial practice and play an increasingly important role in education and training

e. User-centered interfaces and tools to provide easier development, navigation, "mining", and general use of information resources

f. Human resources and education, both to educate the next generation of industrial and academic leaders in information science and technology and to establish a foundation for new learning technologies.

The grand challenge applications, alluded to earlier, represent some fundamental problems in science and engineering, which can only be solved by tremendous computational resources and simulations; and whose solution will have broad economic and scientific impact. Figure 1-8[23] gives an overview of the performance requirements of these and other challenges; as well as a time frame within which these challenges are expected to be solved. The availability of high-speed computers, such as the CRAY-2, as well as the Gigabit data bus for scientific use by laboratories and universities, provides a stimulating environment for R&D and acts as a precursor of the service types one can expect in the future of the NII and GII.

---

[22] From the HPCC Program's home page whose URL is http://www.hpcc.gov/

[23] Information Infrastructure Sourcebook, Version 3.0 Volume 1, pg. 81.

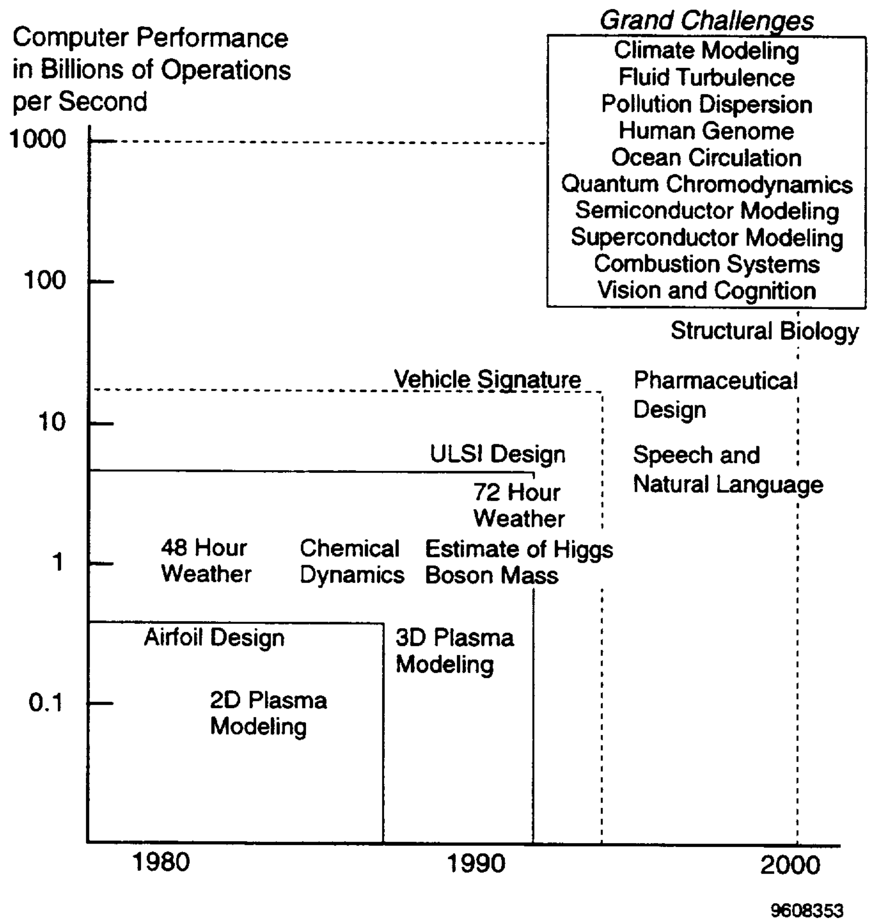


Figure 1-8. Performance Requirements for the Grand Challenge

1.4 SUMMARY OF NII/GII APPLICATIONS AND QoS REQUIREMENTS

The various NII/GII applications discussed in the previous section are summarized in Table 1-7. Specifically, the satellite addressable market of these applications are emphasized. It is clear from this table that the data requirements for these markets are quite varied. A time frame divided into present, evolving and future time periods for the deployment of these applications is given in Figure 1-9 which categorizes these applications as a function of bandwidth, interactivity requirements, and complexity. Clearly, for example, Emergency Services would require a long time-frame to set-up, but its bandwidth requirements are fairly modest. At the other end of the spectrum, electronic banking facilities are already in place in a number of areas and they require very low bandwidth.

Table 1-7. Summary of Satellite Addressable Markets for NII/GII Applications

NII/GII Markets	Satellite Addressable Requirements	Broadband Services needed	Market
Tele-commuting	• Remote access to office workstation using portable Pcs • Extension of corporate LAN using groupware • Remote video conferencing • Remote interactive video	• Two-way voice fixed/mobile • Multiple-way voice teleconferencing • Low/Medium/high speed computer network, fixed/mobile • Video conferencing (low speed) • Video conferencing (broadcast-quality)	Mobile satellite services enable sat. offices to be in non-premium areas. Expected 20 to 30 percent of workforce in next decade will telecommute. Thirteen million telecommuters expected by 1998
Tele-medicine	• Mobile health services (for rural and underdeveloped/underserved regions of world) • Medical examination (2-way video, electronic stethoscope, otoscope, opthalmoscope) • Radiology (x-rays, CT scans, MRI, mammagrams) • Pathology (high resolution, color pathological imagery) • tele-psychiatry (remote video consultations) • Computer-aided diagnosis and treatment • Medical education (interactive audio/video) • Virtual reality (digitized 3D active model of human body for surgical/other training) • Smart card medical data (multimedia patient data)	• Symmetric links of 128 to 384 kb/s for consulation • One-waychannels of 1.54 to 10 Mb/s for medical images, x-rays, etc. • Secure and reliable connections	The market for telemedicine exceeded $25 billion in 1994. Expected to rise above $50 billion by 2000. It is estimated that about 25 percent of data traffic on NII will be health care related
Distance Learning	• Desktop (standalone) computer-based training • Corporate server-based training • Remote training using a PC connected to the training site by a switched or permanent network • Off-the-air TV programming (e.g., weekend medical training, seminars) • Dedicated institutional learning system with multiple "classrooms"	• Point-to-point/point-to-multipoint video (64 kb/s to 45 Mb/s) • Broadband video conferencing (0.38 Mb/s to 45 Mb/s) • High-speed data transfer (10 db/s to 50 Mb/s) • Broadband videotex and video retrieval (30 to 100 Mb/s)	U.S. market for video conferencing was $1 billion in 1993, $1.5 billion in 1994. Projected to grow to $15 billion by 2000. Desktop video-conferencing projected to be $1 billion in 1997. Corporate training market Market is estimated between $80- to $120-billion by year 2000
Multimedia Video Services	• Customizable news, sports and financial information (satellite news gathering) • Interactive games • Satellite software distribution	• Low-speed link from home (< 64 kb/s) • Bursty but large bandwidth link to home (384 kb/s to 10 Mb/s)	Interactive video market to grow to $2.5 billion by 2000 (growth rate approximately 25%)
Disaster Manage-ment	• Monitoring (satellite sensors) • International links from disaster site • Backup for regional and international links • News gathering	• Low-speed links (< 64 kb.s) for data • Broadband video conferencing/Imaging (0.38 to 45 Mb/s)	disaster (both man-made and natural) striking at a higher rate. tremendous role for satellites envisioned
Home Banking/Electronic Commerce	• Wireless backbone for worldwide financial/business information databases/networks	• Mainly low-speed transactions from home (< 64 kb/s) • High-speed for video clips and image services to home (0.384 to 5 Mb/s)	By year 2005, the number of NII based electronic transactions could be in $17- to $24-billion range
Govern-ment Services	• Interconnection of government organizations • Transfer of government documents and forms to public • Networked databases for various government organizations	• Low-speed link in one-direction (<64 kb/s) • Bursty and large channel in other direction (0.384 to 5 Mb/s)	U.S. government will spend approx. $20 billion on civilian and $12 billion on defence related applications of information technology by year 2000
Advanced Networks for R&D	• Supercomputing for Grand Challenge applications • Virtual Environments and simulations • Human Resources and Education	• Full motion video (45 to 300 Mb/s) • Low-speed data (0.064 to 1.544 Mb/s) • Ultra high-speed data (over 1 Gb/s) • Video teleconferencing (at 384 kb/s)	Market data not available.

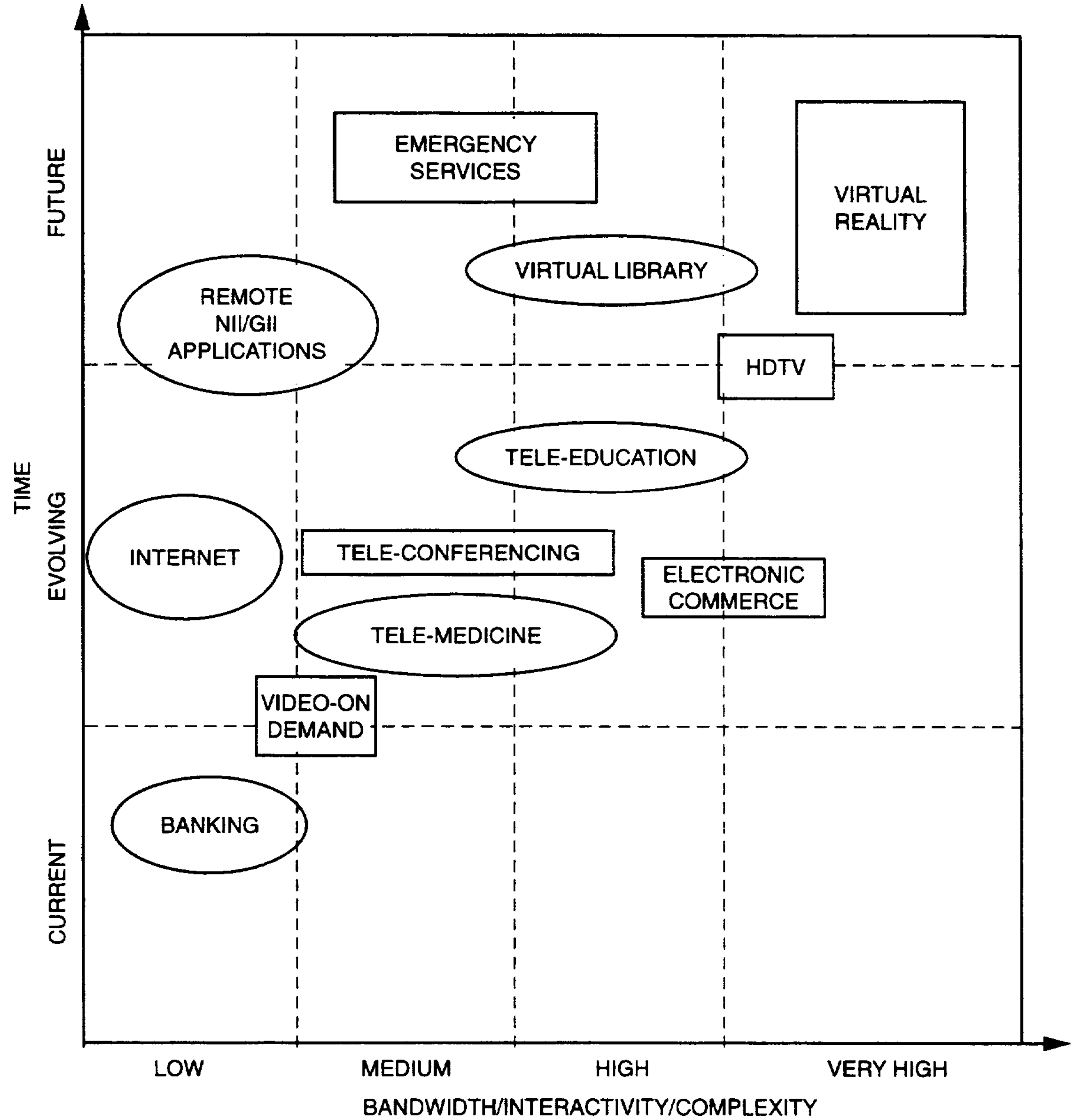


Figure 1-9. Emerging NII/GII Applications and Their Time-Frame for Deployment

Figure 1-10[24] depicts the bandwidth and traffic flow characteristics of the various services that make up the NII/GII applications. The traffic requirements are broadly classified as steady or bursty. For example, image transfers, especially if they are required to be high resolution, take up a very large bandwidth. Data transfer also tends to be bursty. In contrast, a simple file transfer requires very little bandwidth and the traffic rate tends to be more steady.

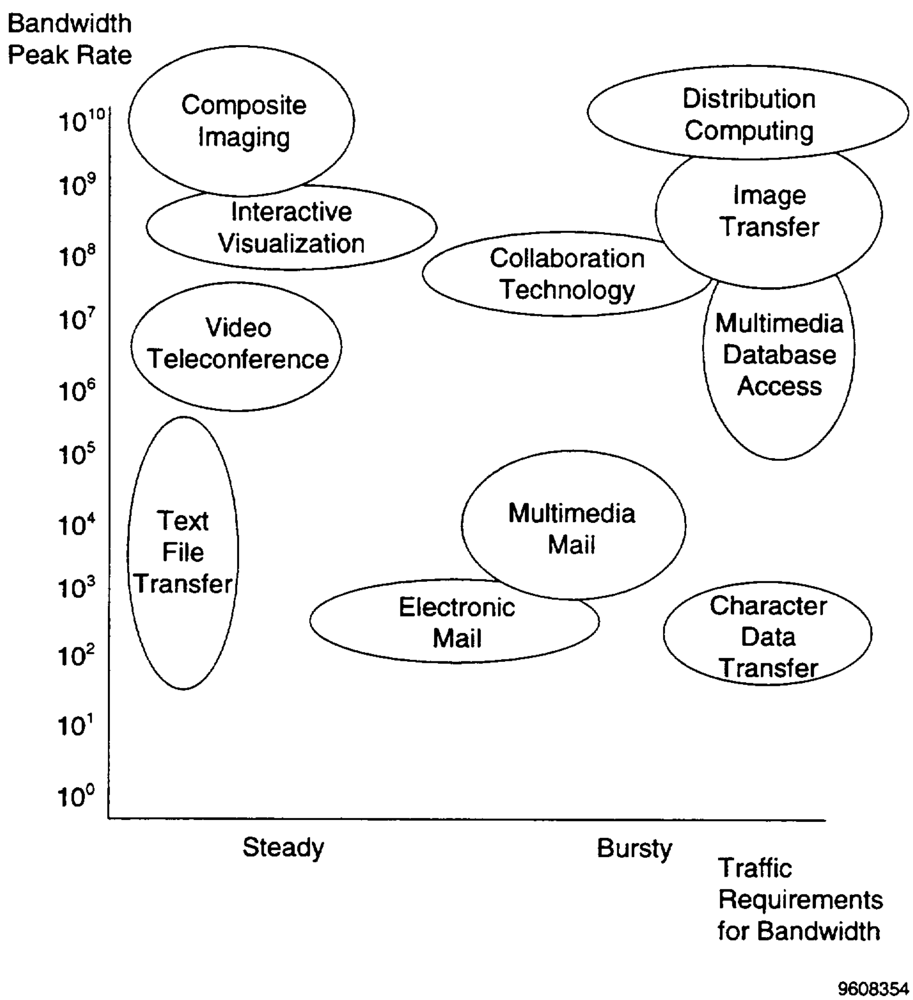


Figure 1-10. NII/GII Applications and Services Bandwidth and Traffic Characteristics

---

[24] Information Infrastructure Sourcebook, Version 3.0, Volume 1, pg. 87.

Figure 1-11[25] depicts the delay requirements of the various services as a function of data transfer size and channel speed. Thus, for example, image transfer may require transfer of file of over 200 MB in size. Transfer of such a file over a T1 link would take over two minutes, whereas the same transfer over a 100 Mb/s FDDI-link would take about two seconds.

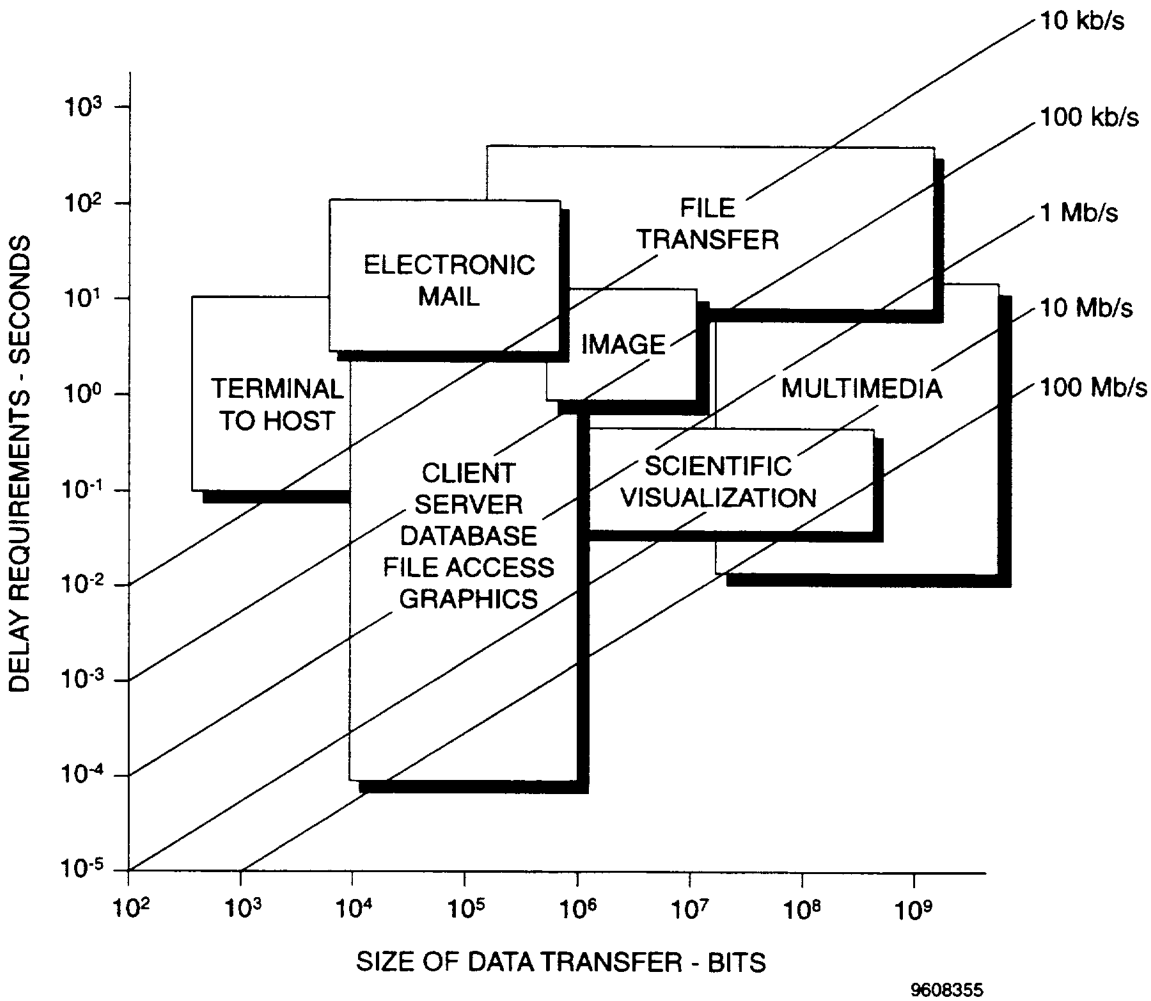


Figure 1-11. Delay Requirements for Various Services

[25] "A Road Map to broadband Networking: Network Evolution within the Utilities Industry", Sham Rane, Telecommunications, Dec. 1994, pg. 34.

For systems where there is need for immediate response, satellite systems, with their large transmission delays may not be well suited. Figure 1-12[26], examines the various NII/GII services from point of view of response time and throughput requirement. The dark area represents services that may not be possible over satellites.

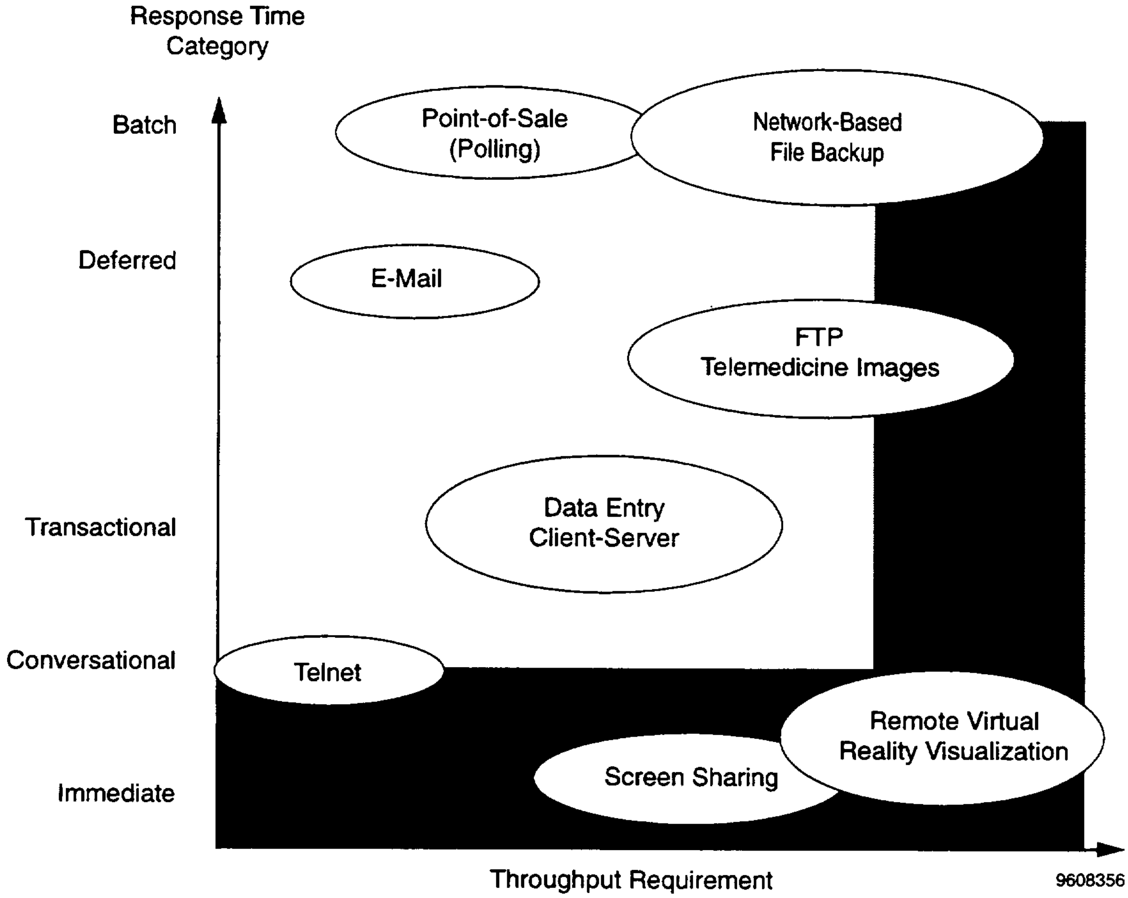


Figure 1-12. Throughput and Response Time Requirements of Various Applications

1.5 FORCES DRIVING THE NII/GII MARKET

Having examined the various NII/GII applications, it is instructive to look at the forces that create a market for these applications. These forces can be broadly categorized as:

a. The proliferation of high performance end devices and software

b. Progress in Networks.

---

[26] "Data Communications Protocol Performance on Geostationary Satellite Links — Lessons learned using ACTS," Hans Kruse, AIAA 16th International Communications Satellite Systems Conference, Feb. 25-29, 1996, pg. 726.

1.5.1 Proliferation of High Performance End Devices and Software

One key to the information boom we experience is the sophisticated set of multimedia devices that are presently available — made possible by the use of integrated-circuit technology which has been successful in reducing its gate density by over seven orders of magnitude (> 10 million) within the last 30 years, as well as bringing down its costs by means of mass production. These devices include the PC hardware, TVs, faxes, and routers. Table 1-8[27] shows the evolution in performance and cost of the PC. Currently, IDC estimates that 37 percent of U.S. households have a PC[28] . This is expected to reach 90 percent by the year 2002. (see Figure 1-13). At that point, the penetration rate will be the same as that of today's television in present-day America.

Table 1-8. The PC, Now and Soon: What $1800 Will Buy

Feature	1995	1996	2000
Processor	Pentium fifth-generation	Pentium fifth-generation	Eighth-generation
Speed	60 MHz	100 MHz	600 MHz
Memory	8 megabytes	16 megabytes	64 megabytes
Storage	420 megabytes	1 gigabyte	8.32 gigabyte
CD ROM drive	Double-speed	Quadruple-speed	Six-speed
Outside link	Fax-modem at 14,400 b/s	Fax-modem at 28,800 b/s	Built-in network connection at up to 100 million b/s

Table 1-9[29] shows the projection of some of these end-devices over the 1993 to 1997 time frame. As can be seen, all of these devices are expected to show healthy growth in sales. Similar growth rates are expected in other regions of the world.

---

27 "The Myth of Multimedia", Wall Street Journal, June 19, 1995.

28 From the Motorola Millenium FCC filing, pg. 19 (Source: IDC Global Home Market Survey, 1995)

29 Hughes Galaxy/Spaceway FCC filing, pg. 33 (Sources: "Surveys, Forecasts and other data bearing on Potential Demand for a proposed Satellite Communication Service," Link Resouces, 20 Jan. 1994 and "1995 SPACEWAY Services Demand Study", Hughes Communications, June 1995).

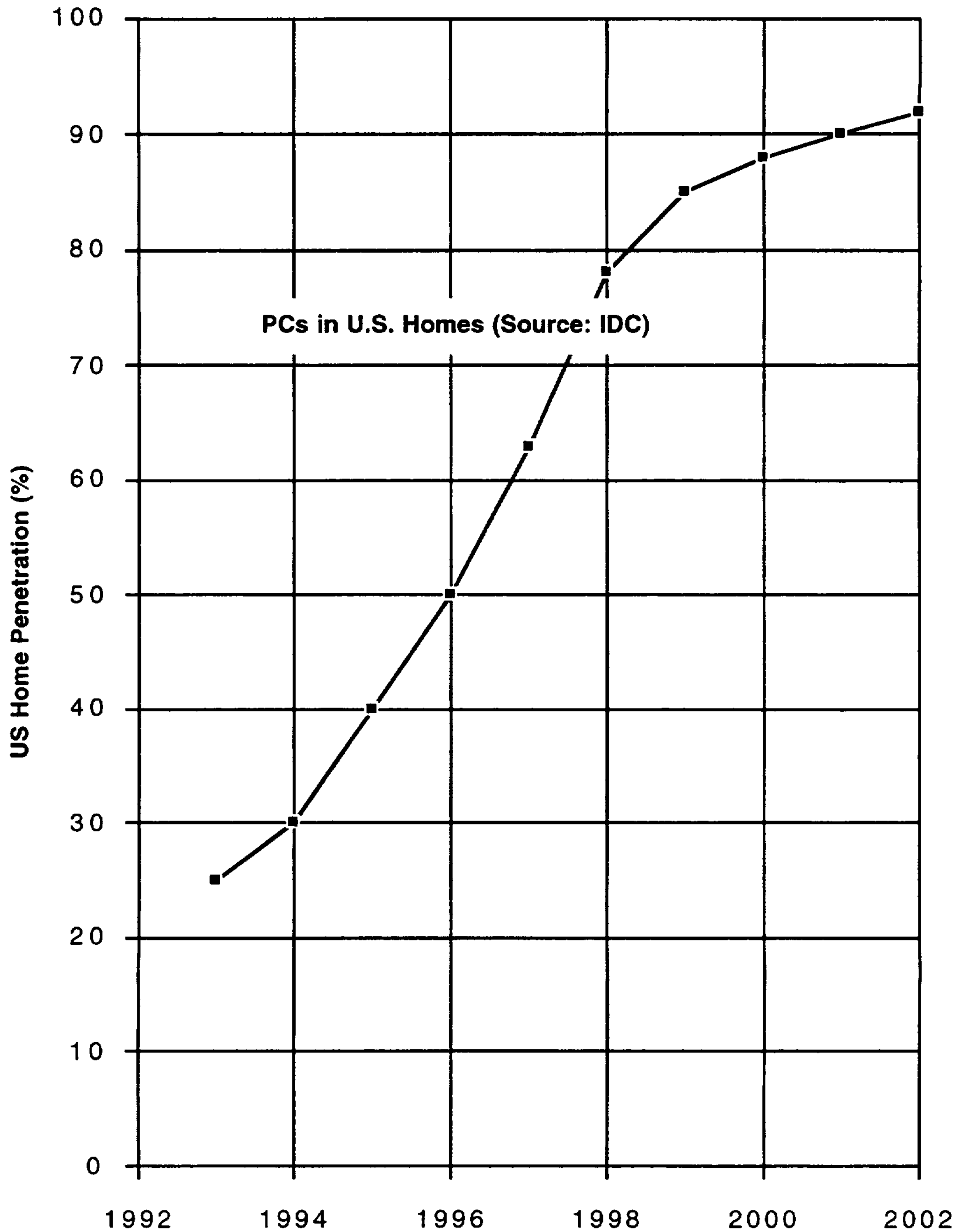


Figure 1-13. Percent Penetration of PCs within U.S. Households

Table 1-9. Projected U.S. Market for Home Computers and On-Line Services

	1993	1995	1997
Home PCs	31 Million	38 Million	45 Million
Home PC/ Modems	13 Million	18 Million	21 Million
On-Line Services Subscribers	4 Million	14 Million	16 Million

The proliferation of PCs and the growing need for interconnectivity results in a similar growth in the bridge/router market, as shown in Figures 1-14 and 1-15[30]. In these figures, the growth rates for both the U.S. and the European market are shown side-by-side. The worldwide hub market is also given in Figures 1-16 and 1-17[31], both in terms of total shipments as well as hub revenues. It can be seen from these figures that United States holds almost half the market share of the hub revenues for the world (for example, in 1995, US sales were $300 million, while in the whole world it was only about $600 million).

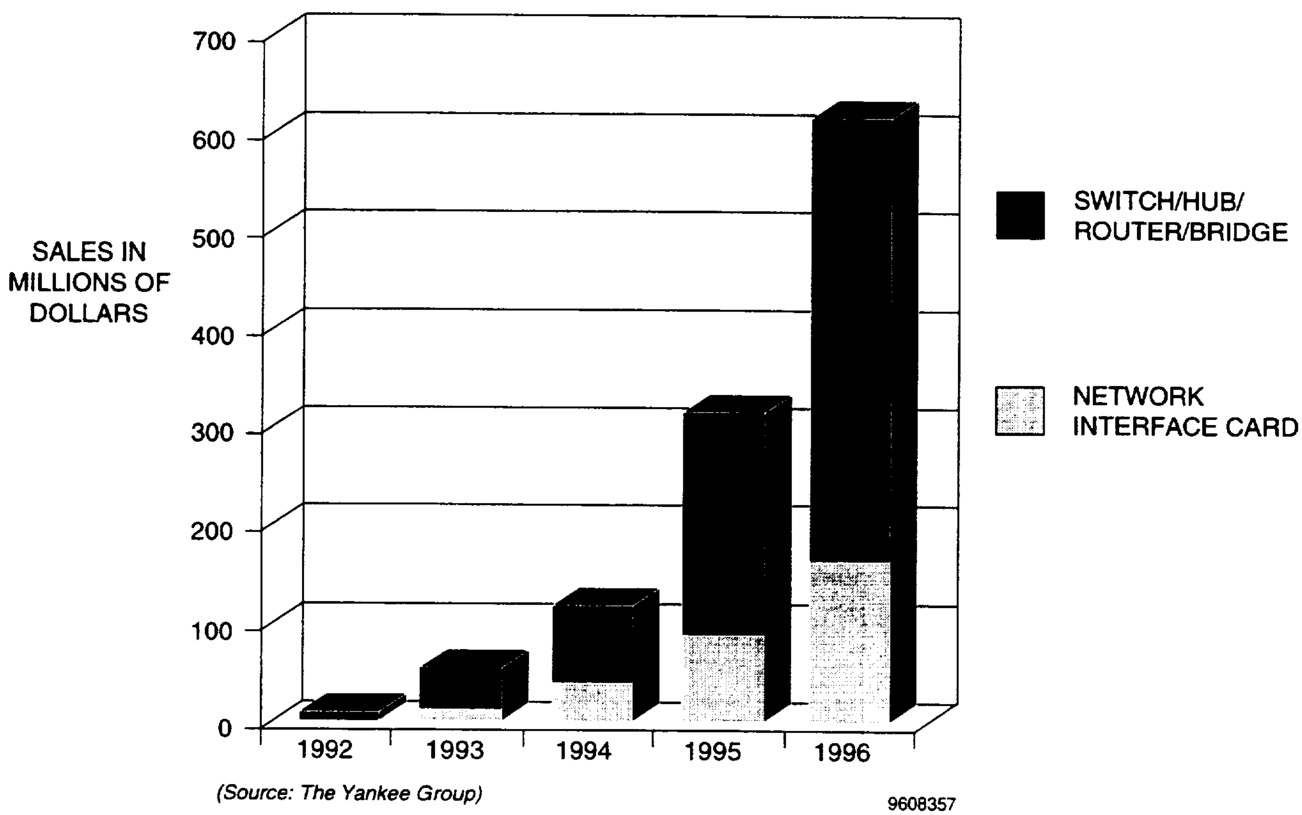
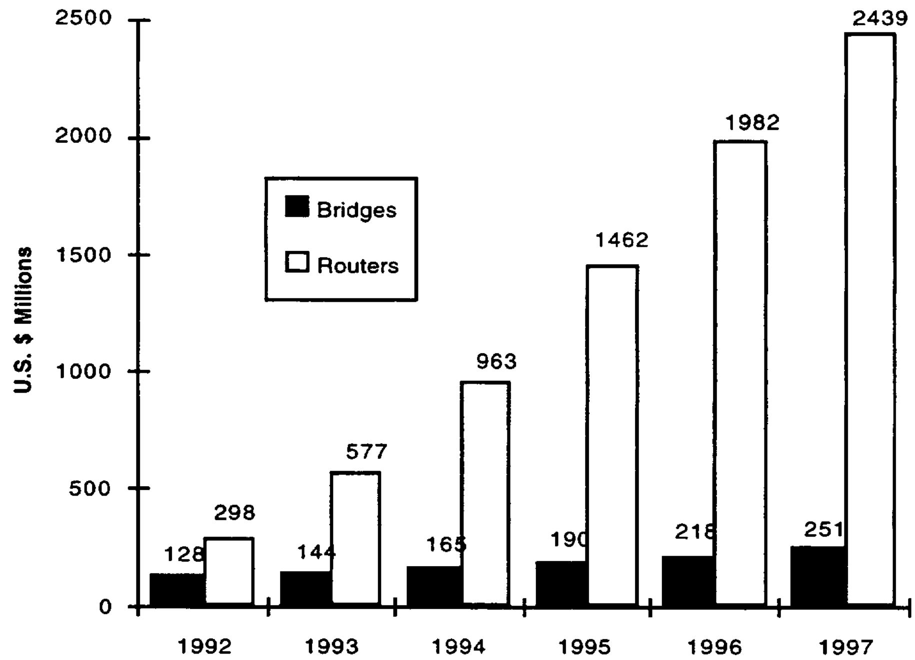


Figure 1-14. Bridge/Router Market Sales for USA

---

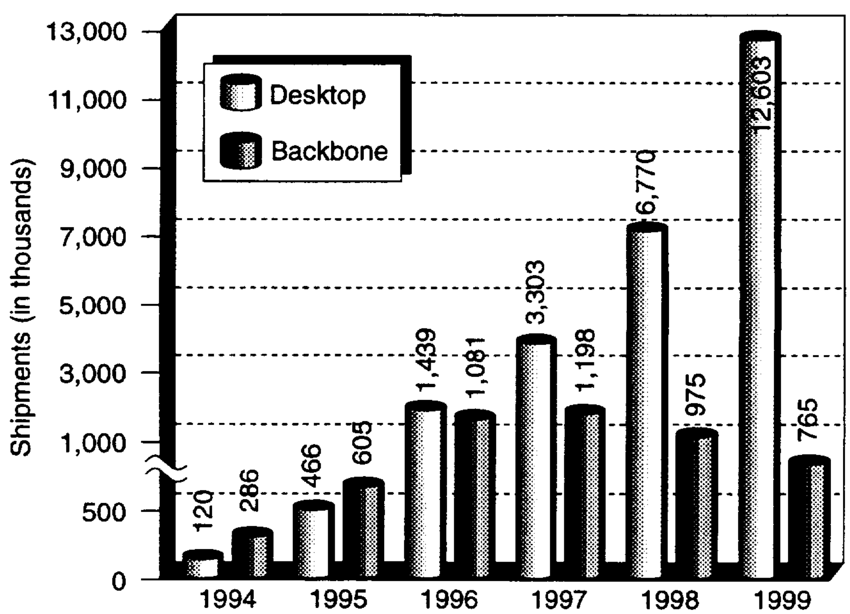
30 Telecommunications, April 1994 (Source: The Yankee Group, Boston) and Telecommunications, June 1994, pg. 18 (Source: Frost & Sullivan)

31 "The future of Networking", LAN, Dec. 1995, pg. 130 (Source: The Yankee Group).



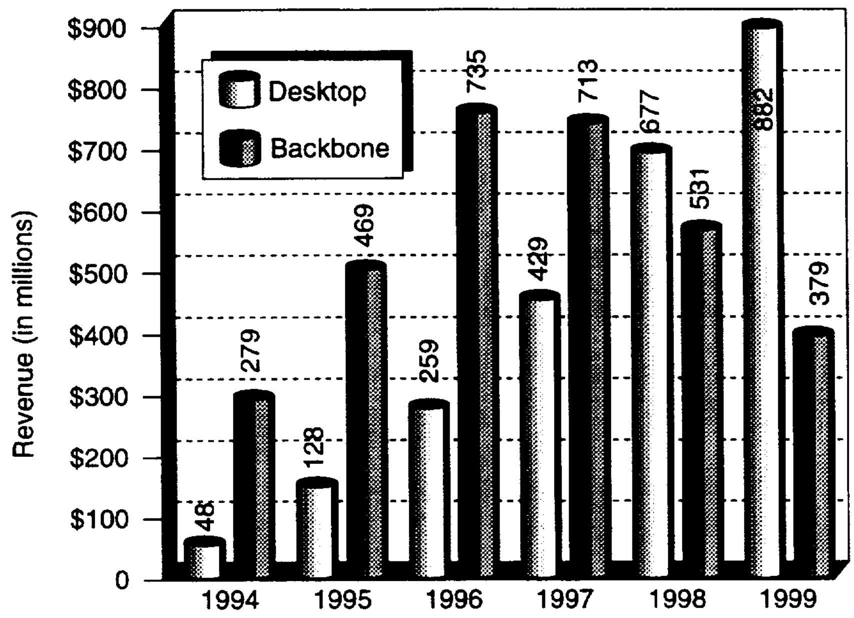
(Source: Frost & Sullivan)

Figure 1-15. The European Market for Bridges and Routers



Source: The Yankee Group 9608358

Figure 1-16. Worldwide Switching Hub Shipments



Source: The Yankee Group 9608359

Figure 1-17. Worldwide Switching Hub Revenue

Alongside growth in hardware has been a steady decrease in information processing costs as shown in Figure 1-18. A major jump in technology from mainframe to distributed has resulted in more than a ten-fold reduction in price per Million Instructions per Second (MIPs) (from about $10/MIPs to $0.2/MIP).

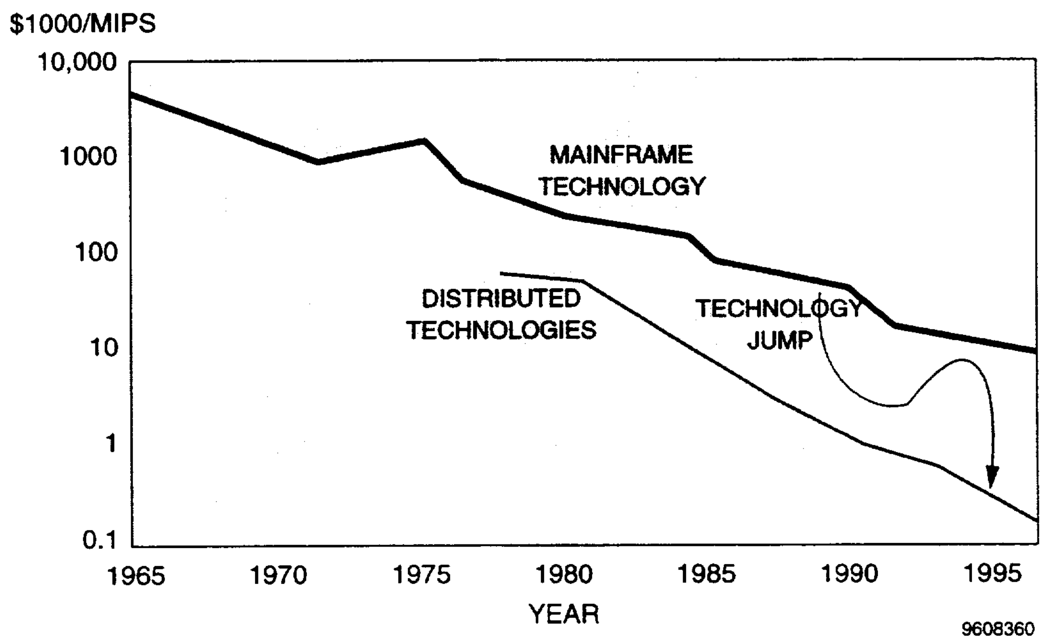


Figure 1-18. Information Processing Price/Performance Improvements

1.5.2 Progress in Networks

The biggest driving force fueling growth in NII/GII applications is the growth of the Internet. Figure 1-19[32] shows the worldwide trend in Internet subscriptions; giving the number of hosts in each region in 1995 and the growth rate between 1994 and 1995. While the United States has the highest number of Internet subscribers, it was Africa which experienced the highest growth rate. The growth rate of Internet connections was 609 percent in U.S. and 587 percent worldwide. Figure 1-20[33] graphically predicts the growth rate in wireless connectivity across the world.

