NTNU
Norwegian University of
Science and Technology

A satellite system for broadband communications to polar areas

Lars Løge

Master of Science in Electronics
Submission date: June 2008
Supervisor: Odd Gutteberg, IET

Norwegian University of Science and Technology
Department of Electronics and Telecommunications

Problem Description

Communications to the merchant fleet is today mainly done with the help of geostationary satellites. Geostationary satellites have poor coverage north of 78-80 degrees northern latitude. As the activity, from oil drilling, ship traffic and an ice free Northwest- and Northeastpassage, increases in polar areas the need for transmission of large amounts of data will increase. Broadband communications to these areas will only be possible with the help of communications satellites in suitable orbits. Narrowband communications to these areas is today possible through the low earth orbiting system Iridium.

The task is to suggest a satellite system suited to provide broadband communications with bit rates in excess of 1 Mbps to the area north of 65 degrees northern latitude.

• Discuss possible carrier frequencies, analyse the propagation issues and find a suitable frequency configuration.

• Assess different satellite orbits and design a constellation that give continuous coverage of the northern polar areas.

• Look at different solutions for satellite antenna design. Create link budgets and find the upper capacity bound.

• Give a brief summary of important tasks for future work.

Assignment given: 24. January 2008
Supervisor: Odd Gutteberg, IET

Abstract

Over the last few years the ship traffic in polar areas have been steadily increasing, especially north of Norway. This is largely due to the growing activity in that area and in northern Russia, mainly from oil and gas exploration and production. All indications suggest that this will continue into the foreseeable future.

In this report a satellite based system for broadband communications to the area north of 65° northern latitude is discussed. Possible carrier frequency configurations and their propagation properties is analysed, and it is found that Ka-band, 20/30 GHz, will give best performance.

Various satellite orbits are then discussed, and a constellation that give continueous coverage and allow for easy handover is designed. It consists of four satellites in Molniya orbits with an eccentricity of approximately 0.72. Each satellite is then operational and quasi-stationary for six hours of every orbit, with two satellites above the coverage area at any time.

Solutions for the satellite antennas are considered, and link budgets are presented. Active phased arrays are found to provide the best performance. A total uplink capacity of 1.6 Gbps is teoretically possible with a user terminal output power of 100 W, but it is not deemed realizable. Instead a configuration with a total capacity on both uplink and downlink of about 1 Gbps is suggested.

At the end of the report a range of issues, related to the realization of the satellite system, requiring future attention is summarized and briefly discussed.

Preface

This report is written as a Master's thesis, and is the final part my Master of Science degree in Electronics from the Norwegian University of Science and Technology.

I would like to thank all the people who have helped me on my way. A special thanks to my supervisor Professor Odd Gutteberg for motivation and help on many technical issues. I also owe gratitude to all my friends and family who have supported me on my path towards wisdom and deeper knowledge.

Lars Løge
Trondheim, 17.06.2008

Contents

Chapter 1

Introduction

Over the last few years the ship traffic in polar areas have been steadily increasing, especially north of Norway. This is largely due to the growing activity in that area and in northern Russia, mainly from oil and gas exploration and production. All indications suggest that this will continue into the foreseeable future.

As the polar ice cap is melting the increased activity will probably also be augmented by larger areas beeing available for exploration. If the reduction in area occupied with pack ice around the north pole continues at the same rate observed over the last ten years, ships travelling between northern Europe and the American west coast can use the Northwest- and Northeastpassage sometime in the near future.

In the modern world people everywhere, regardless of their location, need communications systems. The polar areas are no exeption. However, far north the number and capabilities of possible communications systems are very limited. For ships and other activities far from land the only option is satellite based systems. In large areas, even on land, other alternatives do not exist.

Today the whole world is covered by the Iridium system, but it offers only a bit rate 2.4 kbps. That is enough for voice, but not much more. Inmarsat and similar providers uses geostationary satellites to guarantee its costumers data rates of up to 256 kbps. The problem with geostationary satellites is that they are not visible north of Svalbard at about 80° north. Furthermore the broadband access offered by Inmarsat does not even cover the satellite's whole field of view. Thus, there exist a growing need for broadband communications in an area that is not covered by any adequate systems today.

This report will try to answer some questions regarding how this need can be satisfied in a best possible way. It is not a comprehensive specification of a satellite system, but more a feasibility study that look at some possible solutions and their potential. The intention is to establish a bound for the throughput and capacity that can be expected from a satellite system providing coverage to polar areas north of 65° northen latitude.

In chapter 2 the issue of frequency selection is handled. Frequencies allocated for satellite communications applications by the International Telecommunications Union is used as a basis. The propagation properties of these frequencies are discussed before a suitable frequency configuration is selected.

Chapter 3 addresses the question of satellite orbit. Various types of orbits are discussed with the aim of finding one that maximizes the time possible users can access a satellite. This leads on into chapter 4 where constellation design is given attention. The excact orbit is defined more clearly and a constellation that are optimized for handover is suggested.

The communications system, which for these satellites is the payload, is looked at in chapter 5. The coverage area is discussed, and before presenting link budgets a brief summary of relevant transmission theory is given. Link budgets for three different satellite antenna options is then evaluated and compared with the aim of finding the best solution.

A summary of the chosen solutions can be found in chapter 6. The radio system is then devoloped further and revised link budgets for the preffered solution in chapter 5 is offered to the reader. At the end of the chapter a few problematic issues that can be challenging to solve is briefly discussed.

The last chapter, chapter 7, conclusions are drawn and a brief summary of subjects that should be given attention in the future as the process of developing this satellite system continues is given. Some of the subjects are technical tasks that are important for the realization of the system, while others are oppertunities that can be very beneficial and enhance the value of the system.

To simulate satellite orbits and antenna coverage the computer program Satellite Tool Kit (STK) created by Analytical Graphics Inc. (AGI) has been used. Figures with satellite orbits and field of view have been created in this program.

Chapter 2

Frequency selection

One of the most important parameters in a satellite based communication system is the operating frequency. The choice of frequency have impact on almost every subsystem on the satellite. An increase in frequency allows for smaller antennas without a reduction in antenna gain. Also at high frequencies the available bandwidth is large, thus, a larger data rate is possible. However, atmospheric conditions attenuate the signal more at higher frequencies so the data rates must be reduced to compensate for the degraded signal.

2.1 Allocated frequencies

Since there is only one frequency spectrum there is limitations on frequency use. Without frequency coordination radio communications systems would in some cases use overlapping frequencies, and interfere with each other. Thus, few systems would have optimal performance. Frequency allocation and assignment is coordinated through the International Telecommunication Union (ITU) and national governments. They allocate frequencies for different purposes, including satellite communications. Several large frequency bands for use by present and future satellite based communications systems are already allocated. Thus, the possible frequencies to choose between are within these bands. Allocation tables for the three geographical regions of the world can be found in ITU Radio Regulations [1]. The Radio Regulations are revised every third year by the World Radiocommunication Conference.

The frequency bands allocated for satellite communications up to just bellow 20 GHz are for the most part already in use. Above 20 GHz, however, the availability is better. Around 20 GHz there is allocated almost 3 GHz of bandwidth to satellite communications, primarily for signals from space to

TABLE 1: Possible frequency configurations.

ALTERNATIVE	DOWNLINK BAND [GHz]	UPLINK BAND [GHz]
A	20	30
B	40	45
C	40	51
D	70	70

earth. A bandwidth of 3.5 GHz is allocated around 30 GHz for earth to space communication, while almost 10 GHz is available for satellites communicating in both directions starting at 37.5 GHz. A small band of only 1 GHz exist at 51 GHz. There is also allocated 10 GHz between 66 and 76 GHz [1]. At higher frequencies there are also bands set aside for satellite communications, but they will not be taken into consideration in this study.

These allocations present several possible frequency configurations. For simplicity in the evaluation process only four generalized configurations will be considered. In satellite communications it is customary to use separate frequencies for uplink and downlink. A satellite has a limited amount of power available, thus, the lowest frequency is normally chosen as the downlink frequency. The four frequency configurations that will be considered here are shown in table 1.