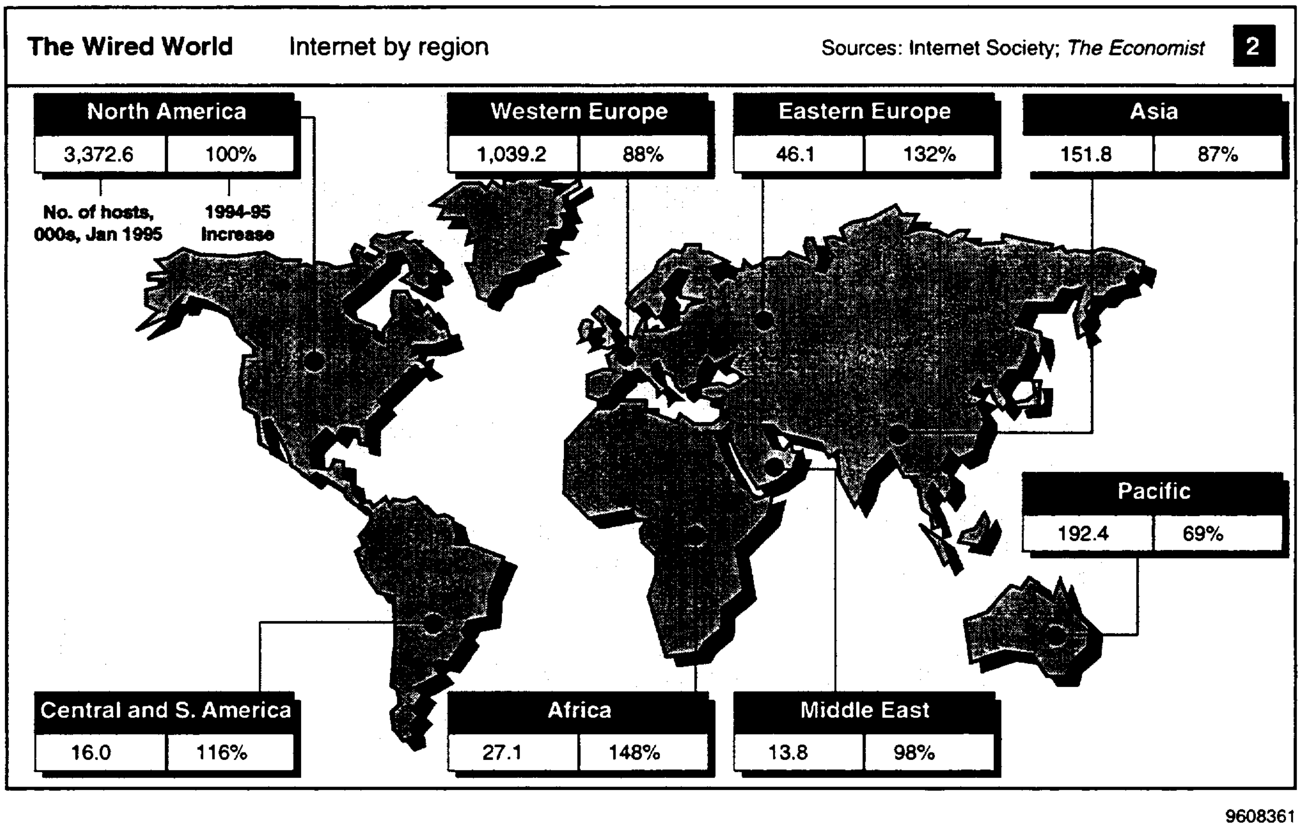


Figure 1-19. Worldwide Internet Subscriber Map with Number of Hosts Shown Per Region

[32] The Economist, July 1, 1995, pg. 4.

[33] Ericsson Advertisement.

In the LAN/WAN arena, there has been a steady increase in the shipments of various Network Interface Cards (NICs). Table 1-10[34] shows the present and projected rates in shipments of Ethernet, FDDI, ATM LANs, etc.

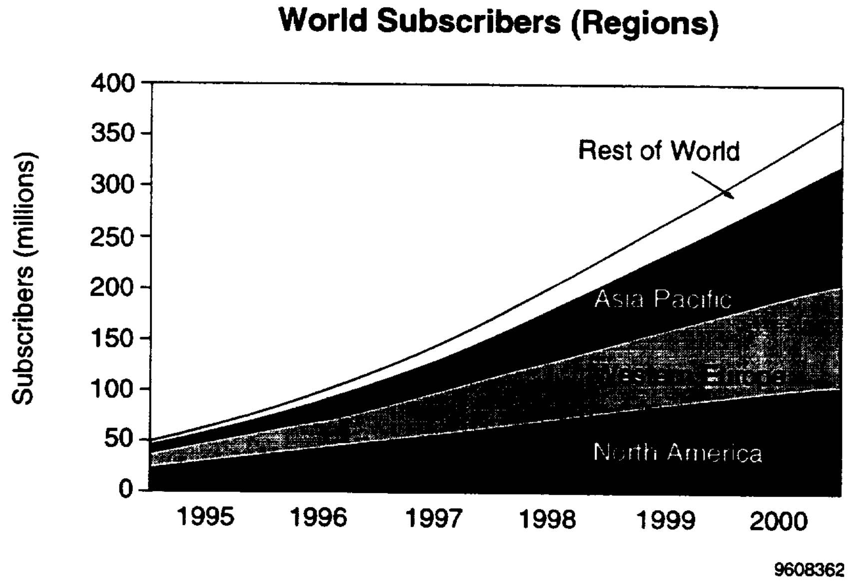


Figure 1-20. Worldwide Wireless Subscriber Growth Predictions

Table 1-10. Number of NICs Shipped and Projected to Ship in Thousands

	Actual 1993	Estimated 1994	Projected 1995	1996	1997	1998	1999
Ethernet	14,555.6	17,175.6	20.267.2	23,915.3	28,220.1	33,299.7	39,293.6
Ethernet 100	N/A	250	600	1,000	1,400	1,700	1,900
Token Ring	3,568.1	4,138.9	4,801.2	5,473.4	6,239.6	7,113.2	8,109
FDDI	18.6	22.344	25.696	28.266	33.919	40.702	43.959
Local Talk	835	653	483	375.4	309.7	289.5	270.62
ATM LAN	1.65	5.2	18.5	66	261	841	1,193
Arcnet	325	289	224.5	169.8	103.2	63.4	38.95

34 LAN Magazine, July 1995, pg. 84 (Source: Computer Intelligence InfoCorp LAN Services).

In addition to the progress in networking, has been the tremendous progress in areas of circuit, packet and cell-based switching. Table 1-11[35] shows the market trend in various switching services and equipment. While X.25 type packet switching showed only a modest growth rate of five percent in United States and eight percent worldwide in 1995, frame relay packet switching technology shows an 200 percent growth rate. The Switched Multimegabit Data System (SMDS) showed a growth rate of approximately 100 percent and (ATM) over 150 percent growth rate. The growth in all these services is overshadowed by the phenomenal growth in internet connections (approximately 600% both in U.S. and abroad).

Table 1-11. Data and Network Services/Product Revenues (in $US Millions)

Service/ Product	U.S. or Worldwide	1994 Revenue	1995 Revenue	1996 Revenue (Projected	1998 Revenue (Prodjected)	1995 Growth Rate	1996 Growth Rate (Proj.)
X.25 Service	U.S	$950	$998	$1030	NA	5%	2%
	Worldwide	$2500	$2700	$2835	NA	8%	5%
Frame Relay Service	U.S.	$183	$554	$1130	NA	203%	104%
	Worldwide	$231	$647	$1261	$3570	180%	95%
Frame Relay Switches and Access Devices	U.S.	$321	$533	$755	NA	66%	42%
	Worldwide	$470	$776	$1095	$1810	65%	41%
SMDS Service	U.S.	$10	$20	$38	NA	100%	90%
	Wordwide	$18	$35	$72	NA	95%	105%
ATM Services	U.S.	$11	$28	$63	NA	155%	125%
	Worldwide	$12	$30	$69	$339	150%	130%
ATM Switches	U.S.	$38	$81	$160	NA	113%	98%
	Worldwide	$47	$110	$244	$1650	134%	122%
ISDN (BRI only)	U.S.	$59	$110	$211	NA	86%	92%
	Worldwide	$234	$608	$1154	NA	160%	90%
Commercial Internet Access	U.S.	$141	$1000	$1799	NA	609%	80%
	Worldwide	$189	$1298	$2194	NA	587%	69%

[35] 1995 NATA Telecom Market Review & Forecast, pg. 164, Table V-1.1)

Of the various switching technologies mentioned above, it is clear that ATM switches will hold the largest market due to falling prices of ATM equipment, the emergence of low-speed ATM interfaces which can use existing cabling, major ATM initiatives by IBM, General Datacom and Newbridge, the continued growth in the deployment of LANs, and the industry's orientation towards switched solutions. Table 1-12[36] forecasts the growth rate of ATM sites both as part of a local area network backbone and as a part of a wide area network. Slower growth is predicted in Europe compared to the U.S. with European users being reported as viewing ATM as a backbone technology that will delay market development until there are suitable wide area services.

Table 1-12. European and U.S. Broadband Growth Forecasts

		1994	1995	1996	1997	1998	1999	2000
ATM LAN Backbone Sites	Europe	NA	0.121	1.06	4.55	12.8	26.7	49.5
	USA	0.159	1.06	4.25	11.5	28.7	50.6	77
ATM LAN Workgroup Switches	Europe	0.095	0.6	2.5	8.61	27.4	76.8	189
	USA	1.17	4.86	14.2	33.1	70	124	199
ATM WAN Backbone Switches	Europe	0.013	0.047	0.155	0.45	1.25	3.15	5.22
	USA	0.088	0.326	0.991	2.43	5.39	11.3	17.6
ATM WAN Service Connections	Europe	1.27	3.81	9.46	20.3	38.9	68.6	113
	USA	4.49	13.5	30.9	58.7	98.4	151	220

1.6 MARKET TRENDS IN THE BITWAYS

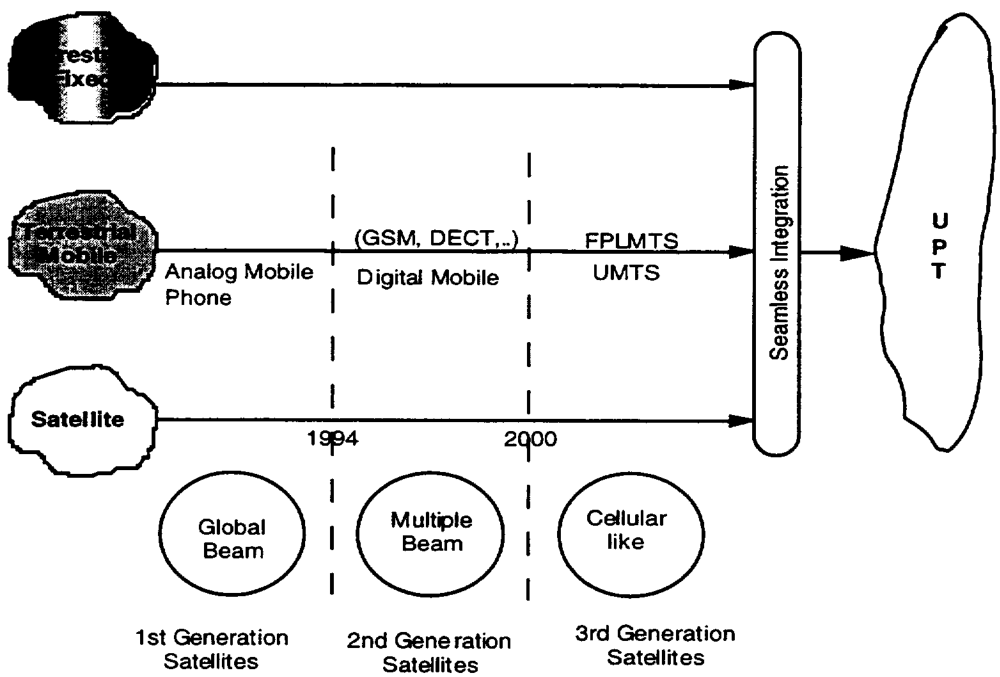
The NII/GII will involve the use of a variety of physical infrastructures or bitways, such as fiber cable, copper lines, wireless PCS and cellular links as well as satellites — all of which will operate seamlessly to provide interconnectivity for the user. Figure 1-21[37] depicts how the various terrestrial, satellite and mobile networks will evolve. We will consider the market trends for each of these physical infrastructures.

---

[36] Telecom*europa*'s Interactive Video Newsletter, Dec. 4th, 1995, pg. 17.

[37] "Satellite and Personal Communication Services: Market, Spectrum and Regulatory Aspects, F. Anaansso, AIAA-96-0991-CP, pg. 159.

Terrestrial links, it is expected, will show major growth in the local loop, where with the introduction of Fiber-to-the-Home (FTTH) or Fiber-to-the-Curb (FTTC), one can bring broadband services to the end user or subscriber. Figure 1-22[38] shows the expected growth in the deployment rates of these various technologies. Both fiber/coax and FTTH/FTTC are very costly solutions to providing broadband services (installation cost approximately $1500 per subscriber). Asynchronous Digital Subscriber Loop (ADSL) has recently received much attention as a possible alternative method to carry broadband signals over the common twisted pair cables and is expected to be much less costly.



FPLMTS: Future Public Land Mobile Telecommunication System
UMTS: Universal Mobile Telecommunication System
UPT: Universal Person Telecommunication

Figure 1-21. Evolution of the Terrestrial and Satellite-Based Communication Networks

[38] "Video Dialtone Yearly Growth", Fiber Optics & Communications New, May 1994, pg. 14.

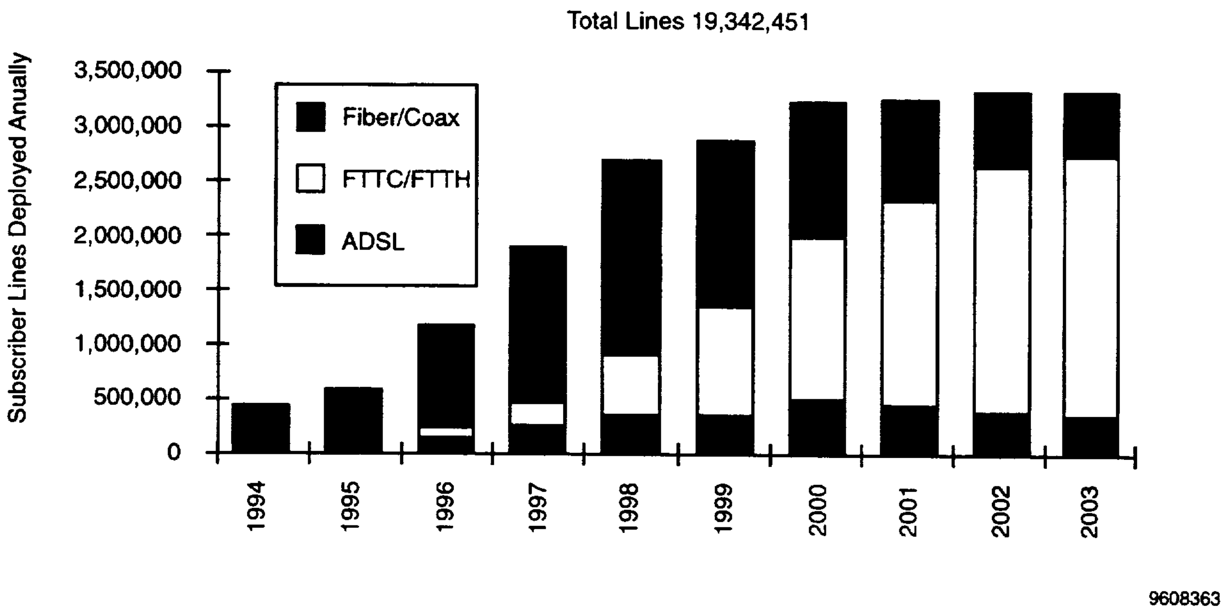


Figure 1-22. Video Dialtone Yearly Growth

In contrast to the terrestrial links, wireless services can provide considerable savings in cost, especially where population in sparse. Wireless systems are one of the fastest growing segments of the telecommunications market. For example, as shown in Figure 1-23,[39] U.S. revenue from wireless data market was close to a $1 billion for 1995. By 2005, it will have exceeded $4 billion. Table 1-12[40] gives a breakdown of the wireless market into the following categories: PCS, Cellular, Satellite, Narrowband paging, Dedicated data, and SMR/ESMR.

---

[39] "Projected US Wireless data market", LAN Magazine, Dec. 1994, pg. 44.

[40] Wireless Technology & the NII, US Office of Technology Assessment, Table 1-2, July 1995, pg. 15.

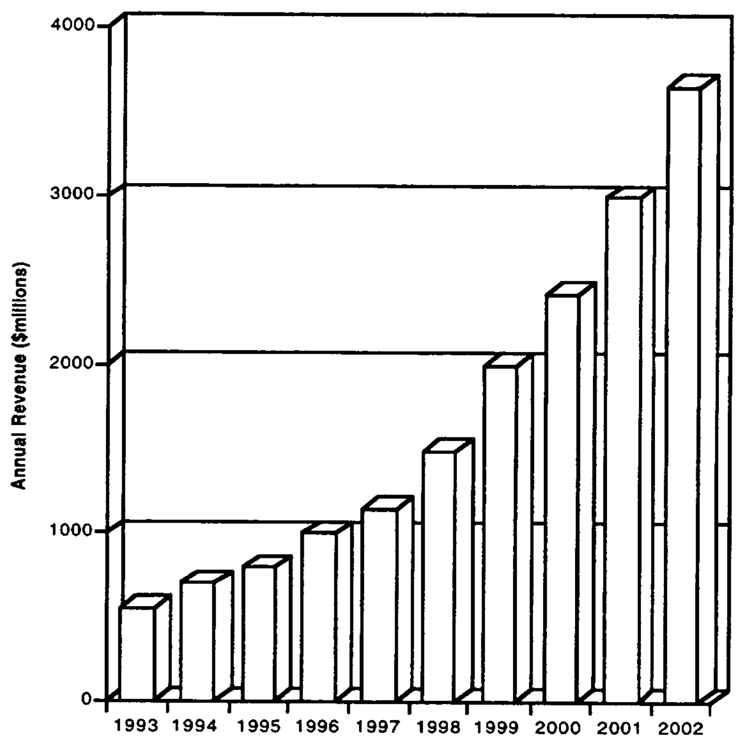


Figure 1-23. Projected U.S. Wireless Data Market

Table 1-13. Wireless Technologies Subscription Forecast

Service	1994		2000		2005	
	Subscriber (millions	Penetratio (percent)	Subscriber (millions	Penetratio (percent)	Subscriber (millions	Penetratio (percent)
New PCS	-	-	14.8	5.4	39.4	13.1
Satellite	0.1	0.0	1.3	0.5	4.1	1.4
Narrowband Paging	24.5	9.0	56.2	20.4	92.2	30.7
Dedicated Data	0.5	0.2	3.4	1.2	5.7	1.9
Cellular	23.0	9.0	46.9	17.0	65.4	21.8
SMR/ESMR*	1.5	0.6	5.2	1.9	9.0	3.0
Total	34.1	13.4	79.7	28.9	136.3	45.4
Total Voice Services	14.6	5.7	48.2	17.5	96.5	32.1

Note: the following U.S. population figures were used: 1994- 255 million. 2000- 275.8 million and 2005- 300.3 million.
* SMR/ESMR = Specialized Mobile Radio/ Enhanced Specialized Mobile Radio.

SOURCES: Personal Communications Industry Association, "1994 PCS Market Demand Forecast", (Washington, DC, Personal Comm. Industry Association, Jan. 1995), Personal Communications Industry Association, 1995 PCS Technologies Market Demand Forecast Update, 1994-2005", (Washington, DC, Personal Comm. Industry Association, Jan. 1995)

The satellite industry, despite its inherent problems with limited transponder capacity and high latency, is in fairly healthy condition. The satellite industry is estimated to bring in revenues of over $15 billion (this includes sale of satellite, launch vehicles, earth terminals and other related equipment). This industry is expected to grow at a healthy rate to double this value in a decade. Figure 1-24[41] is an evaluation of the supply and demand for transponders in the United States, Western Europe and Asia/Africa. As can be seen, there is tremendous potential for growth in Europe, Asia and Africa. Figure 1-25[42] highlights the Asia-Pacific Market for civil geostationary communications satellites from 1996 to 1999.

1.7 CONCLUSIONS

Satellites will play a major role in the national and global information infrastructure and will play at certain times a complementary role and at other times a competitive role with respect to terrestrial communication systems. As can be seen through the extensive market forecasts provided in this report, NII/GII applications are expected to create a tremendous demand for bandwidth-intensive communication links. Future global links will consist of digitized voice, data and video transmitted in packetized cells and transmitted over BISDN and/or high data rate ATM networks. To operate seamlessly with the heterogeneous terrestrial networks, satellites will have to maintain the same quality of service as optical fibers and be able to carry high Synchronous Optical Network (SONET)-rate data signals, in a reliable and clean fashion. Innovative techniques will have to be used to circumvent some of the common problems that plague satellite systems. The fact that there is already a big scramble for the FCC filing of future satellite communication links for Ka-band is a clear indication that there is an awareness in the industry of the tremendous market potential for satellite communication links.

---

[41] Supply & Demand for Transponders, Via Satellite, Jan. 1995, pg. 24.

[42] Satellite Communications, Jan. 1996, pg. 27. (Source: Euroconsult).

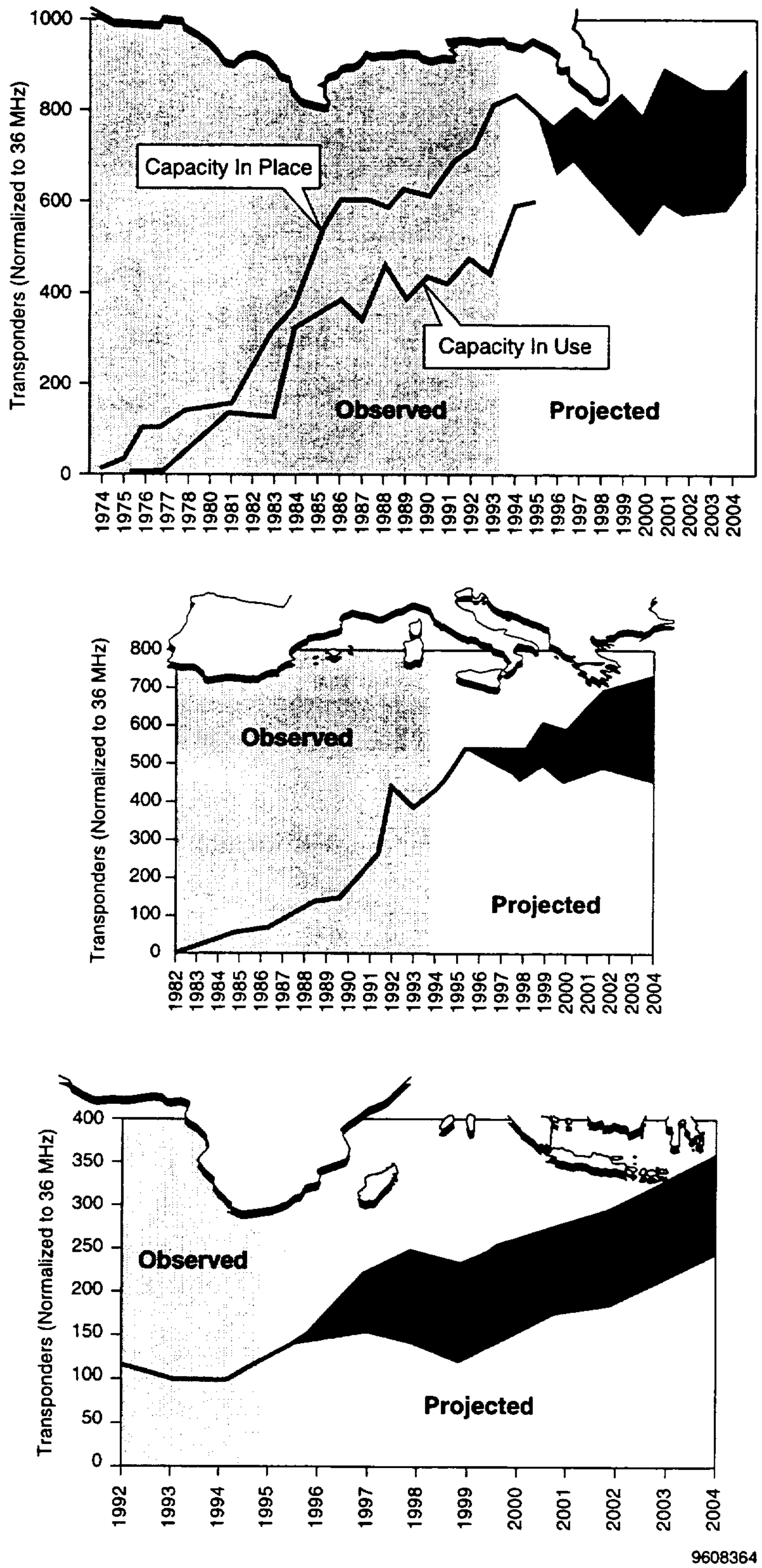
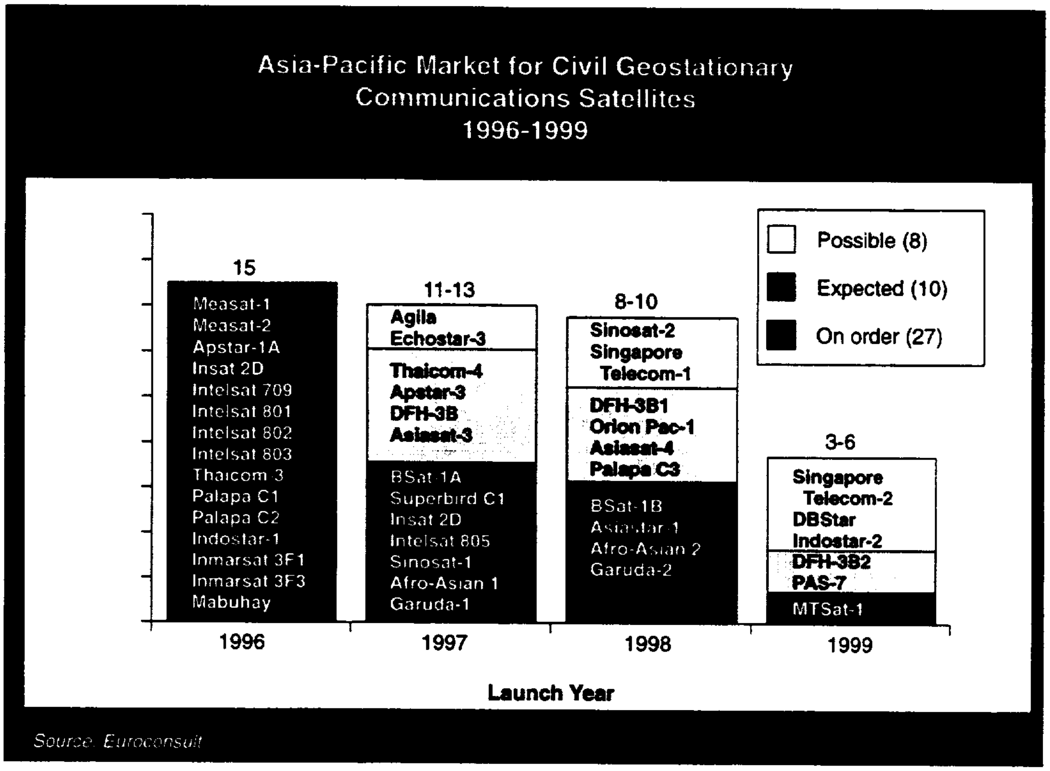


Figure 1-24. Supply and Demand for Transponders in USA, Western Europe and Africa/Asia



9609012

Figure 1-25. Asia-Pacific Market for Civil Geostationary Communications Satellites

SECTION 2 – NETWORK ANALYSIS

2.1 INTRODUCTION

In Section 1, the role of satellites in the national and global information infrastructure and the market for satellite applications were discussed. In this section, the impact of satellite system parameters on end-to-end user services in an interoperable wired and wireless telecommunications infrastructure is evaluated. Specifically, satellite delay, delay variations, BERs, cell loss ratios, transport protocols, traffic volume, and user terminal processing requirements for the proposed Low-Earth Orbit (LEO), Medium-Earth Orbit (MEO) and Geostationary Orbit (GEO) satellite architectures are analyzed. Issues of interoperability of these satellite architectures with terrestrial-based telecommunication networks are discussed and the barriers to interoperability are identified. These analyses will help us to prioritize LEO, MEO and GEO satellite network systems.

Satellite communication systems are expected to play a multifold role in the evolving telecommunications networks and vastly strengthen the capabilities of the NII and GII by reaching remote, inaccessible areas, providing connectivity in areas where the terrestrial infrastructure has been damaged, providing cost effective broadcast/multicast services and providing global connectivity anytime, anywhere. Satellite communications networks, if designed properly, can also be made to operate seamlessly with terrestrial systems. In fact, the key feature of the new networks, i.e., ATM virtual networks, is to provide multimedia services and variable bandwidth when needed; characteristics that are closely matched with the unique strengths of satellite communications. Satellite communications has global reach and can provide flexible Bandwidth-on-Demand (BOD) capability through the use of agile spot beams. These features, coupled with the multipoint/broadcast nature of satellite links can pave an important role for satellite communication systems in the NII/GII.

However, the national and international standards for the ATM layer (as well as other high-speed network systems such as frame relay and SONET) are being specified taking into account the characteristics of optical fiber networks with their low delay and low random bit errors. Unfortunately, satellite systems, for the most part, suffer from burst errors (with the use of channel coding) and large propagation delays. The effect of burst errors causes loss of frame and synchronization at the physical layer, loss of ATM cells at the ATM layer, and undetected errors at the ATM Adaptation Layer (AAL) layers. Similarly, satellite propagation delay causes an adverse effect on congestion control. These, and related problems, will be analyzed in subsequent sections pertaining to future LEO, MEO and GEO orbit satellite networks.

2.2 BROADBAND SERVICE PERFORMANCE REQUIREMENTS

2.2.1 NII/GII Broadband Services

Future NII/GII applications will require the use of a number of services identified earlier under subsection 1.2.1. These services vary greatly in their bit rate requirements, burstiness/information density, holding times, and sensitivity to delay or loss of packet. These terms are defined as follows:

1. *Information density* is defined as the ratio of the average information source rate to the maximum rate. Sometimes the term burstiness is also used to describe broadband service types. Burstiness is the reciprocal of information density. Interactive data tends to be very bursty as compared to voice and video services

2. *Holding time* is defined as the time period during which a circuit is kept busy by a service, i.e., the mean service call duration

3. *Performance criteria* can be either delay sensitive, error sensitive, or both. Services that have human elements at the end, such as voice, are generally sensitive to end-to-end delay as well as delay variations in a cell-based switching network. However, voice services tend to be very tolerant to errors. On the other hand services such as data transfer are very sensitive to bit errors but not too sensitive to delay. Some services, such as interactive video, are sensitive to both delay (because of the human element) and bit errors (depending upon the degree to which compression has been used, the more compressed the greater is its sensitivity to bit error).

The various services, described earlier in subsection 1.2.1, that drive the NII/GII applications of the future are summarized in Tables 2-1 through 2-5 in terms of their data rates, information density, holding times and performance criteria. As seen in this table, services vary greatly in their bit rates (from a few b/s to hundreds of Mb/s), holding times (from a few seconds to hours), and information density (from <0.1 to 1). In addition, there is considerable difference in the delay and error sensitivity of these services. All these varieties of services will need to be transported seamlessly over the future broadband network, both over satellite networks as well as terrestrial networks.

Table 2-1. Characteristics of Sample Conversational Services [1,2]

Sample Services	Bit Rate (Bit/S)	Information Density	Holding Time	Performance Criteria
Alarm	10 - 100	-	-	-
Telemetry	100 - 1K	low - 0.01	short	-
Security Monitoring	100 - 1K	low - 0.05	short	-
Electronic Fund Transfer	300 - high	low - 1.0	short, medium	error rate
Inquiry/ response	300 - 9.6K	< 0.1	short	delay
Voice	3.6K - 64K	0.4 - 0.5	medium	delay
Video telephony	10K - 20K	0.4 - 1.0	medium	delay
High Fidelity Audio	128K - 384K	0.5 - 0.8	medium	delay/ error rate
Facsimile	1.2K - high	0.8 - 1.0	medium, large	error rate
Color facsimile	2M	0.8 - 1.0	medium, large	error rate
Bulk data transfer	1.2K - high	0.8 - 1.0	medium, large	error rate
CAD/CAM	10K - 100M	0.8 - 1.0	medium, large	error rate
File transfer	10K - 50M	0.8 - 1.0	medium, large	error rate
LAN interconnection	1M - 100M	low - 1.0	medium, large	error rate/ delay
High-speed LAN	600M - 1G	low - 1.0	medium, large	error rate/ delay
MAN	155M - 600M	low - 1.0	medium, large	error rate/ delay
Video transmission	30M - 45M	low - 1.0	medium, large	delay
High quality Video conf.	100K- 500K	low - 1.0	medium, large	delay

Table 2-2. Characteristics of Sample Messaging Services [1,3]

Sample Services	Bit Rate (Bit/S)	Information Density	Holding Time	Performance Criteria
Electronic Mail	300 - 1M	low - 1.0	short, medium	-
Voice Mail	64K	0.4 - 0.5	medium	delay
Image Mail	64K	0.8 - 1.0	medium, large	delay/ error rate

---

1 J.C. Kohli et. al., "Emerging Broadband Packet-Switch Technology in Integrated Information Networks," IEEE Network, pp. 37-51, Nov. 1988.

2 N. Jayant, BD. Ackland et. al., "Multimedia: Technology Dimensions and Challenges," AT&T Technical Journal, Sept./Oct. 1995, pp. 14-33.

3 D.J. Wright and M. To, "Telecommunication Applications of the 1990s and their Transport Requirements," IEEE Network Magazine, pp. 34-40, March 1990.

Table 2-3. Characteristics of Sample Retrieval Services [2,3]

Sample Services	Bit Rate (Bit/S)	Information Density	Holding Time	Performance Criteria
Audio (CD quality)	128K - 384K	0.8 - 1.0	medium, large	delay/ error rate
Video information retrieval	30M - 45M	low - 1.0	medium, large	delay
High-resolution X-ray film	50M - 100M	low - 1.0	medium, large	delay/ error rate
Entertainment video	50M - 800M	0.8 - 1.0	medium, large	delay

Table 2-4. Characteristics of Sample Distribution Services without User Individual Presentation Control [2]

Sample Services	Bit Rate (Bit/S)	Information Density	Holding Time	Performance Criteria
Video distribution	1M - 8M	low - 1.0	medium, large	-
NTSC, PAL, SECAM	1M - 5M	low - 1.0	medium, large	-
EQTV	10M - 15M	low - 1.0	medium, large	error rate
HDTV	17M - 32M	low - 1.0	medium, large	error rate

Table 2-5. Characteristics of Sample Distribution Services with User Individual Presentation Control [3]

Sample Services	Bit Rate (Bit/S)	Information Density	Holding Time	Performance Criteria
Cable Image	150M	low	short	delay
Cable text	2M	low	short	delay

Note: (1) Holding time is the mean service call duration: short - seconds to minute, medium - 1 minute to 30 minutes, and large - 30 minutes to hours.

(2) Cable text uses a full digital broadband TV channel for cyclical transmission of text, images, audio and video.

2.2.2 ATM Broadband Services

ATM, as mentioned earlier, has become a *defacto* standard for broadband communication services and is believed to remain so for the forseeable future. Thus, it is important that the future satellite systems adhere as closely as possible to ATM standards to facilitate interoperability between terrestrial and satellite systems. The ATM standard has broadly defined four QoS classes[4] :

4 Shirish S. Sathaye, "ATM Forum Traffic Management Specification Version 4.0", ATM Forum/95-0013R10.

1. Constant Bit Rate (CBR) services are used by connections that request a static amount of bandwidth that is continuously available during the connection's lifetime. Example include CBR applications:

   - Interactive Video (e.g., video-conferencing)
   - Interactive Audio (e.g., telephone)
   - Video Distribution (e.g., television, distributed classroom)
   - Audio Distribution (e.g., radio, audio feed)
   - Video Retrieval (e.g., video on demand)
   - Audio Retrieval (e.g., audio library)
   - Data/text/image transfer

2. Variable Bit Rate (VBR) services which can be subdivided into:

   - Real-time VBR (rt-VBR) services intended for real-time applications (i.e., those which require tight constraints on delay and delay variations, as would be appropriate for voice and video applications)
   - Non-real-time VBR (nrt-VBR) intended for non-real-time applications which have bursty traffic characteristics. Examples include: transaction processing (such as airlines reservations, banking transactions, process monitoring) and frame-relay interworking

3. Unspecified Bit Rate (UBR) service category is intended for Non-real-time applications (i.e., those not requiring tightly constrained delay and delay variation). Examples of such applications are:

   - Interactive Text/Data/Image Transfer (e.g., credit card verification)
   - Text/Data/Image Messaging (e.g., e-mail, telex, fax)
   - Text/Data/Image Distribution (e.g., news feed, weather satellite pictures)
   - Text/Data/Image Retrieval (e.g., file transfer, file browsing)
   - Aggregate LAN (e.g., LAN connection or emulation)
   - Remote Terminal (e.g., telecommuting, telnet)

4. Available Bit Rate (ABR) service category, for which the limiting ATM layer transfer characteristics provided by the network may change subsequent to the establishment of a connection. A flow control mechanism is specified which supports several types of feedback to control the source rate in response to changing ATM layer transfer characteristics. Examples include:

- Any UBR application listed above that can take advantage of the ABR flow-control protocol to achieve a low cell loss ratio
- Critical data transfer (e.g., defense information)
- Supercomputer applications
- Data communications applications requiring better delay behavior, such as remote procedure call, distributed file service (e.g., Network File Service) or computer process swap/paging.

These categories are listed in Table 2-6 along with some of their key attributes.

Table 2-6. ATM Service Categories and Their Key Attributes

Service Class	CBR	Rt-Vbr	Nrt-VBR	UBR	ABR
Timing End-to-End	Required	Not Required	Not Required	Not Required	Not Required
Bit Rate	Constant	Variable	Variable	Variable	Variable
Connection Mode	Connection Oriented	Connection Oriented	Connection Oriented	Connection-Less	Connection-Less
Traffic Control	No	No	Yes	No	Yes

2.2.3 ATM Quality of Service Parameters

The ATM layer QoS is measured by a set of parameters characterizing the performance of the ATM layer connection. Some key parameters of ATM networks are:

a. Maximum Cell Transfer Delay (CTD): The CTD is defined as the elapsed time between a cell exit event at measurement point 1 (MP1) (e.g., at the source UNI) and the corresponding cell entry event at measurement point 2 (MP2) (e.g., at the destination UNI) for a particular connection.

b. Cell Delay Variation (CDV): Two performance parameters associated with CDV have been defined by CCITT. The first parameter, termed as 1-point CDV, describes the variability in the pattern of cell arrival events observed at a single measurement point. The second parameter, termed 2-point CDV, describes the variability in the pattern of the cell arrival events observed at the output of a connection portion (MP2) with reference to the pattern of the corresponding events observed at the input to the connection portion (MP1). The 2-point CDV includes only the delay variability introduced within the connection portion, unlike the 1-point CDV which also includes the variability present at the cell source (i.e., the customer equipment).

c. Cell Loss Ratio (CLR): The CLR of a connection is defined as the ratio of lost cells to the total number of cells that were transmitted. Lost and transmitted cells counted in severely errored cell blocks should be excluded from the cell population in computing CLR.

d. Cell Error Ratio (CER): The CER of a connection is defined as the ratio of errored cells to the sum of the successfully transmitted cells as well as the errored cells. Successfully transferred cells and errored cells contained in cell blocks counted as severely errored cell blocks should be excluded from the population used in calculating CER.

e. Cell Misinsertion Rate (CMR): The CMR of a connection is defined as the total number of misinserted cells observed during a specific time period divided by the time duration interval. Misinserted cells and time intervals associated with cell blocks counted as severely errored block cells are excluded from the calculation of CMR.

f. Severely Errored Cell Block Ratio (SECBR): The SECBR of a connection is defined as the ratio of severely errored cell blocks to the total number of transmitted cell blocks. A cell block is a sequence of N cells transmitted consecutively on a given connection. A severely errored cell block outcome occurs when more than M errored, lost or misinserted cells are observed in a received cell block. The value of N is uniquely determined by the Peak Cell Rate (PCR) of the aggregate flow, i.e.,

$$N = \frac{PCR}{25} \text{ (with N rounded to the next larger power of 2) and}$$

$$M = \frac{N}{32}$$

At the time of call set-up, the following parameters are negotiated: peak-to-peak CDV, maximum CTD, and CLR. Default values are used for other parameters by the network.

The ITU-T has recommended a set of bounds for the key QoS performance parameters. These objectives are given in Table 2-7. For satellite hops, the ITU has assigned a fixed delay allocation of 260 + ∂ ms, independent of the actual path length or processing complexity. The value of ∂ is unspecified but assumed to be less than 60 ms. The satellite link should have a 2-point CDV of less than 1.5 ms to maintain an end-to-end 2-point. CDV of less than 3 ms. In terms of the various block error parameters, satellites have been assigned block allowances of 42 percent for national portions and 35 percent for international portions. The satellite QoS objectives are shown within parenthesis in Table 2-7.

Table 2-7. Provisional QoS Class Definition and Network Performance Objectives[5]

	CTD	2-pt. CDV	CLR	CER	CMR	SECBR
Nature of the network performance objective	Upper bound on the mean CTD	Upper bound on the diff. betn. upper and lower 10^{-8} quantities of CTD	Upper bound on the cell loss probability	Upper bound on the cell error probability	Upper bound on the mean CMR	Upper bound on the SECBR probability
Default objectives	No default	No default	No default	$4*10^{-7}$	1/day	10^{-4}

QoS Classes:

Class 1 (stringent class) CBR, VBR, ABT	400 ms (260+∂) ms	3 ms (1.5 ms)	$3*10^{-7}$ ($\sim 1*10^{-7}$)	$4*10^{-7}$ ($\sim 1.4*10^{-7}$)	1/day	10^{-4} ($\sim 3.5*10^{-5}$)
Class 2 (tolerant class) CBR, VBR, ABT	U	U	10^{-5} / 10^{-6} ($\sim 3.5*10^{-7}$)	$4*10^{-7}$ ($\sim 1.4*10^{-7}$)	1/day	10^{-4} ($\sim 3.5*10^{-5}$)
Class 3 (bi-level class)VBR, ABT	U	U	U	$4*10^{-7}$ ($\sim 1.4*10^{-7}$)	1/day	10^{-4} ($\sim 3.5*10^{-5}$)
U class	U	U	U	$4*10^{-7}$ ($\sim 1.4*10^{-7}$)	1/day	U

Note: U = "Unspecified" and δ= 60 ms.

2.3 IMPACT OF SATELLITE NETWORKS SYSTEMS ON BROADBAND QOS PARAMETERS

There has been considerable analysis, in recent years, of multi-protocol and multi-service broadband systems (such as ATM) for driving NII/GII applications. However, most of these studies assume that the bit-way that will be used for the NII/GII is high-bandwidth optical fibers. While optical fibers (with their inherent high bandwidth, low loss and excellent BER performance) are well-suited to carry broadband services, they are by no means the only choice or even the preferred choice. As mentioned in subsection 2.1, the costs of laying a nationwide terrestrial optical network are staggering. The problem of creating a worldwide broadband fiber infrastructure is magnified further by the minimal terrestrial infrastructure existing in most parts of the world. Satellite systems are expected to lead the way in providing the type of broadband services that are needed in the NII/GII.

5 B-ISDN ATM Layer Cell Transfer Performance," ITU-T Recommendations I.356, Draft 4R, 02/96, pp. 21, 27.

Satellite system designs have, in recent years, seen a shift from the traditional geosynchronous orbit system design used for broadcasting, to systems involving constellations of satellites in the LEO, MEO and GEO which can provide global personal communication and broadband services. There are many considerations that must be taken into account before deciding which satellite network architecture is most suitable for NII/GII applications. These include:

a. Cost of the system (dependent upon the number and complexity of the satellites; number of launch services required; number, size and complexity of the earth stations; cost of the switching center; satellite replacement cost; and the cost of system integration)

b. Availability of service (dependent upon satellite coverage, elevation angles, frequency of operation, and smoothness of handover between satellites)

c. Quality of the service provided (dependent upon satellite delay, delay variations, BER, cell loss ratio, transport protocols used, beam/satellite handover, traffic volume, and user terminal processing).

The cost and availability of service parameters have been explored in great detail in a number of publications[6,7], and are not included here. In this report we will concentrate mainly on the impact of the QoS parameters.

LEO systems consist of satellites operating below the Van Allen Belt (500 to 1500 km) with the number of satellites varying from 36 to 66[8]. Some proposed LEO systems employ a large number of satellites to provide a high minimum elevation angle requirements, e.g. 840 satellites with 40° minimum elevation angle for the Teledesic system. Each satellite is generally simple in architecture, small in size and low in cost, but, because of the high rate of fuel consumption needed to overcome atmospheric drag, their lifetimes are generally short (3 to 8 years). LEO systems require complicated hand-off mechanisms to route a call through the myriad of rapidly moving satellites, within the short period of time (a couple of minutes) for which the LEO satellite remains in view of the end-user. MEO systems consist of 10 to 12 satellites which orbit at an altitude of 10,000 to 16,000 km (between the Van Allen electron and proton belts) 8. Because of their relatively high altitude, MEO satellites have less stringent elevation angle requirements (15° for Ka-band) than LEO systems and will typically be in sight of a user for up to 90 minutes. Satellites in MEO should also last longer (8 to 10 years) because they experience much less drag than do LEO satellites. GEO systems, orbiting at an altitude of 35,786 km., require the least number of

---

6 R. Busch & P. Cress, "Selecting the Best Constellation for Mobile Satellite Services", AIAA-96-1068, pg. 680-687.

7 K.G. Johannsen, "Mobile P-Service Satellite System Comparison," International Journal of Satellite Communications, Vol. 13, pg. 453-471, 1995.

8 R.A. Nelson, "Satellite Constellation Geometry,' Via Satellite, pp. 110-122.

satellites (3 to 4) for global coverage with each satellite expected to have a long lifetime (10 to 15 years). However, each GEO satellite would have to be larger and more complex than those at LEO or MEO in order to support a larger capacity and the large antennas needed to communicate with the small user terminals. The following subsections present evaluations of the LEO, MEO and GEO systems in meeting critical interoperability requirements, such as delay, delay variation, transmission errors, transport protocols, beam/satellite handovers, traffic and congestion controls and user terminal complexity.

2.3.1 Reference System Architectures

To evaluate the QoS parameters for LEO, MEO and GEO systems, a representative system for each architectures has been chosen. The evaluated QoS parameters can be taken to be representative of each architecture.

The Teledesic system, consisting of 840 satellites in 21 orbital planes (20 satellites per plane) is chosen as the reference for the LEO system. The altitude for each satellite is assumed to be 700 km. The ICO system, consisting of 10 satellites in two 45° orbital planes (each with 5 satellites) is used as the baseline for the MEO system. The ICO satellite altitude is 10,355 km. For the GEO system, a constellation of four satellites (at 95°´, 180°E, 105°W and 20°W) at a geosynchronous altitude of 35,786 km. is assumed. The use of Inter-Satellite Links (ISLs) between the four satellites provides a full global coverage. The footprint for this system (Figure 2-1) is shown on the next page, for elevation angles of 5, 10 and 15 degrees.

We have assumed Ka-band operation for all these systems because the Ka-band has large bandwidth to support multimedia services and its space segment and user terminals are more economical. Compared to C- and Ku-bands, Ka-band signals are more susceptible to rain fade and atmospheric anomalies. As a consequence, adaptive fade compensation and higher elevation angles (40° for LEO case) are used.

Elevation Angles: 5°, 10°, and 15°

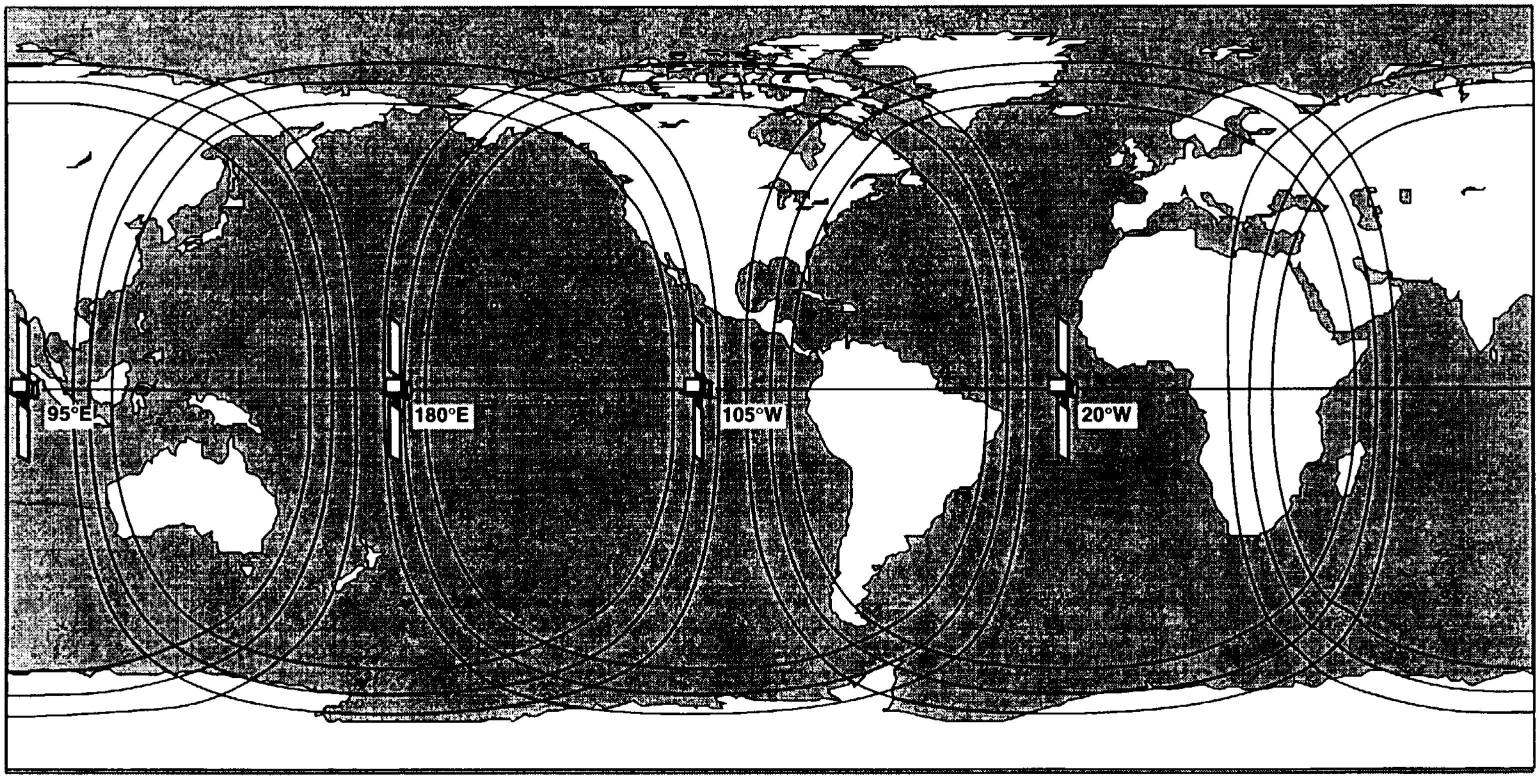


Figure 2-1. Footprint for Sample GEO Architecture

2.3.2 Satellite Delay

Delay and delay variation (jitter) play a very important role in determining the QoS possible for the broadband services used in the NII/GII. Unlike terrestrial networks, where propagation delay normally accounts for only a small fraction of the total end-to-end delay of a system, for satellite systems it can form the most significant portion of the delay. An attempt will be made here to characterize the delay of "typical" LEO, MEO and GEO systems for systems designed to operate at Ka-band.

The geometry used in calculating the distance of the satellite from a fixed point on Earth is given in Figure 2-2. In this figure, the following terms are used:

h is the altitude of the satellite

R_e the radius of Earth (6378 km)

α is the angle subtended at the satellite between the nadir and the user

β is the angle subtended at the center of the earth between the nadir and the user

θ is the elevation angle of the line-of-sight to the satellite

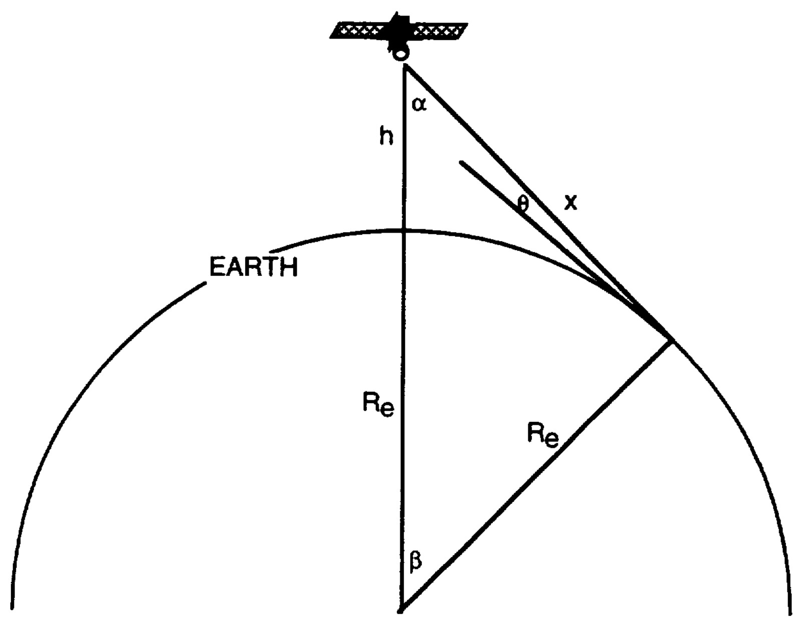


Figure 2-2. Geometry Used to Calculate Satellite Orbit

The value of x can be calculated as a function of Re, h and θ using the following relation:

$$\frac{Sin(\theta + 90°)}{R_e + h} = \frac{Sin(\alpha)}{R_e} = \frac{Sin(\beta)}{x} \qquad \text{E2-1}$$

Using the above relation it can be shown that the value of x is[9] :

$$x = R_e\left\{\gamma\sqrt{1-\left(\frac{Cos(\theta)}{\gamma}\right)^2} - Sin(\theta)\right\} \qquad \text{E2-2}$$

where:

$$\gamma = 1 + \frac{h}{R_e}$$

LEO and MEO System

For both the LEO and MEO systems, the distances and propagation delays between nearest neighbor satellites in adjacent orbits and next nearest neighbor satellites in the same orbit are determined. These are used to estimate the propagation delay and its variation for typical communication routes.

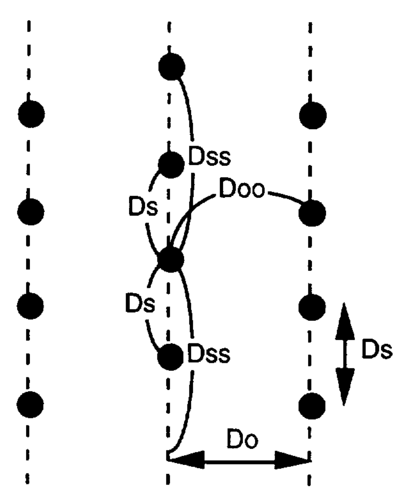


Figure 2-3. Possible ISL Paths and Their Nomenclature

9 C. Michel, D. Rouffet, "Progressive Set-Up of LEI Inclined Constellations,", AIAA-94-1055-CP, pg. 1683 (addendum).

The distances shown in Figure 2-3 are defined as follows:

Ds is the distance between a satellite and its nearest neighbor in the same orbital plane,

D*ss* is the distance to the next satellite in the same orbital plane

Doo is the distance to the nearest satellite in the adjacent plane

Do is the distance between orbits.

These distances can be evaluated using the following formulae: [10]

$$Do = 2(R_e + h)(Sin(\phi / 2))Cos(L)$$
$$Ds = 2(R_e + h)(Sin(v / 2))$$
$$Dss = 2(R_e + h)(Sin(v))$$
$$Doo = \sqrt{\left\{\left(\frac{Ds}{2}\right)^2 + (Do)^2\right\}} \qquad \text{E2-3}$$

where: ϕ is the separation between the orbital planes

L is the Latitude and

v is the separation (in degrees) between satellites within a plane

The system we have used as an example for calculating propagation delay for LEO systems is the Teledesic system which consists of 840 satellites in 21 planes (40 satellites per plane). The angular separation between any two satellites within a plane is 9° (assuming the satellites are equally spaced). The distance, Ds, between two adjacent satellites within a plane is calculated as 1111 km using equation E2-3. The corresponding delay is 3.7 ms. We expect that a maximum of 20 ISLs (with connections between adjacent satellites) will connect any point on the earth to its geographically opposite point. By skipping intermediate satellites in the cross-links (i.e., using Dss) or by connecting with satellites in other planes (i.e., using D*oo*) it should be possible to reduce the maximum number of ISLs to 10.

A similar analysis has been used for calculating the propagation delay of the MEO system, using the ICO system as our baseline. The ICO system consists of ten satellites in two planes (5 each) at 45° two with additional spare satellites, one in each plane. The angular separation between any two satellites within a plane is 72° (assuming the satellites are equally spaced). The distance, Ds, between two adjacent satellites within a plane can be

10 S.J. Campanella, "Transmission over Dynamic LEO Satellite Links", AIAA-96-1042-CP, pp. 458-466.

calculated as 19,671 km using equation E2-3. The corresponding delay is 66 ms. We expect a maximum of three ISLs will be needed to connect to geographically opposite locations on Earth. With links between the planes, the number of inter-satellite links needed could be reduced.

The system parameters and resulting delay values for LEO and MEO systems are shown in Table 2-8.

Table 2-8. Minimum and Maximum Propagation Delays for LEO and MEO Systems

Parameter	LEO (Teledesic)	MEO (ICO)
Altitude, h (km)	700	10,355
Elevation angle, θ, (degrees)	40°	15°
Maximum satellite to ground distance, x (km)	1021.5	13,907
Maximum two-way sat. to ground delay (ms)	6.8	92.7
Number of planes	21	2
Number of satellites per plane	40	5
Separation between adjacent satellites within a plane, ν (degrees)	9°	72°
Distance between satellites in same plane, D_s (km)	1111	19,671
Crosslink delay between satellites in the same plane (ms)	3.7	66
Distance between skipped satellites in same plane, D_{ss} (km)	2214.5	31,828
Crosslink delay between skipped satellites in the same plane (ms)	7.38	106.1
Maximum number of crosslinks	20 (D_s links) or 10 (D_{ss} links)	3
Maximum delay (ms)	81 or 78	291
Minimum delay (ms)	7	93

GEO System

The baseline system chosen as a reference to evaluate GEO architecture consists of four satellites at 20°W, 95°E, 180°E and 105°W, as shown in Figure 2-4. The distances between the satellites are also given in this figure.

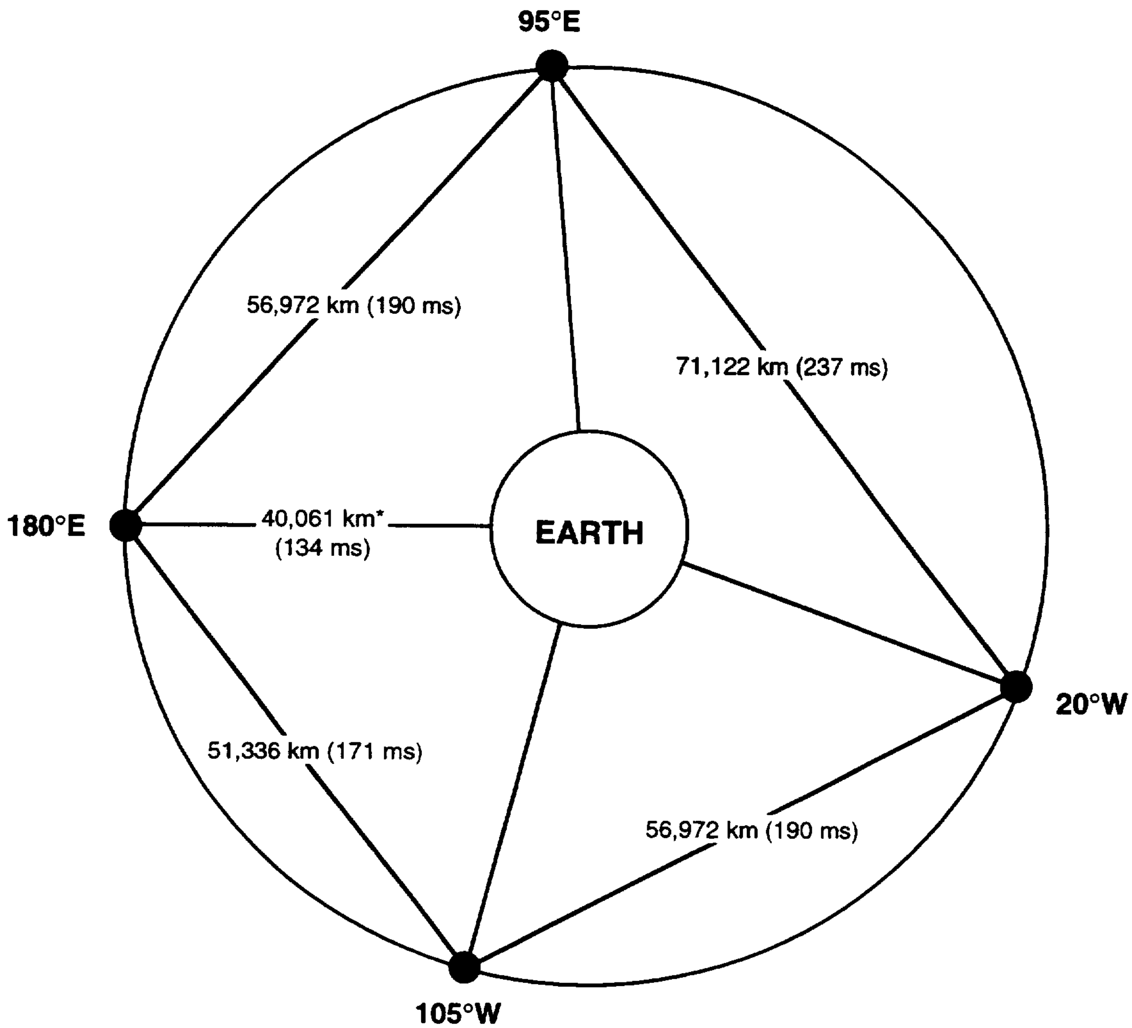


Figure 2-4. Satellite Configuration and Distances Shown for the Chosen GEO System

The maximum and minimum delays for this GEO system has been computed and is given in Table 2-9.

Table 2-9. Minimum and Maximum Propagation Delays for Geo System

	Min (ms)	Elevation Angle 0° Max (ms)	5° Max (ms)	10° Max (ms)	15° Max (ms)
Without ISLs	239	278	274	271	267
With 1 ISL	410	515	511	508	504
With 2 ISLs	600	705	701	698	694

Assuming an elevation angle of 15° the maximum delay without ISL is 267 ms. To have full connectivity one would need at most two ISLs. This increases the delay to 694 ms. There is also a great variation between the minimum and maximum possible delays (239 vs. 267 ms for no ISL, 600 vs. 694 ms for a system with 2 ISLs).

From Tables 2-8 and 2-9, the minimum and maximum delay values for the LEO, MEO and GEO system are summarized in Table 2-10.

Table 2-10. Comparison of Propagation Delays for LEO, MEO and GEO Systems

	LEO	MEO	GEO
Elevation Angle	40°	15°	15°
Maximum Number of Crosslinks	10	3	2
Maximum Delay	78	291	694
Minimum Delay (without ISL)	7	93	267

Clearly, MEO and LEO systems conform to the two-point CDV delay requirements for ATM switched systems. GEO systems can only comply with the Class 1 standards if no ISL is used.

In addition to propagation delays, one may need to add the following processing delays: delays in assembling and disassembling of Time Division Multiplexed (TDM) frames, delays arising out of Forward Error Correction (FEC) coding/interleaving and decoding, and on-board processing and switching delays. The values of these delays will have to be evaluated on a case by case basis. Ranges for these delays are given in Table 2-11.

Table 2-11. Non-Propagation Delays in Satellite Systems [11]

	Delay (ms)	
Function	**Minimum**	**Maximum**
Assemble SVPs, form TDM frames	0	2
FEC Coding, interleaving and modulation	16	80
Switching on Satellite (includes buffers)	2	200
Dwell cycle time (hopping or fixed beam)	1	2
Assemble SVPs, Form TDM frames	0	2
Demodulation, FEC decoding, deinterleaving	16	80
Disassemble TDM Frames and SVPs	0	2
Total delay (ms)	35	368

[11] K. Barker and K. M. Price, "Concept Definition for Satellite-Delivered B-ISDN", AIAA-96-1065-CP, pg. 647-654.

The total sum of all propagation and non-propagation delays will have to be evaluated in designing a specific satellite system. Delay has a marked impact not only on the QoS of the broadband services, but also on the traffic and transport protocols.

2.3.3 Satellite Delay Variation

Along with delay, delay variation is another important parameter in determining the QoS provided by satellite networks. Variations in delay can cause severe traffic problems and result in very poor quality of service. Delay variation is caused by changes in signal propagation time for a given virtual circuit connection due to satellite motion, timing errors, satellite handover, and processing at the satellite and user terminals. Doppler depends on the motion of the satellite relative to the earth's surface and varies with the distance from the nadir. Maximum Doppler occurs on the projected ground track, at the horizon. These parameters are compared for the three systems — GEO, MEO and LEO — in Table 2-12.

Table 2-12. Doppler Shift and Delay Variation as a Result of Satellite Motion

	Elevation	Maximum Doppler Shift	Maximum Doppler Shift at 20 GHz	Maximum Doppler Shift at 30 GHz	Delay Variation
LEO (Teledesic)	40°	2.3×10^{-5}	0.46 MHz	0.69 MHz	1.1 ms
MEO (ICO)	15°	5.2×10^{-6}	104 kHz	156 kHz	11.8 ms
GEO	15°	8×10^{-9}	160 Hz	240 Hz	0.24 ms

Since both Doppler and delay change significantly for the LEO and MEO systems, procedures must be devised to maintain timing and frequency alignment of the traffic channels within a spot beam with the timing and carrier frequency at the satellite. Doppler compensation may be used by the earth stations to correct the frequency of the carrier. Similarly, path delay compensation may be introduced by the earth stations to ensure correct timing to within the 1.5 ms specified for ATM signals. This can be achieved using buffers. Buffers size must be large enough to absorb delay variations caused by other impairments mentioned above. The effect of satellite handovers is discussed in detail in subsection 2.3.6.

2.3.4 Transmission Error Characteristics

Currently, high-data rate services using fast-packet switching technologies (e.g., ATM) and broadband transmission protocols (e.g., SONET) are designed assuming low error rate transmission, such as that available from optical fiber systems. Broadband fiber networks are designed, especially at high data rates, to operate at BERs of 10^{-10} or better. Even better performances than 10^{-10} may be needed for traffic consisting of highly compressed data.

Satellite services will have to be designed to match this level of performance for high-speed networking. This section describes alternate techniques for achieving the performance goals shown in Table 2-7.

Satellite links are inherently noisier than fiber links and also suffer from problems such as rain fade, which become more acute at higher frequencies such as Ka-band. Typically, a good satellite link today operates at a BER of 10^{-6} or 10^{-7}. The standard method to improve the BER performance of digital satellite communications systems is to use various error-correction and coding, techniques (e.g., convolutional coding, concatenated coding and trellis coded modulation schemes). With the use of such coding schemes, one can achieve BERs of better than 10^{-10} or 10^{-11}. However, coded systems alter the generally random nature of bit errors into bit errors occurring in bursts. Bursty errors occur because any error made in the decision of a coded data sequence affects a large number of bits (generally 10 to 40, depending on the code). The bursty nature of errors in such satellite links introduces QoS degradation in broadband integrated digital service networks such as ATM. It affects both the ATM layer and AAL. The QoS parameters that are affected are: CLR, CMR, CER and SECBR.

Effect of Error Bursts on ATM Protocols

The ATM layer is concerned with data transmission between two adjacent network nodes, usually not the end points. The ATM cell consists of 53 bytes, of which 48 contain data and 5 contain overhead. The ATM cell header consists of subfields related to routing information (Virtual Path Identifier [VPI] and Virtual Circuit Identifier [VCI]), the indication of Payload Type (PT), the Cell Loss Priority (CLP), and the Header Error Control (HEC). The 8-bit HEC is used to protect data in all these subfields and to delineate the ATM cells and achieve synchronization at the receiver end. The check bits in the HEC are capable of detecting one or more errors in the cell header, but are capable of correcting only one-bit errors. While the single-bit-error correction is capable of correcting most errors found in fiber links, given the random distribution of errors for fiber links, it is inadequate for the bursty type of errors encountered in satellite links. The HEC is incapable of correcting multiple errors and discards these cells. The ATM cell discard ratio for satellite systems is, therefore, orders of magnitude higher than over links with random errors.

A satellite system, which exhibits burst errors in the transmission link, can still use the 1-bit error correction capability of the ATM HEC by spreading the bits of the error burst amongst many cell headers so that only a 1-bit error is experienced in the ATM cell header. This procedure of spreading the bits in the error burst amongst different cell headers is called bit interleaving. Comsat Laboratories has analyzed and experimented with a 45 Mb/s Intermediate Data Rate (IDR) satellite link with and without the use of bit interleaving. The computed results are reproduced in Figure 2-5[12]. Clearly, the ATM cell discard probability is reduced by orders of magnitude by the use of bit-interleaving.

Effect of Error Bursts on AAL Protocols

The AAL of the ATM protocol adapts higher level data into formats compatible with ATM requirements. This layer is dependent upon the higher layer services being transported, and several different AALs have been defined for services such as data-only transport, voice, video and others. The AAL is an end-to-end process, used only by the two communicating entities to insert and remove data from the ATM layer. This greatly simplifies the processing requirements in an ATM network, in that the switches route independent of the AAL protocol of a connection.

There are four classes of AAL protocols that are currently being standardized, i.e., AAL Type 1 (CBR services), AAL Type 3/4 and AAL Type 5 (VBR services). Each service employs different Cyclic Redundancy Checks (CRCs) on the Segmentation and Reassembly (SAR) header. While all the services have error detection capability, only some have error correction capabilities. Table 2-13 summarizes the impact of burst errors on each service type. As can be seen from this table, Type 5 services which makes use of a 32-bit CRC error checking, should be able to detect burst errors of length 32 or less, making them fairly impervious to the type of burst errors arising from satellite coding. On the other hand, Type 1 services, which employ only 3-bit CRC, are ineffective against burst errors and can result in loss of framing.

---

12 D.M. Chitre et. al., "Asynchronous Transfer Mode (ATM) Operation via Satellite: Issues, Challenges and Resolutions," International Journal of Satellite Communications, Vol. 12, 1994, pp. 211-222.

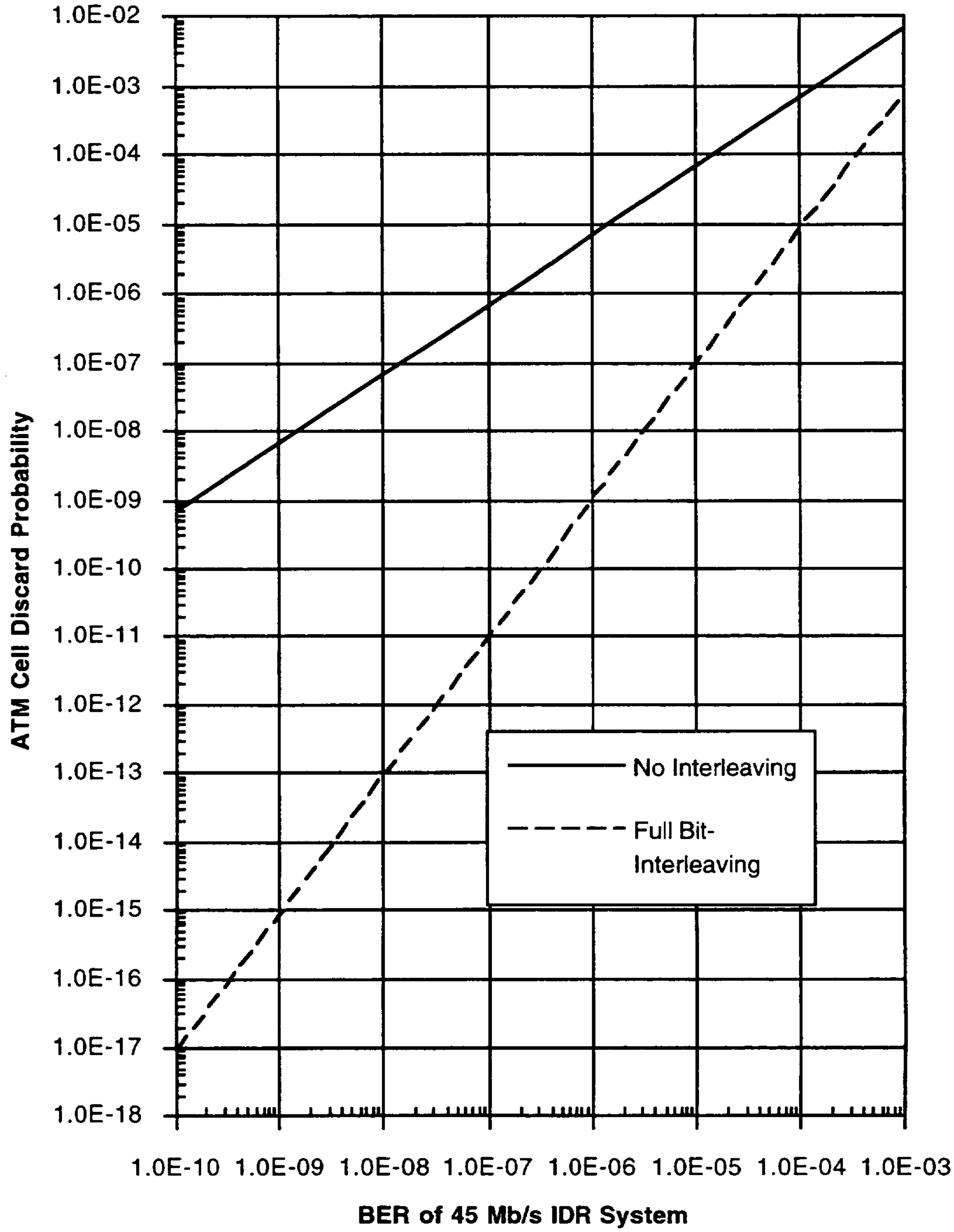


Figure 2-5. Computed ATM Cell Discard Probability for a 45 Mb/s IDR System With and Without Bit-Interleaving

Table 2-13. Burst-Error Effects On Various AAL Service Types

Defined AAL Type	Service Type	CRC-Type	Error Detection/ Correction Capability	Result of Burst Errors
Type 1	Constant bit rate (e.g., voice and video)	3-bit CRC	Both	Possible loss of frame synchronization.
Type 3	Variable source bit rate-timing sensitive (e.g., packet video)	10-bit CRC	Only error detection	Detects all single bursts of length 10 or less and burst of double and triple errors.
Type 4	Variable source bit rate-connection oriented (e.g., file transfer)	10-bit CRC	Only error detection	Detects all single bursts of length 10 or less and burst of double and triple errors.
Type 5	Same as layer 3/4 but with no multiplexing	32-bit CRC	NA	Expected to be impervious to most burst errors.

Bit-interleaving (of the 48-byte cell payload) can also be used to improve the CLR and CER for those AAL services that have error correction capabilities. If the AAL has only error detection capability, interleaving over multiple cell payloads may slightly degrade the CER performance by spreading burst errors. However, this is offset by the orders of magnitude improvement in the CLR.

In general, bit interleaving can be classified into two categories: intra-cell and inter-cell bit interleaving. In either case, the interleaving size or the interleaving depth (i.e., the number of cells being interleaved) should be large enough to successfully correct burst errors; with the objective of having at most 1-bit error in the cell header. To perform bit-interleaving, temporary storage of cells is necessary, which results in cell queuing delay. The interleaving depth (number of cells) is constrained by the cell transfer delay requirement. To reduce the interleaving delay, a high information transmission rate should be used; hence, it is advantageous to perform bit-interleaving after the cell streams are multiplexed.

Although bit-interleaving alleviates burst error problems, it might create new ones if a periodic noise source is encountered in the transmission links. To avoid the occurrence of new error bursts resulting from a periodic noise source, one can use pseudo-random interleaving (for example, by means of scrambling the data).

2.3.5 Transport Protocols

The Transmission Control Protocol (TCP) is one of the most commonly used transport protocols worldwide and is used extensively on the internet. It seems inevitable that TCP will be the most commonly used transport protocol on the NII/GII. With the rapid growth of the internet and the widespread acceptance of ATM as the technology of choice for future multimedia networks, it is important to assess how well TCP operates in an ATM network that includes a satellite link. Experiments performed using TCP/IP protocol suite over the ACTS geostationary satellite link show a very poor throughput performance[13]. There are essentially three major problems with operating TCP over satellite links, especially when encapsulated over ATM. These are:

1. Window size limit
2. Slow start and congestion avoidance procedure
3. Lack of selective acknowledgments.

The theoretical throughput of a TCP session has been shown, using queuing theory[14], to be:

$$\gamma = c\frac{1-\rho^{L}}{1-\rho^{L+1}} \qquad \text{E2-4}$$

where: c = offered load to the session (packets/s) = L/τ

τ = round trip latency between the peer TCP entities (s)

L = window size (packets)

ρ = link utilization = c/v

v = transmission rate (packets/s)

For the case where link utilization is much less than 1, the maximum throughput reduces to:

$$\gamma = \frac{L}{\tau} \qquad \text{E2-5}$$

The round trip latency, τ, is the sum of the propagation delay of the satellite link, transmission delay of ATM cells, SAR delay of ATM cells, and the processing delay within the TCP stack. The last three terms are negligible compared to the first (i.e., propagation delay of the satellite link). The maximum TCP packet size is specified to be 2^{16} bytes (i.e., 65 kbytes).

Figure 2-6 depicts the TCP throughput as a function of a round-trip delay for link utilization of 0, 0.4 and 0.8. The maximum throughput significantly decreases as the delay

---

13 H. Kruse, "Data Communications Protocol Performance on Geostationary Satellite Links. Lessons Learned Using ACTS", AIAA-96-1072-CP, pp. 722-725

14 A.A. Lazar, "The Throughput Time Delay Function of an M/M/1 Queue," IEEE Trans. on Information Theory, Vol. IT-29, No. 6, Nov. 1983.

increases. This performance degradation has a severe impact on not only satellite services but also terrestrial transmission of TCP-based traffic. For example, a coast-to-coast terrestrial TCP connection encounters about 32 ms of delay without including processing delay at switching nodes and non-ideal transmission media (transmission speed < speed of light). The theoretical throughput is about 16 and 9 Mb/s for link utilization of 0 and 0.8, respectively. Thus, a high-speed transmission link operating at 155/622 Mb/s has no use for high-throughput TCP services.

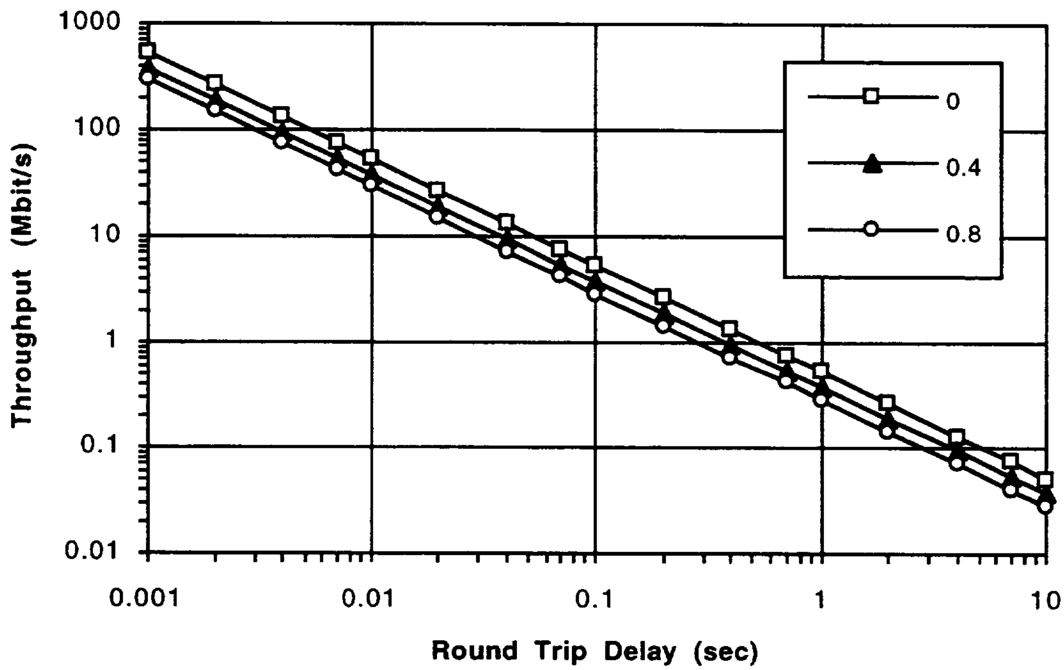


Figure 2-6. TCP Throughput

The impact of delay is more significant for satellite links. The maximum throughput for the LEO, MEO and GEO systems for typical round trip delays is given in Table 2-14. The values in the table will be about 33 percent lower for 50-percent link utilization. From this table it is seen that the impact of satellite delay is significant, in particular in the MEO and GEO systems.

Table 2-14. Maximum TCP Throughput For Various Satellite Network Systems

	LEO	MEO	GEO
Maximum round-trip delay	156 ms	582 ms	1.388 s
Minimum round trip delay (assuming no ISLs)	14 ms	186 ms	534 ms
Maximum throughput (assuming maximum delay)	3.33 Mb/s	893 kb/s	375 kb/s
Maximum throughput (assuming minimum delay)	37 Mb/s	2.8 Mb/s	974 kb/s

For example, the maximum achievable throughput in GEO systems, even without the use of ISL, is a bare 974 kb/s. The worst case MEO system, which assumes ten double-hops to achieve worldwide interconnectivity, also suffers from poor performance due to the limited window size. In actual implementations, the TCP window size is generally made smaller (e.g., 24 kbyte) than the proposed 65 kbyte, reducing the TCP throughput further. A proposed "window scaling" option has been specified for TCP, but few systems currently support this option.

The throughput limitation is severely compounded by traffic control mechanisms that have been included in TCP. To stabilize the internet and prevent congestion collapse, an algorithm was incorporated into TCP called "slow start." In this algorithm, a new connection, or a connection which has just experienced packet loss, ramps up by sending one packet during the first round-trip interval, two packets the next round-trip interval, and so on with an incremental build. This slow start mechanism is used in conjunction with a "congestion window" which cuts the receiver window in half every time a packet loss is detected (which implicitly signifies congestion). The performance of these mechanisms is dependent upon the round trip interval, and for geosynchronous satellites, it will take many round-trip intervals (several seconds) before full throughput can be achieved upon startup or upon each packet loss.

In addition to the degradation experienced due to the congestion control mechanism, the loss of a packet is also detrimental to throughput over a satellite link due to its retransmission procedure. A proposed modification of TCP could be to use a selective acknowledgment scheme, but no mechanism to do this has been proposed at this time.

Although the performance of TCP is poor in the satellite environment, it is also considered poor even in a LAN environment operating over ATM, and several papers[15] have been written proposing that TCP be modified to operate in the ATM environment. It is likely

15 Keung and Siu, "Degradation of TCP Performance under Cell Loss," ATM Forum contribution 94-0490.

that the performance of TCP over satellite-routed ATM connections will improve if TCP is modified for high bandwidth-delay connections.

A protocol that was originally designed for the Broadband ATM Adaptation Layer, but is applicable to any environment with a large bandwidth-delay product, is the Service Specific Connection Oriented Protocol (SSCOP)[16]. SSCOP incorporates a special selective repeat protocol for error recovery that prevents unnecessary retransmissions. The error recovery process in SSCOP makes use of sequentially numbered data packets. If the receiver determines, through the examination of received packet numbers, that a packet is missing, it explicitly requests a retransmission of the missing packet. There are no "time outs" in this protocol and new packets can be transmitted simultaneously during the error recovery process. The receiver must allocate a resequencing buffer to hold correctly received packets until the missing packets are retrieved and resequenced. SSCOP implements flow control through the use of an adjustable sliding window. The receiver grants a "credit" window to the transmitter that allows it to transmit a certain number of packets. If a window's worth of packet is outstanding on the connection, the transmitter must stop transmitting until the receiver returns one or more acknowledgments. The credit can be dynamically reduced or increased by the receiver to control the rate of transmission.

In SSCOP, the transmitter periodically requests status information from the peer by sending a POLL message. This poll can be triggered either through the expiration of a timer or after a certain number of packets have been sent. The receiver, upon receipt of a POLL message, sends a status (STAT) message to the transmitter which conveys the current operating window, the packets acknowledged, and the sequence number for any packets that have not yet been received. The receiver also sends unsolicited status (USTAT) messages whenever an error or loss is detected. The USTAT mechanism allows the transmitter to increase time between POLLs while still achieving rapid recovery from errors.