2.2 Free space loss

In general a radio system link can be regarded as two antennas, one transmitting and one receiving, with a distance, D, between them. An omnidirectional antenna will radiate isotropically, and thus divide its transmitted power equally in all directions. As a result the power density, S, in Watts per meter at a distance D from the antenna is:

$$S = \frac{P_t}{4\pi D^2} \tag{1}$$

where P_t is the transmitted power [2]. Now if the transmitting antenna has a gain G_t, and the receiving antenna has an effective area A_r the received power P_r in Watts is

$$P_r = \frac{P_t G_t A_r}{4\pi D^2} \tag{2}$$

The part bellow the fraction line in equation 2 is the actual free space loss, and it is independent of the frequency. However, in radio link design it is customary to define the antenna gain as the ratio of its effective aperature area to the effective area of a hypothetical isotropic antenna [3]. The result is that the antenna gain can be expressed as

$$G = \frac{4\pi A_e}{\lambda^2} \tag{3}$$

where G is the gain, A_e is the effective area and λ is the wavelength [4]. When this is used to find an expression for A_r the result is:

$$A_r = \frac{G_r \lambda^2}{4\pi} \tag{4}$$

Substituting (4) into (2) gives:

$$P_r = P_t G_t G_r \left(\frac{\lambda}{4\pi D}\right)^2 = \frac{P_t G_t G_r}{L_S} \tag{5}$$

where L_S equalls the normal definition of the free space loss [5].

The wavelength is given by the speed of light divided by the frequency. Thus, the free space loss, L_S, can be expressed as

$$L_S = \left(\frac{4\pi D}{\lambda}\right)^2 = \left(\frac{4\pi f D}{c}\right)^2 \tag{6}$$

where c is the speed of light and f is the frequency. In order to ease calculations (6) can be converted into dB. When the distance, D, is given in km and the frequency, f, is given in GHz, the space loss in dB is:

$$L_S = 92.4 + 20\log(D) + 20\log(f) \tag{7}$$

From (7) and (6) it is evident that the an increase in the frequency will increase the free space loss. To be more precise, a doubling in frequency entails a 6 dB larger free space loss. In table 2 the free space loss for the three different downlink frequencies given in table 1 is presented with 20 GHz as the reference frequency. The free space loss for the four uplink frequencies is shown in table 3 with 30 GHz as the reference frequency. This suggests that the alternative with the lowest frequencies should be selected for minimized free space loss.

However, an increase in frequency will result in a higher antenna gain for the same aperture area, as stipulated by equation 3. A frequency increase can therefore not necessarily be said to have the adverse effect suggested by the free space loss. This can be confirmed by examining equation 2 which is, as discussed above, independant of frequency.

TABLE 2: Space loss for the downlink frequencies.

FREQUENCY [GHz]	SPACE LOSS [dB]
20	0.0
40	6.0
70	10.9

TABLE 3: Space loss for the uplink frequencies.

FREQUENCY [GHz]	SPACE LOSS [dB]
30	0.0
45	3.5
51	4.6
70	7.4

2.3 Propagation loss

On its way between the satellite and earth, or vice versa, a signal will, in addition to normal free space loss, be attenuated due to various effects and phenomena. Some effects vary with time while others are more or less constant. In the atmosphere some of the signal power are absorbed, mainly by water vapour and oxygen. For a given satellite altitude and ground station elevation angle this is a fairly constant effect.

The time variable effects include tropospheric scintillation due to changes in the refractive index, and Farady Rotation and scintillation in the Ionosphere. Clouds in the signal path will also degrade the signal, and for frequencies above 10 GHz rain attenuation must not be forgotten. Depolarization due to ice crystals and aerosols is another source of propagation loss [4]. The variable effects will normally apear as fades with different amplitude depending on the effect and its severity.

All effects contributing to propagation loss, be it constant effects or variable effects, are frequency dependant. Thus, the choice of frequency must be evaluated against all these effects. In the following sections the various losses will be discussed with focus on the four frequency configurations given in table 1.

2.3.1 Atmospheric absorption

Signal attenutation due absorption in the atmosphere are mainly caused by non-symetrical molecules and molecules with a permanent magnetic moment. Non-symetrical molecules, also called polar molecules, have a preferred orientation when placed in a electromagnetic field. If the electromagnetic field changes direction a polar molecule will attempt to realign itself with the new field direction. The energy needed to achieve this movement the molecule get from the electromagnetic field. A radiowave carrying information from a satellite is an electromagnetic field, and in the atmosphere there are polar molecules. Thus, a signal traveling through the atmosphere is attenuated. The absorption is dependant on frequency and type of molecules. At frequencies bellow 70 GHz water and water vapour are dominant. In addition to background attenuation water vapour is resonant at 22.3 GHz [6].

Oxygen has a permanent magnetic moment giving rise to several resonant absorption lines aroung 60 GHz. These resonance lines are close to each other, and under normal conditions they are smeared in to one broad absorption peak. The resonant absorption peaks of oxygen are in addition to the general background attenuation oxygen causes. The background absorption of both water vapour and oxygen increase with the frequency [6].

All polar and paramagnetic molecules in the atmosphere attenuates the signal, but only water vapour and oxygen have observable effects. Absorption from other gases is negligble in the frequency range of interest in this study [7].

In [8] formulas for calculation of the attenuation caused by atmospheric gases are given. There also the total zenith attenuation in a standard atmosphere is found. A standard atmosphere is a mean annual global reference where the surface pressure is 1 013 hPa, the surface temperature is 15 °C and the surface water vapour density is 7.5 g/m^3. Since the attenuation is dependant on frequency a plot like the one in figure 1 can be produced.

Using figure 1 an estimate of the zenith attenuation in the frequency bands of interest from sea level can be found. Up to 40 GHz the attenuation is bellow 0.5 dB. After 40 GHz the attenuation increases rapidly up to the oxygen absorption peaks around 60 GHz before it deacreases again towards 80 and 90 GHz. Table 4 lists the zenith attenuation for the frequencies of interest given in table 1.

In high latitude areas the height of the atmosphere is smaller than it is around equator. The temperature is also lower which has the effect of decreasing the water vapour density. Thus, it should be possible to assume that the results extracted from figure 1 is the worst case scenario for atmospheric attenuation in the geographic area of interest in this study.