Figure 2-7 illustrates the general model of the protocol operation. A transmitter, a receiver, and a data link (lower protocol layer) are illustrated. At the transmitter, a user generates packets that are queued until they can be sent to the peer. Once packets are sent, they are placed in a transmitting buffer until explicitly acknowledged, If the receiver requests a packet for retransmission, it is placed in a retransmission queue, where it has higher priority than the queue for new transmissions. At the receiver, a resequencing buffer is needed, and in-sequence data are delivered to the user when available. The forward direc-

[16] Draft ITU-T Recommendation Q.2110, "B-ISDN-ATM Adaptation Layer-Service Specific Connection Oriented Protocol (SSCOP)".

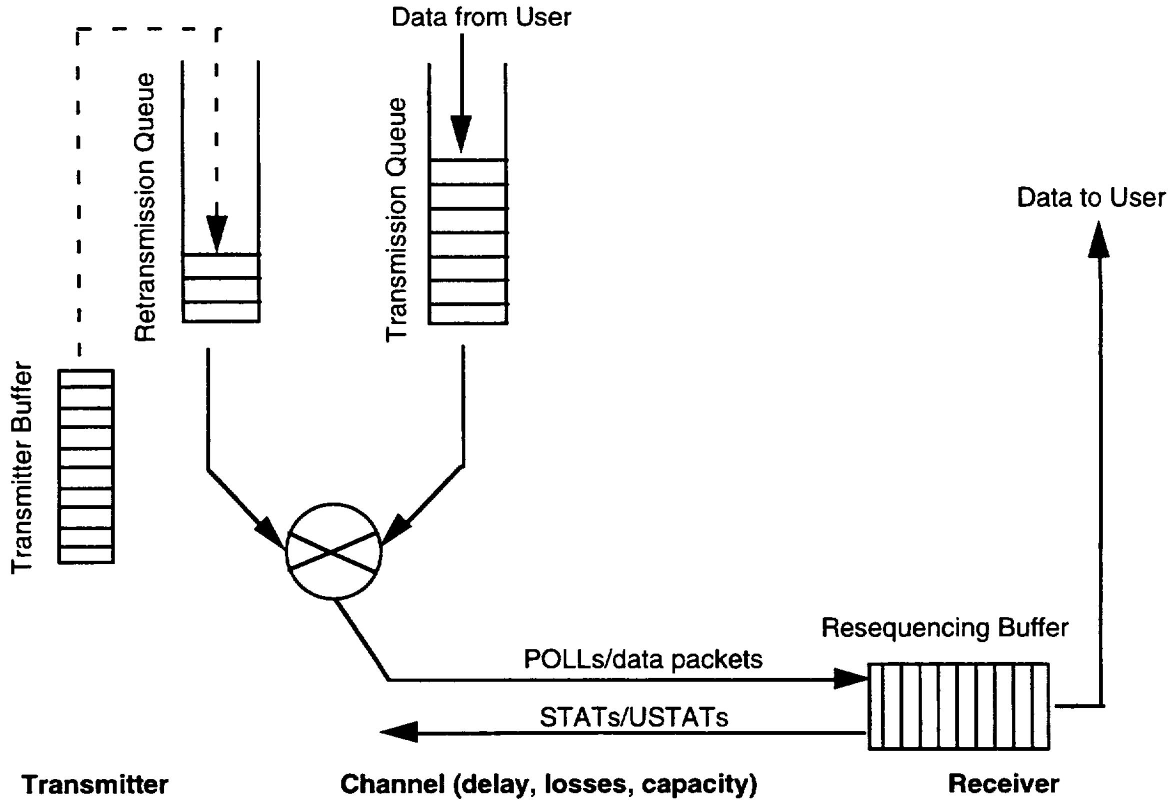


Figure 2-7. A General Model of SSCOP (simplex operation)

tion of the channel contains data packets and POLL messages, while the backward direction allows for STAT and USTAT messages to be sent to the transmitter.

The transmitter driven control flow of SSCOP is beneficial for high-speed implementation, since per-packet overhead is reduced to a sequence number, the overall processing overhead is reduced, and parallel processing between the transmitter and receiver portions of the implementation can be used. In addition, the protocol provides a large enough value of the sequencing number (up to 24 bits in length) to accommodate gigabit-per-second data rates over multiple satellite hops.

2.3.6 Beam/Satellite Handover

The term handover (or handoff) is used in terrestrial wireless services as well as mobile satellite systems and refers to the process in which a call that is being routed through a particular path is re-routed via another path to maintain an uninterrupted connection. Handovers are used extensively in current terrestrial mobile systems for switching between cells and switching between mobile switching centers. For satellite systems, there are essentially two types of handover mechanisms; intra-satellite beam-to-beam handovers and satellite-to-satellite handovers. In the former case, the user terminal moves from one spot beam to another within the same satellite and its circuit connection has to be re-routed by switching to the new beam. In the former case resulting from either satellite or user terminal roaming, the user terminal has to switch connections from the coverage of one satellite to another.

Satellite-to-satellite handovers can be further sub-divided into systems using Satellite-Oriented Frequency Assignment (SOFA) and those using Region-Oriented Frequency Assignment (ROFA). Figure 2-8 illustrates these concepts.

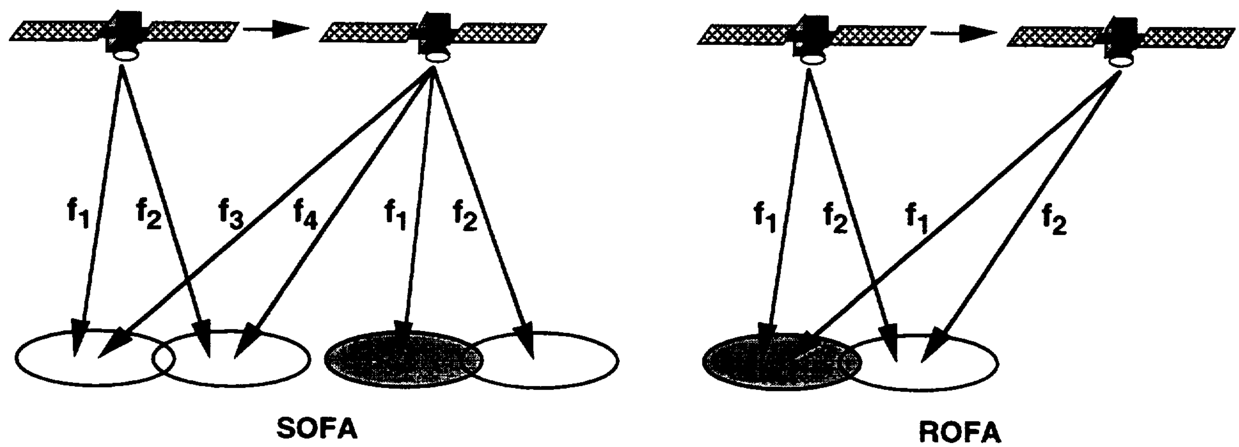


Figure 2-8. SOFA and ROFA Concepts

In the case of SOFA, the frequencies used in each satellite beam remain fairly constant as the satellite moves in the orbit. At the time of handover, the user terminals are required to change frequency. In the case of ROFA, the user terminal maintains a fixed frequency with respect to its geographical location. At the time of handover, it is the satellite which has to switch frequencies. ROFA requires a more sophisticated satellite payload than does SOFA. However, the use of ROFA makes the design of the user terminal much simpler, thereby reducing its cost.

Normally when a user terminal can communicate with two or mores satellites, diversity channels will be allocated on two different satellites, subject to resources being available. In this case, satellite-to-satellite handover is embedded in the diversity channel allocations. When a user terminal, in communication with one satellite also comes in sight of a second satellite it will be allocated a second 'diversity' channel on that satellite. Only when the visibility of the first satellite is lost will the channel through it be released. Satellite handover is normally a make-before-break process where the duration of the overlap is equal to the time of visibility of the two satellites.

In comparing the LEO, MEO and GEO systems with respect to handovers, the key parameters of importance are the motion of the satellite with respect to the Earth and the number of beams. If satellites appear to move rapidly, the handover process is frequent. The period between handovers depends upon the number of beams in each satellite and the relative rate of motion.

Satellites in GEO orbit appear fixed with respect to the earth and handover is generally not needed. However, for LEO and MEO orbits, satellites move fairly rapidly with respect to the earth, with speeds up to 25,000 km/hour in LEO orbits. Figure 2-9 shows the apparent rate of motion of a satellite with respect to the earth, as a function of the altitude of the satellite[17]. While MEO satellite systems have an apparent rate of motion of approximately 1° per minute, the LEO systems have rates of 18° to 35° per minute at zenith (with just slightly slower rates at 40° above the horizon). The maximum time-in-view of a LEO satellite by a stationary ground user terminal, assuming 40° elevation, is about 3.5 minutes. The corresponding time-in-view of a MEO system (with 15° elevation) is about an hour. Additionally, due to the difference in their altitudes, a larger number of beams are required to cover the same area on Earth by LEO systems as compared to MEO. Consequently, both beam-to-beam and satellite-to-satellite handovers will occur more frequency in LEO systems than MEO. These results are compared and summarized in Table 2-15.

---

[17] R, Rusch and P. Cress, "Selecting the Best Constellation for Mobile Satellite Services," AIAA-96-1068-CP, pp. 680-687.

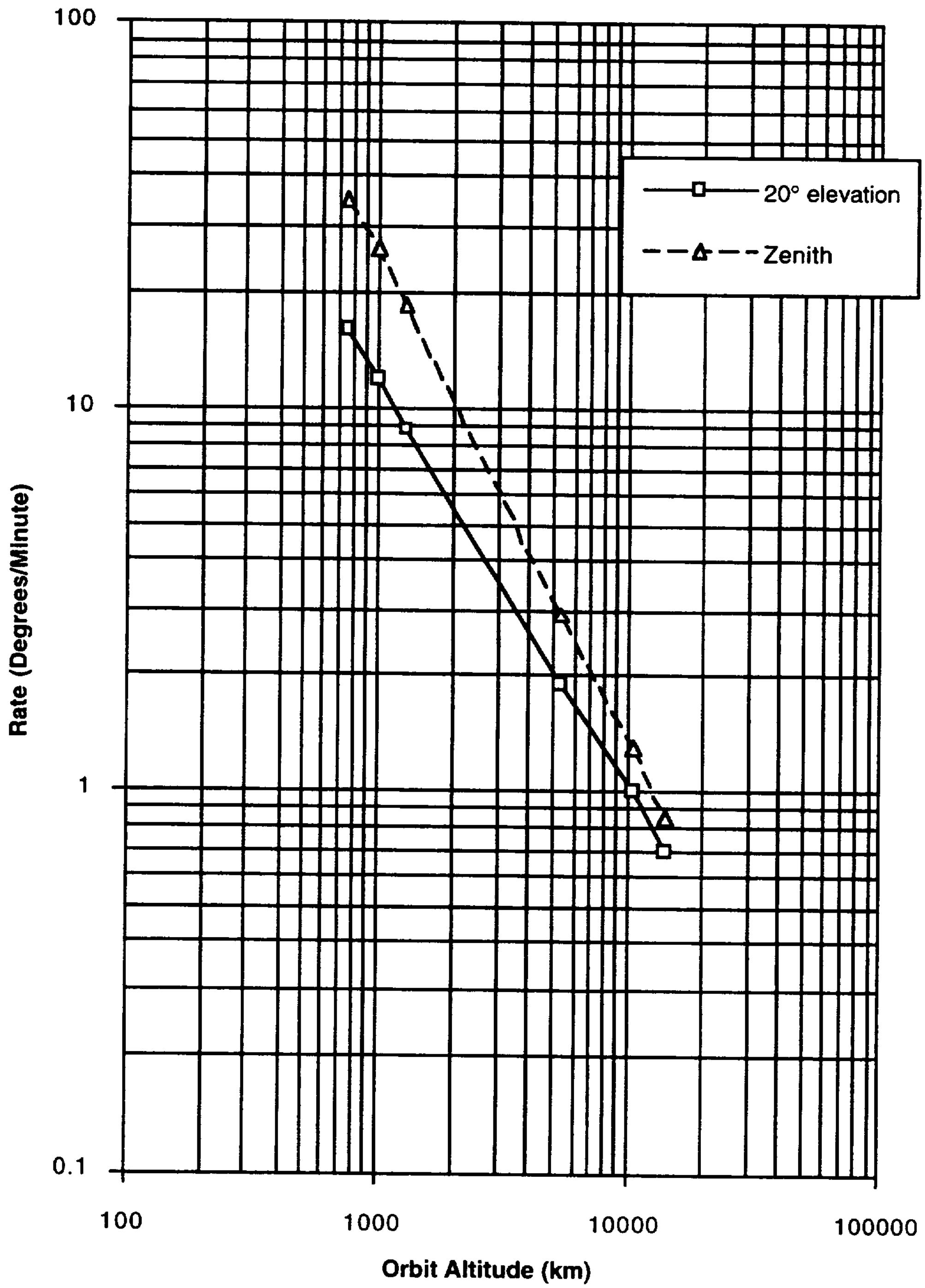


Figure 2-9. Satellite Rate of Motion with Respect to Earth

Table 2-15. Comparisons of Satellite Networks for Handover

	LEO (e.g., Teledesic)	MEO (e.g., ICO)
Elevation	40°	15°
Rate of Satellite motion with respect to Earth	18° to 35° per minute	1° per minute
Maximum Satellite time-in-view	3.5 minutes	approximately 1 hour
Average time for beam-to-beam handover	< 1 minute	5 minutes

The impact of satellite and beam handovers is a potential traffic loss during a handover process. To prevent this event, a precise timing coordination is required among satellite baseband processor routing, ISL re-routing, and user terminal transmit and receive processing. An abrupt change in path delay must be absorbed by user terminal buffers such that there is virtually no delay variation at the terrestrial interfaces. Also, user terminals operating in the LEO or MEO systems must be equipped with tracking antennas. In contrast, none of these features are required for GEO system terminals.

2.3.7 Traffic and Congestion Control

Satellite systems, even those operating at the higher microwave frequency bands (such as the Ka-band), are far more limited in their bandwidth capacity than terrestrial fiber systems. The scarce bandwidth resources of a satellite system must be wisely allocated amongst the various traffic sources so that bandwidth efficiency is maintained while, at the same time, providing users with an acceptable QoS. The situation is made further complicated by the fact that, as shown in subsection 2.2.1, NII/GII applications make use of a variety of voice, video and data services; each differing widely in their data rates, burstiness characteristics and QoS requirements.

To avoid overloading the satellite network, and yet efficiently carry the different types of user traffic, one could use deterministic multiplexing, i.e., allocating bandwidth based on the peak rate of all the connections. This would ensure that congestion would never occur, but such a method under utilizes the bandwidth of the system. A more efficient solution is to use statistical multiplexing, according to which users are allocated bandwidth that is lower than the peak transmission rate of that particular connection. This improves the bandwidth efficiency, but can lead to congestion based on the finite probability that at some link section, and for some time period, the sum of all the peak rates of all the link connections exceeds the total bandwidth. If the situation persists, cells are discarded due to overflowing buffers within the network and the network is deemed congested.

Two broad categories of congestion control mechanisms have been proposed in recent years; namely preventive control (such as leaky bucket) and reactive control (such as sliding window). Traditional packet-based networks, such as X.25, use windows-based congestion control mechanisms, relying heavily on the end-to-end exchange of control messages. As transmission speeds have increased to the Gb/s range, the propagation delays have started to dominate over other delay components, such as switching and buffering delays. As a result of the large propagation delays, the windows-based congestion control mechanism, which depends heavily upon feedback information, cannot be updated fast enough to react to a congestion situation. The introduction of multiplexing different types of traffic with varying delay and cell loss requirements, has also created the need for new congestion control mechanisms to replace the traditional window-based schemes.

Preventive Control

The most common preventive control is selective call admission whose function decides whether to accept new calls into the network depending upon the availability of bandwidth resources. Call admission can be carried out at the call setup time, when the user requests a service from the network by providing parameters such as peak and average bit rates and the maximum burst length. The network determines, on the basis of these parameters and available network resources, whether the connection should be accepted. To enforce the contract between the user and the network, a policing function needs to be added. For example, one could use various queuing buffers on each Virtual Circuit (VC), such as a peak rate controller (to ensure that the connection does not exceed the agreed peak rate) and leaky bucket (which ensures that the average bit rate is maintained). This approach helps greatly in controlling traffic and thereby preventing congestion, but is not always practicable. Many sources, such as data transfer, are unable to predict these parameters in advance.

Another method, commonly employed in ATM networks, is to prioritize the cells by using a single bit in the header, called the Cell Loss Priority (CLP) bit, to define high (CLP=0) and low priority (CLP=1) cells. The implication of CLP=1 is to indicate that the respective cell should be dropped first in the event of buffer overflow in the network. In addition, a traffic policing function may be introduced which has the option of changing the CLP of non-compliant cells from CLP=0 to CLP=1 instead of simply discarding the cell. Thus, an attempt is made to pass non-compliant cells first, before discarding them.

While preventive controls can help to reduce congestion problems before they occur, some degree of congestion will have to be tolerated to use satellite bandwidth efficiently. Consequently, there is a need for fast congestion reaction mechanisms which can quickly relieve congestion problems whenever they occur.

Reactive Controls

Currently defined congestion control mechanisms for ATM and other multimedia networks are very poorly defined and most are unsuitable for systems with large delays, such as satellite networks. To be effective, congestion control must satisfy the following criteria:

a. They must operate on very short time scales to be successful, since at broadband rates, finite sized queues may quickly overflow during congested conditions

b. They must allow high network efficiency without sacrificing QoS parameters such as CLR

c. They must be fair (i.e., they should not favor high rate users over low rate users)

d. They should be easy to implement.

Reactive congestion control operates in three phases:

1. Detection of the congestion condition
2. Carrying of congestion information to the user
3. Source reaction to congestion information.

Congestion detection mechanisms have not been subject to standardization, since such methods are implementation dependent. A simple congestion detection mechanism, implemented in most current ATM switches, considers the link to be congested when a cell is discarded at the transmit link buffer.

Once congestion has been detected at a link in the satellite network, the information must be conveyed to the sources so that they can then react to bring the traffic load to an acceptable level. Two mechanisms, currently proposed, to accomplish the information delivery are the Explicit Forward Congestion Notification (EFCN) and the Explicit Backward Congestion Notification (EBCN) methods. EFCN relies upon the use of a single bit on the cell header, termed Explicit Forward Congestion Indication (EFCI) bit, to convey congestion information. Whenever a network element determines that it is congested, it sets the EFCI bit on each cell of a particular virtual channel contributing to the congestion. At the destination, this indication is sent to the higher protocol layer, which instructs its peer protocol entity to reduce its traffic load. As currently defined, EFCN is inadequate for effective reactive congestion control. First, no semantics have yet been specified for EFCN. Consequently, the end user, receiving this notification, cannot be sure of the true state of congestion in the network. Second, EFCN is unenforceable; no mechanism exists for the higher-level protocols to act on this indication and regulate the flow. Finally, and most critical to satellite communications, EFCN necessarily incurs a round-trip propagation delay in notifying the source.

EBCN, on the other hand, sends notification of congestion in the reverse direction to the congested path (i.e., towards the source). This notification could be some type of a performance-management type cell. EBCN results in better congestion control, especially if congestion is detected close to the source, and for systems where the propagation delays are large, as in a satellite network. The sources are able to receive the congestion information in less time, and therefore react to a situation faster than with EFCN.

To compare LEO, MEO and GEO systems for their responsiveness to congestion control, a parameter termed as bit-length is used. Bit-length is the product of the end-to-end propagation delay of a network with the link bandwidth, and is a measure of the responsiveness between two end points to congestion control. For example, a 100 km long, 1 Gb/s link has a bit-length of approximately 1.6 Mbits. EBCN controls would require, on an average, 1 bit-length for information to reach a user. EFCN, on the other hand, would typically require two bit-lengths for the congestion information to reach the user. Table 2-16 compares the bit-length for the various types of satellite systems. One can see from this table that for satellite systems operating in GEO, one can have bit lengths of approximately 42 million bits, even for the case where we have considered minimum propagation delay (i.e., single-hop, no ISL). By the time EFCI notification would reach the user, about 84 million bits would have already been sent, thereby exacerbating the congesting condition.

Table 2-16. Bit-Length Comparisons for LEO, MEO and GEO Systems

	Delay (ms)		Bit-Length			
			STS-1 rate (51.84 Mb/s)		STS-3 rate (155.52 Mb/s)	
	minimum	maximum	minimum	maximum	minimum	maximum
LEO	7	78	363 kbits	4.04 Mbits	1.088 Mbits	12.13 Mbits
MEO	93	291	4.82 Mbits	15.1 Mbits	14.46 Mbits	45.26 Mbits
GEO	267	694	13.84 Mbits	35.98 Mbits	41.52 Mbits	107.9 Mbits

Clearly, congestion control mechanisms become a major problem at high data rates for satellite networks with a large amount of delay.

2.3.8 User Terminal Requirements and Complexity

This section describes the issues involved in the design of a user terminal which transmits and receives traffic data from other users via the connectivity provided by the satellite network. A block diagram of a typical user terminal is given in Figure 2-10.

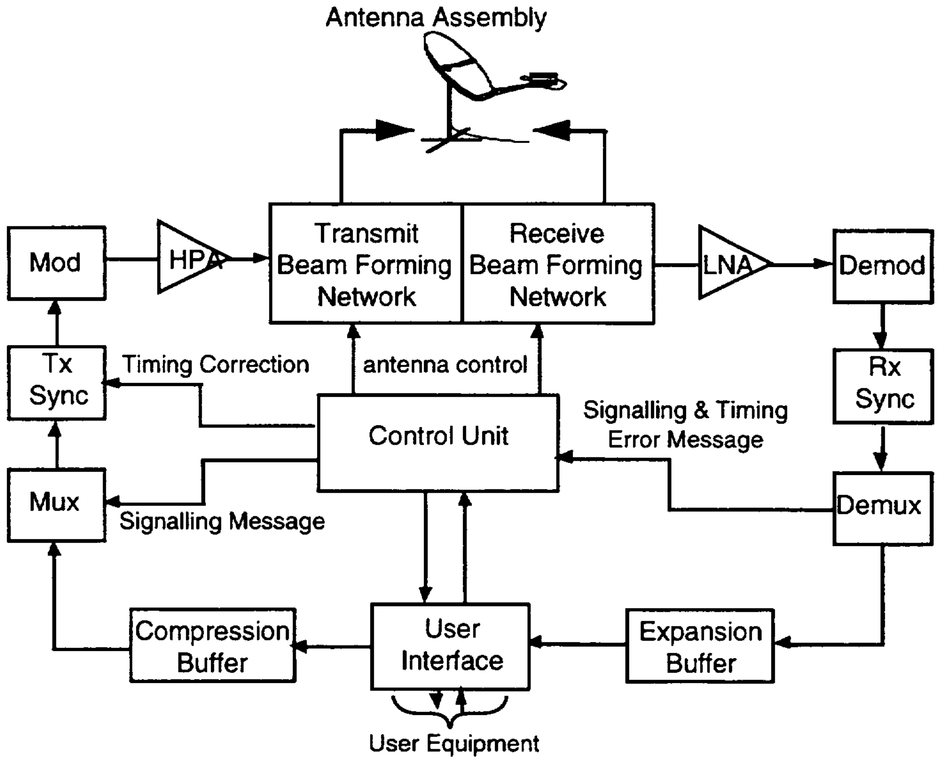


Figure 2-10. User Terminal Block Diagram

To describe the variations in the user terminal functions of various satellite systems (LEO, MEO and GEO), it helps to review the functions of the user terminal. These are:

a. Perform downlink frame acquisition and establish synchronization. Start decoding signaling messages from the Network Control Center (NCC)

b. Transmit an initial acquisition burst based on the receive frame timing and the satellite position information (or an equivalent message received in the NCC signaling channel)

c. Perform transmit timing correction according to the timing correction message received from the satellite (timing error is measured by the satellite like in the ACTS system or at the NCC)

d. Upon establishment of signaling links with the satellite and NCC, send traffic capacity requests

e. Program a demultiplexer, multiplexer, and user interface control maps according to the assigned transmit and receive satellite time slots and user interface channel formats

f. Transmit and receive traffic in the assigned time slots

g. Maintain transmit and receive synchronization, perform data demultiplexing and multiplexing, and process signaling messages

h. Perform (coordinated) traffic reconfiguration as needed to dynamically reallocate satellite capacity.

The user terminal consists of the following subsystems:

a. An RF assembly consisting of a solid state HPA, an LNA, a steerable antenna/feed, upconverter and downconverter

b. Uplink processor consisting of:

   1. A transmit synchronizer to acquire and synchronize to the satellite timing

   2. A burst traffic/ packet formatter and a compression buffer which inputs user data from the Data Terminal and produces uplink burst data at the transmission rate

   3. A burst modulator (typically QPSK). This subsystem normally includes an FEC encoder, interleaver and block encoder (e.g., Reed-Solomon type)

c. Downlink processor consisting of:

   1. A burst demodulator which includes an FEC decoder, deinterleaver and block decoder

   2. A receive synchronizer to acquire and synchronize to the downlink reference timing

   3. A time division demultiplexer and expansion buffer which extracts the downlink data and signalling messages provides the user with continuous data to the data terminal

d. User terminal interfaces to data equipment (e.g., DTE) or terrestrial networks (T1/E1, T2/E2, T3/E3, ATM, SDH/SONET, ISDN and LANs)

e. Control unit to process signalling messages and control and monitor terminal operation.

The various satellite networks differ greatly in their user terminal cost and complexity. GEO satellites have a 31 to 34 dB greater propagation loss as compared to LEO satellites. Similarly, MEO satellites have an increase of between 20 to 26 dB propagation loss as compared to the LEO satellites. The larger propagation losses of MEO and GEO satellites translate into larger antenna diameters either at the user terminal or on-board the satellite (GEO/MEO). This advantage of LEO satellites has to be weighed against the increased complexity in the user terminal for LEO because of the need to communicate with a larger number of fast-moving satellites. The GEO systems, because of their fixed position relative to earth and their small number of satellites (3 to 4 for worldwide interconnection), result in greatly simplified user terminals. Nor do GEO systems have to grapple with complex handover problems, as do LEO and MEO systems. These problems are especially severe for LEO systems, where each satellite is visible for approximately 3.5 minutes, and routing has to be dynamically assigned in short periods of time over many satellite links. Problems in timing and sequencing of data are possible in cases where the user switches from one satellite to another which may require either sophisticated timing coordination or the addition of extra demodulators and demultiplexers in the user terminal. This would add to the cost of the terminal as well as increase its size and weight.

2.4 BARRIERS TO INTEROPERABILITY OF SATELLITE AND TERRESTRIAL NETWORKS

Seamless integration of terrestrial and satellite networks can be addressed in terms of QoS and cost. First, the QoS that the end user expects from a fully terrestrial network should not be degraded when the end-to-end connection involves a satellite link. Second, the services provided via the hybrid (satellite/terrestrial) network must be cost competitive. The effect of satellite systems on QoS parameters have been discussed in great detail in subsection 2.3. The main distinguishing characteristics of satellite networks with respect to terrestrial networks are their delay, delay variations, poorer link quality and the need for beam/satellite handover. These characteristics were shown to vary quite widely for the proposed LEO, MEO and GEO satellite systems. Each characteristic has a profound effect on both the QoS parameters of the end-to-end system as well as cost.

The barriers against seamless interoperability of satellite and terrestrial networks can be categorized into the following components:

a. A major administrative hurdle is the lack of awareness amongst vendor and the terrestrial carrier community of the parameters and protocols necessary for efficient operation of satellite links. For example, while there is provision for the use of a large window size in the TCP protocol, most vendors design TCP with smaller windows. While this is adequate for terrestrial networks, it is totally inadequate (as shown in subsection 2.3.4) for those satellite network systems that have large propagation delays.

b. Important broadband standards, such as ATM and SONET, as well as data protocols (X.25/X.75) and transport protocols (such as TCP/IP) have, in the past, been designed assuming the use of terrestrial links. Similarly, congestion control schemes, too, have been designed with the implicit assumption of terrestrial networks. In subsection 2.3 we have shown how these protocols behave poorly when the network involves large propagation delays (such as in GEO satellite networks). Many of these standards can be greatly improved, as we have pointed out in subsection 2.3, by a few slight changes to the protocol.

c. Bandwidth efficiency, which is so crucial in satellite networks, has been given very low priority in standardizing the next-generation digital transport architecture (e.g. ATM/SDH). From very early on, a fundamental decision was made to base ATM and SDH on the characteristics of fiber optic media. Many characteristics of these networks result in very high percentage of overhead. For example, ATM CBR speech will require over 70 kb/s satellite bearer channel for carrying one voice channel. This is in direct contrast to currently designed satellite networks which attempt to maximize the use of the limited satellite bandwidth by applying advanced signal processing techniques to reduce the bandwidth requirement of each service (for example, voice services are being designed to as low a rate as 3.6 kb/s).

While both technical and administrative barriers do exist towards a seamless integration of satellite and terrestrial networks, there are ways to overcome these barriers. For one, not all the satellite networks have the same degree of interoperability problems with terrestrial networks. For example, LEO systems might often experience less delay than intercontinental fiber links, and their QoS may not differ too much from such terrestrial systems. Another welcome trend in reducing the interoperability barrier, is the fact that some of the problems experienced by satellite networks are also being experienced in other areas, such as inter-connection of high-speed LANs. Thus, new standards (such as SSCOP) are being formulated which are attempting to address these as well as satellite concerns. More participation by satellite service providers, equipment vendors, satellite manufacturers and satellite service users is needed in the communications standards organizations (such as: ITU-T, ITU-R, ATM Forum and ANSI). Finally, there is growing awareness amongst the communications industry about the need and importance of satellite networks. The meteoric rise of wireless network connections and the realization of the enormous costs involved in introducing fiber to the home, or even fiber to the curb, is renewing interest in global satellite networks. The clamor amongst the satellite companies to file Ka-band broadband service networks, attests to the market reality of such networks. This market reality is bound to affect the broadband standardization process and force the terrestrial and satellite networks to work together in a seamless fashion.

SECTION 3 – CONCEPTUAL ARCHITECTURES

3.1 INTRODUCTION

From the time satellite systems were first introduced, they have steadily evolved in terms of their capabilities and services. The first generation of satellites were purely analog, and acted as "bent-pipe" systems making possible international telephone connections as well as providing global broadcast of television signals. Second generation satellite systems, currently under development, have shifted to digital designs that make use of sophisticated digital processing techniques with on-board switching capabilities to allow the interconnection of a large number of spot beams. The third generation system will have even more digital processing capabilities and in addition, will be able to operate seamlessly with terrestrial systems.

In this section, conceptual architectures for the third generation of satellite networks are developed. These architectures are essentially designed to be able to support the NII/GII applications identified in Section 1 and to meet the QoS performance analyzed in Section 2. An additional design goal aspires to achieve seamless interoperability with terrestrial networks.

This report proceeds as follows: First, the requirements for the third-generation satellite systems are enumerated followed by a description of conceptual third-generation satellite network designs, including a brief overview of the emerging Ka-band filings for satellite multimedia communications services. A tradeoff comparison of LEO, MEO and GEO systems is made and common system issues affecting the design of each architecture is examined. LEO and MEO systems differ radically from GEO systems in that the satellites in these orbits move with respect to the earth. Consequently, LEO/MEO system architectures require different considerations than GEO systems. For example, handover issues which are important for both LEO and MEO systems do not have to be considered in GEO systems. In this report, conceptual architectures for GEO systems are explored first. Some architectural issues discussed are: satellite constellation geometry, data rates, number of beams, QoS parameters, protocols, access techniques, spacecraft antenna size and type and Inter-Satellite Links (ISL) requirements. LEO/MEO system architectures are also developed in a similar fashion. Finally, the key technologies required to enable these architectures are identified and some of the important research work in these technologies are discussed.

3.2 REQUIREMENTS FOR SPACE-BASED INFORMATION INFRASTRUCTURES

The various NII/GII applications as well as the market for these applications by the year 2005 were described in Section 1. These applications are being spurred on by the reduced costs and increasing multimedia capabilities of personal computers, the phenomenal growth in internet usage, and the availability of new mobile communication services to users. It was shown in Section 1 that services that support the NII/GII applications differ widely in their data rates and in their sensitivity to errors or delay. Future third-generation satellite systems, as well as terrestrial wired and wireless systems, are expected to carry these widely-differing set of services through a common network. This is in sharp contrast to systems of the past which consisted of separate voice, computer (data) and video networks. To facilitate this common integrated system, a packet-based network system, using a fixed 53-byte cell structure has been devised. This system, called ATM, was examined in Section 2. Important QoS parameters for ATM services were also defined and examined in Section 2. Third-generation satellite systems are likely to use ATM as the physical and logical mode of transport for the services that support the NII/GII applications. Using common standards, such as ATM, will help in achieving interoperability between terrestrial and satellite systems.

Clearly defining the requirements for these systems helps to build a framework for the design of these third-generation satellite system architectures. These requirements can be summarized as follows:

a. Enough transponder bandwidth to support the various types of services and data rates required by the NII/GII applications

b. Support the QoS requirements described in Section 2 of this report

c. Interoperable with terrestrial wired and wireless systems

d. User terminal size of one meter or less so as to reduce costs and allow mass production

e. Cost effective.

The first two requirements have been briefly examined in Sections 1 and 2. In Section 1, the services that support the NII/GII applications were introduced and the markets for these services were described. In Section 2, various attributes of these services, such as bandwidth requirement and QoS parameters, were enumerated. These services and their attributes are recast in Table 3-1. The QoS parameters and the data rates for each service are given in this table, as are the NII/GII applications that make use of these services. Applications that belong to the stringent class of ATM services (i.e., Type 1), have a

Table 3-1. Service Requirements for NII/GII Satellite Architectures

ATM Categories	CBR Services			rt-VBR		nrt-VBR		UBR					ABR		
Services	Video/ Audio Retrieval	Inter-active Video	Multimedia File Transfer	Real-time Voice	Real-time Video	Transaction Processing	LAN /WAN	Text/ Data/ Image Messaging	Text/ Data/ Image Distribution	Text/Data/ Image Retrieval	Remote Terminal	Critical Data Transfer	Super-computing	Distributed File Service
Examples	Video-on-demand, audio library	Tele-conferenc-ing	Data/text/ image transfer	Telephone	NTSC/ PAL/ HDTV	Airline res., banking trans., process monitoring	Frame relay inter-working	e-mail, telex, fax	news feed, weather satellite pictures	file transfer, file browsing	telnet	defense infor-mation	NREN, HPCC	Network File Service
Data Rates	100 Kb/s to 45 Mb/s	100 to 500 Kb/s	10 Kb/s to 50 Mb/s	3.6 to 64 Kb/s	1 to 32 Mb/s	300 b/s to 9.6 Kb/s	64 Kb/s to 100 Mb/s	64 Kb/s to 1 Mb/s	1 to 8 Mb/s	64 Kb/s to 45 Mb/s	28.8 to 64 Kb/s	1.2 Kb/s to 1.5 Mb/s	100 Mb/s to 1.2 Gb/s	64 Kb/s to 45 Mb/s
NII/GII Applications Supported	Telecommut ing, Tele-medicine, Disaster Mngmnt, Gov't. Svcs	Telemedi-cine, Dist. Learning, Govt. Ser-vices, R&D	Telecommut-ing, Tele-medicine, Disaster mgmt, Dist. Learning, Gov't. Svcs, R&D	Telecomm uting, Tele-medicine, Disaster Mngmnt, Govt. Ser-vices	Telemedi-cine, Dis-aster mgmt, Distance Learning	Telecommuting, Electronic commerce, Air-line Reservations	Telecommu ting, Dist. Learning, Govt. Services, R&D	Telecommut-ing, Tele-medicine, Disaster mgmt, Distance learning, Gov't. svcs, R&D	Telecommuting, Telemedicine, Disaster mgmt, Distance learning, Govt. services, R&D	Telecommut-ing, Telemedi-cine, Dis-aster mgmt, Distance learning, Govt. svcs, R&D	Telecommuti ng, Tele-med., Dis-tance lear-ning, Gov. Services, R&D	Telemedici ne, Disast-er mngmnt elect. com-merce, gov't svcs	Telecommu-ting, Tele-medicine, R&D	Telemedi-cine, Multi-media svcs, Dist. Learning, Gov't. svcs, R&D
Max user-to-user delay (ATM strin-gent type)	400 ms	400 ms	Can be >400 ms	400 ms	400 ms	Can be >400 ms	Can be >400 ms	Can be >400 ms	Can be >400 ms	Can be >400 ms	Can be >400 ms	Can be >400 ms	Can be >400 ms	Can be >400 ms
Max. CER of Link (ATM)	1.40E-07	1.40E-07	1.40E-07	1.40E-07	1.40E-07	1.40E-07	1.40E-07	1.40E-07	1.40E-07	1.40E-07	1.40E-07	1.40E-07	1.40E-07	1.40E-07
Max. SECBR of Link (ATM)	3.50E-05	3.50E-05	3.50E-05	3.50E-05	3.50E-05	3.50E-05	3.50E-05	3.50E-05	3.50E-05	3.50E-05	3.50E-05	3.50E-05	3.50E-05	3.50E-05
Required BER of Link	<1E-10	<1E-10	<1E-10	<1E-7	<1E-10	<1E-10	<1E-10	<1E-10	<1E-10	<1E-10	<1E-10	<1E-10	<1E-10	<1E-10
Link Availibility (%)	99.5	99.5	99.5	99.9	99.9	99.5	99.99	99.5	99.5	99.5	99.5	99.5	99.5	99.5

maximum bound on cell transfer delay of 400 ms (see subsection 2.2, Table 2-3). Such a boundary precludes ATM Type 1 service class from being offered in GEO systems employing multiple hops or ISLs. However, single-hop GEO systems as well as LEO and MEO systems can support Type 1 ATM service.

In terms of cell error ratio, ATM specifies an upper limit of 1.4×10^{-7} for satellite links. However, an end-to-end BER performance of 10^{-10} or better is required in most systems. As described in Section 2, to achieve these error rates, satellite systems require added error coding and correction circuitry at the baseband level.

For satellite links, the maximum Severely Errored Cell Bit Rate (SECBR) allowed by ATM for all its classes of services (except the unspecified class) is 3.5×10^{-5}. Satellite systems suffer from outages created by rain fade (e.g., during thunderstorms in temperate zones). These outages typically last less than 30 minutes and are seasonal. Taking into account the loss of signal due to the rain outages, satellite systems are designed to have a link availability of 99.5 percent or better. A link availability of 99.5 percent is equivalent to 44 hours per year of outage, whereas a link availability of 99.9 percent (required for low-speed voice services) is equivalent to 9 hours of outage per year. When higher link availability (>99.99%) is needed, site diversity (with two earth terminals separated by at least 10 km) may be employed. The link availability parameter impacts the selection of the minimum elevation angle.

To satisfy requirement 3, i.e., interoperability of satellite and terrestrial systems, it is necessary that the satellite systems be designed to carry narrowband ISDN (N-ISDN) services as well as broadband ISDN (B-ISDN) services at various SONET rates. The N-ISDN service is composed of two B channels (at 64 kb/s) and one control channel (at 16 kb/s). The two B channels can be aggregated together to form super channels at rates of 384, 1536 and 1920 kb/s. Table 3-2 gives the key telecom signal hierarchies supported in North America as well as the rest of the world.

Requirement 4, which limits the size of a typical user terminal to less than one meter (i.e., approximately three feet), is necessitated by cost considerations. User terminal size should be briefcase size or smaller to be mass producible and thereby cost-effective. This limit on the user terminal size affects the uplink and downlink power budgets, and consequently, uplink/downlink data rate capacities and/or link availability.

The last requirement, i.e., of cost effectiveness, is very important and can often be the deciding factor in determining the viability of a satellite network architecture. Because of its importance, the cost issues will be dealt with separately in greater detail in Section 4.

Table 3-2. Telecom Signal Hierarchies in North America and Abroad

STANDARD	INFORMATION STRUCTURE	BIT RATE
N-ISDN		
	2B + D	144 kb/s
B is Bearer Channel (64 kb/s) which carries information and	23 B + D (N. America)	1.544 Mb/s
D is data channel (16 kb/s) used for carrying signaling info.	30 B + D (outside N. America)	2.048 Mb/s
Digital Signaling Hierarchy DS-1/E-1	N. America & Japan (24 voice channels)	1.544 Mb/s
	Europe (32 voice channels)	2.048 Mb/s
DS-2/E-2	N. America & Japan (96 voice channels)	6.312 Mb/s
	Europe (128 voice channels)	8.448 Mb/s
DS-3/E-3	N. America (672 voice channels)	44.736 Mb/s
	Japan (480 voice channels)	32.064 Mb/s
	Europe (512 voice channels)	34.364 Mb/s
SONET/B-ISDN/SDH		
	STS-1/OC-1	51.84 Mb/s
	STS-3/STM-1/OC-3	155.52 Mb/s
	STS-12/STM-4/OC-12	622.08 Mb/s
	STS-24/STM-8/OC-24	1244.16 Mb/s
	STS-48/STM-16/OC-48	2488.32 Mb/s

3.3 COMPARISON OF LEO, MEO AND GEO ARCHITECTURES FOR USE IN THIRD-GENERATION SATELLITE NETWORKS

3.3.1 Common System Issues for GEO, MEO and LEO Systems

On the basis of the set of requirements discussed in the previous section, various conceptual architectures for the third-generation satellite networks can be designed. Common architectural issues that affect the design of all systems are discussed in this section. These include:

a. Determination of frequency band of operation

b. Elevation angles

c. Earth coverage

d. Number of satellites

e. Radiation issues

f. Lifetime

g. Time of sight

h. Transmission delay

i. Antenna size

j. Free-space loss.

Frequency Band of Operation

Choosing a frequency band of operation is generally a compromise between bandwidth, component availability and rain fade susceptibility. For example, while the C-bands (6/4 GHz) and Ku-bands (14/12 GHz) represent a fairly mature level of technology growth and have readily available low cost components, most of the spectrum allocations have already been made in these bands and only limited slots for broadband applications are available.

The Ka-band spectrum (30/20 GHz) provides about 2.5 GHz of total bandwidth (in each band). The Ka-band frequency will have to be shared among various filed multimedia satellites (over 14 companies have filed for allocations with the FCC as shown in Section 3.3.3), LEO feeder links (e.g., Iridium and Odyssey), Local Multipoint Distribution System (LMDS) links, and NASA's ACTS links. Recently, the WARC-95 committee has accepted segmentation of the Ka-band as shown in Figure 3-1. Ka-band technology is reaching maturity and components in this band are readily available. A drawback with Ka-band in comparison to Ku- and C-bands is that it suffers from greater attenuation due to rain.