All ground stations not positioned at the sub satellite point will have

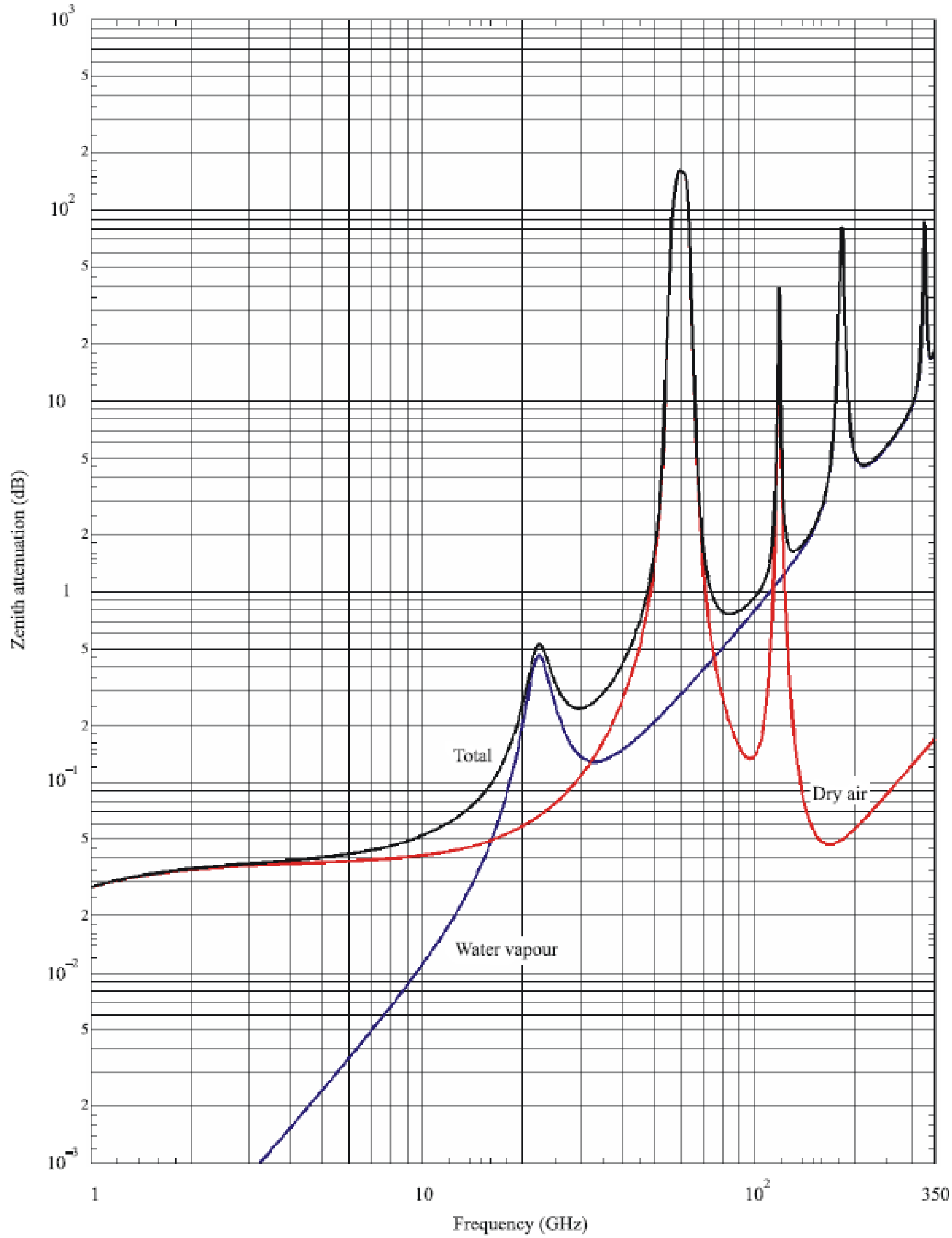


FIGURE 1: Total, dry air and water vapour zenith attenuation from sea level when a standard atmosphere is assumed [8].

TABLE 4: Atmospheric attenuation from sea level when a standard atmosphere is assumed.

FREQUENCY [GHz]	ZENITH [dB]	60° ELEVATION [dB]
20	0.3	0.3
30	0.3	0.3
40	0.4	0.5
45	0.7	0.8
51	2.0	2.3
70	1.9	2.2

a longer path through the atmosphere. This will increase the attenuation compared to the zenith values. For elevation angles between 5° and 90° the attenuation caused by atmospheric absorption can be found using the zenith attenuation and the cosecant law [8]:

$$A = \frac{A_{zenith}}{\sin \varepsilon} \tag{8}$$

where A_{zenith} is the zenith attenuation and ε is the elevation angle. In table 4 the approximate attenuation due to atmospheric absorption at an elevation angle of 60° are shown for the various frequencies of interest.

Table 4 show that the lowest frequencies have the lowest loss. With an elevation angle of 60° both frequency configuration A and B have a one way atmospheric attenuation bellow 1 dB. Because of the uncertainties in these estimates alternative A and B can be regarded as almost equall in terms of atmospheric attenuation. Alternative C and D both have an attuation above 2 dB on the uplink frequency. However, since alternative C uses the same downlink band as alternative B, it outperforms alternative D.

2.3.2 Rain attenuation

When a radiowave hits a drop of water there will be absorption and scattering. This of course cause the signal to be attenuated. Since rain is a collection of drops of water a radio signal passing through rain will deteriorate. At low frequencies the wavelength is large compared to the raindrops. Attenuation due to scattering as well as absorption is then negligible [6]. This changes at higher frequencies, and for frequencies above 10 GHz rain is the dominant propagation phenomenon [4].

In order to calculate the rain attenuation some information is needed about the location of the ground stations. That is the rain height and the

TABLE 5: Rain attenuation from sea level with 60° elevation.

FREQUENCY [GHz]	0.01 % [dB]	0.1 % [dB]	1 % [dB]	2 % [dB]
20	11.0	3.7	0.9	0.5
30	21.6	7.8	2.0	1.2
40	32.4	12.3	3.3	2.1
45	37.4	14.3	3.9	2.4
51	42.7	16.6	4.6	2.9
70	51.9	20.6	5.8	3.7

rainfall rate in the coverage area. North of 70° latitude the rain height can be assumed to be somewhere between 1.5 km and 2.5 km [9]. The polar areas are almost like deserts in terms of rain, but in coastal areas around 70° latitude the rainfall rate can almost reach 30 mm/h for more than 0.01 % of the time [10].

Using procedures given in [11] the rain attenuation exceeded for a given percentage of time can be found. This procedure is strictly speaking only valid up to 55 GHz, but lacking other estimation methods it will be used here for the 70 GHz band.

Table 5 show the estimated rain attenuation exceeded for different percentages of time for an elevation angle of 60° when situated on the northern coast of Norway, around 70° north. These values are averaged over a year. Thus, the attenuation levels given for 0.01 % will be exceeded for a total time of under one hour during a year. 0.1 % equalls just under nine hours, while 1 % is a little more than three and a half day during a year.

When evaluating the rain attenuation the service probability comes into play. If a service probability of 99.99 % is demanded table 5 show that huge margins for rain attenuation is needed in all the frequency bands. That will probably not be worth the cost, so all of the proposed frequencies are out of question if a service probability of 99.99 % is needed. If a service probability of 99.9 % is tolerated, frequency alternative A probably has acceptable rain margin requirements. The other frequency alternatives is most likely not usable unless an outage time of 1 % can be accepted.

These attenuation values may look very discouraging, but a few things must be kept in mind. These figures can be assumed to be the worst case. Thus, in real life the neccessary margin should by all accounts be smaller. Another thing to remember is that the precipitation events causing these fades will never extend over the complete coverage area of the satellite. At any time some of the users may lose contact with the satellite for a short time due to local heavy rain, but the mayority of the users will not be affected by that in any negative way.

2.3.3 Attenuation due to snow

The effect of snow in the propagation path have received limited attention in the research community. This is probably because snow have little effect on the frequencies traditionally used, and it is a rare phenomenon in most parts of the world.

Experiments done in Norway indicates that dry snow have a negligible effect on frequencies bellow 30 GHz. Wet snow, or sleet, can on the other hand have a crippling effect on a satellite link. Larger attenuation than what the equivalent rain rate would give have been measured [12]. For frequencies above 30 GHz the available research data is scarce.

In the core of a flake of wet snow is a crystal structure that suspends the outer edges of water. As a result flakes of wet snow looks like super sized raindrops. This increased size will give more scattering than the equivalent rainfall rate cause. From this it should be possible to conclude that attenuation due to sleet or wet snow is larger, but proportional to the equivalent rainfall rate [13]. Thus, attenuation due to wet snow should have more or less the same impact on frequency selection as attenuation from rain.

2.3.4 Cloud and fog attenuation

Rain and snow are not the only hydrometeors that can be present in the signal path between earth and space. Clouds and fog can also attenuate a signal. Rain has by far the most devestating effect on the communication links, but clouds are present more often and must not be forgotten at higher frequencies in low margin systems.

Fog usually do not extend more than 150 m above ground. With such thin layers of fog the attenuation on earth-space paths are very low. Thus, up to 100 GHz attenuation due to fog can be ignored [7].

Equation 9 give the cloud attenuation in dB exceeded for a certain percentage of time. It is valid for elevation angles between 5° and 90° [14, 15].

$$A = \frac{LK_l}{\sin \varepsilon} \tag{9}$$

Here L is the total columnar content of liquid water, K_l is the specific attenuation by water droplets and ε is the elevation angle.

Values of L for various time percentages can be found in [14]. For example, 20 % of the year the total columnar content of cloud liquid water will exceed

TABLE 6: Cloud attenuation from sea level with 60° elevation.

FREQUENCY [GHz]	1 % [dB]	5 % [dB]	10 % [dB]	20 % [dB]
20	0.1	0.1	0.0	0.0
30	0.2	0.1	0.1	0.0
40	0.3	0.2	0.1	0.1
45	0.4	0.3	0.1	0.1
51	0.5	0.4	0.2	0.1
70	1.0	0.7	0.3	0.2

0.1 kg/m^2 in parts of the area north of 70° latitude. The specific attenuation by water droplets is temperature dependant, but in clouds the temperature will in general lie around 0 °C. The specific attenuation can be read from figure 1 in [14], or calculated from formulas given in [14] and [15].

The estimated cloud attenuation for 1, 5, 10 and 20 % is shown in table 6. It is evident that clouds have a smaller impact on a satellite link than rain, but at 70 GHz the cloud attentuation exceeds 1 dB for a total of around three and a half day during a normal year. However, since the communication link must be designed with a margin substantial enough to handle the rain attenuation, which will be several dB as shown in section 2.3.2, the cloud attenuation should have little impact on the frequency selection.

2.3.5 Depolarization loss

The earth's atmosphere can under certain conditions change the polarization of a radiowave. When this occur the antenna will have a polarization mismatch with the received wave. This is normally reffered to as depolarization loss. In the lower part of the atmosphere depolarization is usually caused by rain or ice crystals [7].