The next available band (50/40 GHz) after Ka is in the millimeter wavelength region. A major concern about millimeter-wave technology is that of increased rain-fade. While millimeter waves, no doubt, suffer from increased attenuation due to rain, this loss is not insurmountable. Prior analysis has shown that elevation angles of greater than 30° are sufficient for reasonably reliable millimeter wave communications[1]. The Milstar program at 40 GHz is currently working in this band and Motorola has filed with the FCC to use this band for its M-Star system. The FCC has allocated a 3-GHz downlink bandwidth in the 37.5 to 40.5 GHz band. The allocated FCC uplink bandwidth is 2 GHz, spread over two bands (42.5 to 43.5 GHz and 49.2 to 50.2 GHz). While the total bandwidth available is far more than that in the Ka band, millimeter wave technology is still in its infancy and involves greater technology development risk. There has been major research by Japan in the millimeter-wave technology in recent years and commercial products may soon come to the market[2].

---

1 A.H. Jackson and P. Christopher, "A LEO Concept for Millimeter Wave Satellite Communication," International Mobile Satellite Conference, Ottawa, IMSC '95, pp. 185-192.

2 Y. Takimoto and M. Kotaki, "Recent Development of MM-Wave Applications in Japan", Microwave Journal, May 1996, pp., 214-226.

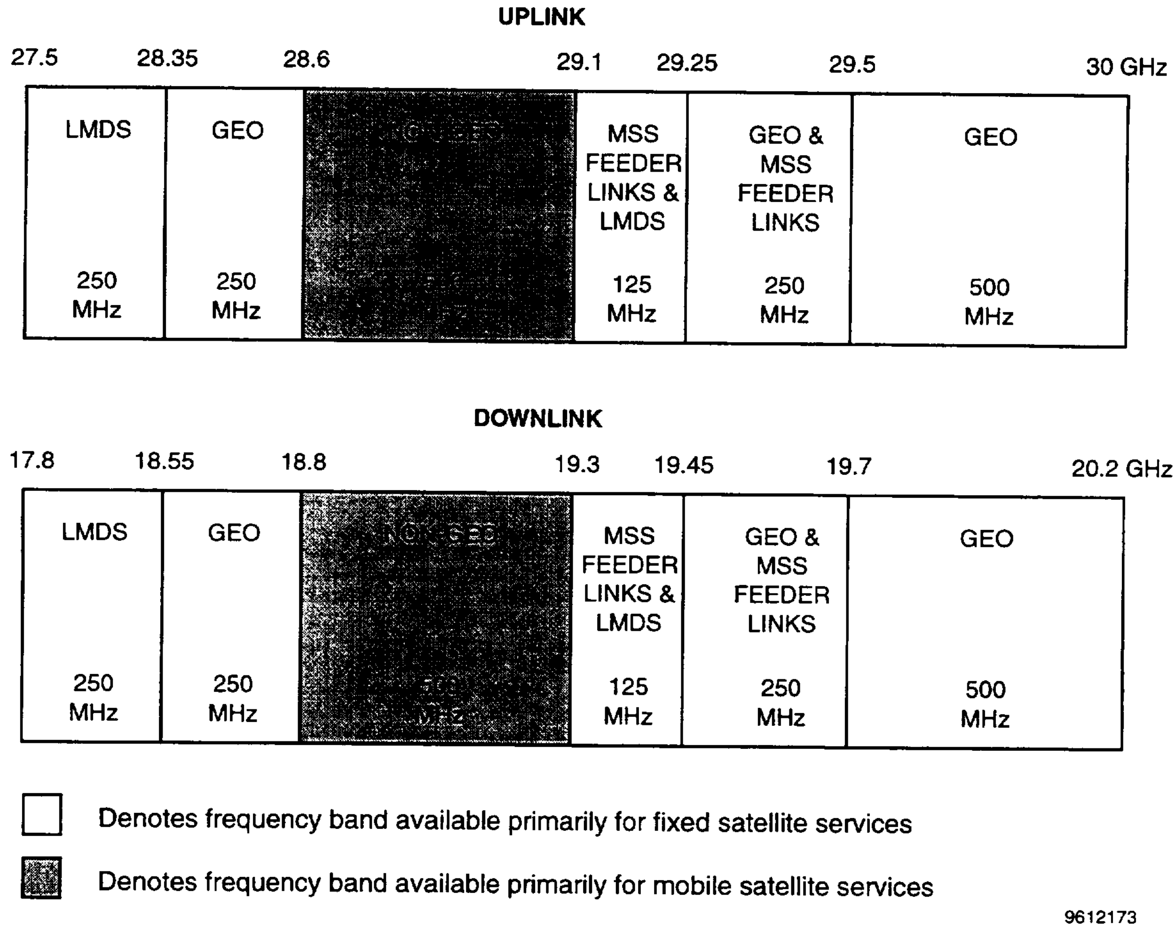


Figure 3-1. Ka-Band Allocation of Frequencies Approved by WARC-95

While the millimeter-wave region is preferable from a point of view of bandwidth availability, its relatively immature technology and its development risk preclude it from further consideration for the third-generation satellite systems in the 2005 time frame. Consequently, the Ka-band is chosen for these systems. Ka-band frequencies have already been allocated for both fixed GEO satellite systems as well as for MEO/LEO systems. There is a total of about 1 GHz band primarily available for GEO systems and about 500 MHz for LEO/MEO systems in each uplink and downlink frequency band. It is possible for multiple GEO systems to use the same 1 GHz bandwidth without interference, provided the satellites are spaced two degrees apart. This is possible because the GEO satellites remain fixed with respect to the ground stations. Multiple LEO/MEO systems, on the other hand, would require extensive cooperation between the different service providers to use the full bandwidth for each system without causing interference. Even if such coordination is possible, it would greatly increase the complexity and cost of the satellite design.

Elevation Angle

Satellite network systems are designed to support a fixed minimum elevation angle. The minimum elevation angle is an important system parameter that affects the availability of the link as well as determining the Field of View (FOV) for each satellite and thereby total number of required satellites for a global system.

Elevation angle affects the satellite architecture differently for the GEO system as compared to LEO and MEO systems. In the GEO system, the satellite is fixed with respect to the earth. The choice of the minimum elevation angle affects the coverage area of each satellite. One can use Equation E2-1 given in subsection 2.3.2 to derive the coverage latitude (β) as a function of the elevation angle (θ).

$$\beta = 90 - \left[\theta + \sin^{-1}\left[\frac{R_e}{h + R_e}\sin(\theta + 90)\right]\right] \qquad \text{E3-1}$$

This relationship is plotted in Figure 3-2. Latitudes of approximately ±60° encompass the majority of the world's population, with most of USA (except for Alaska) at latitudes lower than 50°. An elevation angle of 20° is sufficient to cover these populated regions.

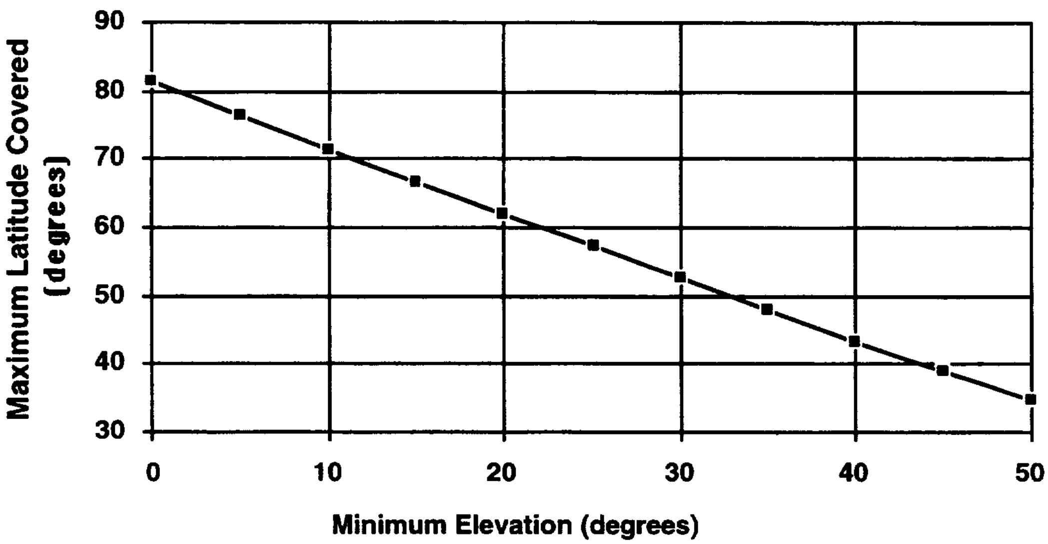
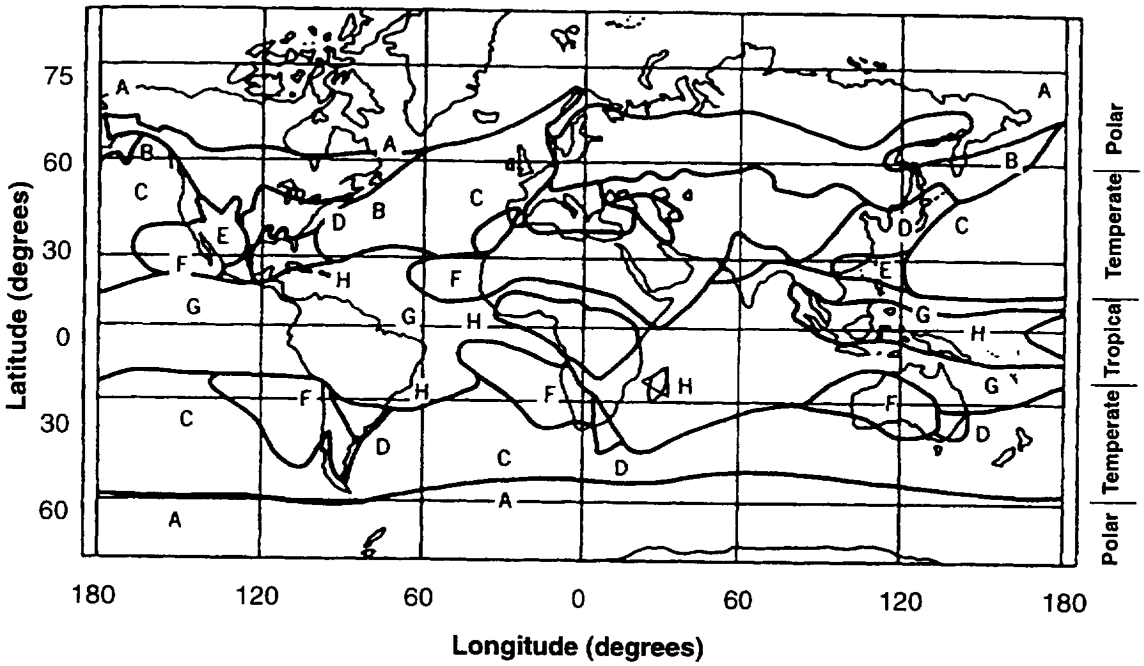


Figure 3-2. Maximum Latitude Covered by a GEO Satellite

Since the satellites in GEO system remain fixed with respect to earth, each user location has a fixed relative elevation angle with respect to the satellites. Users located at higher latitudes have low elevation angles and the signals from them to the satellite requires traversing over a longer path of atmosphere, thereby being more susceptible to rain fade. Figure 3-3 shows the various climatic conditions across the world. Some regions in the world with the most rainfall (region E) are fortuitously at low latitudes and have high elevation angles with respect to the GEO satellites.



Rain rate climate regions

Polar	A	Thundra (dry)	Subtropical	E	Wet
	B	Tiaga (moderate)		F	Arid
Temperate	C	Maritime	Tropical	G	Moderate
	D	Continental		H	Wet

9609013

Figure 3-3. Crane Model of Climatic World Conditions

In contrast to the GEO system, LEO and MEO systems' satellites move relative to the Earth. Thus, each user location sees the satellite traverse over the full range of elevation angles before the satellite switches its signal to an adjacent satellite. This variation of elevation angle translates into a larger variation in rain fade for each user terminal. However, the maximum rain fade value is still determined by the minimum elevation angle (as was the case in GEO).

Earth Coverage

The coverage diameter of a satellite is a function of both the minimum elevation angle and satellite altitude. The coverage diameter of each satellite, measured along the chord between opposite edges of the FOV, can be shown to be:

$$D = 2R_e \sin\left(90^\circ - \sin^{-1}\left[\frac{R_e}{h+R_e}\cos(\theta)\right] - \theta\right) \qquad \text{E3-2}$$

Or, as an explicit function of FOV:

$$D = 2x\sin(\alpha) = 2x\sin\left(\frac{FOV}{2}\right) \qquad \text{E3-3}$$

where, from the derivation in Appendix C,

$$x = (R_e + h)\cos\left(\frac{FOV}{2}\right) - R_e\sqrt{1 - \left[\gamma\sin\left(\frac{FOV}{2}\right)\right]^2} \qquad \text{E3-4}$$

The coverage diameter is plotted in Figure 3-4 as a function of satellite altitude and minimum elevation angle. The coverage diameter varies from about 706 km for a satellite at altitude 700 km (LEO-Teledesic) to about 8756 km (GEO) assuming a fixed minimum elevation of 40°. The coverage diameter also changes with elevation angle; the lower the minimum elevation angle, the larger the coverage diameter. As can be seen in Figure 3-4, this variation is more dramatic for GEO than for LEO satellites.

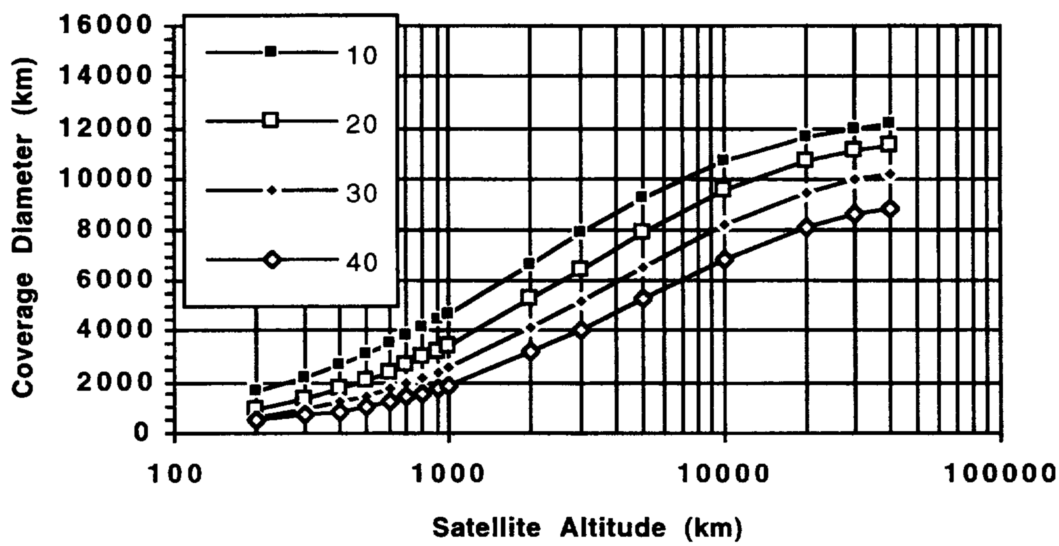


Figure 3-4. Earth Coverage Diameter of a Satellite

A key aspect of earth coverage by the various of satellite constellations is the coverage over populated areas. While GEO systems can position each satellite to cover a large population area and then set satellite resources to continuously cover high traffic areas, such a scheme cannot be implemented in LEO and MEO systems where the satellite constantly moves with respect to the earth. The satellite resources are essentially wasted during the periods in which LEO/MEO satellites traverse through ocean areas where there is little or no traffic. The same satellites, on the other-hand, are often "overworked" in terms of system utilization and capacity when traversing over high traffic areas. This problem of unequal utilization of satellite resources is more pronounced in LEO satellites which pass over the user terminal's horizon in just a few minutes (approximately 3 to 5 minutes) than for MEO satellites where the satellites are in view for about an hour.

Number of Satellites

The number of satellites needed for full global coverage is also a function of both the minimum elevation angle supported by the system and satellite altitude. To compute the number of satellites needed for full global coverage, the footprints/FOV of each satellite in a LEO/MEO or GEO orbit is taken as a snapshot in time.

Assuming a slight overlap in each satellite's coverage (an overlap factor of $\sqrt{2}$ is generally taken) one can easily compute the number of satellites needed to cover Earth's entire circumference. One can also compute the number of planes (and spacing between the planes) needed to cover the earth's surface. The total number of satellites needed to cover the entire earth surface is the product of these two numbers, i.e., the number of satellites per plane and the number of planes. This number has been computed as a function of elevation angle and altitude and is shown in Figure 3-5.

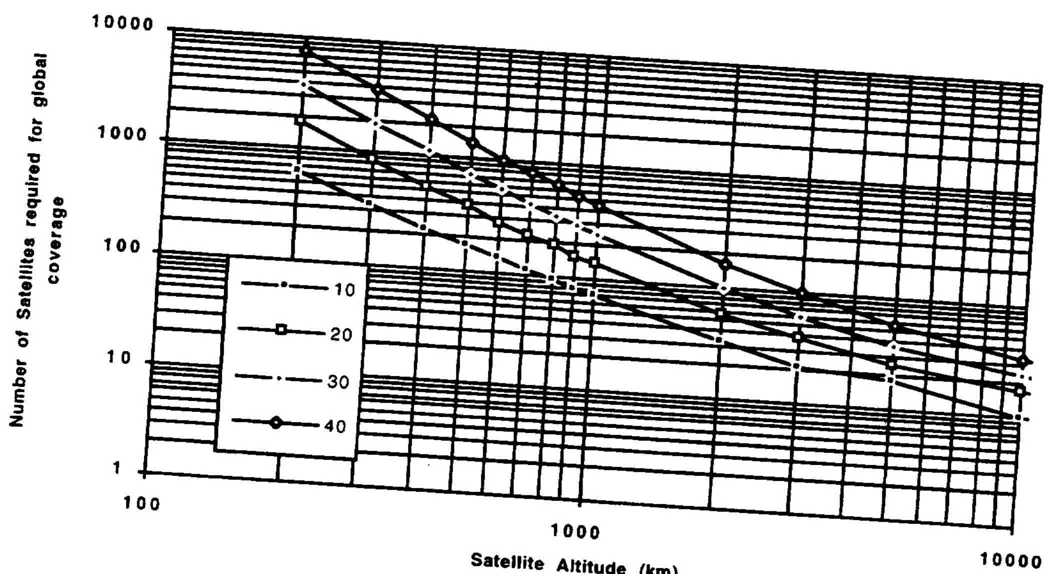


Figure 3-5. Number of Satellites Needed for Global Coverage

It is quite clear from Figure 3-5 that the choice of elevation angle and satellite altitude has a dramatic effect in determining the number of satellites needed for global coverage. For example a 40° elevation angle and 700 km satellite elevation, as chosen by Teledesic, requires a minimum of over 800 satellites. Changing the elevation angle to 20° reduces this number to about 210. Similarly, keeping the elevation angle the same (40°), but changing the altitude from 700 to 1500 km reduces the satellite count from 800 to about 200.

Radiation

While it is desirable to increase the LEO satellite altitude so as to limit the number of satellites needed for global coverage, another phenomenon limits the altitude of LEO satellites to less than 1500 km. This phenomenon is the presence of the Van Allen radiation belt, caused by the trapping of charged particles around the Earth's magnetic field. Figure 3-6[3] shows the Van Allen belts (as a function of distance given in multiples of the Earth radii). Satellites traversing these belts can suffer from high frequency of electronic anomalies and malfunction. Repeated motion in these orbits results in considerably reduced satellite lifespan.

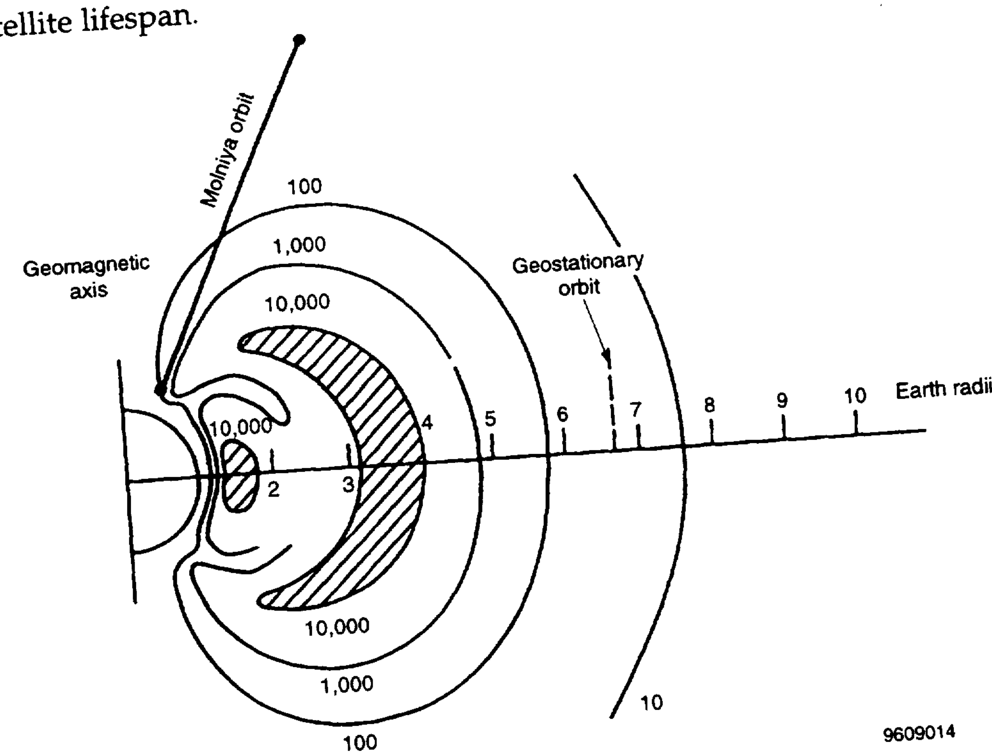


Figure 3-6. Van Allen Radiation Belt

Geostationary orbits have almost negligible dosage of radiation from the Van Allen belt, but both the LEO and MEO orbits are quite susceptible to these belts. Figure 3-7 shows the radiation dose rates (in radians/year) as a function of altitude for varying shield

---

3 Walter Morgan and Gary Gordon, "Communications Satellite Handbook", Wiley Interscience, 1989, pg. 548.

thicknesses[4]. It can be seen from this figure that there is a dramatic increase in rate dosage above 700 km; increasing to a maximum at around 4000 km and then gradually reducing to lower rates at over 10,000 km (MEO).

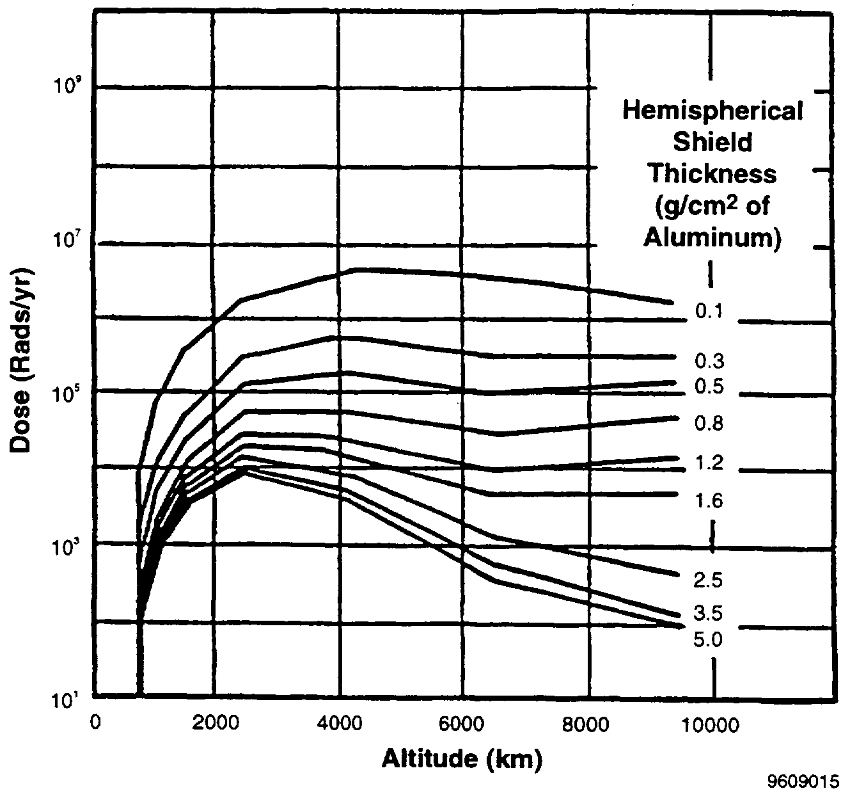


Figure 3-7. Radiation Dose Rates

Lifetime

The lifetime of LEO satellites is considerably reduced from those at MEO and GEO primarily because of increased atmospheric drag due to higher relative motion between the satellite and earth (LEO speeds are about 25,000 km/hr as compared to 17,600 km/hr at MEO and 0 km/hr at GEO). In addition to atmospheric drag, LEO systems also have much larger number of sun eclipses thereby requiring more thermal and power cycles. It is estimated that satellites in MEO have five times (satellites in LEO have 50 times) the number of thermal cycles than GEO satellites have. These increased thermal/power cycles create a greater load on the satellite power requirements and thereby limit their lifetime.

---

4 W. Larson & J. Wertz, "Space Mission Analysis and Design, "2nd Edition, Microcosm, Inc. 1991, page 202.

Time in Sight

The amount of time the satellite is in view of the ground station is an important parameter because it determines how often handovers will be required and how fast the beam steering will have to be to maintain the user terminal in sight of the satellite.

Figure 3-8 depicts the geometry used in calculating the arc length, S_t, traversed by the satellite while in view of a fixed ground station. Using simple geometrical formulas one can calculate S_t to be:

$$S_t = 2 \cdot (R_e + h) \cdot \left[\frac{\pi}{2} - \theta - \sin^{-1}\left\{ \left(\frac{R_e}{R_e + h} \right) \sin\left(\theta + \frac{\pi}{2} \right) \right\} \right] \qquad \text{E3-5}$$

where θ is the minimum elevation angle (radians), R_e is the radius of earth (6378 km) and h is the altitude of the satellite. Satellite velocity at a fixed altitude, h, is given by:

$$v_{sat} = \sqrt{\frac{G \cdot M}{(R_e + h)}} = \sqrt{\frac{398600.5}{(R_e + h)}} \quad (\text{km / sec}) \qquad \text{E3-6}$$

Using equations E3-5 and E3-6 one can determine the time period for which the satellite is visible for a fixed earth terminal. This time period has been computed and the results are plotted in Figure 3-9 as a function of the minimum elevation angle.

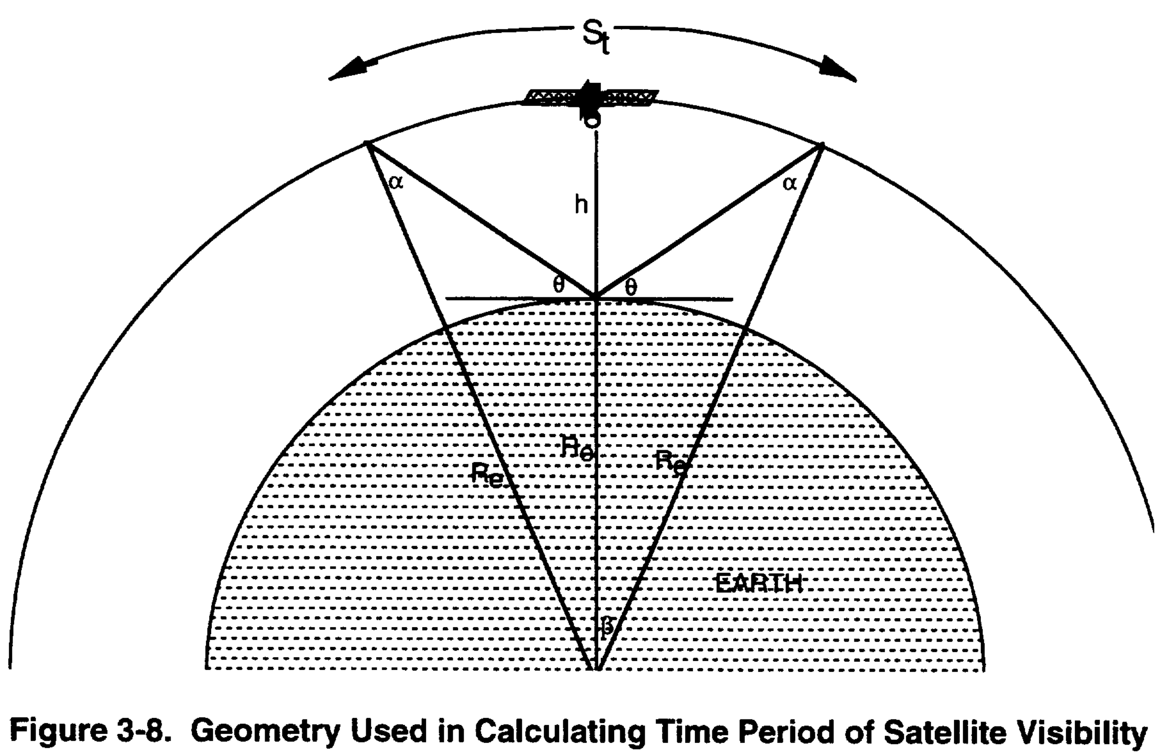


Figure 3-8. Geometry Used in Calculating Time Period of Satellite Visibility

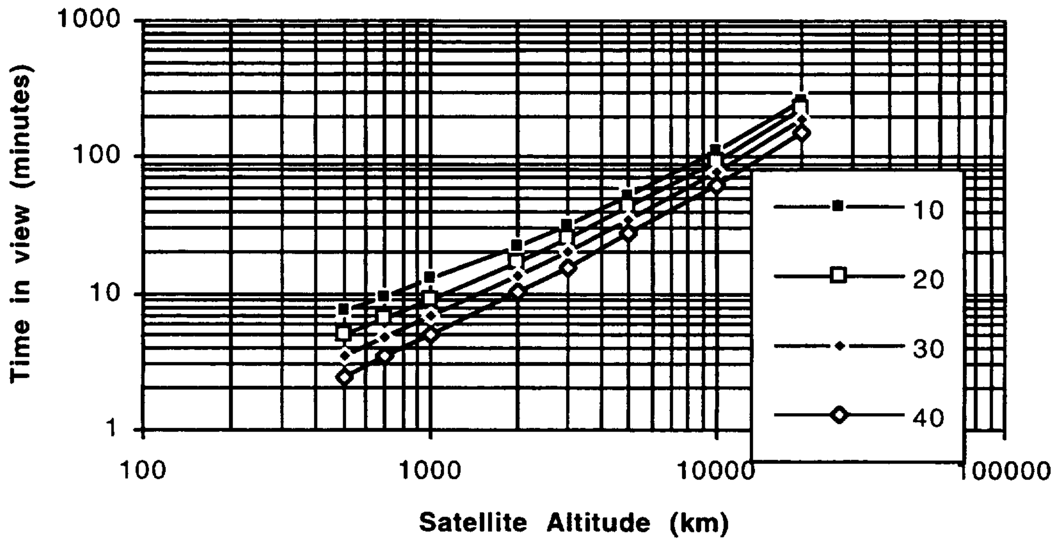


Figure 3-9. Satellite Visibility Time

From Figure 3-9 one realizes that LEO satellite viewing is very short (about 3.5 minutes for a satellite at an altitude of 700 km and assuming 40° elevation angle). MEO systems (10,000 to 15,000 km) have a maximum time in view of about 1.5 to 2 hours. This makes the complexity of satellite to earth-terminal routing much simpler for MEO systems than for LEO systems; both in terms of handover and beam-steering requirements.

Transmission Delay

The effect of transmission delay has been dealt with in great detail in Section 2 and will not be repeated here. Suffice is to say that LEO and MEO systems experience a transmission delay comparable to terrestrial systems, whereas GEO systems can have fairly large delays (>240 ms). The delay of a single-hop GEO system has been given consideration by the ATM standards body and such a system can satisfy the QoS requirements for delay for critical Class 1 services. However, GEO systems with multiple hops or using ISLs will not be able to satisfy ATM Class 1 service requirements.

Antenna Size

Antenna diameter is a very important design parameter for any satellite system. Its choice influences the beam size, antenna gain, spacecraft stability and satellite cost. In general, one would like to use the largest possible antenna diameter to increase the antenna gain and thereby increase the power level reaching the user terminal. The benefit of doing so is obvious; one can reduce the user terminal size.

There are, however, a number of constraints that limit the antenna diameter. First and foremost is the fact that larger antennas imply larger satellites and therefore higher costs in manufacturing and deployment. For LEO satellites, large size antennas create increased atmospheric drag which puts a strain on the fuel consumption and reduces the satellite lifetime. Another problem with large diameter antennas is their small beamwidths (the larger the antenna, the smaller the beam width). Most satellites are designed for a pointing accuracy of ±0.05°. However, this may not be accurate enough for beams that are very small in size and thus the losses from pointing error can become substantially larger. Overcoming these losses would require improved pointing accuracy, which implies the use of autotracking on an uplink pilot tone for each beam, thereby adding considerable complexity to the satellite system.

Another indirect constraint on large diameter antennas is the fact that the resultant smaller size beams necessitate a larger number of satellite beams to cover the FOV of the satellite. The number of beams required to cover the FOV for typical LEO, MEO and GEO systems is plotted in Figure 3-10. This figure has been derived using the parameters given in Table 3-3.

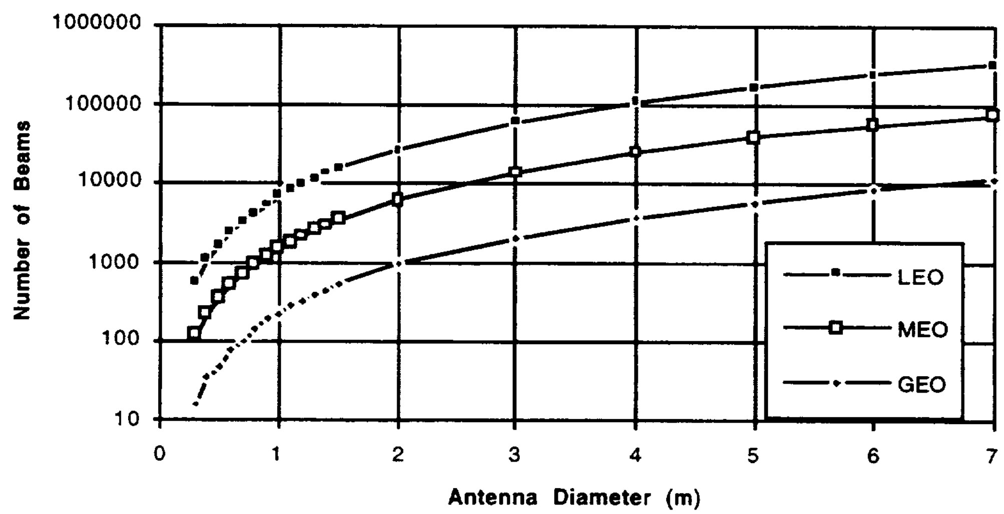


Figure 3-10. Number of Beams Required to Cover Each Satellite's FOV in Typical LEO, MEO and GEO Constellation Systems

Table 3-3. Parameters Chosen in Calculating Number of Beams in FOV of a Satellite

Parameter	LEO	MEO	GEO
Satellite altitude	700 km	10,355 km	35,786 km
Minimum elevation	40°	20°	20°
Satellite FOV	87.3°	41.98°	16.34°
HPBW (degree) F= Frequency (GHz) and D = satellite antenna diameter (m)	$HPBW = \frac{21}{FD}$	$HPBW = \frac{21}{FD}$	$HPBW = \frac{21}{FD}$
Number of Beams	$(FOV/HPBW)^2$	$(FOV/HPBW)^2$	$(FOV/HPBW)^2$

Using a larger size satellite antenna and its resultant larger number of beams required to cover the FOV of each satellite puts indirect constraints on the satellite's on-board processing requirements. As can be seen in Figure 3-10, the number of beams needed to illuminate the complete FOV of a satellite increases exponentially with increasing satellite antenna diameter. In comparing LEO, MEO and GEO systems, the lower the satellite altitude, the larger the number of beams in the FOV of a fixed diameter satellite antenna. A 0.5-meter antenna would require as much as 1681, 361 and 49 beams in LEO, MEO and GEO satellites, respectively, to cover the FOV of a satellite. It should be noted that in the analysis above, no overlapping of beams has been assumed. If overlapping beams are used (with an overlapping factor, k) the number of beams increases by a factor of k^2. Typically, an overlapping factor of $k=\sqrt{2}$ is used, which results in a doubling of the number of beams within the satellite FOV.

A Beam-Forming Network (BFN) is used in LEO and MEO systems will have to be able to create all the beams in the FOV of a given satellite; though at any moment in time many of these beams may be wasted over ocean regions with scarce traffic. GEO system satellites, because of their relative fixed position with respect to earth, can be designed to create a smaller number of beams than required to fill the whole FOV of each satellite. Thus areas with little or no traffic need not have any spot beams directed towards them. A scanning beam can be used for the rare cases where a ship or boat needs to communicate with other ships or with a terrestrial network.

On the basis of this analysis, one can conclude that LEO and MEO systems should use as small an antenna diameter as possible to avoid building large BFNs. Assuming, from practical considerations, a BFN limited in size to 500 non-overlapping beams, the satellite antenna diameter at 20 GHz is limited to about 0.25, 0.6 and 1.5 meters for LEO, MEO and GEO systems, respectively. GEO system satellites can use antenna diameters greater than 1.5 meters, (with the same BFN) by not "illuminating" all the beams in FOV of the satellite at 20 GHz. Most current GEO systems are designed for 1° to 2° HPBWs, which implies satellite antenna diameters of 0.5 to 1 meter at 20 GHz. An exception is NetSat 28, which assumes 0.2° beams. This requires approximately 5.5 meters (i.e., 18 ft.) diameter antennas.

Free-Space Loss

Free-Space Loss is dependent up satellite altitude and satellite elevation. Free-space loss varies directly in proportion to $4\pi L^2$ (where L is the path length between satellite and user terminal). This loss is plotted in Figure 3-11 as a function of both satellite altitude and elevation angle.

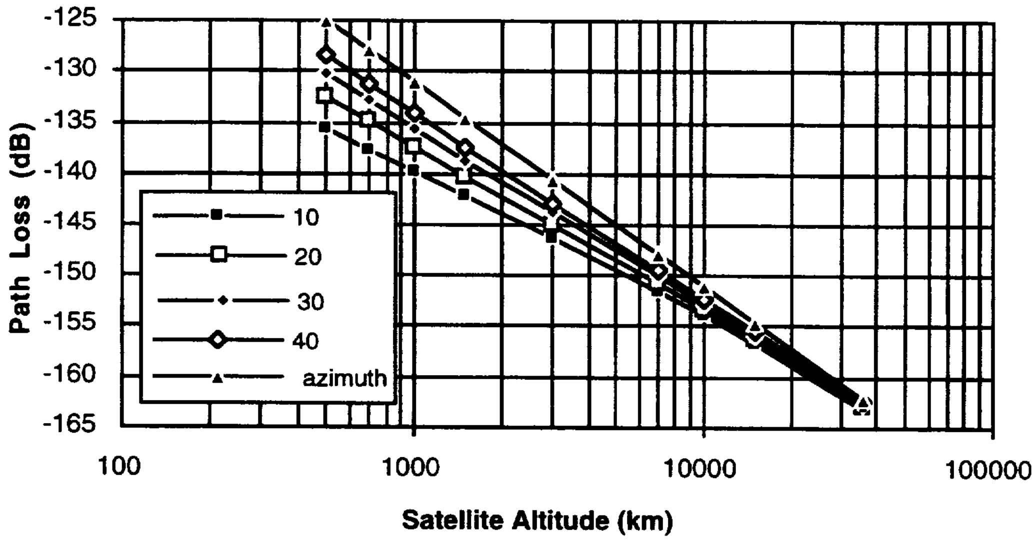


Figure 3-11. Free-Space Loss

The effect of elevation angle on free-space loss is more pronounced at the lower altitudes (such as LEO) than at the higher altitudes (such as GEO). The difference in path loss between LEO and GEO systems varies from about 25 to 33 dB (depending on choice of elevation angle for LEO system). Similarly, the path loss difference between MEO and GEO systems is about 10 dB. This loss can be made up for by increasing the satellite antenna gain through increasing its diameter (gain is proportional to the area of antenna). A GEO satellite will require a diameter 3 times larger than MEO and 30 times larger than LEO to maintain the same power level at the receiver (assuming all other losses remain the same). There is a limit to the size satellite antennas can be constructed. Often, one adjusts for the free-space loss by increasing the receiver antenna size or by reducing the fade margin.

3.3.2 Summary of Design Issue Comparisons of LEO, MEO and GEO Systems

The previous section described in detail some issues considered in designing satellite architectures. Quite often, the choice of one parameter has repercussions on various other parameters. For example, the choice of minimum elevation angle not only affects the satellite availability, but also the satellite FOV, the number of satellites needed for coverage, the transmission delay, and the free-space loss. It has also been pointed out that the choice of each design parameter has, often, totally different effects for GEO and MEO or LEO systems. Table 3-4 attempts to summarize these parameters of importance and their effects on LEO, MEO and GEO systems.

Table 3-4. Comparison of LEO, MEO and GEO Architectures

Parameter	LEO	MEO	GEO	Comment
Frequency Band of Operation	Ka	Ka	Ka	The Ka-band, as allocated by FCC, will be able to allow only 1 LEO/MEO system. Multiple GEO systems possible as long as satellites are spaced 2° apart.
Satellite Altitude (km)	500 to 1500	10,000-15,000	35,786	Satellite altitudes primarily determined by Van Allen radiation belts. Low orbits suffer from increased drag.
Satellite Lifetime (years)	3 to 8	8 to 10	10 to 15	The lower lifetime for lower orbit systems is because of greater atmospheric drag and greater radiation dosage.
Number of thermal/ power cycles	50 X GEO	5 X GEO	GEO	LEO/MEO systems have far more eclipses and power cycles than GEO systems. This limits power availability as well as satellite lifetime.
Elevation Angle	40°	20°	20°	Elevation angle for GEO determines satellite FOV. Choice based up on covering major traffic areas of the world. LEO/MEO systems experience full range of elevation angles at each earth location, as compared to GEO where the elevation angle is fixed.
Number of satellites needed for global coverage	200 to 850	10 to 12	3 to 4	This parameter is dependent up on the satellite altitude as well as elevation angle chosen.
Earth coverage per satellite (at 10° elevation)	2 to 5%	24 to 28%	34%	The further the satellite is from the earth, the larger is its FOV.
Free-Space Loss (dB)	128 to 138	151 to 154	163	The variations in loss are due to differences in elevation.
Inc. in antenna diameter needed to offset increase in loss wrt LEO	LEO	3 X LEO	30 X LEO	Antenna gain is proportional to the square of the antenna diameter
Time Delay (ms)	7 to 80	93 to 291	267 to 694	MEO/LEO systems will be able to support all ATM classes of service. GEO may not be able to support Type 1 QoS Class, but can support all others.
Satellite speed wrt earth (km/hr)	≈25,000	≈17,600	0	The rapid motion of LEO/MEO systems with respect to earth causes greater drag, more thermal/power cycles and more handovers.
Satellite period of rotation	≈ 100 minutes	≈ 3.5 to 8 hours	24 hrs	See above comment
Satellite time of view	5 to 10 minute	1 to 2 hours	24 hrs	LEO/MEO systems require much more rapid and sophisticated handover systems than do GEO systems.