Small raindrops are spherical since the surface tension is the dominant force acting on the raindrop. As the size increase external forces such as aerodynamic drag will affect the shape of the raindrop turning it into an oblate spheroid. A radiowave crossing an area containing nonspherical raindrops will experience different attenuation and phase shift for the two characteristic polarizations. Ice crystals can cause a differential phase shift in a radiowave when the ice crystals are not randomly oriented, but have a preffered direction. In both cases the result is a change in the polarization of the radiowave [7].

For frequencies above 35 GHz and elevation angles above 60° very little reasearch have been done regarding the depolarization caused by rain and ice crystals. However, it has been established that rain is the prominent cause of depolarization on signal paths between earth and space [7].

Using the rain attenuation figure for 30 GHz given in table 5 and procedures given in [11] it can be shown that the cross polarization discrimination will be more than 42 dB in 99.9 % of the time. Circular polarization and an elevation angle of 60° is then assumed. This is a substantial higher cross polarization discrimination than what has been measured at Kjeller and Svalbard towards geostationary satellites earlier [12], but the increased elevation angle accounts for most of the difference. This result suggest that in the dry climate of the polar areas depolarization losses are negligible for the frequencies discussed in this study.

2.3.6 Faraday rotation

The ionosphere is an inhomogenous and anisotropic plasma. A radiowave entering the ionosphere can be divided into two characteristic polarizations. The two characteristic waves experience different phase shift and attenuation. Thus, when the two waves are recombined the polarization of the total field has changed. This phenomenon is known as Faraday Rotation, and can be a source of depolarization loss [4].

Faraday rotation is inverse proportional with the square of the frequency. The result is that a radio wave with a high frequency will have very little polarization rotation. At 20 GHz the worst case is a rotation of a few tenths of a degree [4]. Thus, it is evident that the polarization loss caused by Faraday rotation is negligible at all frequencies relevant for this study.

2.3.7 Scintillation

Scintillation is rapid fluctuations around the mean of the radiosignal, either in phase or amplitude [6]. These fluctuations can be produced in both the ionosphere and the troposphere. In the ionosphere irregularities in the electron density can effect signals with frequencies up to 6 GHz [7]. Thus, ionospheric scintillation do not affect the frequencies discussed here.

Tropospheric scintillation are produced by rapid changes in the refractive index within 4 km of the earth's surface. These changes occur when air from different humidity and temperature layers are mixed [4]. This effect have been observed on frequencies from 3 to 30 GHz [12]. However, it is only at low elevation angles that tropospheric scintillation have any considerable impact. For elevation angles above 20° it can be more or less ignored [7].

2.4 Radio noise

All materials that absorbs radiation will also emit radiation as long as their physical temperature is above 0° Kelvin. Thus, everything that surrounds us acts as a black body radiator to some extent. Up to 300 GHz this radiation can be ragarded as additive white gaussian noise. For a radiosystem with a noisebandwidth B the noise power from a black body radiator is given as

$$N = kT_aB \tag{10}$$

where k is Boltzmann's constant and T_a is the equivalent noise temperature [4]. The equivalent noise temperature of a medium is directly connected to the amount of energy the medium absorbs. Thus, the noise temperature can be expressed as a function of the attenuation caused by the medium:

$$T_a = T_m \left(1 - 10^{-\frac{A}{10}}\right) \tag{11}$$

where T_m is the average physical temperature of the medium in Kelvin and A is the attenuation in dB [7].

The equivalent noise temperature of an absorbing medium is added to the antenna noise temperature. Thus, all the absorptive effects discussed in section 2.3 will in addition to attenuate the signal also increase the noise. However, large variations in radio noise will only be experienced in the downlink. The receiving antenna on the satellite will normally only see a part of the earth's surface. A satellite antenna fixed on the northern part of the world will observe a more or less constant noise temperature around 280 K. As a result the frequency has only impact on the noise temperature of the downlink [7]. Thus, only the downlink frequencies will be discussed in the following sections.

2.4.1 Noise from atmospheric gases

Using equation 11 and the attenuation values from table 4 the noise temperature of the gases in the atmosphere can be found. In table 7 the equivalent noise temperature of the atmosphere is given for the frequencies in question. The estimates are based on a mean path temperature, T_m, of 270 K at an elevation angle of 60°.

From table 7 it is evident that increased frequency also give an increase in noise from atmospheric gases. This corresponds well with the findings in section 2.3.1, and strenghtens the assessments made there.

TABLE 7: Noise temperature due to atmospheric gases in standard atmosphere with an elevation of 60°.

FREQUENCY [GHz]	NOISE TEMPERATURE [K]
20	18.0
40	29.4
70	107.3

TABLE 8: Noise temperature due to rain with an elevation of 60°.

FREQUENCY [GHz]	0.01 % [K]	0.1 % [K]	1 % [K]	2 % [K]
20	253	158	51	30
40	274	259	146	105
70	275	273	203	158

2.4.2 Noise from rain

In the same way as for atmospheric gases the increase in noise temperature due to rain can be found. The best prediction of the mean physical temperature of rain is 275 K [7]. If the attenuation levels given in table 5 are used the equivalent noise temperature of the rain will be as shown in table 8 for an elevation angle of 60°.

Table 8 show that in addition to deep fades rain will substantially increase the antenna noise temperature. Rain therefore have an even larger impact on the received signal than the attenuation levels suggests when a low noise receiver is used. However, it should be noted that for rain attenuation levels above 10 dB the noise temperature increases asymptotic towards the physical temperature. Thus, if a service probability better than 99.99 % is needed the choice of frequency will have little impact on the possible increase in noise temperature due to rain.

2.4.3 Noise from clouds

Table 6 indicates the attenuation caused by clouds are marginal compared to the possible rain attenuation levels. Then it is natural to asume that the increased noise from clouds will be well below the noise produced by rain. Thus, minimizing the noise due to clouds should not be prioritized when selecting the frequency.

2.4.4 Cross polarization noise

When two ortoghonal polarizations are utilized the effects discussed in section 2.3.5 can cause the two signals to interfere with each other. If the two signals are independent of each other it is customary to assume the interference behaves as additive white gaussian noise.

This interference noise can in many cases degrade the signal environment more than the attenuation caused by depolarization. However, as discussed in section 2.3.5 the cross polarization discrimination is very high. Thus, cross polarization noise should not be an issue here.

2.5 Evaluation

The two way loss for the four frequency configurations can now be found by adding the relevant losses. As discussed in section 2.3 that is the atmospheric attenuation and the rain attenuation. Attenuation due to wet snow or sleet can, as discussed in section 2.3.3, in some cases be larger than rain attenuation. However, since attenuation due to snow is proportional to rain attenuation they have more or less the same impact on the frequency selection. Rain attenuation can be predicted more accurately and will hence be used for evaluation here. The rest of the propagation losses originating in the atmosphere are too small to have a significant impact.

Under normal conditions only one of the signal directions will be influenced by rain. Since the satellite will have limited ability to compensate for increased attenuation the worst case scenario is that the downlink experiences rain. The estimated two way attenuation for the four frequency configurations are given in table 9.

In addition to the attenuation values given in table 9 the free space loss will be larger for higher frequencies as discussed in section 2.2. This can however be slightly misleading as a higher frequency allows for smaller antennas without reduction in gain. In section 2.2 it is mentioned that the actual free space loss is independant of frequency, thus, the free space loss will not be given much weight when deciding the best frequency to use.

The result shown in table 9 clearly indicates that the best frequency configuration in terms of attenuation is alternative A. In terms of radio noise alternative A is also the best choice. The differences between the alternatives are not as large, but even without rain alternative A is more than 2 dB better than alternatives B and C, and almost 8 dB better than alternative D. However, with regards to the total system noise temperature of the receiver the difference between the four alternatives are smaller.

Table 9: Total worst case attenuation from sea level with 60° elevation.

Alternative	0.01 % [dB]	0.1 % [dB]	1 % [dB]	2 % [dB]
A	11.6	4.3	1.5	1.1
B	33.7	13.6	4.6	3.4
C	35.2	15.1	6.1	4.9
D	56.3	25.0	10.2	8.1

The size, and then in turn complexity, of a communications satellite is more or less scaled by the needed transmit power [3]. By minimizing the signal attenuation and noise the needed transmit power is also minimized. Thus, in general terms it can be concluded that frequency alternative A will most likely give the most cost effective satellite system, and should therefore be chosen here.

Chapter 3

Satellite orbit

The choice of orbit is crucial in a satellite system. A space based communications system will only bring in revenue when the users are within the coverage area of a satellite. Thus, the time a satellite spend with its users inside the field of view should be maximized.

The orbits that normally are used for communication purposes can be divided into three generall categories. Those are the geostationary earth orbit (GEO), low earth orbit (LEO) and highly elliptical orbit (HEO). In the following sections the coverage performance in polar areas and assumed cost efficiency of each category will be assesed.

3.1 Geostationary earth orbit

Conventionally high capacity communications systems have utilized satellites in a geostationary earth orbit. A satellite in geostationary earth orbit appears to be stationary over a fixed point on earth. As a result only a single satellite is needed for continuous coverage of a large area [4].

There is however some limits to which areas a geostationary satellite can cover. A geostationary satellite orbits around the equator at an alltitude of 35 786 km. Thus, areas north of about 80° north can not be covered by a geostationary satellite [16]. This is illustrated in figure 2 which show the view towards earth from a geostationary satellite positioned at 10° east.

In figure 2 it is possible to see all of Norway, Iceland, parts of Greenland and just barely Svaldbard. This suggest that it is possible to use geostationary satellites for communications as far north as Svalbard [12]. However there are issues with the use of a geostationary satellite that make this a less attractive option. The main issue is the low elevation angle at the ground stations. At low elevation angles the atmosphere have increased effect on the

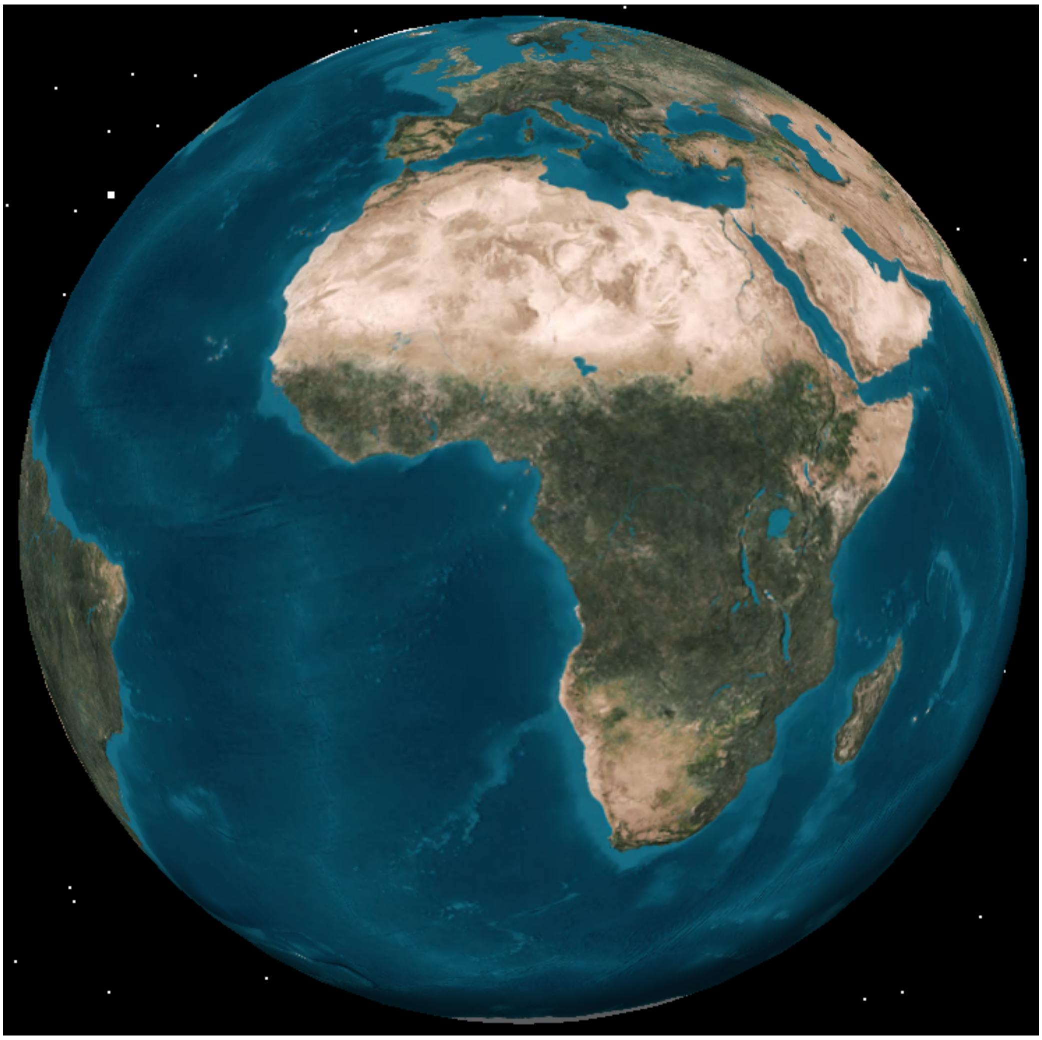


FIGURE 2: View of the earth from a geostationary satellite positioned at 10° east.

radio signals. Atmospheric attenuation increases and both tropospheric and ionospheric scintillation cause large variations in signal strength. Thus, it is not desirable to use a satellite for communications when the elevation angle drops bellow 10° [11].

Figure 3 show a standard two dimensional map with a selection of elevation angles between 0 and 25° for a geostationary satellite positioned at 10° east. The northern tip of Norway are just north of the 10° elevation contour, thus, indicating that satellite communcations north of 70° lattitude should be done with non-geostationary satellites.

Moving south the elevation angle towards a geostationary satellite in-

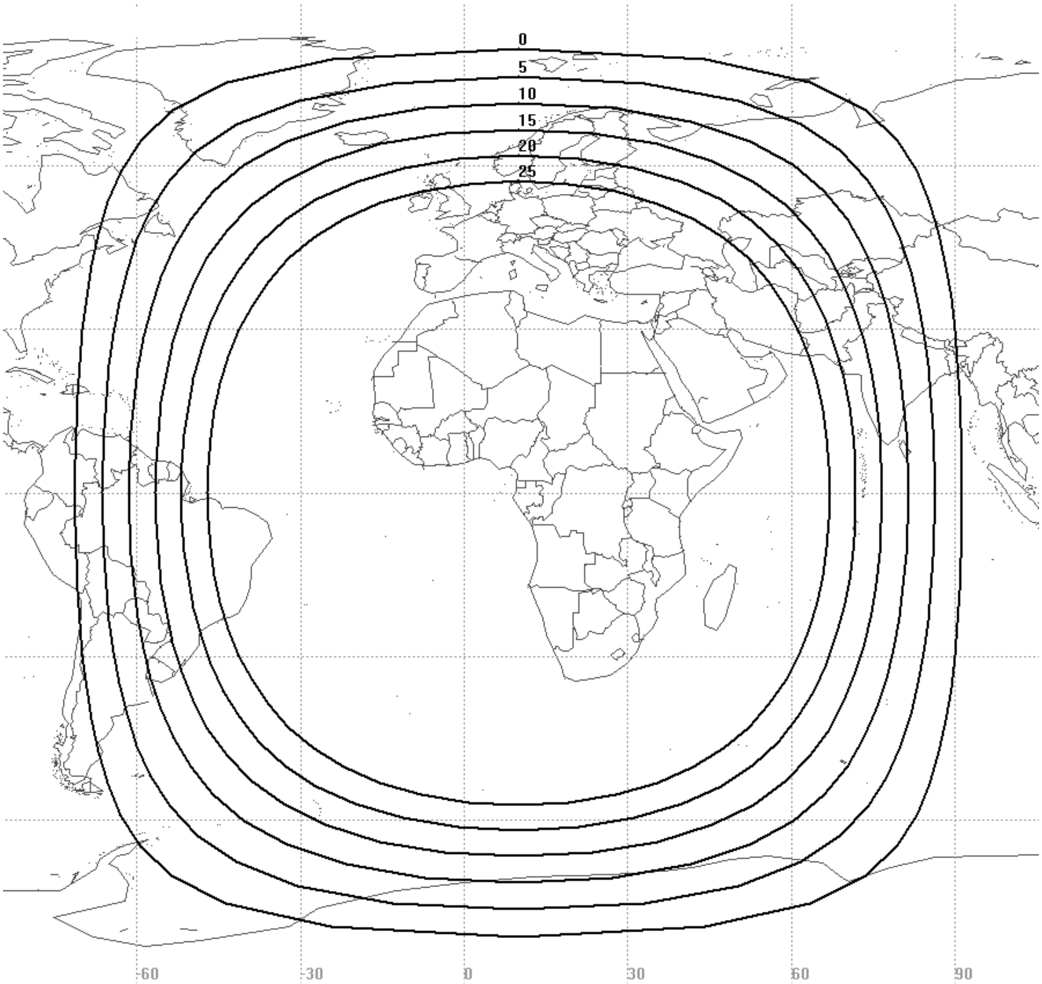


FIGURE 3: Two dimensional map with contours denoting the elevation angles seen from a geostationary satellite positioned at 10° east.

creases slowly, and at the southern tip of Norway it is still bellow 25°. Ships and other users of satellite communication systems in these areas can also be better served by a satellite system providing a higher elevation angle.