3.3.3 Proposed Ka-Band Multimedia Satellite Systems

In 1995, fourteen organizations requested authorization from FCC to construct and launch Ka-band satellites for the 1998-2000 time frame so as to provide multimedia communications services. These systems and their attributes are summarized in Table 3-5. It is interesting to note that, with the exception of Teledesic which has been designed to operate in LEO, all other filings are specified for GEO. Many of the services that are envisioned by these systems are essentially the same as those of the NII/GII applications discussed in Section 1. User data rates vary from 16 kb/s to 8.448 Mb/s for low to medium rate services and up to 310 Mb/s for high speed systems. The BER target for most systems is 10-10 and link availability is in the range 99.2 to 99.9%.

As can be seen from Table 3-5, Astrolink, CyberStar, Galaxy/Spaceway, GE*Star, VoiceSpan and Orion provide global coverage using 3 to 15 satellites in GEO orbital locations. Teledesic will use 840 satellites (connected by means of ISLs) in LEO orbital locations to provide global coverage. Domestic and regional systems include EchoStar, KaStar, Millennium, NetSat 28, NorStar and PanAmSat. The per-satellite capacity of each system is typically 1 to 9 Gb/s. A notable exception to this is NetSat 28 which consists of a single GEO satellite with a total capacity of 772 Gb/s. NetSat 28 uses a proprietary on-board optical switch that allows 500 simultaneous users operating at 1.544 Mb/s to connect to 1000 beam locations thereby achieving its 772 Gb/s capacity (500x1000x1.544 Mb/s). The total system capacity varies from 1 to 90 Gb/s, exceptions being NetSat 28 which has 772 Gb/s capacity and Teledesic which has a total capacity of 4200 Gb/s (840 satellites X 5 Gb/s per satellite).

All these systems use a large number of narrow spot beams (each 1° to 3°) which are isolated using frequency, polarization, and space diversity techniques. The number of spot beams varies from 7 (Morning Star) to 576 (Teledesic). Most of the systems use multi-frequency TDMA and digital FDMA on the upstream and TDM/TDMA on the downstream. An exception to this is AT&T's Voice-Span which uses CDMA on the uplink and TDM on the downlink. Almost all of these systems perform on-board processing and switching (exceptions are Morning Star, GE*Star, NorStar and PanAmSat). Inter-satellite links, operating in the 60-GHz band, are used by Astrolink (1 Gb/s), CyberStar (two 1-Gb/s links), Galaxy/Spaceway (1 Gb/s), Teledesic (eight 155.52 Mb/s) and EchoStar (120 MHz).

Table 3-5. Emerging Ka-Band Satellite Systems Providing Multimedia Communications Services

	Astro-link	Cyber-Star	Galaxy/Spaceway	GE*Star	Morning Star	Teledesic	Voice-Span
Proponent	Lockheed Martin	Loral Aerospace Holding	Hughes Communi-cations GALAXY	GE American Communications	Morning Star Satellite Company	Teledesic Corp.	AT&T Corp.
Services	Videoconferencing, tele-commuting, wireless LANs, dist. learning, telemedicine, corporate training, monitor-ing/data retrieval, WAN/LAN & private interconnection	High-speed data, video telephony, medical & technical imaging, CAD/CAM data, image tx and private VSAT network services, tele-medicine/commuting	Voice, high-speed data, video conferencing telephony, global video distb., wideband inter-active services, VSAT, tele-imaging	High-speed data, video, audio, video telephony/conferenc-ing, database & on-line service access, DTH & multimedia	Telecom infrastructure for less-developed countries, VSAT-type services for home businesses, telecom-muting, disaster & emergency services, tele-medicine	Provide global info infrastructure in under-developed & rural com-munities, disaster & emergency service, dist. learning, tele-medicine, video con-ferencing, interactive multimedia	PC-based applicat-ions, tele-medicine, interactive video ser-vices, multimedia messaging & mail-boxes, information & database access, dist. learning, electronic sware distb.
Orbit Class	GEO	GEO	GEO	GEO	GEO	LEO	GEO
Number of Satellites	9	3	20	9	4	840 + up to 84 spare	12 + 4 partial ground spares
Number of Planes	5	3	15	5	4	21	7
Coverage	Global. No coverage for poles & ocean. 1° beams.	Global (high population density areas). No coverage for poles & ocean. 1° to 2° beams.	Global. 1° to 3° beams.	"Global"- not including Canada, Brazil, Africa, China & Russia. <1° beams.	"Global"- including Americas, Europe, Middle East, Asia, Australia & NZ. 1.6° beams.	Global (including poles and ocean)	Global, most but not all land masses are covered. 1° beams.
Operating Frequency	Ka-band	Ka-band	Ka- and Ku-bands	Ka-band	Hybrid Ka- and Ku-band	Ka-band	Ka-band
Satellite Capacity (each)	8 Gb/s (or 4 Gb/s duplex) BW of 9 GHz	4.9 Gb/s (or 2.5 Gb/s full duplex)	4.4 Gb/s (or 2.2 Gb/s full duplex)	Approx 6 Gb/s (44 times 243 MHz)	Over 150 compressed video (515 Mb/s)	5 Gb/s (or 2.5 Gb/s full duplex)	5.9 Gb/s (or 2.95 Gb/s full duplex)
# of Transponders	68	54	48 (Ka) & 24 (Ku)	44	16	64	32 or 64
Transponder BW	125 MHz	125 MHz	125 MHz (Ka) & 36 MHz (Ku)	243	50 MHz (Ka) & 24 MHz (Ku)	396 MHz	120 MHz
Total # beams	192 spot beams (64 HBs) +3 FB (gateway)+ 1 SB (gateway)	27 spot beams (x2 polarization)	48	44 spot beams (4 bands of 243 MHz, with 11 re-use)	7-spot beams (limited global)	576 (64 HB)	32 or 64
Onboard Processing	Regeneration and cell switching	Regeneration and baseband switching	Regeneration and BB switching/routing.	Bent-pipe	Bent-pipe	Regeneration and fast packet switching	Message and beam switching, ATM switching/reouting

Table 3-5. Emerging Ka-Band Satellite Systems Providing Multimedia Communications Services (Continued)

	Astro-link	Cyber-Star	Galaxy/Spaceway	GE*Star	Morning Star	Teledesic	Voice-Span
ISL frequencies (GHz)	Two 1 Gb/s ISLs at 60 GHz (ISL dist betn. 37,000 to 62,000 km)	Two 1 Gb/s ISLs at 60 GHz	1 Gb/s	NA	NA	Eight 155.52 Mb/s ISLs (60 GHz)	NA
Access Schemes	•Uplink: High BW TDMA/ FDMA @128/768 Kb/s, 4.096 Mb/s & 8.448 Mb/s •Downlink: High BW TDMA @130 Mb/s. •Gateway: 310 Mb/s TDMA uplink/ TDM downlink	Uplink: FDM/TDMA at 384 kb/s, 1.544 Mb/s & 3.088 Mb/s Downlink: TDM at 92 Mb/s	DAMA (bandwidth on demand)	uplink: FDMA downlink: TDMA	ALOHA used for requesting assignment of video channels	Uplink: hopping beam FDMA at 224 kb/s Downlink: hopping beam TDMA at 324 Mb/s	Uplink: CDMA Downlink: TDM
EIRP	56 dBW	57 dBW	54 dBW	58 dBW	53 dBW	50 dBW	56 dBW
User Data Rates	16 kb/s to 8.448 Mb/s, 310 Mb/s (gateway)	16 kb/s to 3.088 Mb/s	16 kb/s to 1.5 Mb/s, as well as asymmetric 6 Mb/s	384 kb/s to 40 Mb/s	30 Mb/s used for 10 compressed video channels	16 kb/s to 2.048 Mb/s; 155.52 Mb/s to 1.244 Gb/s (special applications)	32 kb/s voice, 144 kb/s to 1.544 Mb/s multimedia
Performance	10E-10 BER. 99.5%-99.9% link availability	10E-10 BER. 99.5% link availability	10E-10 BER. 99.7% link availability	NA	10E-06 BER. Link availability not specified	10E-09 BER. 99.9% link availability	10E-10 BER. 99.7% link availability
Terminal Size	65 cm to 1.2 m, 2.4 m to 4.5 m (gateway)	70 cm to 3 m	45 cm to 1.2 m (66 cm at Ka-band & 45 cm at Ku-band)	75 cm to 1.5 m	60 cm, used for both bands (retrofitting existing Ku-band receive antennas to operate in Ka-band)	16 cm to 1.6 m	66 cm to 1.8 m
Dry Mass	2,185 kg	1,900 kg	2,000 kg	1,768 kg	1,360 kg	747 kg	2,200 kg
Satellite Life	10-12 yrs	12 years	15 years	15 years	12 years	10 years	12 years
Capital Expenditure	$4 billion	$1.2 billion	$5.1 billion	NA	$823 million	$9 billion	NA
Capital Cost*	$7,100	$10,200	$7,400	NA	Not applicable for 1-way VOD	$270	NA
Date of Operation	2000 to 2001	2000	1998 to 2004	44 months after authorization	Around 2000	2001	2000 to 2003
References	FCC File Numbers: 187-SAT-AMEND-95, 188/189-SAT-P/LA-95	FCC File# 109-SAT-P/LA-95 & 110-SAT-P/LA-95	FCC File # 174-SAT-P/LA-95 through181-SAT-P/LA-95	FCC File # 169-SAT-P/LA-95 through173-SAT-P/LA-95	FCC File # 190-SAT-P/LA-95 through193-SAT-P/LA-95	FCC File # 22-DSS-P/LA-(840), 43-SAT-AMEND-95, 127-SAT-AMEND-95	FCC File # 156-SAT-P/LA-95 through162-SAT-P/LA-95

Table 3-5. Emerging Ka-Band Satellite Systems Providing Multimedia Communications Services (Continued)

	Echostar	KaStar	Millenium	NetSat 28	NorStar	Orion	PanAm-Sat
Proponent	Echostar Satellite Corp.	KaStar Satellite Comm. Corp.	Comm. Inc., (subsidiary of Motorola)	NetSat 28 Company	Norris Satellite Comm.	Orion Network Systems	PanAmSat Licensee Corp.
Services	Tie with DBS service, videophone, video conferencing, VSAT-services, dist . learning, telecommuting, digital messaging, CAD/ CAM & sat. news gathering	Terrestrial & undersea cable restoration, remote medical diagnostics/treatment, dist. learning, video-conferencing, DARS, air traffic control com., VSAT markets, video-phone, disaster & emergency services	ATM services, telecommuting, telemedicine, PPV TV, video games, home health, home shopping, multimedia video, business training, WANs, internet access, collaborative services	Video, data, pay-per view movies, VOD, home banking/shopping, games, video-conferencing, health care, dist. learning, telecommuting, telemedicine, govt. services, alarm reporting	Not Available	Digital audio, compressed video, SNG, voice & multimedia services	High-quality video, voice, data, digital audio, VSAT (at 64 & 128 kbps) and other specialized services
Orbit Class	GEO	GEO	GEO	GEO	GEO	GEO	GEO
Number of Satellites	2 (one co-located with DBS sat.)	2	4	1	1	3	2
Number of Planes	2	2	4	1	1	3	2
Coverage	US domestic (CONUS plus Alaska, Hawaii and Puerto Rico). 1° beams.	US domestic (CONUS Alaska & Hawaii) plus part of Mexico and Caribbean). 1° beams.	Full coverage of US & Latin America. 1° beams.	CONUS. Approx. 0.2° beams.	USA	"Global", 2 satellites with CONUS coverage & 1 with coverage of populated regions of IOR. 1° beams.	Americas, Europe and West Africa (AOR coverage). 1°/2°/3° beams.
Operating Frequency	Ka-band	Ka-band	Ka-band	Ka-band	Ka-band	Ka-band	Ka-band
Satellite Capacity	Approx. 6 GHz (or 5.76 Gb/s)	Approx. 6 GHz (or 7.46 Gb/s)	Approx. 7.125 GHz (or 4 Gb/s)	772 Gb/s (386 Gb/s full duplex)	576 MHz (24 transponders each 24 MHz)	2.88 Gb/s (32 transponders with 114 MHz BW each)	1.24 Gb/s (24 transponders of 54 MHz each)
# of Transponders	48	48	57	1000	24	32	24
Transponder BW	120 MHz	125 MHz	250 MHz	150 MHz	24 MHz	114 MHz	54 MHz
Total # beams	24 dual pol. narrow spot beams + 1 elliptical CONUS for service request	24 dual pol. fixed spot beams, 2 SWB & 1 elliptical CONUS for service request	25 dual pol. fixed spot beams + 7 single pol spots (using 57 transponders)	1000 spot beams	8 spot beams	25 fixed and 2 steerable spot beams	12 steerable spots (4 with 1°, 4 with 2° & 4 with 3°)
Onboard Processing	Regeneration, switching & routing	Dynamic processing, switching and routing (BB proc. for low-speed & ATM for broadband)	"ATM-like" on-board switching/routing	Optical crossbar switching of digital FDMA carriers	Bent-pipe	Regeneration and dynamic routing on packet-by-packet basis	Bent-pipe

Table 3-5. Emerging Ka-Band Satellite Systems Providing Multimedia Communications Services (Continued)

	Echostar	KaStar	Millenium	NetSat 28	NorStar	Orion	PanAm-Sat
ISL frequencies (GHz)	120 MHz ISL at 60 GHz	120 MHz ISL at 60 GHz	59 to 64 GHz band	None	None	May consider	No
Access Schemes	Uplink: FDMA Downlink: TDM	Uplink: FDMA Downlink: TDM	Uplink: FDMA Downlink: TDM	FDMA/TDM (500 T1s per beam)	NA	Uplink: FDMA Downlink: TDMA	Varied, depends upon application
EIRP	57 dBW	56 dBW	58 dBW	60 dBW	NA	56 dBW	49-58 dBW
User Data Rates	384 kb/s to 1.544 Mb/s, as well as 155.52 Mb/s	384 kb/s to 1.544 Mb/s, as well as 155.52 Mb/s	384 kb/s to 51.84 Mb/s	64 kb/s to 1.544 Mb/s	NA	384 kb/s to 3.088 Mb/s	64 kb/s to 45 Mb/s
Performance	10E-10 BER. 99.5 to 99.9% link availability	10E-10 BER. 99.2 to 99.9% link availability	10E-10 BER. 99.5 to 99.9% link availability	10E-6 to 10E-10 BER. 99.2% link availability	NA	10E-10 BER. 99.5% link availability	10E-10 BER. 99.3 to 99.8% link availability
Terminal Size	NA	66 cm to 2.0 m	70 cm and larger	30 cm to 1.5 m	90 cm to 4 m	70 cm to 2.5 m	Depends upon application
Dry Mass	1,860 kg	1,690 kg	1,800 kg	1,700 kg	NA	1,550 kg	NA
Satellite Life	12 years	10 years	10 years	12 years	10 years	13 years	15 years
Capital Expenditure	$340 million	$645 million	$2.3 billion	$250 million	NA	Approx. $725 million	Approx. $409 million
Capital Cost for a 64-kbps full duplex	$1,900	$2,800	$9,200	$40	NA	Approx. $5,400	Approx. $10,500
Date of Circuit Operation	46 to 48 months after authorization	54 to 66 months after authorization	1998 to 2001	2000	NA	2000	1999
References	FCC File # 167-SAT-P/LA-95 &168-SAT-P/LA-95	FCC File # 128-SAT-P/LA-95	FCC File # 163-SAT-P/LA-95 through 166-SAT-P/LA-95	FCC File #: 194-SAT-P/LA-95	FCC Application: July 16 1990	FCC File # 195-SAT-P/LA-95 through 197-SAT-P/LA-95	FCC File # 198-SAT-P/LA-95 through 199-SAT-P/LA-95

All the FCC filings specify a user terminal with an antenna diameter varying between 30 cm to about one meter. Larger diameter terminals (up to 3 meters) are available for gateway stations. The output power from the user terminal is limited to between 1 to 10 W and for gateway terminals to between 10 to 200 W.

Of the various architectures that have been filed, Astrolink and Teledesic systems are the only global systems with high enough system capacity (>70 Gb/s) to be able to carry the traffic load of future NII/GII applications and services. The NetSat 28 system, though not global (it only supports traffic across CONUS), needs also to be given consideration based on its sheer system capacity (772 Gb/s). The architectures of these three systems are briefly examined here to provide a starting point for designing future third-generation systems.

Astrolink

The Astrolink system makes use of three orbital locations to provide maximized coverage over the various land masses, and two orbital locations over the oceans (one over Atlantic and the other over Oceania). The five orbital locations are chosen to maximize traffic over each satellite and minimize ISL use to only those cases requiring connectivity over separate continents. There are a total of nine satellites, with each orbital location (except the Oceania region) having a pair of co-located satellites. Astrolink uses a mixture of space diversity (64 hopping spot beams with 192 beam positions), frequency diversity (it splits the available Ka-band frequencies into 4 bands of frequencies and reuses these bands across the satellite FOV) and polarization diversity (each satellite uses a separate polarization to effectively double available bandwidth). Astrolink is designed to provide a minimum of 20° elevation angle across most of the populated regions of the world (located within ± 60° latitudes). Each Astrolink satellite uses three beams to switch between 64 active signal paths to achieve up to 192 contiguous spot beams, each with 1° half-power beam width, to synthesize its primary coverage area for user terminals. Each Astrolink satellite supports a total of more than 10,000 simultaneous full duplex 384 kb/s circuits, or equivalent number of other data rates. On the user terminal uplinks, a hybrid TDMA/FDMA access scheme is employed with user channel bursts of 128 kb/s, 384 kb/s, 768 kb/s, 4.096 Mb/s and 8.448 Mb/s available, depending on user requirements. The user terminal downlink employs a broadband 130 Mb/s TDM signal. In addition to the 192 position user hopping beams, Astrolink also provides three fixed gateway spot beams and one steerable gateway beam. Gateway terminals, located at major fiber trunk locations around the world, will provide connectivity between any ground terminal user and terrestrial fiber networks through the use of these high bandwidth gateway beams. Each gateway beam can support up to 310 Mb/s (equivalent to two OC-3s) data rates. Astrolink uses on-board processing and fast packet switching to provide required user connectivity and to improve overall link performance by isolating the effects of uplink and downlink propagation impairments. Astrolink is designed for user terminal sizes between 65 cm to 1.2 meters. Gateway terminals are larger and can be between 2.4 to 4.5 meters. Astrolink is designed to operate around 2001, with a lifetime per satellite of approximately 10 to 12 years.

Teledesic

The Teledesic system operates in LEO (at an altitude between 695 and 705 km) and makes use of 21 orbital planes with 40 satellites per plane for a total of 840 satellites. Four spare satellites are kept per plane (for a total of 84 spares) to be activated in case of primary satellite malfunction. The Teledesic system is based on the non-hierarchical "geodesic" or mesh network, which is increasingly becoming the model for terrestrial ATM-based communication network systems. Thus, each satellite in the Teledesic system is linked with four satellites on each plane (2 in front, 2 behind) and with two satellites in each adjacent plane for an overall total of eight links. Such a mesh system is designed to be tolerant to faults and local congestion. The Teledesic system maps the surface of the earth into a grid of 20,000 "supercells". Each supercell is a square of 160 km and is composed of 9 cells arranged in a 3x3 grid. The satellite system is designed for very high availability of 99.9 percent over most of the United States. This requires a terminal elevation angle of 40° or greater. The field of view (or footprint) of each Teledesic satellite, assuming a minimum of 40° elevation, is a circle of radius 706 km. The satellite footprint encompasses a maximum of 64 supercells or 576 cells. To avoid increasing the complexity of the user terminal, Teledesic uses beam-steering (through the use of phase array antennas) to keep the beam fixed to the user terminal cell. The beam steering compensates for both the satellite's, as well as the earth's motion. Handover is achieved using Region Orientated Frequency Assignment (ROFA) techniques described earlier in subsection 2.3.6. Channel resources (i.e., frequencies and time slots) are associated with each cell and are managed by the current "serving" satellite. As long as a terminal remains within the same cell, it maintains the same channel assignment for the duration of the call. At any instant, each fixed supercell is served by one of 64 transmit and one of 64 receive beams on one of the Teledesic satellites. Each scanning beam supports up to 1440 16-kb/s channels (or multiples of this up to 2.048 Mb/s, and for special applications from 155.52 Mb/s to 1.24416 Gb/s). FDMA is used for the uplink and asynchronous TDMA is used for the downlink. The on-board processor uses fast packet switching (each cell is 64 bytes or 512 bits long) based on an ATM-like structure. The user terminals range in diameters from 16 cm (8 cm for Mobile terminals) to 1.8 meters determined by the terminal's maximum channel rate.

NetSat 28

The NetSat 28 system consists of a single geostationary Ka-band satellite located at 103°W, designed to provide multimedia services across the regional CONUS area. The system can be best envisioned as a gigantic switch in the sky capable of connecting over 500 simultaneous users operating at up to T1 rates (1.544 Mb/s) through 1000 beams. The satellites use a Gregorian antenna (20x20 ft.) with a focal feed array consisting of 1000 feed elements (arranged in an array of 22x23 elements) used to create 1000 spot beams covering the continental United States. Each beam averages 70 miles in diameter (i.e., 0.2° beam width).

NetSat 28 will provide data rate services from 64 kb/s to 1.544 Mb/s with bandwidth assigned on demand and user terminals varying from 0.3 meters (for home and mobile users) to 1.5 meters (for business and movie-quality video). The uplink access is through FDMA, while the downlink uses TDM. The heart of the system is a proprietary 1000 input to 1000 output optical crossbar switch. Uplink signals from 500 beams are received by each antenna element array, filtered (to separate uplink and downlink frequency spectrum) and then converted to an optical signal. The resulting 1000 optical signals, each containing 500 FDMA signals, are switched and re-routed to the appropriate beam feed. These optical output beams are converted back to Ka-band signals, amplified and transmitted through the specific antenna feed for which it was destined.

With a few modifications, the NetSat 28 system, with its strong optical processing capability, can be used to create a global system. This would require a larger number of satellites, interconnected by means of ISLs. Because the switch is optical, it could very well be used to set up optical ISLs.

3.4 CONCEPTUAL ARCHITECTURES FOR THIRD-GENERATION SATELLITE NETWORKS

In this section, conceptual architectures for third-generation systems will be formulated based on the requirements enumerated in subsection 3.2 and the set of system parameters evaluated in subsection 3.3. Due to their radically differing modes of operation, GEO and MEO/LEO architecture designs are considered separately.

3.4.1 GEO Systems

GEO systems, with their greater spectrum availability (satellites can share the same bands provided they are spatially 2° apart), their lower number of satellites required for global coverage, their reduced number of beams and their simpler network control (no tracking or handovers needed), are the preferred method of choice for third-generation satellite networks. The only key disadvantages to the GEO systems are their increased propagation loss and higher delays. The increased propagation loss can be offset somewhat by increasing satellite antenna diameter. The satellite delay is a problem for only the Class 1 (stringent) ATM services; and that too, only for the specific conditions when multiple-hops and/or ISLs are employed by the GEO system. As discussed in Section 2, the problems of delay experienced by satellite users are similar to problems encountered in LAN systems operating over terrestrial ATM links. Many new standards, such as COMSAT's SSCOP, are currently being formulated that are designed to be more tolerant to large delays. Once these standards become accepted, GEO-satellite systems (including those with ISLs and multi-hops), will be able to offer the QoS needed for even the most stringent services used in the future NII/GII applications.

A number of design architectures are possible for GEO systems. Two designs that have been discussed earlier (subsection 3.3.3) are those of Astrolink and NetSat 28. The conceptual design that we propose here uses a mixture of the on-board processing capabilities of the NetSat 28 design with the global coverage of the Astrolink design. In addition, some advanced capabilities for satellite systems will be assumed. These include: greater on-board processing capability which will allow the processing of larger number of beams, more advanced ISL and Phased-Array Antenna (PAA) capabilities, and availability of dual-polarization capability on each satellite. The conceptual GEO architecture's frequency of operation, bandwidth allocation, satellite geometry, data rates, spacecraft antenna size and access techniques are elaborated in the following sections.

Frequency Band of Operation

As discussed in subsection 3.3.1, only the Ka-band has both the bandwidth and component maturity to support third-generation satellite systems that will be used for carrying NII/GII applications within the 2000 to 2005 time-frame. Figure 3-12 shows the Ka-band frequency allocation proposed for the GEO concept design. This frequency plan closely follows the WARC-95 guidelines and is assumed identical for all satellites in this concept design.

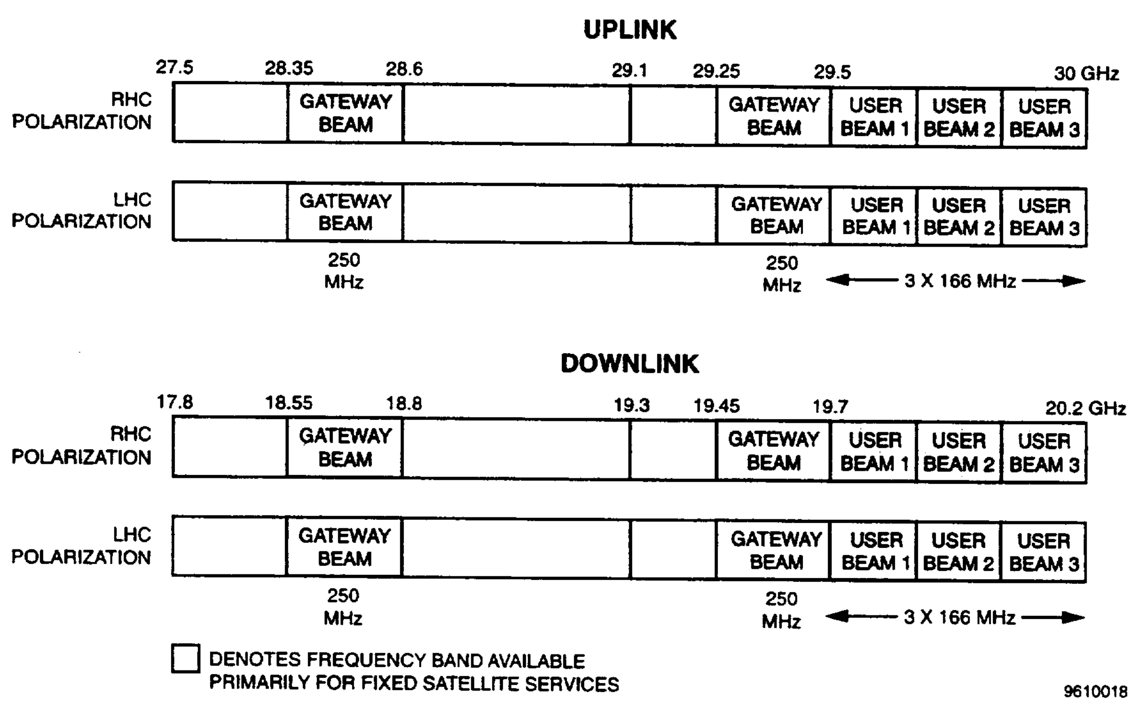


Figure 3-12. Frequency Allocation for Ka-Band GEO Satellite Architecture

Each satellite will have different antennas for uplink and downlink and will be able to operate with two orthogonal polarizations on each antenna. This provides a total transmission bandwidth of 2 GHz for uplink and 2 GHz for downlink. Of this bandwidth, the two lower sets of frequencies in each band (i.e., 28.35 to 28.6 GHz and 29.25 to 29.5 GHz in the uplink band and 18.55 to 18.8 GHz and 19.45-19.7 GHz in the downlink band) are required to be shared with other Fixed Satellite Services (FSS) and/or Mobile Satellite Services (MSS) feeds. The sharing requirements of these bands make them unsuitable for being assigned to user terminals and are, instead, assigned to larger gateway terminals. The frequency coordination for the gateway terminals with the other FSS or MSS terminals should be relatively simple. Ten steerable gateway beams are employed, each with 1-GHz bandwidth (2 polarizations x 2 x 250 MHz/polarization).

The 500-MHz band in the frequency range 29.5 to 30 GHz (uplink) and 19.7 to 20.2 GHz (downlink) does not require to be shared with the fixed or mobile services. This eliminates the need for difficult frequency coordination of the ground terminal with the terrestrial microwave links and is, therefore, perfect for assigning to user terminals. The user terminals are connected by means of fixed spot beams which "illuminate" the user location. To avoid interference between adjacent spot beams, the 500-MHz user band (1 GHz with polarization) is subdivided into three channels, each with approximately 166 MHz bandwidth. The 166-MHz user terminal channels are allocated to the fixed spot beams using a three-cell frequency re-use pattern as shown in Figure 3-13. The numbers within each beam indicate which of the 166 MHz channels has been assigned to that beam. Note that no adjacent beams have the same channels allocated to them. This achieves a reasonably good beam-to-beam isolation.

Finally, in addition to the user and gateway channels, ISLs will be required to carry 10 GHz bandwidth. The ISLs will operate in the 50/60 GHz frequency band to avoid any interference with the Ka-band channels.

The Satellite Constellation Geometry

While three satellites are generally sufficient to provide global coverage for GEO systems, using more satellites can increase the availability of a connection by providing path diversity to the user terminal. There seems to be, however, a diminishing rate of return as the number of satellites increases beyond five[5]. The Astrolink design is optimally

---

[5] FCC Filing 187-SAT-AMEND-95, Astrolink Communications Satellite System, submitted Sept. 27, 1995.

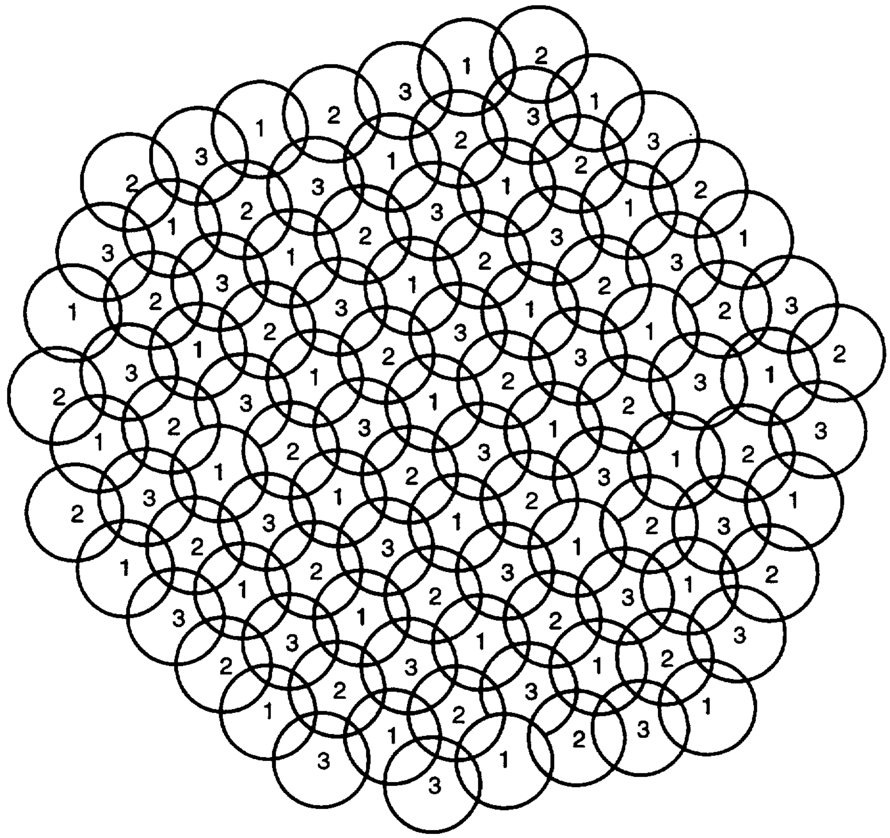


Figure 3-13. Frequency Re-Use Pattern

configured in this respect, and a similar design is suggested for use in the conceptual architecture. The Astrolink system consists of five GEO satellites located at:

1. 96°W (Americas)
2. 37°E (Europe, Africa, West Asia)
3. 115° E (East Asia, Australasia)
4. 29° W (Atlantic Ocean region)
5. 168° E (Oceania).

Some areas, such as Eastern U.S. and the increasingly important regions of Asia, have double coverage. These locations can use added coverage to either increase their communications bandwidth capacity or to increase their system availability by providing diversity channels. Another major advantage of this geometry is that this allows each satellite to dedicate its resources to providing regional/countrywide services. Only the infrequent "calls" that are international need be routed through ISLs. Limiting ISL use to a minimum is important from the point of view of satisfying the stringent class (i.e., Type 1) of ATM services, which does not allow for more than 370 ms of satellite delay (i.e., only a single hop).

Most proposed Ka-band satellite networks have designed their systems to produce 1° to 2° spot beams. Assuming a minimum elevation angle of 20° (see subsection 3.2), this limits the number of beams within the FOV of the satellite to be approximately 267 (for 1°) or 67 (for 2°). As an example, Astrolink satellites are designed to provide 1° beams, thus setting the number of beams within its FOV to 267. By avoiding the beam locations where there is little or no traffic (such as the ocean areas) Astrolink limits the maximum number of beams to 192. A further simplification by Astrolink is to limit the number of active beams to 64 and use beam-hopping between three positions to cover all the 192 spots.

While the Astrolink design is reasonably good, it provides only 9 GHz of bandwidth per satellite. This may not be enough to carry all the high-bandwidth applications required for the future NII/GII traffic. We believe that there is sufficient improvement in on-board processing technology to be able to provide a four-fold increase in the number of active beams. This can be achieved by reducing the beam-width to half of Astrolink's design (i.e., 0.5°). A 0.5° HPBW would allow as many as 1068 beams within the FOV of the GEO satellite. Assuming a similar coverage as Astrolink, of these 1068 beams only about 768 need be used. By hopping between three positions, 256 active beams can cover 768 beam locations. Many of the satellites (for example, the satellite over Oceania) will actually require far fewer spot beams to illuminate traffic regions and might require only hopping between two beam positions. Using 256 active spot beams (each with 332 MHz bandwidth) along with the ten gateway beams (each with 1-GHz bandwidth) provides 95 GHz of transmission bandwidth per satellite. This is approximately a ten-fold increase compared to the Astrolink design and should be sufficient to carry future NII/GII application traffic.

Access Techniques

The aggregate bandwidth available for each user terminal is 332 MHz which should be able to support two STM-1 (i.e., STS-3) signals at a combined 311.04 Mb/s data rate. ISDN rates of 144 kb/s with multiples up to 1.544 Mb/s, as well as standard telecom signal hierarchies such as 2.048 Mb/s, 6.312 Mb/s up to the STS-1 rate of 51.84 Mb/s can be made available to the user on demand. User terminals can access the allocated uplink channel in TDMA format, receiving the burst time plan information from the satellite to synchronize and control their data bursts within the beam switched TDMA frame. Depending upon its required information rate, each user terminal can be allocated the required number of TDMA time slots. The TDMA burst plan is dynamically variable to provide the greatest flexibility to changing user traffic. Higher data rates than 51.84 Mb/s can be implemented at a user uplink site by accessing multiple-uplink channels in parallel, using an appropriate data multiplexer. The ten steerable gateway beams (each with 500 MHz bandwidth) can be used to carry even higher data rate signals of up to 622 Mb/s (STS-12 or STM-4) data rates. Fixed gateway terminals with larger size antennas will be able to accommodate the higher data rates.

Each downlink spot beam will have the full 322 MHz bandwidth available to be shared between user terminals by means of a simple TDM scheme. All ground terminals receive 311.04 Mb/s data stream, and extract and buffer only the data which is addressed to them.

All communication links will use advanced modulation techniques (such as QPSK and 8-PSK) and coding (Viterbi and Reed-Solomon) as well interleaving techniques to support BER of better than 10^{-10}. Such a BER performance goes beyond ATM specifications and satisfies even the most stringent class of services.

Antenna Size

The antenna size is determined mainly by the HPBW (see Table 3-3). To create 0.5° beams for the 20-GHz downlink, the antenna size has to be 2.1 meters in diameter. The resultant 768 spot beams per satellite can be created in two ways; through (a) a focal-plane fed array or (b) phased-array antenna. Using the focal-plane fed array method, antenna elements are positioned at the focal point of a reflector, such that each antenna element produces a separate beam in a fixed direction. Thus, each spot beam would require a separate antenna and separate associated Traveling Wave Tube Amplifier (TWTA). The phased-array antenna approach, on the other hand, makes use of a beam-forming network (which could be made up of either active or passive elements), and connects each beam to all the antenna elements. The beam direction is determined by the phase taper created by the beam-forming network. The phased-array antenna approach has the following advantages over the focal-plane method:

a. Since each beam is shared by many antenna elements, each antenna element can radiate at lower power levels. This reduces the complexity, mass, power and volume of the satellite

b. The Phased-Array Antenna (PAA) approach allows the formation of shaped or contoured beams. This can be of importance in cases where there is too much interference from other systems

c. The PAA approach allows graceful degradation of components. In other words, a single antenna element failure does not result in catastrophic failure (unlike the focal plane fed approach, where a single antenna element failure can result in the loss of a beam).

The only major drawback of the PAA approach is its increased complexity. With the considerable technological progress currently underway in PAA designs, its use on-board third generation satellite systems seems reasonable. One key design parameter of the PAA approach is the number of antenna elements required. Given the antenna's beam diameter, the number of antenna elements will depend on the spacing required between the elements. To avoid grating lobe problems, the antenna elements need to be spaced ≤ 0.5 λ apart[6] (λ is the wavelength of the RF signal) For a 20-GHz downlink, this mean that the antenna elements should be spaced no more than 0.75 cm apart. With this spacing about 267 elements can be fitted in length along a 2.1-meter diameter; requiring a total of around 61,000 elements to cover the full antenna FOV. Designing a beam former to connect 768 spot beams to a PAA of size 61,000 is beyond the capability of present day systems and so will not be further considered. However, it is possible to design a focal-plane fed system with 768 antenna elements and non-uniform excitation of the elements.

Communications Payload

The complete communications payload is shown in block diagram format in Figure 3-14.

The receive section consists of receive antennas, 256 switches (3x1), filters, downconverters and demodulators. The transmit section is essentially a mirror image of the receive section and consists of modulators, upconverters, High Power Amplifiers (HPAs), and 1x3 switches connected to transmit antennas. There are essentially three sets of signals shown in Figure 3-14. These are:

1. The user signals from the user terminals connected to the satellite through fixed spot beams
2. The gateway signals from the gateway terminals connected to the satellite through steered beams
3. The ISL signals which provide connections to other satellites.

The user signals from the 768 beams are detected by the antenna elements and a 3X1 switch selects one beam out of every three (in satellites with less number of beams, 2X1 switches may be used). Each of the 256 selected signals are filtered, amplified (through Low-Noise Amplifiers [LNAs]), downconverted, demodulated and routed through the On-Board Processor (OBP) to the required output port. The whole sequence of events is reversed at the transmitter section.

The ten gateway signals are directly filtered, downconverted and demodulated without having to pass through a BFN or switch. Similarly, on the transmit side, the gateway signals are modulated, upconverted and amplified (through HPAs) before radiating to the earth from a separate antenna.

---

6 Eli Brookner, "Practical Phased Array Antenna Systems," Artech House, 1991, pp. 2-4.

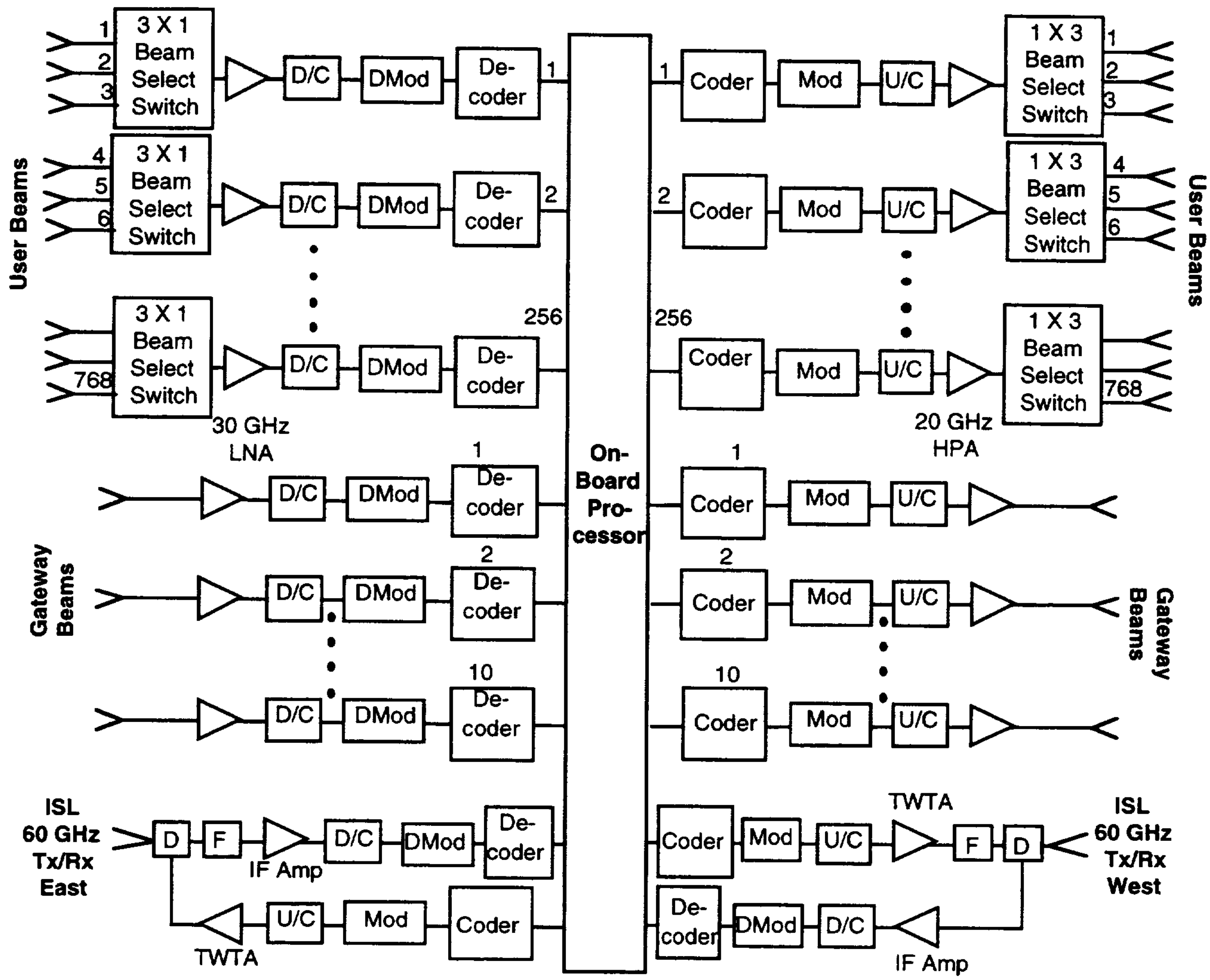


Figure 3-14. Functional Block Diagram of a GEO Communications Payload

Two ISLs are assumed per satellite, thus providing each satellite with connectivity to two other satellites. A common transmit/receive antenna is used for the ISLs and the signals are routed through diplexers to the proper transmit or receive channels.