3.2 Low earth orbit

A satellite is generally said to be in low earth orbit if the orbit is fairly circular and has an altitude between 500 km and 1 500 km. This definition is dictated by atmospheric drag in the lower end and the inner Van Allen belt in the upper end [4]. Continuous covarage of any area is not possible with a single LEO satellite. The number of satellites needed depends largely on the orbital height and size of the system's coverage area.

The Iridium system has already for several years been providing global coverage, but Iridium is only designed for narrowband communications such as voice and low rate data as it provides a bit rate of 2.4 kbps [17]. A similar system providing wideband communication capabilities, or even an upgrade of the Iridium system, will by all accounts be doomed to bankruptcy. This assumption is based on the fact that the Iridium system never has been profitable, and after a bankruptcy the United States Department of Defence now owns the system. An uppgraded wideband system will be even more expensive, thus, other smaller and more specialized constellations must be contemplated.

3.3 Highly elliptical orbit

The degree of ellipticity in an orbit is refered to as the orbit's eccentricity, which is defined as:

$$e = \frac{c}{a} \tag{12}$$

where c is the distance between the center of the ellipse and the center of earth and a is the semimajor axis of the ellipse. Thus, for a circular orbit e is equall to zero, and if e is equall to one the ellipes collapses into a straight line between the two focal points [18].

A satellite in an orbit with high eccentricity will have large differences between its apogee height and its perigee height. Kepler's second law states that a line drawn between the center of earth and an orbiting satellite will sweep equal areas in equal time. As a result a satellite in a highly elliptical orbit will spend a majority of its time around the apogee. From earth this will make the satellite look almost stationary for long periods of time [16].

As mentioned above it is desirable to maximize the time a satellite spend above the coverage area of interest. This is achieved by placing the apogee above the coverage area. If a repeating ground track orbit is chosen the apogee can regularly be above the same spot on earth. A satellite that completes an integer number of orbit periods in an integer number of days is in a repeating ground track orbit. In order to reduce the number of satellites needed each satellite should pass over the coverage area at least once each day. Thus, the maximum orbit period is one sidereal day. For the same reason the dwell time around apogee should be maximized. Repeating ground track orbits with orbit periods below a half sidereal day have short maximum dwell time because the low orbit altitude limits the eccentricity. The result is that there is two usable orbit periods, one sidereal day and a half sidereal day.

However, even though the ground track is repeating the apogee is not necessarily above the same spot on earth orbit after orbit. This is due to the fact that the earth is not a perfect sphere, but slightly oblate. Because of this oblateness a difference in the gravitational force between equator and the poles is generated. The gravitational force is strongest around equator causing a satellite in a low inclination orbit to curve faster than the perfect Keplerian orbit. Because of the slight increase in speed the perigee will be rotated in the direction of motion. With a high inclination the perigee will rotate in the opposite direction because of the weaker gravitational force around the poles [16].

Thus, in order to keep the apogee above the same spot on earth the inclination must be chosen with care. The perigee drift as a function of the inclination, i, is

$$\frac{d\omega}{dt} \approx \frac{5\left(5\cos^2(i) - 1\right)}{\left(\frac{a}{R_e}\right)^{\frac{7}{2}}\left(1 - e^2\right)^2} \tag{13}$$

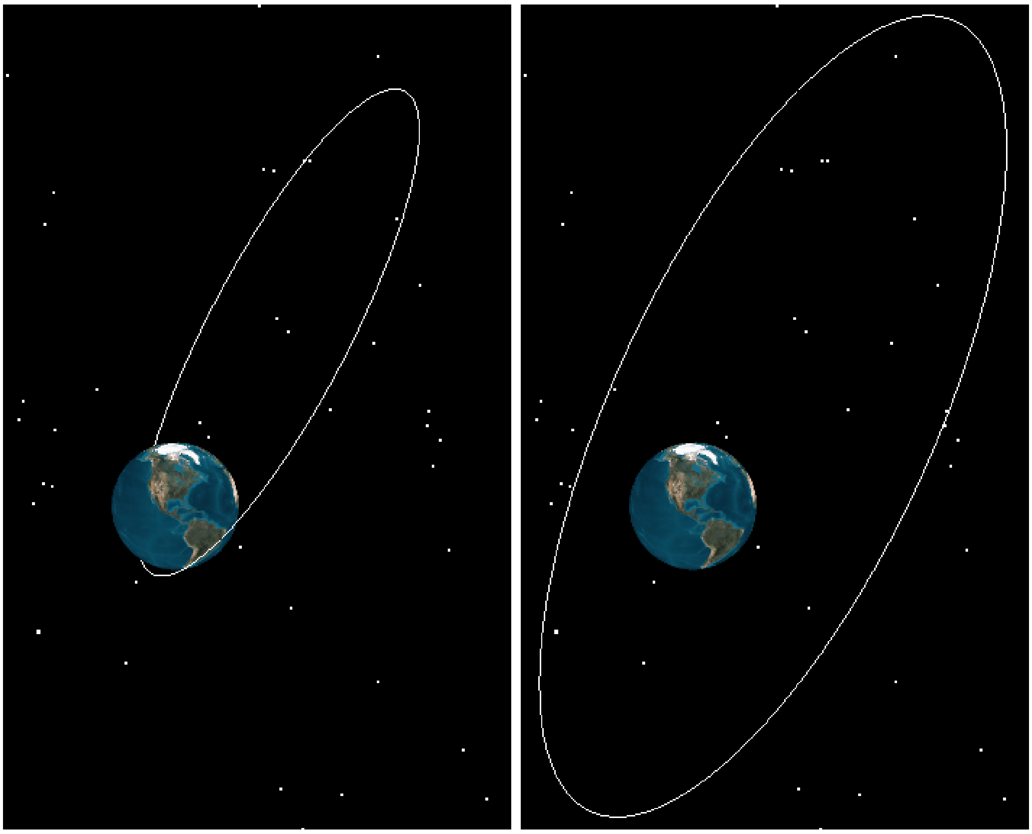
where a is the semimajor axis, R_e is the earth radius and e is the eccentricity [18]. The apogee will be stationary if the perigee drift is zero. If equation 13 is set equall to zero the necessary inclination can be found.

$$\begin{aligned} 5\cos^2(i) - 1 &= 0 \\ i &= \underline{\underline{\begin{cases} 63.4^\circ \\ 116.6^\circ \end{cases}}} \end{aligned} \tag{14}$$

A satellite orbit with an inclination of 63.4° or 116.4° is often referred to as critically inclined. The inclination of 63.4° is the one used by the Russian, and former Soviet, Molniya satellites. The Molniya satellites have an orbit period of a half sidereal day, which is almost twelve hours. A corresponding orbit with a period of one sidereal day was named Tundra by the Soviet. Historically there has been few satellites in Tundra orbits, but today the American satellite radio system Sirius uses Tundra orbits.

Figure 4 illustrates basic Molniya and Tundra orbits. The inclination is of course the same, but the difference in orbit period and eccentricity result in two distinct orbits.

The view of earth from apogee is very close to the same for Molniya and Tundra orbits when the apogee is placed on the same longitude. Figure 5 show the large areas visible from a satellite in a Molniya orbit at apogee.



(a) Molniya orbit (b) Tundra orbit

FIGURE 4: A basic Molniya and Tundra orbit.

Since the Molniya orbit has a period equal to half a sidereal day it will pass through two apogees each day, while a Tundra satellite only is at apogee once a day. The Molniya orbit used in figure 5 has apogee above 15° east and 165° west. From 15° east the area above the solid black line is visible, and from 165° west the area above the dotted line is visible.

Both the Monliya and Tundra orbit allow a single satellite to see huge parts of the world for long periods of time. Thus, it appears that for the latitudes of interest in this study a critically inclined highly elliptical satellite orbit will give the most cost effective system. The long dwell times caused by a high eccentricity will give close to geostationary conditions with high elevation angles, even far north. Also the number of satellites needed for continuous coverage are drastically reduced due to these long dwell times.

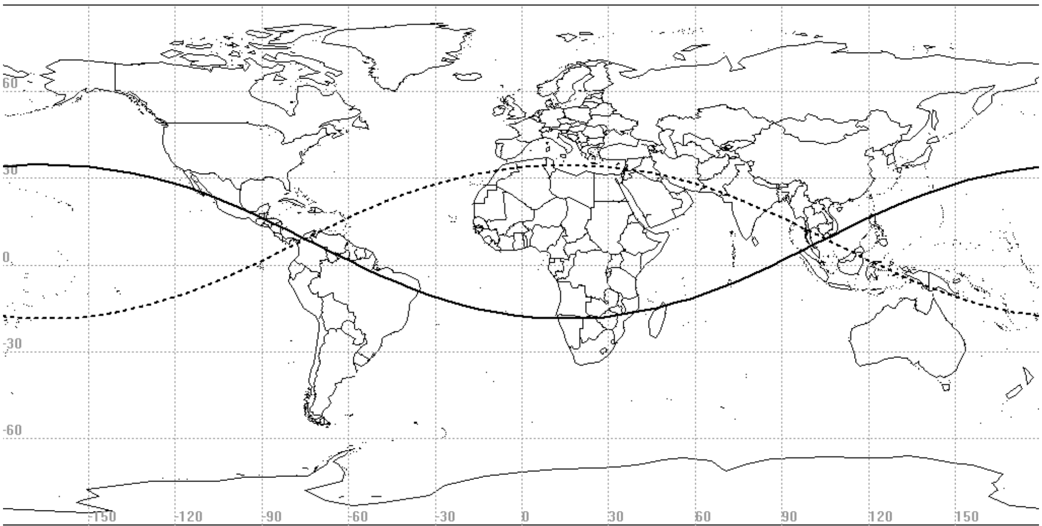


FIGURE 5: Area visible from the apogee of a Molniya orbit. The area to the north of the solid black line is visible from apogee at 15° east, while the area above the dotted black line is visible from apogee at 165° west.

Chapter 4

Constellation design

In chapter 3 possible orbits were discussed, and it was concluded that the best suited orbits for satellite communications in polar areas are the Molniya and Tundra orbits. However, neither of those provide continuous coverage with a single satellite. Thus, a constelltion of multiple satellites is needed.

When multiple satellites are used to provide continuous coverage of an area the handover problem must be adressed. Movement of traffic from one satellite to another should not be noticeable to the user. The easiest solution to this problem is to place both the incoming and outgoing satellite within the antenna beam of the user. Then handover can be performed by simply switching the outgoing satellite off at the same time the incoming satellite is turned on, and the user terminals only need one antenna.

This solution puts constraints on how the satellites must be positioned in relation to each other. How severe these constraints are is depentant on the ground station antennas. If antennas with fairly wide beams are used the two satellites can be far from each other when the switch is made. On the other hand, tracking antennas with narrow beams demand that the satellites are more or less above the same spot on earth at the time of handover. Thus, the simplest antenna solution allows for more freedom in the constellation design, but the extra gain available in a tracking antenna may be needed.

4.1 Tundra orbit

As mentioned in section 3.3 the Tundra orbit has an orbit period equal to one sidereal day. That is the same orbit period as a geostationary satellite, but because of the inclination and the eccentricity the Tundra orbit is only geosynchronous. If a satellite is launched into a critically inclined circular geosynchronous orbit the ground track will form a figure eight as shown in

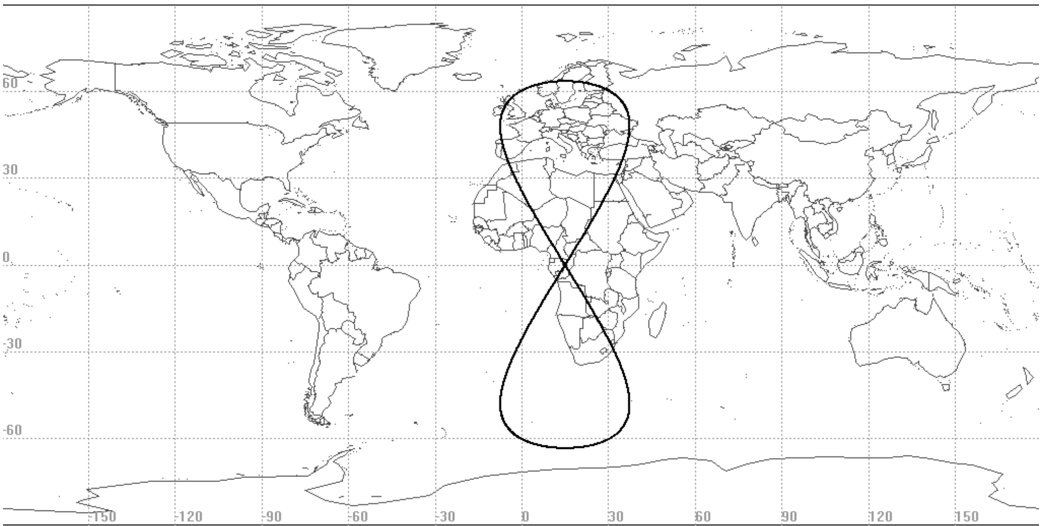


FIGURE 6: The ground track of a critically inclined circular geosynchronous orbit with the ascending node at 15° east.

figure 6. Since the eccentricity is zero the two loops will have the same size.

If the orbit is changed to become slightly elliptic the two loops will no longer be symetrical. Because of the inclination of 63.4° the northern loop will shrink as the eccentricity is increased. Should the eccentricity increase above 0.42 the northern loop will disappear and the ground track will become an oval circle. Thus, it is possible to control the time a satellite spend in a loop by adjusting the eccentricity. This control possibility makes the ground track intersection point ideal for seamless handover between incoming and outgoing satellites.

When handover between satellites happens at the ground track intersection point the number of satellites in the constellation dictates the time a satellite must spend in the small loop. In table 10 the time spent in the northern loop is given along with the corresponding eccentricity and apogee height for three to eight satellites. Higher eccentricity give a smaller loop, thus, the more satellites employed the more stationary will the system apear to the user.

Figure 7 show the ground track for three different loop sizes, eight, six and four hours. The four hour loop is so small that a satellite is almost stationary relative to the ground when it is in the loop. The eigth hour loop, on the other hand, traverse over a large area, and are far from stationary. Nevertheless the orbit is visible at high elevation angles, even far north, for the full eight hours, and should therefore not be disregarded.

TABLE 10: Time spent in the northern loop with the corresponding eccentricity and apogee height for possible number of satellites in Tundra orbit.

SATELLITES	TIME [hours]	ECCENTRICITY	APOGEE HEIGHT [km]
3	8.0	0.265	46 960
4	6.0	0.343	50 248
5	4.8	0.374	51 555
6	4.0	0.390	52 230
7	3.4	0.400	52 652
8	3.0	0.404	52 820

FIGURE 7: The ground track of three Tundra orbits with different eccentricity. A satellite will spend eight hours in the black loop, six in the grey and four in the white. The ground track intersection point is placed fairly arbitrarily at 15° east.

Even though the ground track loop created by the Tundra orbit solves the handover problem for antennas with tracking capabilities, it can still be of interest to use non-tracking antennas in the system. Non-tracking antennas need a high degree of stationarity. Thus, an increase in the number of satellites will give an increase in performance for non-tracking systems. However, the cost of a satellite system scales with the number of satellites needed [16]. As a result the number of satellites should be minimized and traded with other system parameters to meet the desired performance level.

It is difficult to pinpoint which eccentricity that give the most cost effective orbit. As stated above the number of satellites should be minimized. However, if adding an extra satellite or two to the constellation reduces the complexity, size and cost of a user terminal, the number of possible users may increase drastically. Another factor is that during operations the satellite antenna must constantly be adjusted to make sure it is pointed towards the users. A more stationary orbit requires smaller adjustments, thus, reducing the complexity of the satellites, and in turn the costs.