The OBP digitally filters and demodulates the individual uplink receive channels and then separates the individual bit-stream cell packets according to their destination (downlink, gateway or ISL) beam. The demodulated gateway uplink and ISL incoming data streams are also separated by the OBP into cell packets according to destination. These data cell packet streams are then combined into the appropriate transmit channel and digitally modulated onto the appropriate Intermediate Frequency (IF) carrier at the OBP output.

User Terminals

Three major constraints define the user terminal. These are:

1. Its size should not exceed one meter (so as to be portable)
2. Its output power should not exceed 5 W (to limit the size and weight of the terminal)
3. The receive power flux density should not exceed -105 dBW/m^2 per MHz (set by the ITU).

Detailed power budget calculations must be performed to determine if the Carrier-to-Noise Ratio (CNR) of the system is sufficient to provide the BER performance required. Preliminary calculations indicate that a one-meter user terminal antenna diameter can provide enough CNR for burst uplink data rates up to 32 Mb/s in light rain. These one-meter terminals will support downlink burst data rates up to 32 Mb/s in moderate rain. Users who wish to maintain low bit error rates in during heavier rainfall may purchase terminals with higher EIRP (e.g., 10-W transmitter and 1.5-meter antenna diameter). Other rain-fade mitigation techniques may also be available, e.g., as implemented on an experimental basis on the NASA ACTS program. Achieving higher data rates is possible by means of accessing multiple uplink channels through a multiplexer. The downlink power budget is also estimated to allow transmission of up to 622 Mb/s with 10^{-10} BER.

3.4.2 MEO/LEO Systems

While both LEO and MEO systems are similar in that they involve non-stationary satellites, there are a number of factors which favor the MEO design over LEO. These are:

a. The LEO system requires large number of satellites (200 to 840). A comparable MEO system requires only ten to twelve satellites

b. The LEO satellites, because of their proximity to earth's atmosphere, have reduced lifetimes (< 5 years). MEO systems suffer very little atmospheric drag and have lifetimes of about eight to ten years

c. LEO systems require large number of high bandwidth crosslinks. This puts enormous strain on the satellite payload design

d. LEO satellites travel over a user terminal's horizon in a matter of a few minutes. Sophisticated high-speed PAAs are required to keep the user connected to the satellite during this brief period

e. LEO systems require complicated hands-off techniques and dynamic switching to keep the user terminal connected to the network

f. LEO systems require multiple satellites per launch to keep costs at comparable level to MEO systems.

As a consequence of all these drawbacks, only the MEO system is considered for the detailed architecture design.

Frequency Band of Operation

The Ka-band is chosen as the frequency band of operation for the same reasons that it was chosen for GEO system. However, unlike the GEO system, LEO/MEO systems have only 500 MHz of available bandwidth (28.6 to 29.1 GHz for uplink and 18.8 to 19.3 GHz for the downlink). Also, unlike the GEO systems, where each service provider could use the full available bandwidth by keeping its satellites at least 2° apart from other satellites, MEO/LEO systems will have to share their available frequency band. This leaves LEO/MEO system providers with two options; either each MEO/LEO system

a. Uses a separate and smaller slice of the bandwidth or

b. Uses the full 500 MHz spectrum by collaborating with all the other system providers.

The second option, while desirable, may be very difficult to achieve and may involve costly additions to each satellite. A further complication arises from the fact that there is already one LEO system filed, i.e. Teledesic, which has asked for allotment of almost the whole of the available bandwidth. In light of this, it may be very difficult for any new LEO/MEO system to get approval from FCC for this band. The only possible way for a new LEO/MEO filing to get approval is if it coordinates its satellites frequencies and movements to avoid conflict with the Teledesic system. We will assume this to be the case for all our subsequent analysis.

Satellite Constellation Geometry

For reasons explained earlier, an MEO constellation is chosen for the MEO/LEO concept design. At MEO altitudes of 10,000 km, the number of satellites in polar orbit required to provide global coverage (assuming 20° elevation) is approximately 17 (see Figure 3-5). ICO, through its studies, showed that changing the orbital inclination to 45° can reduce the number of satellites required to provide global coverage to a total of 10. We will assume here the same satellite geometry as that of ICO, i.e., two orbital planes each at 45° inclination and each containing five equi-spaced satellites per orbit. Each satellite is at an orbital altitude of approximately 10,355 km.

The FOV from each satellite at this altitude, and with 20° minimum elevation angle, subtends an angle of approximately 42° which corresponds to an FOV footprint chord diameter of about 9629 km. The number of beams supported by each satellite and the HPBW of each beam is determined by the satellite antenna diameter. If we limit the number of beams to 196, the HPBW of each beam will be about 3°. This corresponds to a nadir beam footprint (chord) diameter of approximately 542 km. The earth surface can be mapped into fixed cells (total of approximately 3000 fixed cells needed to map the whole surface) of this size (542 km) and each satellite passing over such cells will be programmed to provide a beam of fixed frequency and polarization to each cell. Adjacent cells can be assigned opposite polarizations, in a checker-board pattern, so as to minimize interference between them.

Access Techniques

A mixture of FDMA and TDMA is proposed for the uplink and TDM on the downlink. On the uplink, each user terminal is assigned one or more frequency slots for the call's duration. The number of slots assigned determine the maximum available rate. One slot corresponds to a standard 16 kb/s basic channel. Since 500 MHz bandwidth is available one could provide an equivalent of up to 3 STS-3 signals. Multiple users located within the same cell unit can each access the satellite using TDMA. The terminal downlink uses the packet's header rather than a fixed assignment of time slots to address terminals. Packets are delimited by a unique bit pattern and a terminal selects those addressed to it by examining each packet's address field.

Handover

Unlike GEO systems where the satellites remain at a fixed position with respect to the earth, MEO/LEO systems involve motion of satellites relative to the earth. This motion creates the need for handovers during which a call from one satellite is switched to another to provide continuous and uninterrupted service for the duration of the call. Two handover techniques were discussed in Section 2; i.e. SOFA and ROFA. At an altitude of 10,355 km, each satellite is traveling at a speed of 17,600 km/hr with respect to the earth. At these speeds a satellite traverses over a cell-site in a few minutes. The SOFA technique would require sophisticated circuitry to keep track of the satellite as it traverses across the horizon as well as to switch to a new satellite when the old one has swept past the horizon. Such sophisticated circuitry translates into a larger size and more costly user terminal.

Instead, the ROFA technique for handover is advisable, with each satellite changing its frequency and beam direction to maintain connection to the cell site. Handovers are performed automatically by the satellite system, whenever a satellite goes out of the 20° elevation angle range of a cell site. Thus, as long as a terminal remains within the same earth-fixed cell, it maintains the same channel assignment for the duration of a call, regardless of how many satellites (and beams) are involved. Channel re-assignment becomes the exception rather than the normal case, thus eliminating much of the frequency management and handover overhead. A database will have to be maintained in each satellite defining the frequency and polarization allowed for each earth-fixed cell.

Antenna Size

The antenna size is dependent upon the HPBW chosen for the system. With a HPBW of 3°, the antenna diameter is fixed at about 0.35 meter. To create the required 196 steerable spot beams for each satellite, a PAA system will be assumed. To avoid the problem of grating lobes the antenna elements must be spaced no more than 0.75 cm apart (or 0.5λ). This implies the need for a PAA of size 50x50 elements. A beam-forming network to switch the 196 beams to all the 50x50 elements with proper amplitude and phase, will also be needed.

Communications Payload

A functional diagram of the communications payload is given in Figure 3-15. As can be seen, the MEO communications payload is similar to that of the GEO system except in the following respects: the PAA is smaller in size (50x50); there are no gateway beams, there are four ISLs instead of two; and the on-board processor is less complex and involves the routing of 196 instead of 256 beams. The four ISLs include communications with neighboring satellites in the same orbital plane, as well as with the neighboring satellites in the adjacent plane. Interference between these inter-satellite links can be minimized by setting them about 2 GHz apart in the 60-GHz band.

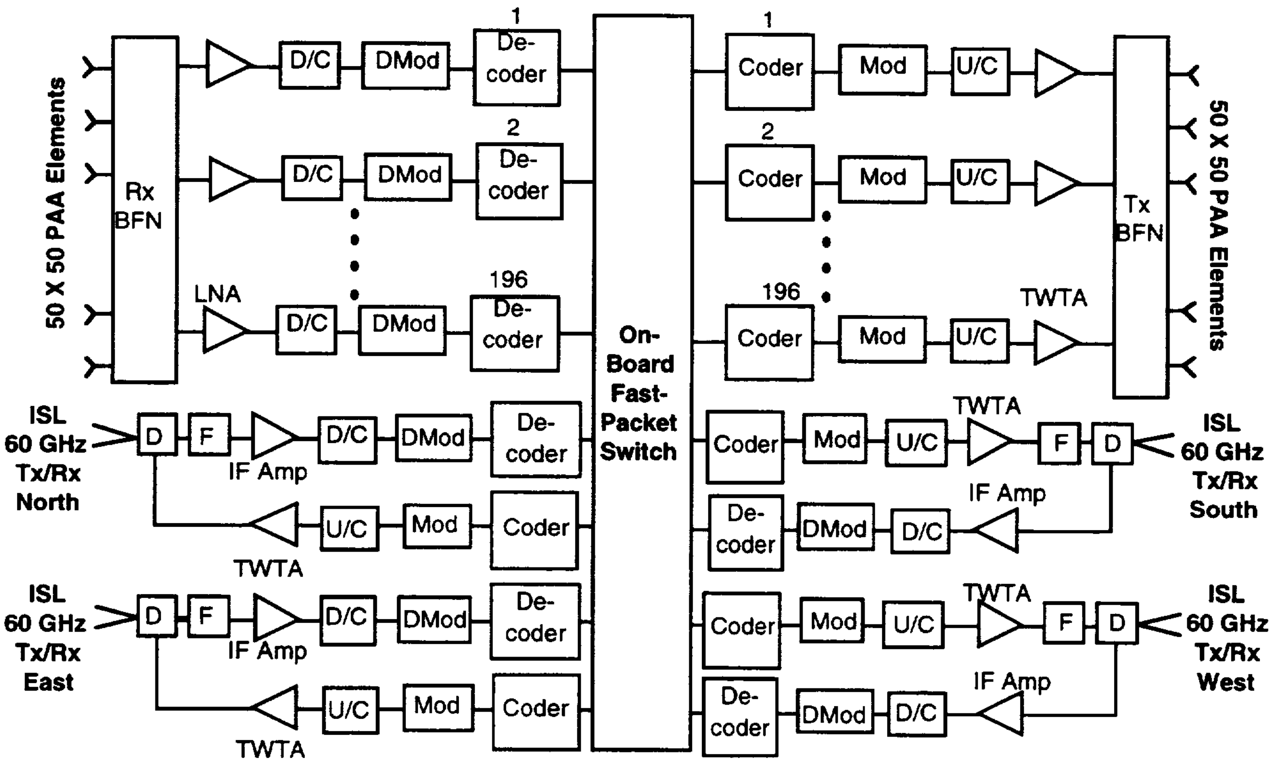


Figure 3-15. Functional Block Diagram of Communications Payload for Conceptual MEO/LEO Design

This system will have an overall bandwidth per satellite of approximately 98 GHz (196 beams x 500 MHz per beam). Such a bandwidth should easily support all the data needs for future NII/GII applications.

3.5 KEY TECHNOLOGIES REQUIRED

To realize the suggested third-generation satellite architectures suitable for carrying NII/GII applications, a number of key technologies needs to be developed. These include:

a. The design of large size PAAs and BFNs with low mass, power and volume

b. Development of high speed on-board processing and switching

c. High data rate and reliable ISLs

d. Standards for interoperability between satellite and terrestrial networks.

These technologies are briefly summarized in the following sections.

Antenna Designs

Third-generation satellite systems will require the formation of large number of spot beams, both fixed and steerable. The two methods to achieve these spot beams were discussed in subsection 3.4.1, i.e., the use of focal plane fed arrays and phased arrays.

In the focal-plane fed array systems, antenna elements are placed at the focal point of a single reflector. In its simplest form, each element, when excited, can form a single beam in a fixed direction. Multiple antenna elements can be used to create a single beam with sophisticated beam patterns and finer resolution by employing beam-forming networks that include attenuators and phase shifters in the waveguide feeds which excite the antenna elements. Switched feed horn arrays can also incorporate high performance ferrite switches to implement beam scanning or hopping. Extensive research has gone into reducing the size and weight of electronically-controlled multiple beam antenna assemblies to support rapid beam scanning over a wide field of view by rapidly switching selectable clusters of antenna elements[7].

The phased-array antenna, in contrast to the focal-plane fed array, uses all antenna elements to create each beam. Using phased-array antennas reduces the power requirement per antenna element, permits rapid beam steering, and allows for graceful degradation of components, as previously discussed in subsection 3.3.1. The increasing need for rapid forming and steering of antenna beams has prompted significant research and development efforts in phased array technology. The major task is to provide a controlled feed of correct amplitude and phase for the individual elements in the array, thus forming a specified beam shape and direction. Traditionally, RF beam forming and steering are performed by individually controlled phase shifters and variable gain amplifiers in the

---

7 E. Georg et.al., "Advanced Satellite Communications: Towards an Integrated Infrastructure," AIAA-94-0970-CP

signal path of each array element. Advances in Monolithic Microwave Integrated Circuit (MMIC) technology now make it possible to fabricate compact arrays of individual solid state amplifiers at each radiating element which are also power efficient and operable at high frequencies. As part of the active antenna program, several agencies such as Intelsat, ESA and NASA have sponsored development of flexible Ku- and Ka-band multi-element multibeam active antennas for advanced communication satellites. Such microwave phased-array antennas can produce multiple fixed, hopping and/or scanning beams with dwell times as small as microseconds. However, for the large phased-array antennas needed in large-capacity, multi-beam satellites to provide dynamic connectivity between corresponding locations of the communications network, the hardware necessary to perform such beam forming becomes prohibitively complex and heavy for the payload carried on-board. A recent trend in PAA research is to develop improved or new Beam-Forming Networks (BFNs) by using optimized architectures, novel concepts and/or alternative technology. Such systems include digital beam forming with neural networks[8] and photonics technology.

Use of photonics technology, for example, promises high-frequency, low side-lobe beam forming and steering by reliable lightweight BFNs[9]. Additionally, the potential for real-time optical processing capability has no parallel in other technologies, which makes its development very attractive and timely for communications satellites requiring on-board processing. Most proposals for optical BFN have centered around the concept of replacing the conventional harness of coaxial cables and waveguides by low weight and density fiber cables. While such methods do reduce the mass, power, and volume of the BFN, they do not fully use the processing capabilities of optics. Comsat Laboratories has pioneered a new technique, called the Coherent Optical Processor technique[10,11], which uses a two-dimensional Fourier transformer technique to derive, in real time, the phase and amplitude of the electromagnetic fields required at the antenna inputs to create a specific far-field beam pattern. Such techniques should make possible very large size PAA systems with large numbers of beams.

---

8 H. Steyskal, "Digital Beam-Forming at Rome laboratory, "Microwave Journal, Feb. 1996, pp. 100-126.

9 D. Paul, R. Razdan and B. Markey, "Optical Beam Forming and Steering Technologies for Satellite Phased Array Antennas,", AIAA-94-1113-CP, pp. 1332-1341.

10 G.Koepf, "Optical Processor for Phased-Array Antenna Beam Formation," Optical Technology for Microwave Applications, Proceedings of the SPIE, Vol. 477, S.K. Yao, editor, 1984, pp. 75-81.

11 R. Razdan et.al., "Communications Performance of a Multi-beam Multi-carrier Photonic Beam-Forming and Beam-Steering Feed Network for Satcom Phased Array Antenna," AIAA-96-1160-CP, pp. 1392-1396.

Inter-Satellite Links

Inter-Satellite Links are a relatively recent phenomenon, with the first demonstrated crosslink provided by MIT Lincoln Laboratory Experimental Satellite (LES) 8 & 9, launched in 1978[12]. Since 1983, ISLs have been used regularly within the Tracking and Data Relay Satellites (TDRS) system. The Milstar satellites also use multiple crosslinks[13]. Inter-satellite links are expected to be an integral part of LEO/MEO designs which try to emulate the terrestrial mesh networks. In such systems, each satellite can be viewed as a switching node connected to numerous other such nodes through such inter-satellite links. GEO systems, on the other hand, use ISLs only for the special case where direct single-hop transmission is not possible.

For all these third-generation satellite systems, the requirements for the inter-satellite link can be very demanding. Inter-satellite links are expected to carry very high-speed data, formed as an aggregate of multiple channels. While some proposed multimedia Ka-band satellite systems plan to use the 60-GHz band to provide the inter-satellite link, its feasibility[14] is questioned in providing high-speed data connections. An alternative is to use laser communications for the inter-satellite links. Such links promise to be smaller in size and volume, with less power consumption, lower cost, free from EMI/RFI and capable of modulation to very high data rates. MIT Lincoln Laboratory, for example, has recently reported a Lasercom breadboard capable of communicating at 10 Gb/s with $\leq 10^{-9}$ BER over geosynchronous link distances[15]. The breadboard was made from off-the-shelf commercial components at 1550 nm, and many of the components are already qualified for transoceanic fiber applications (which are often more severe than space applications).

Recent developments in optical technology, such as the creation of the Master Oscillator, Power Amplifier (MOPA) which allows delivery of 1 W of single-mode laser output, and the development of Erbium-doped fiber amplifiers, promise to provide higher-speed links with even less noise than that provided by the MIT design.

---

12 William Brandon, "Commercial and Military Experimental Systems in the New Era," 15th AIAA International Communications Satellite Systems Conference, Plenary Session, Feb. 1994.

13 L.F. Kwiatowski, et.al., "The MILSTAR System," 15th AIAA International Communications Satellite Systems Conference, AIAA-94-1013-CP, March 1994, pp. 744-748.

14 J.E. Freiedell, et.al., "Why Laser Communication Crosslinks Can't be Ignored," AIAA-96-1060-CP, pp. 1397-1409.

15 J.C. Livas, "High Data Rate Error Correcting Coding," Presented at SPIE International Symposium, Proc. SPIE 2123, paper no. 37, 1994.

On-Board Processing

On-Board Processing (OBP) and routing is, by far, the most crucial part of the third generation satellite system. Using OBP allows the formation of multiple switchable beams with each beam having only the data addressed to the terminals within the footprint of the beam. This allows frequency reuse over separate beams thereby improving spectral efficiency. Additionally, OBP allows decoupling the uplink and downlink error statistics, thereby improving the overall BER.

In the past, OBP technology has been slow to develop because of its added mass and power consumption and its greater cost. Advanced Communications Technology Satellite (ACTS) and Milstar were amongst the first systems to use OBP technology. Since then some of the new LEO systems (such as Teledesic and Iridium) as well as the recent Ka-band filings (such as NetSat 28 and Astrolink) propose to use advanced OBP technology to provide multimedia connectivity over multiple switched beams.

The OBP can be based upon either a circuit-switched system or a fast packet switched system. From point of view of compatibility with ATM networks, the fast packet-switched architecture is clearly better suited. The on-board fast packet switch routes the traffic coming from either fixed spot beams or the scanning beams, by means of the header information. To get to the header information, it must first downconvert the received signal and then demodulate it. The routed signal or signals are modulated, upconverted and passed to the relevant beam. The on-board fast packet switches will have to be able to handle and route the various packets at very fast rates to avoid congestion. Speed and size requirements of the OBP can put a severe strain on the MPV budget. Optical technology can be used to reduce these strains. Comsat Laboratories has developed and tested a 1-Gb/s fiber-optic distribution system for on-board baseband processing[16]. The optical components used in this development were all available off-the-shelf. Implementing an optical switch matrix on-board the satellite will require various design trade-offs between space-division, time-division and Wavelength-Division Multiple Access (WDMA) systems[17]. An optical fast packet switch, based on a proprietary scheme, forms the heart of the NetSat 28 system which will provide connectivity to 500 users, with each user having data rates up to 1.544 Mb/s. Such a system is well on the way to the creation of an all-optical system, i.e., where the beam-forming network, on-board processor, and inter satellite link are all operating in the optical band.

---

16 B. Pontano et.al., "COMSAT Laboratories' on-board Baseband Switch Development", Second Space Communications Technology Conference, Cleveland, OH, Nov. 1991, pp. 207-214, NASA Conf. Pub. 3132.

17 "Optical Packet Switched Networks", Collection of papers, IEEE Journal of Lightwave Technology, Vol. 11, Nos. 5/6, May/June 1993.

Interoperability Standards

Future third-generation systems are envisioned having seamless operation between terrestrial and satellite systems. This is based on the fact that both systems are expected to grow tremendously. Terrestrial systems will be able to provide very high bandwidth optical links, but is an unsuitable choice for providing connectivity to remote locations with poor communications infrastructure. Satellite systems, while somewhat limited in bandwidth (compared to optical systems), are ideally suited for setting links between remote and poorly connected systems. In addition, satellite systems will be able to provide mobile communication links. Recognizing the future role of both satellite and terrestrial networks, it becomes important that standards allowing interoperation between the two systems are formulated.

The various standards bodies (such as ITU) have realized that the future networks will be all digital in nature and will involve traffic of varying data types and rates with differing burstiness characteristics. A cell-based communications standard, termed as ATM, was created to be able to handle this variety of services. Satellite systems, too, are moving towards full digitization of traffic; and the use of ATM for satellite networks would greatly increase interoperability with terrestrial systems.

Unfortunately, the standards bodies have, in the past, ignored satellite networks in formulating communication standards such as ATM. However, through the efforts of COMSAT Labs and other organizations, satellite concerns are now being addressed in these standard bodies. The key differences between the terrestrial and satellite links are the greater delays and poorer communication links for satellite systems. Fortuitously, some of these problems, such as large delays, are also experienced by terrestrial systems resulting in reformulation of the terrestrial standards. It is expected that with the growing awareness of satellite systems, and their unique characteristics, more general standards will evolve. Such standards will make interconnection between satellite systems and terrestrial systems relatively simple and cheap.

In addition to interoperability between terrestrial and satellite links there is concern about communications between satellite networks[18]. For example, with the current allocation of frequency in Ka-band there is possibility of only one satellite system operating in LEO/MEO orbit. Standards may have to be developed to allow multiple satellite systems to operate without interference.

---

[18] E. Georg et.al. "Advanced Satellite Communications: Towards an Integrated Information Infrastructure,", AIAA-94-0970-CP, pp. 463-473.

3.6 CONCLUSIONS

In the preceding sections, various conceptual architectures for space-based information infrastructures were formulated. These architectures are designed to support future NII/GII applications and operate seamlessly with terrestrial networks as a part of the third-generation communication systems. The relative merits/demerits for each architecture were examined. Based on the results of this study, LEO/MEO architectures are rejected for NII/GII applications. The Ka-band can, currently, only support one LEO/MEO satellite system. Since the Teledesic system is already proposed, that leaves no place for any other LEO/MEO system. The GEO-based architecture, on the other hand, can utilize the full available bandwidth, without interfering with other GEO satellite systems, provided the satellite systems are spaced two degrees apart. The GEO architecture suggested here has a 80 Gb/s capacity per satellite. With a total of five satellites, the system capability is 400 Gb/s, which is sufficient to handle the type of traffic dictated by future NII/GII applications.

SECTION 4 – ECONOMIC ASSESSMENT

This section develops cost estimates for the GEO conceptual architecture proposed in Subsection 3.4. A MEO system was also discussed, but was not defined in sufficient detail to allow costing. The MEO Ka-band system that was discussed also suffered from the problem of having to share the 500 MHz of non-GEO spectrum with the Teledesic system. The discussion is organized as follows:

4.1 Cost Methodology

4.2 Cost Assumptions

4.3 Cost Estimates

4.4 Major Cost Drivers

4.1 COST METHODOLOGY

The methodology is to use a satellite system financial spreadsheet model which spreads revenues and expenses over the system lifetime, and calculates resultant circuit wholesale cost to realize the desired return on investment. The model results and sample spreadsheet pages are given in subsection 4.3. The assumptions going into the model are described in subsection 4.2.

User terminal costs are not estimated per se, but are rather given as marketing goals. In other words, if user equipment cost is too high, small users are discouraged from using the system. This may be desirable in that the costs of servicing many small users is disproportionately higher than that for fewer large users.

The system model is that of a service provider selling wholesale capacity to a number of service providers such as public telephone companies. This is a model similar to that for the Globalstar mobile satellite service system. Globalstar provides the satellite system with satellite and network control via a few gateways. The local service providers furnish local gateways and billing for customers in their area. The users purchase their own terminals which may come in several sizes depending on user rain region and desired quality of service. The user cost per minute estimated is the wholesale cost, which is typically doubled or tripled by the local service provider, to which a monthly service or billing charge is added.

Tables 4-1 through 4-4 summarize the parameters of the two conceptual architectures costed in subsection 4.3 and described in Section 3:

a. GEO system (reference subsection 3.4.1)

b. MEO system (reference subsection 3.4.2) (partial parameters given)

Table 4-1. Satellite System Parameters

Parameter	MEO	GEO
Number of satellites (on-orbit spares)	10 (2)	5 (1)
Satellite lifetime (yr)	10	12
Satellite dry mass (kg)		2,000
Satellite end-of-life power (kW)		30
Satellite altitude (km)	10,355	35,786
Minimum elevation (degrees)	20	20
Satellite field of view (degrees)	42.0	16.3
Uplink frequency - users (GHz)	28.6 – 29.1	29.25 – 30.0
Uplink frequency - gateways (GHz)		28.35 – 28.6
Downlink frequency - users (GHz)	18.8 –19.3	19.45 – 20.2
Downlink frequency - gateways (GHz)		18.55 – 18.8
Intersatellite link frequency (GHz)		59.0 – 64.0
User bandwidth, up/down (GHz) (6 x 166 MHz with polarization reuse)		1.0
Gateway bandwidth, up/down (GHz) (20 x 250 MHz with polarization reuse)		5.0
Intersatellite link bandwidth (GHz)		2.0
Number of intersatellite links	4	2
Number of simultaneous uplink and downlink user beams	196	256
Downlink channel size (Mb/s)		311
Total downlink capacity (Gb/s)		80

Table 4-2. Satellite Bus Parameters

	MEO	GEO
Number of satellites (on-orbit spares)	10 (2)	5 (1)
Satellite altitude (km)	10,355	35,786
Satellite dry mass (kg)		2,000
Satellite end-of-life power (kW)		30
Launch vehicle	Atlas IIAS	Atlas IIAS
Satellites per launch vehicle	2	1
Bus power (kW)		30.0
Available payload power (kW)		25.5
Processor power (kW)		5.1
Remaining payload dc power (kW)		20.4
Available RF power @ 60% (kW)		12.2

Table 4-3. Satellite Communication Link Parameters

Parameter	MEO	GEO
Satellite uplink antenna diameter (m)	0.24	1.40
Satellite uplink antenna beamwidth (°)	3.0	0.50
Beam size on earth (at nadir) (km)	542	315
Number of beam positions in field of view	196	1,068
Number of simultaneous user beams	196	256
Uplink bandwidth per user beam (MHz)	500	332
Total user uplink bandwidth per satellite (GHz)	98	85
Uplink data rate per user beam (Mb/s)		311
Total user uplink data rate per satellite (Gb/s)		80
User uplink burst rate (Mb/s)		32
Number of simultaneous users per uplink beam		10
Max. number of simultaneous 32-Mb/s uplink users per sat.		2,500
Number of simultaneous gateway scanning beams	0	10
Bandwidth per gateway beam (MHz)		500
Total gateway bandwidth per satellite (GHz)		5
Data rate per gateway beam (Mb/s)		622
Total gateway data rate per satellite (Gb/s)		6.2
Satellite downlink antenna diameter (m)	0.36	2.10
Satellite downlink antenna HPBW (°)	3.0	0.50
Beam size on earth (nadir) (km)	542	315
Number of beams in field of view	196	1,068
RF power per downlink channel (W)		47
Number of simultaneous user beams	196	256
Downlink bandwidth per user beam (MHz)		332
Total user downlink bandwidth per satellite (GHz)		85
Downlink data rate per user beam (Mb/s)		311
Total user downlink data rate per satellite (Gb/s)		80

Table 4-4. Terminal Parameters

Parameter	MEO	GEO
User terminal size (m)		1.0
User terminal power per channel (W)		5
User uplink burst rate (Mb/s)		32
Downlink burst rate to multiple users (Mb/s)		311
Max. number of simultaneous 32-Mb/s bursts per satellite		2,500
Max. number of simultaneous 2-Mb/s users per satellite		40,000
Max. number simultaneous 2-Mb/s users (worldwide)		200,000
Estimated number of user terminals		5,000,000
Gateway terminal size (m)		3 to 5
Gateway terminal power per channel (W)		10 to 100
Estimated number of fixed gateway terminals worldwide		120

Detailed satellite designs with mass and power have not been performed for these concepts. Their proposed parameters are based on engineering judgment for second generation systems; e. g., initial operational capacity in 2010, after that of the first round of Ka-band filers such as Astrolink, Cyberstar, Spaceways (GEO Ka-band). The assumed technology level is thus more aggressive than that of the current FCC filings. The result is that the same size (and cost) spacecraft has more capacity and hence a lower cost per circuit minute. In particular, the GEO system satellite dry mass is assumed to be 2000 kg, and uses an equivalent Atlas IIAS class launch vehicle.

4.2 COST ASSUMPTIONS

The cost assumptions are divided into four categories:

a. Global Assumptions

b. Space Segment Assumptions

c. Terminal Segment Assumptions

d. Control Segment Assumptions

4.2.1 Global Assumptions

Global cost assumptions are as follows:

a. Costs are adjusted to reflect the year 2005 level of technology achievements

b. An initial operating capability is assumed in the year 2010 and a full operating capability in 2012.

c. All costs are expressed in the year 1996 economics

d. No learning curve factors were assumed for multiple production unit cost

e. Commercial acquisition practices are assumed: fixed price contract with on-orbit delivery; limited oversight, minimum CDRLs and other deliverables

f. Technologies are mature prior to development phase and investment of advanced contractor participation

g. Block buy and build

h. Modular spacecraft design where the payload and bus are assembled and tested in parallel as independent modules that are integrated after the functional performance of each has been properly verified

i. Use Protoflight spacecraft for flight unit

j. Commercial equipment qualification philosophy

k. Circuit costs are derived from a monthly, simplex, E-1 (2.048 Mb/s) circuit cost produced by a COMSAT cost model (see Table 4-5). Input factors to model are as follows:

1. Number and capacity (Gb/s) of satellites
2. Lease cost (wholesale) per duplex (2 Mb/s) circuit (adjusted until IRR equals target rate, nominally 30 percent)
3. 30 percent internal rate of return (IRR) (baseline)
4. 10 percent cost of capital
5. 30 percent income tax, linear depreciation
6. Capital investments (see Table 4-6)
7. O&M expenses assumed at 30 percent of the revenue for the year
8. System fill factor equals percent of constellation in place times assumed percent of available circuits purchased wholesale
6. Assumed percent of available circuits purchased wholesale starts at 15 percent the first year and rises to 35% in the fifth and subsequent years. Wholesale means commitments by retailer to sell this much capacity at the market price, which is typically two or three times the wholesale cost.

4.2.2 Space Segment Assumptions

The space cost assumptions are as follows:

a. The life cycle encompasses the period between 2006 and 2023 with full service over a 10-year period from 2012 through 2021.

b. System fill factor is 35 percent of the full constellation capacity (wholesale assumption)

c. The international landing rights cost per year are not included in the cost calculation

d The transmission cost per minute reflects a wholesale cost due to a bulk buy

e. No learning curve was used when calculating the total production cost of satellites

f. The constellation includes an initial satellite system and no replenishment system

g. The initial spacecraft is delivered 36 months After Receipt of Order (ARO) with subsequent deliveries on four-month centers

h. Space segment cost breakdown and spread with time is given in Table 5-6:

1. Satellite cost is $120 M for each of six (5 operational and one on-orbit spare)

2. Launch cost is $60 M, an assumed reduction of fifty percent from the 1996 price for an Atlas IIAS launch vehicle.
3. Nonrecurring space cost totals $100 M
4. Insurance rate is 16% of total space (launched cost)

4.2.3 Terminal Segment Assumptions

Terminal cost assumptions are as follows:

a. User terminal nonrecurring costs are included, but recurring costs are not included. The cost goal for the small user terminal is <$2,000 (worldwide quantity is 5 M).

b. Gateway and TT&C terminals for network control are included in the wholesale cost.

c. Gateway terminals for service providers (retailers) are not included in the wholesale cost.

d. O&M is calculated at 12 percent of the total hardware cost per year

e. Hardware and software costs are estimated using engineering judgment based on similar present terminals with similar functions

   1. Ground R&D (nonrecurring) is $120 M and recurring is $200 M

4.2.4 Control Segment Assumptions

The control cost assumptions are as follows:

a. Network control costs are estimated at $50 M.

b. Spacecraft control costs are estimated at $2 M per year.

b. Software maintenance is calculated at eight percent of the non-recurring cost per year

c. Hardware maintenance is calculated at five percent of the total hardware cost per year

d. Staffing is individually assigned per center type based on man-loading by grade and quantities

e. Hardware and software costs are estimated based on similar centers with similar functions

4.3 COST ESTIMATES

The cost estimate calculation is given in Tables 4-5 and 4-6. The result, for a 30 percent return-on-investment (IRR), is a wholesale lease cost of $6580 per month for a full duplex E-1 (2.048 Mb/s) circuit. This reduces to 7.5 cents per minute for a simplex 2-Mb/s connection. The retail price for this circuit would be two or three times higher, or approximately 19 cents per simplex circuit minute. If proportioned to data rate, the retail price of a 28.8 kb/s simplex circuit becomes 0.26 cents per minute.

This can be compared with the present day Internet access charge of $20 per month for five hours connection, which is equals seven cents per minute, for 14.4 to 28.8 kb/s.

4.4 MAJOR COST DRIVERS

The major cost drivers are interrelated and include the following:

a. Capital cost of system per circuit

b. User terminal costs and O&M (if any)

c. Capacity of satellite in terms of total data rate or number of circuits. The Ka-band multi-media systems being proposed today have around 10 times less capacity, and thus their cost per circuit minute would be 10 times higher.

d. Effective utilization of the available circuits. The financial model has used utilization figures based on circuits being sold wholesale to service providers such as PTTs. The packet nature of the proposed services using ATM protocols may allow further lowering of cost by charging the user for the bits transmitted, not simply the connection time.

e. Allocation of rain margin, or minimization of static, "available but unused" margin. Depending on frequency band utilized, the rain margins may be 10 dB or more. "Smart" rain margin allocations to the regions and circuits needing margin to keep quality of service can turn unused margin into additional capacity.

Table 4-5. Financial Analysis Worksheet

Wholesale circuits to telecommunications providers (2010 launch)

Satellite full simplex capacity	80	Gb/s	Input Number
No. of Satellite	5		Input Number
System capacity	400	Gb/s	Sat Capacity * # Sats
No. of 2 Mb/s duplex circuits	100,000		Total capacity (Gb/s) * 1000 / 2 Mb/s /2
Capital cost/duplex circuit	$17,228		Total capital invest / number of 2-Mb/s duplex circuits
Lease cost per duplex circuit	$6,580	/month	Input number
Lease cost per simplex circuit	$0.0751	/min	Calculated as ((duplex lease cost per month)/2)/((365.25/12)*24*60)
IRR	30.00%		Total Net cash flow + PV capital

Income Tax Rate	30%	Input Number
Cost of Capital	10%	Input Number

Capital investment multiplier = 1.0000 0.58038 1.00000 2.32153 Choice for capital investments of $1B, $1.723B (baseline), and $4B

(x $1,000,000)

Year	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	Total
1. Capital Investment (Input #)	($60)	($200)	($342)	($390)	($348)	($255)	($128)										($1,723)
2. Revenues (# ckt * ckt $ * sys fill ftr)					$355	$1,105	$1,974	$2,369	$2,764	$2,764	$2,764	$2,764	$2,764	$2,764	$2,764	$2,764	$27,912
3. O&M cost (30% revenue)	($5)	($10)	($20)	($50)	($107)	($332)	($592)	($711)	($829)	($829)	($829)	($829)	($829)	($829)	($829)	($829)	($8,459)
4. Other costs					$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0
5. Depreciation (Tot capital invest/12 yr)					($144)	($144)	($144)	($144)	($144)	($144)	($144)	($144)	($144)	($144)	($144)	($144)	($1,723)
6. Pre-Tax Profit		Rev + O&M + Oth + Depre			$105	$630	$1,238	$1,515	$1,791	$1,791	$1,791	$1,791	$1,791	$1,791	$1,791	$1,791	$17,816
7. Taxes(Pre-tax profit*tax rate)					($32)	($189)	($371)	($454)	($537)	($537)	($537)	($537)	($537)	($537)	($537)	($537)	($5,345)
8. Profit After Taxes		Pre-tax profit - taxes			$74	$441	$867	$1,060	$1,254	$1,254	$1,254	$1,254	$1,254	$1,254	$1,254	$1,254	$12,471
9. Non-Cash Charges (depreciation)					$144	$144	$144	$144	$144	$144	$144	$144	$144	$144	$144	$144	$1,723
10. Net Cash Flow		Profit + Depreciation			$217	$585	$1,010	$1,204	$1,397	$1,397	$1,397	$1,397	$1,397	$1,397	$1,397	$1,397	$14,194
11. With PV Capital	($1,286)	$0	$0	$0	$217	$585	$1,010	$1,204	$1,397	$1,397	$1,397	$1,397	$1,397	$1,397	$1,397	$1,397	$12,908
12. System Fill Factor					5%	14%	25%	30%	35%	35%	35%	35%	35%	35%	35%	35%	
13.		% available circuits leased			15%	20%	25%	30%	35%	35%	35%	35%	35%	35%	35%	35%	
14.	($1,226)	Constellation in place			30%	70%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	

=C13+D13/(1+$O$5)+E13/(1+$O$5)^2+F13/(1+$O$5)^3+G13/(1+$O$5)^4+H13/(1+$O$5)^5+I13/(1+$O$5)^6+J13/(1+$O$5)^7

	C	D	E	F	G	H	I	J
		($200)	($342)	($390)	($348)	($255)	($128)	$0
		1	1	1	1	1	1	1
		10%	10%	10%	10%	10%	10%	10%
		110%	110%	110%	110%	110%	110%	110%
		110%	121%	133%	146%	161%	177%	195%
($1,226)		($181)	($283)	($293)	($238)	($158)	($72)	$0

SPACE SYSTEMS LORAL

Table 4-6. Capital Costs Spread with Time

Year	-3	-3	-2	-2	-1	-1	0	0	1	1	2	2	3	3	
6-month period	1	2	1	2	1	2	1	2	1	2	1	2	1	2	Totals
Launches									*	*	*	*	*	*	($ M)
Ground R&D			$30	$30	$30	$30									$120
Ground recurring					$50	$50	$50	$50							$200
Total ground	$0	$0	$30	$30	$80	$80	$50	$50							$320
Control segment	$5	$5	$10	$10	$10	$10									$50
Space R&D	$25	$25	$25	$25											$100
Satellite 1			$20	$20	$20	$20	$20	$20							
Satellite 2				$20	$20	$20	$20	$20	$20						
Satellite 3					$20	$20	$20	$20	$20	$20					
Satellite 4						$20	$20	$20	$20	$20	$20				
Satellite 5							$20	$20	$20	$20	$20	$20			
Satellite 6								$20	$20	$20	$20	$20	$20		
Total satellite recurring (6)			$20	$40	$60	$80	$100	$120	$100	$80	$60	$40	$20		$720
Launch 1								$30	$30						
Launch 2									$30	$30					
Launch 3										$30	$30				
Launch 4											$30	$30			
Launch 5												$30	$30		
Launch 6													$30	$30	
Launch recurring (6)								$30	$60	$60	$60	$60	$60	$30	$360
Total space recurring	$0	$0	$20	$40	$60	$80	$100	$150	$160	$140	$120	$100	$80	$30	$1,080
Total space nonrecurring	$25	$25	$25	$25	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$100
Total space capital	$25	$25	$45	$65	$60	$80	$100	$150	$160	$140	$120	$100	$80	$30	$1,180

Year	-3	-2	-1	0	1	2	3	
Total space recurring	$0	$60	$140	$250	$300	$220	$110	$1,080
Insurance 16%	$0	$10	$22	$40	$48	$35	$18	$173
Total space nonrecurring	$50	$50	$0	$0	$0	$0	$0	$100
Total space	$50	$120	$162	$290	$348	$255	$128	$1,353
Total ground	$10	$80	$180	$100	$0	$0	$0	$370
Total capital investment	$60	$200	$342	$390	$348	$255	$128	$1,723

Satellites cost $120 M
Launch cost (50%) $60 M — Assumes cost reduction of 50% from 1996 to 2010

The sensitivity of the model to variations in input parameters is shown in Table 4-7. For example, the utilization of the satellite constellation resource is varied from 10% to 50%. The actual utilization will likely fall in between these limits. The baseline shown is a 35% utilization. This is the utilization achieved two full years after the full complement of five active satellites is in place. The buildup in utilization is shown on line 13, % of available circuits leased, of Table 4-5.

Cost per minute in the table is based on a 7-day, 24-hour/day week. Since few people would be likely to use it that way, the metric may be judged unrealistically optimistic. Supppose, therefore, a usage of 5 days/week, 12 hours per business day with two 2-hour peaks at 90% utilization and the remaining 8 hours per day at 25% utilization. For these assumptions, the retail price per minute of a 2 Mb/s duplex circuit for the baseline case increases from 19¢/min to $1.14/min. (The retail price per minute of a 28.8 kb/s simplex circuit for the baseline case increases from 0.26¢/min to 1.6¢/min). This usage assumption may, itself, be judged unrealistically pessimistic; i.e., the associated utilization of the leased circuits is only 16%. In addition, this usage assumption implies that no traffic is carried at night or on weekends. Certainly, creative pricing strategies by the seller could induce a large usage. One example would be the transfer of large data files which are not time sensitive.

Table 4-7. Sensitivity Analysis

Parameter Varied	Monthly Lease Cost – 2 Mb/s Duplex Circuit	Proportional per Minute Lease Cost – 28.8 kb/s Duplex Circuit
IRR = 20%	$3,222	0.13¢
IRR = 30% (baseline)	$6,580	0.26¢
IRR = 40%	$12,000	0.48¢
IRR = 50%	$20,200	0.81¢
Utilization = 10%	$18,470	0.74¢
Utilization = 20%	$9,460	0.38¢
Utilization = 35% (baseline)	$6,580	0.26¢
Utilization = 60%	$3,700	0.15¢
Satellite capacity = 20 Gb/s	$26,330	1.06¢
Satellite capacity = 50 Gb/s	$10,530	0.42¢
Satellite capacity = 80 Gb/s (baseline)	$6,580	0.26¢
Satellite capacity = 400 Gb/s	$1,320	0.05¢
Capital investment = $1 B	$3,820	0.15¢
Capital investment = $1.7 B (baseline)	$6,580	0.26¢
Capital investment = $4 B	$15,270	0.61¢

APPENDIX A – TECHNOLOGIES TO ENABLE LOW COST SATELLITE COMMUNICATIONS

Abstract

This appendix presents the results a study[1] to identify the most promising technologies implemented that could be used to reduce the cost of satellite communications. The economics of satellite communications is first summarized, and then the system design and technologies to reduce user costs are discussed. Finally, specific recommendations for technologies are made for fixed, broadcast, and mobile satellite service systems.

A1. INTRODUCTION

As illustrated in Figure A-1, technical feasibility, available financing, and existence of a market must be present for a business to be viable. This information is generally captured in the "business plan" for the proposed enterprise (usually a proprietary document). The FCC filing for a proposed communications system is a public document, and must contain enough of this information to allow a licensing decision. As the name "business plan" suggests, the primary metric used to evaluate the viability of a system is the total cost over its life, which translates into the cost to the user. Additional metrics may be considered or mandated by the government, such as universal access, protection of privacy, and usability by the government.

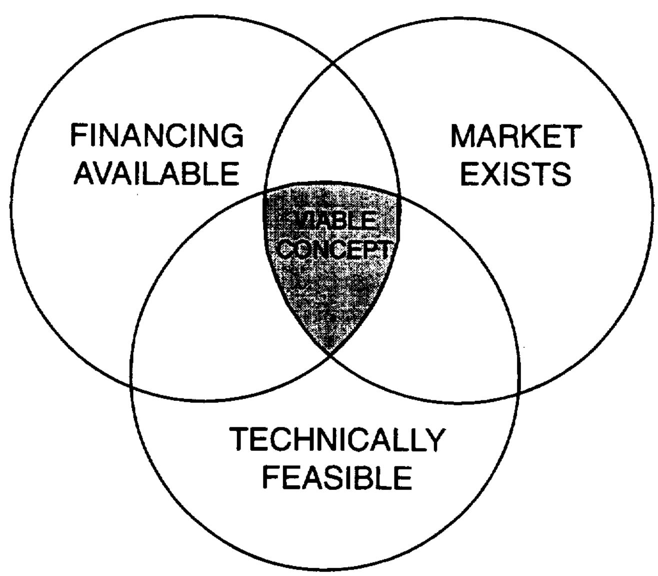


Figure A-1. Conditions for a Viable Concept

---

[1] This work was sponsored by NASA Lewis Research Center under contract NAS3-25934.

A2. ECONOMICS OF SATELLITE COMMUNICATIONS

Satellite system economics is discussed so that the impact of different technologies on user cost can be assessed. The discussion is presented in the following subsections:

A2.1 Breakdown of LCC by System Segment

Table A-1 gives the breakdown of Life Cycle Cost (LCC) for a proposed government FSS satellite system. The LCC is dominated by the ground terminal costs. The satellite payload cost is approximately 7 percent of the total LCC. The total space segment cost (satellite, spacecraft control, and launch) is 36 percent. Higher payoff technologies will be those that reduce ground terminal costs, even at the expense of increased satellite costs. The relative importance of ground segment costs is expected to be similar for a variety of systems that have large numbers of ground terminals.

a. Direct broadcast satellite systems such as DirecTV and PrimeStar are each expected to have more than three million user terminals costing $700 (initial) or $500 average. User terminals consist of 0.45-meter (18-inch) antennas, RF receive-only electronics, and digital codecs.

b. Worldwide MSS systems such as Globalstar, Iridium, Odyssey, and Inmarsat P are expected to each have millions of handsets at a cost from $500 (Globalstar) to $3000 (Iridium). User handsets are likely to be dual or triple mode, operating with satellite or terrestrial wireless systems. User handsets will resemble cellular phones.

Table A-1. Relative Segment Costs of Satellite Communications System

System Segment	Percent of Total Cost	Space Segment	Ground Segment
Satellite (bus and payload)	18%	18%	
Launch (of satellites)	13%	13%	
Control (spacecraft and network)	10%	5%	5%
Ground terminal development and production	35%		35%
Ground terminal operation and maintenance	24%		24%
Totals	100%	36%	64%

c. Worldwide Ka-band FSS systems such as Spaceway and CyberStar suggest millions of users per system (FCC filings, 9/95). User terminals (0.7- to 2-meter diameter) could cost from $1000 to $5000.

In all of these cases, the total price of ground terminals or handsets is several billions of dollars, and is equal to or greater than space segment costs (satellites plus launches). Ground terminal operation and maintenance and user billing costs will be reduced by increased automation.

A2.2 User Costs Categories

User costs can be divided into two categories:

a. User equipment costs (handset or ground terminal). There may be other costs implicit in the supply of the service (e.g., the existence of electricity, terrestrial phone lines, personal computer, or television receiver). These other items could be most significant in undeveloped regions.

b. System use charges (cost per minute, cost per packet, etc.) for a given type of service. Service could be priced according to data rate, number of packets sent, time of day, quality of service, etc. Each user also has a monthly charge for accounting services.

Sometimes the system provider may supply the user equipment free (e.g., cellular phones) in return for a commitment to a specified amount of system use. For simplicity, this discussion assumes that costs are separate.

Table A-2 shows how the number of users is expected to increase as the user equipment cost decreases. The numbers are based on costs ($1995) for U. S. consumers. A number of conclusions can be made. (It is assumed that reasonable system use charges or the unavailability of alternate systems provide the incentive for user purchases.)

a. Purchases of user equipment rise dramatically as user equipment costs drops below $500. The reason is that even small users with small telecommunications expenditures can justify the purchase. As seen with cellular phones and to a lesser extent with DBS receivers, a fad effect occurs in which everyone in a particular group must have one.

b. The acceptable cost of user equipment is related to the user's anticipated system use cost. User equipment costs exceeding one or two year's anticipated usage charges are unlikely to be acceptable, except in the case of mitigating factors such as for personal safety, plant security, and primary communications system backup.

Table A-2. Number of User Terminals vs. User Equipment Cost and System Use

	Small Users (<$300/yr) (Number = 90X)	Medium Users ($500–$1000/yr) (Number = 9X)	Large Users (>$2,000/yr) (Number = 1X)	Relative Number of User Terminals
Low cost user equipment (<$300)	Small users will buy.	Medium users will buy.	Large users will buy.	100X
Medium cost user equipment ($500 to $1000)	Few small users will buy	Medium users will buy.	Large users will buy.	20X
High cost user equipment (>$2000)	Small users will not buy	Few medium users will buy.	Large users will buy.	2X

The system architect tries to minimize user costs — both user equipment costs and system use charges. This will most probably require a tradeoff between increasing user equipment size and cost to increase system capacity and, hence, lower system use charges. Alternately, the satellite EIRP could be increased, resulting in less system capacity, but reducing user ground terminal size and cost. The existence of large numbers of users pushes the result towards lowering the cost of the ground terminals at the expense of higher system use charge.

The optimum solution can be calculated using minimum LCC as the performance criteria, if user parameters are known (number of users, user data rates, system utilization, etc.).

a. For the MSS systems currently under construction, FCC filings have given estimates of numbers of users and expected financial performance of the systems. It is interesting that financial viability is demonstrated for systems with a range of user equipment and user access charges (Globalstar, $0.60/min and $700 handset; Iridium, $3.00/min and $3,000 handset).

   The key design constraint for MSS systems is the user handset omni-directional antenna. These systems are designed to accommodate millions of users worldwide. Their economic success depends on the number of actual users. MSS systems have a system design limit to the number of users over a given geographic region; for example, 5000 4.8-kb/s duplex circuits over the continental United States.

b. DBS (BSS) systems do not limit the number of users in a given geographic region, since DBS is receive only (with perhaps provision for a low capacity order wire).

Thus the cost tradeoff favors lowering the cost of user equipment at the expense of increased space segment cost.

For example, DirecTV user equipment (0.45-meter dish plus receiver and codec) is currently priced at $500. (The user also needs a conventional television set which costs $200 or more, but most users in developed countries already have a TV.) The break-even point for DirecTV is 3,000,000 users in the United States, and they are significantly ahead of schedule in reaching this figure.

c. For Ka-band FSS systems such as Spaceway and CyberStar, it will be harder to optimize the relationship between user equipment cost and system use charges since the business plan may encompass users with different traffic volumes (i.e., business with 40 hr/week communications versus home with 1 hour/week. A flexible system design could allow for a mix of ground terminal sizes for different data rate users, but still has to target a particular traffic volume user.

   A typical Atlas IIAS class GEO FSS satellite may have 8 Gb/s simplex capacity. This allows 5200 1.5-Mb/s or 125,000 64-kb/s simultaneous one-way circuits. Depending on use patterns, 10 to 100 times as many users could share this capacity.

A2.3 Factors that Influence User Costs

Table A-3 lists a number of factors that influence user costs. Factors can be classified as system, user preference, technical, and cost.

System factors are functions of the system architecture. An example is "system utilization" which depends on the network control system, user access method, and communication protocols. Using reconfigurable antennas could allow the beams to follow-the-sun and improve utilization. Using switching on the satellite could avoid double hop connections. Using ATM protocols could more efficiently fill channels with packets.

User preference factors depend on user preferences, and may have different impacts than with terrestrial systems. Examples are data rate, availability, mobility, and location.

Technical factors relate to the technical details of the system implementation. A given of this study is to consider satellite systems. Usually technical factors are of no interest to the user except insofar as they influence user costs. (An exception can be user equipment size for mobile or transportable systems.) Examples of technical factors are antenna sizes and transmitter powers.

Cost factors are the cost of selected pieces of the satellite system. Examples are launch costs and user equipment cost. The LCC breakdown given in Table A-1 showed that user equipment cost was the largest component of the LCC.

Table A-3. Factors that Influence User Costs

	Influence of Factor on:	
Factor	**User Equipment Cost**	**System Use Charge**
System capacity (for a fixed size and cost satellite)	None	Directly related
System utilization	None	Directly related
User data rate	Increases somewhat for higher data rates.	Directly related
Service availability requested by user	Large increase for high rain regions.	Large increase for high rain regions.
Satellite antenna gain (transmit and receive)	Decreases somewhat with increased antenna gain.	Decreases somewhat with increased antenna gain.
Satellite transmit power	Decreases with increased satellite transmit power.	Increases somewhat with higher transmit power.
Ground terminal size (power and diameter)	Directly related (significant for larger sizes)	Inversely related.
Mass of satellite hardware (bus and payload).	None	Directly related since capacity is directly related to mass
Cost of launch service	None	Small increase with increasing launch cost.
Frequency band of operation	Varies with service	Higher for UHF, L-band, and EHF.
User mobility	Lower for smaller user antennas. Higher for transportable antennas.	Higher for lower gain user antennas; much higher for omni user antenna.
User geographical location	Higher for high rain regions or high latitudes using GEO satellites.	Higher for high latitude locations using GEO satellites.
Standards and protocols for interoperability	Complexity of user equipment interfaces.	Overhead traffic. May improve utilization.

A2.4 Conclusions of Satcom Economics

Satellites are favorable for developing nations, enabling the rapid establishment of a communications infrastructure without having to lay copper wire, cable, or fiber optics to a dispersed population. Thus the satellite communications scenario might consist of the following cases:

FSS Many (100,000) small users (64 kb/s or less) interconnected with duplex circuits; or a moderate amount (3000) of villages interconnected at medium rates (2 Mb/s).

BSS DBS (6 Mb/s TV) and/or DAR (384 kb/s audio) direct broadcast to the home or village; possibly with order wire (<1 kb/s). Millions of simultaneous users are envisioned.

MSS Regional mobile service to handhelds and village phone booths (4.8 kb/s to 64 kb/s duplex links). 100,000 simultaneous circuits shared by several million users is possible.

Within this service context, the satcom economics conclusions are as follows.

a. Technology applications should be judged on a systems level, with impact on user equipment cost and system use charge.

b. Ground terminal costs are the largest component of the overall system LCC. In a developing nations context, additional ground support costs may be required (e.g., electricity supply, computer, television.).

c. Relatively large, high power satellites are favored to reduce user terminal costs.

d. Technologies that increase system capacity or system utilization have a direct impact on the system use charge.

e. Satellite technologies are constrained by mass, in addition to cost and performance. Mass is typically a "zero sum" game on the spacecraft, considering use of a given size launch vehicle.

f. Satellites must be able to interconnect, without adverse cost impacts, with terrestrial wired and wireless networks. International standards and protocols should accommodate satellite as well as terrestrial communication links without prejudice to the satellite user equipment or system architecture.

A.3 SYSTEM DESIGN TO REDUCE USER COSTS

There are a number of ways to reduce user costs, ranging from incremental improvements to complete redesign of system function.

a. Reduce cost and mass of existing hardware by incorporating evolutionary technology advances. In the time period from 1996 to 2000, reductions of 50 percent may be feasible for the space segment, which could translate into 25 percent user cost savings.

b. Improve system utilization (MSS and FSS systems). The typical satellite communications system design is based on meeting peak traffic requirements, and achieves a 15 percent overall utilization (actual average traffic versus theoretical maximum traffic). Since utilization is directly related to user cost, there is considerable potential for cost reduction (50 percent user cost savings if utilization is doubled). Systems with flexible architectures (e.g., reconfigurable antennas, on board switching) are more likely to achieve high utilization.

c. Improve transport efficiency by minimizing overhead bits. This is unlikely to occur since the trend is to put more intelligence in the network, which results in more signaling and control bits being embedded in the communications traffic. This trend may indirectly improve utilization, but by itself slightly increases user costs.

d. Transmission capacity improvement has the greatest potential to reduce user costs for satellite communication systems. Spacecraft power is the limiting resource, and the user equipment needs a minimum W/m^2 to close the link. If the coverage area (e.g., CONUS) is fixed, the required transmit power is fixed. This is typically the case for broadcast systems with many users. However, for point-to-point communications, power can be reduced by shortening the range (GEO to MEO to LEO) or using a larger gain antenna (larger size or higher frequency). Potential cost reduction could be 90 percent on the satellite via a ten-times increase in capacity. The uplink from user to satellite also requires a large antenna on the satellite.

e. User equipment (ground terminal) cost can be reduced by using small handsets with omnidirectional antennas. As noted in point (d) above, this requires high power and/or large antennas on the satellite, and/or shortened range. The new LEO MSS systems (e.g., Globalstar, Iridium, Odyssey) use L-/S-bands, and are able to service a small number (600 per satellite) of 4.8 kb/s circuits by using LEO orbits.

   Small user terminals at higher frequencies could have autotracking phased arrays to improve link performance on the forward and return links. The satellite also would have a large receive antenna (high G/T) to enhance the user to satellite link.

New system architectures can be envisioned that use a very-high-gain steerable beam to transmit individual user traffic at high data rates, or multiple steerable beams at lower data rates. Using Ka-band is one approach, and using optical frequencies would be even better for making a limited number of point-to-point connections. Digital beam forming could be used to obtain multiple beams from one aperture.

A4. TECHNOLOGIES THAT INFLUENCE USER COSTS

User costs have been differentiated by user equipment costs and system use charges. This subsection discusses the trends in satellite communication systems, services, and technologies. Finally, the technologies that are expected to have the most influence on user costs are identified.

A4.1 Trends in Satellite Communication Systems

Several trends have developed through the 1970s and 1980s such as using larger, more powerful GEO satellites. However, the 1990s have brought some paradigm shifts such as using proliferated, smaller satellites. Trends are listed below.

a. GEO orbit has been the preferred location for communications satellites due to the advantages of:

   1. Continuous coverage of a large region from a single satellite

   2. The ability to use non-tracking ground terminals.

   Even worldwide systems such as Intelsat and Inmarsat used the GEO orbit. Only in the 1990s have Non-GEO Orbits (NGOs) such as LEO and MEO been proposed for worldwide systems. An exception has been the Russian use of highly elliptical orbits (Molniya orbits) for satellite communications to high latitude locations. The trend is now to choose the orbit that gives the best and lowest cost service to users.

b. Larger and more powerful satellites are being used in order to increase capacity and minimize impact on ground terminals. Larger satellites can support larger antennas and higher power transmitters. Economies of scale also come into play. However, size is constrained by capacity of launch vehicles, about 1900 kg wet mass for a GEO satellite with the Atlas IIAS or Ariane.

c. Constellations of smaller satellites are being proposed for worldwide systems in LEO or MEO orbits. The minimum size of these satellites is constrained by antenna size requirements and traffic capacity requirements. Within budget limits, required traffic capacity, and available spectrum, these satellites may also push the launch vehicle size envelope.

d. Payload mass fraction (proportion of communications payload mass to total satellite mass) is rising because of advances in materials and component designs. This allows smaller satellites with the same capacity or more capacity from the same size design.

e. Satellites power capacities are increasing. Current DBS satellites (1996 launch) have prime power capacities up to 8 kW. Planned satellites (Atlas IIAS class) have up to 15 kW capacity. The combination of high power and digital compression allows a DBS satellite to have enough capacity (100 channels) to be economically viable.

f. Antenna sizes increasing within the constraints of launch vehicle envelope size. Designs have been made for large, unfurlable antennas such as on TDRS and AMSC.

g. Processors speeds continue to increase, while costs and size decrease. The existence of high performance processors allows automatic control of ground terminals, automatic network control, and digital processing of communication signals (e.g., coding, compression, beam forming, and switching). All of this reduces costs and allows for new services.

h. Networks are using higher speeds (e.g., terrestrial networks using ATM/BISDN). Switching is being proposed on the satellite in order to enhance connectivity (Iridium MSS system under construction, and Spaceway and CyberStar proposed FSS systems).

i. User equipment has tended to become smaller and less costly (VSATs), even collapsing into handsets for low data rate MSS service (Globalstar, Iridium, Odyssey, Inmarsat P).

A4.2 Trends in Satellite Communication Services

Communication satellites are:

a. Carriers of the services offered to users by terrestrial communication networks

b. Suppliers of separate "bypass" or backup communication networks

c. Suppliers of unique services such as GPS time and position service and GOES meteorological mapping service.

In the 1970s and early 1980s, communication satellites were simply used to relay terrestrial communications via large trunking terminals. Thousands of voice circuits were multiplexed together and relayed from one office switch to another. Single television channels used entire transponders and CONUS coverage for distribution of network feeds to remote affiliate stations. There were few individual users with enough traffic to occupy an entire transponder and rich enough to afford 10-meter ground terminals (e.g., Dow Jones used satellites to distribute the "Wall Street Journal" to remote printing sites).

In the 1980s, VSATs (very small aperture terminals used with a large hub) allowed private networks to be established independent of the terrestrial communications infrastructure. These spread very rapidly due to favorable economics. Even though the VSAT equipment cost $5,000 to $10,000 or more, the circuit costs (lease of transponder) were greatly reduced compared to terrestrial for a network of many nodes over a large geographical area. Examples of private networks include banks (ATM transactions) and stores for inventory control (Walmart). The TV Receive-Only (TVRO) video pirates emerged in this time frame, with millions of C-band dishes being sold to those people beyond the reach of terrestrial TV, and to people avoiding local cable television fees. Worldwide mobile satellite service emerged via the Inmarsat system, using GEO satellites and small steerable antennas on ships.

In the 1990s, several satellite-unique services have emerged in addition to the previously existing services.

a. GPS supplies time and position information to users via $500 handsets.

b. GOES takes pictures of Earth and relays processed weather maps to users.

c. Worldwide MSS has been improved via smaller "briefcase size" Inmarsat terminals that are easily transportable and set up in minutes.

d. Direct broadcast satellite service (BSS) was established with DirecTV to provide television service direct to the user via $500 ground terminals. DirecTV (and Primestar) are designed to provide TV-on-demand and pay-per-view direct to the user.

e. Regional paging systems using satellites cover large areas such as CONUS or Europe.

Another trend is the use of terrestrial communications standards and protocols to allow interoperability between satellite and terrestrial systems, and to transport traffic conforming to such standards. Examples include MPEG video standards and ATM/B-ISDN network standards. A significant problem is that these standards have not necessarily been developed to accommodate satellite transmission, and may need to be modified to allow efficient use by satellite transmission networks.

The future holds more innovations, such as worldwide mobile satellite service via handheld sets that are dual mode, using the satellite system when no terrestrial cellular relay exists (Globalstar, Iridium, and Odyssey systems are under construction with goals of 1998 service). The proposed Orbcomm system would supply worldwide message service. The proposed GEO Ka-band Spaceway and Cyberstar systems will supply switched, fractional T-1 service worldwide for business services such as videoconferencing and document transfer, or connection to the Internet The proposed Teledesic system would provide 16 kb/s and higher links for e-mail, Internet connection, and other interactive services. Another need is for a regional and worldwide (including polar) air traffic control and routing system.

A4.3 Trends in Satcom Technologies

Satellite communications have benefited from the general advance of terrestrial as well as aerospace technologies. Specific items relating to the communications payload and user equipment are discussed below.

a. Antenna sizes are limited by the frequency band of operation and the required coverage area and gain. Satellite manufacturers have expended large amounts of effort to minimize antenna losses, costs, and mass, with considerable success. Current efforts are focusing on active antenna designs such as direct radiating phased arrays with active elements (phase shifters, power dividers, amplifiers, LNAs) in the antenna assembly. The promises of such designs are ease of integration and test (reduces cost) and operational flexibility (allows greater system utilization, which reduces cost, or allows new services).

b. Power amplifiers have advanced in capacity and efficiency, and followed the trend to higher frequencies (C- to Ku- to Ka-band, and even 60 GHz for ISLs). Efficiencies have made dramatic advances, particularly Traveling Wave Tube Amplifiers (TWTAs), and power conditioner units which now exceed 60 percent at lower frequencies. MMIC device efficiencies have also been improving, but are typically half that of TWTAs for the devices available. This is adversely affected the development of active transmit phased array antennas.

c. On-board processing technology has been slow to develop for satellite applications because of its mass and power consumption, which detract from payload capacity. ACTS and Milstar have taken first steps. Some commercial systems (Iridium, Spaceway, Teledesic) are now proposing to use on board processing (demodulation, digital switching, remodulation) to facilitate networking.

d. ISLs have not been used (except on some military systems), because of their impact on satellite capacity. (The mass of the ISL reduces the mass available for the other communications payload.) Some new systems (e.g., Iridium, Teledesic) propose to use ISLs to facilitate networking.

e. User equipment (ground terminals) has been reduced in size, primarily due to higher EIRP and G/T on the satellites. More capable, lower cost processors have allowed the automation of user terminal functions, thus reducing operations and maintenance costs, and hence user equipment cost.

f. Processors are being used in transmitters and receivers for source coding (compression and security) and channel coding (forward error correction). This allows cost reduction for supplying the service and supplying new services (pay-per-view) by eliminating signal piracy and password theft.

g. Packaging and design for manufacturability is becoming important for cost reduction of multiple satellite systems such as Globalstar (48 + 8 satellites) and Iridium (66 + 11 satellites).

A4.4 Technologies that Influence User Costs

This subsection takes the trends discussed above and lists technologies, roughly in order of impact on economics, that have the potential to influence user costs (user equipment and system use charges). There are three classifications: most important, important, and other. Technologies in the "other" category are believed significant, but relative could not be assessed.

Most Important Technology Areas

a. Standards and protocols that are "satellite friendly" and allow seamless, low-cost interoperability with terrestrial wired and wireless networks.

b. New system architectures to meet user needs via new services. System architectures that lower user costs. Hybrid space–terrestrial communication system architectures.

c. On-board processing and switching with low mass and power consumption.

d. Active antenna technology for spacecraft, gateways, and user equipment. Spacecraft active antennas require high efficiency MMIC HPAs, packaging to reduce cost, and integrated thermal control. Gateways for LEO systems require multiple beams (at least two beams for hand-offs) with wide field-of-view (down to 10° elevation angle). User equipment (at Ka-band) could use a small phased array to track the satellite(s) (to obtain higher gain).

e. User equipment (ground terminals and handsets) that is compact, transportable, and/or unobtrusive. Lower cost user equipment is required.

Important Technology Areas

f. Higher EIRP and G/T on the satellite to decrease user terminal size, or to increase user data rate for a given user terminal size. Larger satellite antennas, higher frequency bands, and higher power amplifiers are possible technologies.

g. Power generation and storage that is more efficient in terms of mass required for a given capacity.

h. Launch vehicles with lower cost and higher reliability.

i. Network control technologies to manage a number of satellites with interconnected communications services (e.g., Spaceway with 10 or more GEO spacecraft; Globalstar with 48 LEO spacecraft interconnected through gateways; Iridium with 66 LEO spacecraft interconnected via intersatellite links; and Teledesic with even more links).

Other Technology Areas

a. Laser links offer the highest EIRPs and hence potentially the best performance. How to use laser links for satcom through the Earth's atmosphere?

b. Intersatellite links with minimum mass and power impact on host satellite.

c. Technical support for U.S. positions on international spectrum and orbital assignments. Tools to enable timely analysis and resolution of inter-system interference issues.

d. Security of transmissions. U.S. source of encryption devices for foreign use.

e. Design for manufacturing of satellites, for systems with many satellites.

f. Technologies for efficient orbit and spectrum utilization.

A5. SPECIFIC RECOMMENDATIONS FOR TECHNOLOGIES

This section gives specific recommendations for technologies that have significant potential to lower user costs (user equipment cost or system use cost). Three tables are presented for FSS, BSS, and MSS satellite communication systems, including hybrid space-ground systems in these categories.

Table A-4 lists recommended technology developments for fixed satellite service systems. The listing is not in any prioritized order. Items not on the list are judged to be less important, or else are being developed as a matter of common practice by Industry. The FSS satellite for the year 2002 is envisioned to have a Beginning-of-Life (BOL) mass of 2000 kg and power of 12 kW, with 700 kg payload capacity. Major challenges are to have a flexible system architecture.

Table A-5 lists recommended technology developments for BSS systems. The BSS satellite for the year 2002 is envisioned to have beginning-of-life mass of 2000 kg and power of 15 kW, with 600 kg payload capacity. Major challenges are high power capacity; tranmitting many channels at one time; establishing an order wire for remote users; and low-cost user terminals.

Table A-6 lists recommended technology developments for MSS systems. These systems have many satellites (e.g., 12 MEO; 48 or 66 LEO) that function in a coordinated manner to provide communication services to a wide area. A challenge is the controlling interference, and the ability to operate in an interference environment (e.g., MSS to MSS; LEO to MEO to GEO).

Table A-4. Technology Developments for FSS Systems

Techmology	Rationale
Standards and protocols for interoperability with terrestrial wired and wireless.	Satellite must operate with terrestrial networks.
Flexible system architecture in assignment of resources (beams, bandwidth, power).	Improve utilization of satellite capacity.
High EIRP and G/T satellites	Allows higher capacity systems or lower cost ground terminals.
Optical communications between satellite and Earth (>500 Mb/s) (far-term application)	Allows high-data-rate communication links to be established via satellite.
Active phased array antennas with low cost and good efficiency (satellite).	Flexibility for reconfiguration on-orbit. Potential cost savings in integration and test.
High-efficiency power devices for active array antennas.	Low efficiency of active phased arrays is factor limiting their use.
Thermal control of phased arrays	Critical technology for feasibility.
High-power, high-efficiency, low-mass SSPAs (L-band to Ka-band).	Multiple carrier systems require linear power amplifiers.
High-data-rate (1 Gb/s) optical cross link.	Network of GEO or MEO satellites.
High-data-rate (1.2 Gb/s) up/down links.	Relay of high data rate terrestrial links.
On-board switching with low mass and power consumption.	Interconnections among multiple beams without double hop time delay.
Autotracking Ka-band user handsets.	Low-cost, transportable user equipment.
Position location of FSS users.	Security and control of service.
Wide angle electronic scanned user terminal. 1.5 Mb/s and 6 Mb/s.	Fixed and transportable (SNG) applications.

Table A-5. Technology Developments for BSS Systems

Technology	Rationale
Low-cost implementation of low-rate order wire (<1 kb/s) with position location.	Remote user can order pay-for-view video. Position is determined for content control.
Reconfigurable beam shaping of coverage area.	Flexibility in positioning and use of satellite.
Provision for local advertising to be inserted into nationwide programming (TV and radio).	Greatly increases value and appeal of service to resellers. Potentially reduces user cost.
Technologies for low-cost user equipment.	Lower cost of user equipment.
Provision for multiple language "dubbing" of video presentations.	Wider customer appeal.

Table A-6. Technology Developments for MSS Systems

Technology	Rationale
20-m deployable spacecraft antenna for high power use at L or S bands.	High EIRP required for GEO satellite to service handheld user terminals.
Moderate data rate (5 Mb/s) ISLs with low mass and low power consumption.	Each MSS satellite has 4 or more ISLs (intersatellite links)
Method to synchronize transmissions from user handsets	Allow more efficient use of spectrum with synchronous multiple-access schemes
Multicarrier demux and demod for CDMA and FDMA signals (low power & mass)	Allows switching on satellite with minimum mass and power penalties.
Baseband switch (64x64) with 10 Mb/s capacity. Minimize mass and power.	Allows switching on satellite for improved connectivity.
Wide-angle, multiple-beam scanning antenna for gateway (C-band and Ka-band).	Easier hand-offs
Wide angle electronic scan antenna for user (Ka-band MSS only)	Higher gain allows higher data rate user links. Mobile antennas require constant steering.
Digital beam forming of multiple beams	Each user can have optimized beam.
Standards and protocols for interconnection with terrestrial wired and wireless.	Interoperability with terrestrial networks.
Analysis and control of inter-system interference.	Self-interference limits capacity.
Bandwidth efficient systems.	Limited MSS bandwidth.
Design for manufacturing of multiple satellites.	Cost reduction. Competitiveness with terrestrial wireless services.

APPENDIX B – STATEMENT OF WORK

This appendix provides the NASA Statement of Work for this task order, "Space-Based Information Infrastructure Architecture for Broadband Services" and is divided as follows:

B.1 Background

B.2 Tasks

B.3 Reporting

The principles of the NII and GII as stated in the U.S. Government's document "National Information Infrastructure Agenda for Action" are as follows:

a. Provide open access to the network for information provider and users

b. Promote competition

c. Create a flexible regulatory environment that can keep pace with rapid technological and market changes

d. Encourage private sector investment

e. Ensure universal service

This document states that the private sector will lead deployment of the NII. Clearly, the private sector will develop many wide area networks for NII/GII applications and services where economic gains are the highest. The availability of NII/GII services will be difficult to achieve beyond these wide area networks and may not respond to the key NII/GII principle of "open access" to the public. Therefore, there are compellingly strong reasons for the U.S. Government to ensure that interoperability exists amongst the various telecommunication providers.

B.1 BACKGROUND

Satellite communications play an important role in providing universal service and accessibility at the national and global level. Satellite communication providers have offered, or plan to offer, a variety of NII/GII services. Other telecommunication providers are also aggressively pursuing NII/GII services. It becomes essential that terrestrial fiber, wireless, and cable networks be interoperable with satellite networks to provide open access and universal service for emerging NII/GII applications. The appliances used in providing NII/GII services also have to be interoperable within these networks.

NII/GII spans communication providers, network service providers, information content providers, and above all users (public, industry, nonprofit organizations, and government). As a result, technical challenges to satellite communication providers are numerous and go beyond the technological challenges of developing satellite technology. The Satellite Industry Task Force (SITF) which met early this year identified several challenges and presented them at the White House Forum. The technical challenges are standards, protocols and interoperability, and technology. The policy challenges are access to spectrum, space and markets.

To achieve interoperability, propagation time delay is the major contributor to the total time delay when satellite and terrestrial networks are compared. Delay becomes a problem for high-speed data communications serving low latency applications. There are also drastic reductions in throughput data rate when existing protocols are used for satellite networks. Satellite communications also have a higher channel Bit-Error Rate (BER) which affects Quality of Service (QoS) for multimedia applications.

Despite delay problems, satellite networks will play an important role in the emerging NII/GII. However, the role of the GEO, MEO, and LEO satellite networks may differ due to the nature of this problem. It is critical that quantitative analyses network performed to assess their performance in relation to satellite orbit location.

The vision of NII/GII is to create an open network accessible to users and service providers alike. To achieve this objective, a four-layer architecture for the Open Data Network (ODN) has been recommended by various NII policy-making committees. At the lowest level is the transport layer, bit way; the upper level is the application layer. Between the two levels are the network layer (the second layer) and the protocol layer (the third layer). Satellite communications play a role in the transport layer and have to interface with the other networks.

The technological challenge for satellite communication providers is to provide NII/GII services at same QoS level as the terrestrial communication providers. This will require hardware and software technologies working symbiotically to provide seamless integration of terrestrial and satellite networks.

Satellite communications perform very well in the present day TV-centric NII/GII; however, they lag behind in the future planned PC-centric NII/GII due to the nature of interactive applications. At present, TV is used in 96% of U.S. households whereas PCs are at a 30% level. As PC use increases, there will be a need to achieve interoperability of PC-based appliances over the networks. Similar developments are taking place globally.

Another issue is for the satellite communication providers to interface with the network layer, for example, ATM. Several key issues have to be resolved in this area. The SITE SIS committee identified several issues of immediate concern including:

a. ATM speech

b. ATM Traffic Management and Congestion Control

c. ATM QoS over satellite networks.

To provide NII services satellite communication providers also have to interface with a variety of appliances such as TVs, PCs, fax machines, telephones, video cameras, etc.. Interoperation of such appliances with the satellite network as well as other networks is another challenge. Undoubtedly, many issues will be solved by industry through technology and standard development activities. However, to achieve NII/GII objectives, a global approach must be taken to achieve interoperability of networks and appliances that entertains high risk, high payoff approaches and is supported by NASA and other government agencies.

Satellite designs and architecture are highly need-driven and are planned far in advance to maximize performance and fully utilize limited resources. This is in contrast to terrestrial fiber networks where high bandwidth capabilities are provided for developing NII/GII applications. The emerging services in the NII/GII require high bandwidth, mobility, transparency of transport layer, interoperability of appliances and networks, security, and low cost. The satellite technology and architectures need to be explored far in advance to meet NII/GII requirements.

The success of second and third generation satellite network systems will depend on their ability to provide interoperable emerging NII/GII applications. The spectrum and technology requirements for these systems need to be identified as soon as possible to achieve the objectives of NII/GII. As a first step toward achieving this goal, four subtasks have been defined below. It is expected that previous studies performed on NASA is behalf will be fully utilized during the execution of these subtasks.

B.2 TASKS

The four subtasks as listed and discussed below:

1. Identify satellite-addressable information infrastructure markets
2. Perform network analysis for space-based information infrastructure
3. Develop conceptual architectures
4. Economic assessment

Subtask 1: Identify Satellite-Addressable Information Infrastructure Markets

Provide an assessment of current, planned and emerging NII/GII applications and user markets and identify where satellites can play a significant complementary and/or competitive role in the emerging information infrastructure. This assessment level will be to the extent necessary to infer requirements for development of reference designs of second and third generation satellite systems. For reference purposes, first-generation systems are considered analog based; second generation systems are considered digital based and have beam-forming capability; and third generation systems will have high degree of seamless interoperability with terrestrial networks.

Subtask 2: Perform Network Analysis for Space-Based Information Infrastructure

Identify and prioritize satellite networks in LEO, MEO and GEO orbits requiring interoperations with wireline and wireless communication and information systems to provide universal services for the applications identified in Subtask 1. Analyze and assess the time delay and transport protocol issues pertaining to LEO, MEO and GEO systems. Define, in particular, user traffic characteristics and maximum simultaneous traffic requirements for interoperable satellite networks. Identify barriers to interoperability and the effect of standards and protocols on capacity, QoS, and user terminal cost.

Subtask 3: Develop Conceptual Architectures

Examine potential space-based information infrastructure architectures. Develop conceptual architectures for third generation satellite networks which could be interoperable with terrestrial networks including ground segment, addressing applications identified in Subtask 1 and issues in Subtask 2. Identify key technologies required for the conceptual architectures.

Subtask 4: Economic Assessment

Perform a preliminary, top-level economic assessment of the architecture concepts to judge business viability. Use this assessment in conjunction with spacecraft platform and ground terminal constraints to provide performance criteria for key technologies. Such criteria could include cost, mass, power, size, capacity, efficiency, etc.

B.3 REPORTING

The output of the subtasks will be detailed in the formal task Final Report. A formal Final Briefing at NASA-LeRC summarizing the output of the subtasks will be delivered in a viewgraph presentation along with 25 paper copies of the briefing as well one hard copy of the viewgraphs. Monthly technical, financial, and management status reports will be provided in accordance with Section F of the contract. Interim briefings at the NASA-LeRC on interim status and results will be provided at the direction of NASA Technical Manager.

APPENDIX C – DERIVATION OF SATELLITE-TO-EARTH STATION DISTANCE

Subsection 3.3.1 held the satellite-to-earth station distance as an explicit function of α, where $\alpha = FOV/2$, to be:

$$x = (R_e + h)\cos(\alpha) - R_e\sqrt{1 - [\gamma \sin(\alpha)]^2}$$

A derivation of this equation follows. Figure C-1 illustrates the angles and measurements discussed.

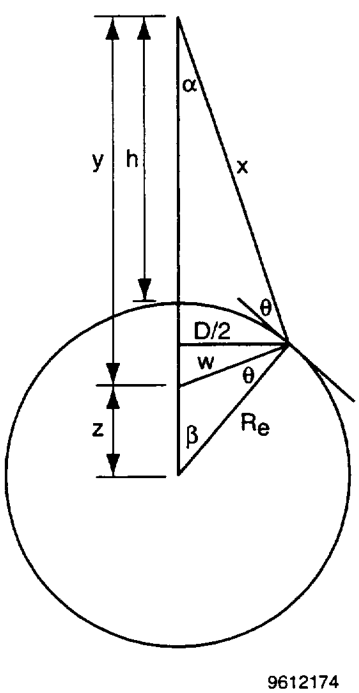


Figure C-1. x as a Function of α

$$x = \frac{w}{\tan(\alpha)} = \frac{w\cos(\alpha)}{\sin(\alpha)}$$

Also,

$$w = \frac{z\sin(\beta)}{\sin(\theta)} = \frac{(R_e + h - y)\sin(\beta)}{\sin(\theta)}$$

By substitution,

$$x = \frac{(R_e + h - y)\sin(\beta)\cos(\alpha)}{\sin(\theta)\sin(\alpha)}$$

Using the relation

$$y = \frac{x}{\cos(\alpha)}$$

we have:

$$x = \left[R_e + h - \frac{x}{\cos(\alpha)}\right]\frac{\sin(\beta)\cos(\alpha)}{\sin(\theta)\sin(\alpha)}$$

Solving this equation for x, one obtains:

$$x = (R_e + h)\frac{\sin(\beta)\cos(\alpha)}{\sin(\theta)\sin(\alpha) + \sin(\beta)}$$

Now, from E2-1 in subsection 2.3.2

$$\cos(\theta) = \gamma\sin(\alpha)$$

or

$$\sin(\theta) = \sqrt{1 - [\gamma\sin(\alpha)]^2}$$

and

$$\sin(\beta) = \frac{x}{R_e}\sin(\alpha)$$

where

$$\gamma = \frac{R_e + h}{R_e}$$

Substituting these last equations into the above equation for x yields:

$$x = \frac{\gamma x \sin(\alpha)\cos(\alpha)}{\sin(\alpha)\sqrt{1-[\gamma\sin(\alpha)]^2} + \frac{x\sin(\alpha)}{R_e}}$$

Again, solving for x in this equation yields the final desired result with x as a function only of α, the half-angle FOV.

$$x = (R_e + h)\cos(\alpha) - R_e\sqrt{1-[\gamma\sin(\alpha)]^2}$$

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)

2. REPORT DATE
December 1996

3. REPORT TYPE AND DATES COVERED
Final Contractor Report

4. TITLE AND SUBTITLE
Space-Based Information Infrastructure Architecture for Broadband Services

5. FUNDING NUMBERS
WU–625–50–5F
C–NAS3–27559

6. AUTHOR(S)
Kent M. Price, Tom Inukai, Rajendev Razdan, and Yvonne M. Lazeav

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Space Systems/LORAL
3825 Fabian Way
Palo Alto, California 94303–4606

8. PERFORMING ORGANIZATION REPORT NUMBER
TR01223

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Lewis Research Center
Cleveland, Ohio 44135–3191

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NASA CR–202310

11. SUPPLEMENTARY NOTES
Project Manager, Clifford A. Arth, Space Communications Office, NASA Lewis Research Center, organization code 6150, (216) 433–3460.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified - Unlimited
Subject Category 17

This publication is available from the NASA Center for AeroSpace Information, (301) 621–0390.

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 words)***
This report concludes that satellites will have a major role in the national and global information infrastructure, requiring seamless integration between terrestrial and satellite networks. The proposed LEO, MEO, and GEO satellite systems have satellite characteristics that vary widely. They include delay, delay variations, poorer link quality and beam/satellite handover. The barriers against seamless interoperability between satellite and terrestrial networks are discussed. These barriers are the lack of compatible parameters, standards and protocols, which are presently being evaluated and reduced.

14. SUBJECT TERMS
Satellite communications

15. NUMBER OF PAGES
170

16. PRICE CODE
A08

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102