The constellation with the lowest deployment cost should by all accounts be the three satellites option. It is the least stationary solution. However, it has the lowest space loss, and the elevation angle at the satellite station at Svalbard are never bellow about 52° as long as the satellite is in the loop. When the satellite is at apogee the elevation angle is a little more than 73°. As a result a non-tracking antenna at a user terminal must have a beamwidth of more than 21° for continuous satellite contact. Given the 3 dB beamwidth, $\theta_{3\text{dB}}$, of an antenna the generalized gain can be found using [4]:

$$G \approx 10 \log \left(\frac{33\,000}{(\theta_{3\text{dB}})^2} \right) \text{ dB} \tag{15}$$

That corresponds to an antenna gain of less than 19 dB, which will probably not be enough for high data rates. Thus, if only three satellites are employed tracking antennas must be used at the user terminals.

As figure 7 show the addidtion of one satellite reduces the size of the loop drastically. The apogee height is increased leading to a 0.6 dB higher space loss, but elevation angle experienced from Svalbard is never bellow 63°. With a maximum elevation angle of 73° a beamwidt of 10° is necessary in a non-tracking antenna. Using equation 15 the corresponding antenna gain of 25 dB is found to be markedly better than for the three satellite constellation, but by all accounts not high enough to give the desired bit rate. Since the user terminals must be equiped with tracking antennas in both the three and four satellite option, the three satellite constellation will probably be the most cost effective of those.

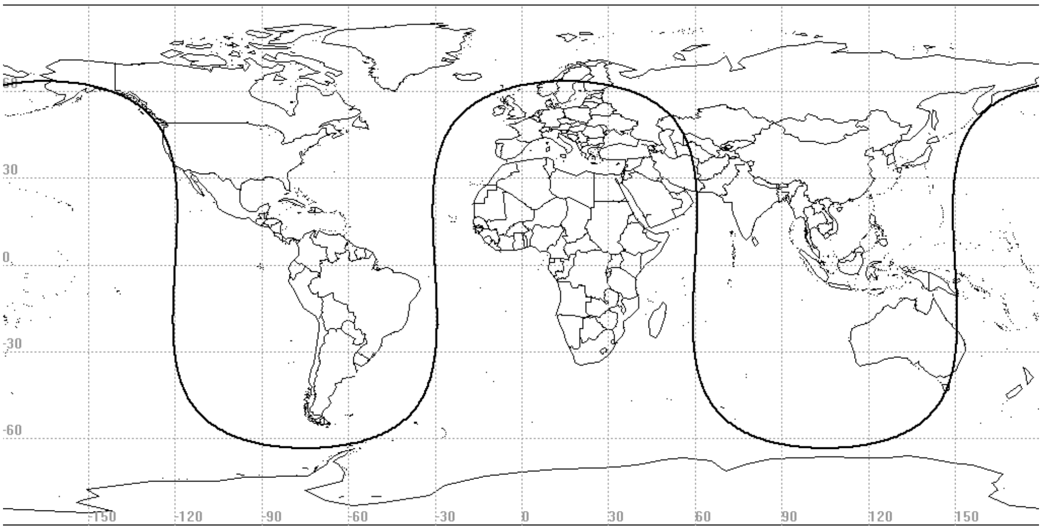


Figure 8: The ground track of a critically inclined circular twelve hour period orbit.

If a six satellite constellation is chosen the elevation angle observed from Svalbard vary between 69.6° and 73.4°. Employing again equation 15 a 4° beamwidth antenna will have a gain of a little more than 33 dB. That may be enough to support high data rates even with a non-tracking antenna. At apogee the space loss will be 0.9 dB higher than for the three satellites option, but it should still be feasible to use a non-tracking antenna. The possible use of a simpler antenna may attract enough additional costumers to justify a constellation consisting of six satellites.

4.2 Molniya orbit

The Molniya orbit is, as discussed in section 3.3, a critically inclined highly elliptical orbit with an orbit period of close to twelve hours. With such an orbit period a satellite in a Molniya orbit will be at apogee two times a day, but they will not be above the same place on earth.

A circular Molniya orbit will have a ground track like the one shown in figure 8. If the orbit is made eccentric the satellite will spend more time around the apogee. As a result the ground track around the apogee will move closer together. As discussed in section 4.1 the ground track of a Tundra orbit will form a closed loop as long as the eccentricity is bellow 0.42. Due to the shorter orbit period the ground track of a Molniya orbit will only draw a closed loop when the eccentricity is above 0.712. However, if the eccentricity lie between 0.712 and 0.738 the ground track will intersect two times around

Table 11: Time spent in the northern loop with the corresponding eccentricity and apogee height for possible number of satellites in Molniya orbit.

Satellites	Time [hours]	Eccentricity	Apogee height [km]
3	8.0	0.7126	39 098
4	6.0	0.7199	39 293
5	4.8	0.7255	39 442
6	4.0	0.7289	39 532
7	3.4	0.7310	39 587
8	3.0	0.7323	39 621

apogee.

Both points can be used for handover, but it is the one closest to the apogee that is the most interesting. The time a satellite spend in that loop can be controlled in a similar matter as for the Tundra orbit, namely by adjusting the eccentricity. Each satellite in the constellation must use a certain time between two consecutive passes over that intersection point. That time decides the necessary eccentricity and corresponding apogee height. For constellation sizes between 3 and 8 satellites these values are given in table 11.

Figure 9 show a map over Europe with the ground track of three Molniya orbits drawn in. Notice that the ground track intersection points are more or less at the same spots as for the Tundra orbits shown in figure 7. As a result the changes in elevation will be in the same order of magnitude as for the corresponding Tundra orbit, but slightly larger for the Molniya orbits due to the shorter distance between the ground and the satellite. On the other hand, the azimuth angle towards the Molniya satellites changes much less.

At the time of handover in a constellation with three satellites the elevation angle from the satellite station at Svalbard will be about 48.5°. With an elevation angle of close to 73° at apogee an antenna with a beamwidth of more than 24° is needed if the user antenna is non-tracking. Using equation 15 that corresponds to a gain of less than 17.5 dB which is not good enough for broadband communications. As a result tracking antennas must be employed if a constellation of three satellites is chosen. However, because the aximuth angle changes less than 4° a 33 dB antenna can be used at the ground stations reducing the necessary tracking to elevation only for fixed stations.

Addition of one satellite reduces also here the loop drastically, but as for the equivalent Tundra orbit the variation in elevation angle observed

FIGURE 9: The ground track of three Molniya orbits with different eccentricity. When the most northern intersection point is used for handover a satellite will spend eight hours in the black loop, six in the grey and four in the white. The ground track intersection point is placed fairly arbitrarily at 15° east.

from Svalbard is still above 10°. Thus, again the users must utilize tracking antennas for high speed data transmission to be possible. The azimuth angle vary less than for the option with three satellites, and single axis tracking can be done with an antenna having a beamwidth of 2.5°. According to equation 15 that give a gain of more than 37 dB. The increased eccentricity results, as table 11 show, in a higher apogee height, but the difference is so small that the additional space loss is negligible.

A Molniya constellation with six satellites results in a variation in elevation angle observed from Svalbard of a little more than 4°. Thus, a 32 dB non-tracking antenna at the user terminals should be able to provide a bit rate that is good enough. If tracking antennas are employed at the user terminals they need a beamwidth of less than 1° for single axis tracking since the azimuth angle are within 0.7°. The result is antennas with a possible gain of more than 45 dB.

4.3 Evaluation

4.3.1 Molniya versus Tundra

Molniya and Tundra orbits are very similar, and they have many of the same features. As figures 7 and 9 show they provide equal coverage with equal number of satellites, and users experience the same range of elevation angles for both orbits. However, due to smaller changes in azimuth angle a Molniya system allows more directional antennas with reduced complexity.

The biggest advantage the Tundra orbits have over the Molniya orbits are the reduced radiation the satellites will be exposed to. With a perigee radius of more than 19 000 km the Tundra orbits go clear of the Van Allen belts. A Molniya orbit will on the other hand cross the Van Allen belt four times each day, two times for each orbit. This extreme radiation environment will give a Molniya system significantly reduced operational life time when compared to a Tundra system [4, 16]. The reduced life time can be compensated by added shielding, but that makes the satellites more costly.

On the other hand the higher orbit altitude of a Tundra system have two major disadvantages. One of them is increased space loss. In equation 7 the distance component of the spaceloss is given as $20\log(D)$. Using that and the apogee heights from tables 10 and 11 it can be shown that the difference in space loss at apogee are between 1.6 dB and 2.5 dB depending on the number of satellites employed.

The other major disadvantage caused by the higher altitude of the Tundra orbits is the increased launch costs. It is difficult to find the exact cost difference with so few system parameters defined. However, some assumptions can be made. A Molniya orbit can be compared to a geostationary transfer orbit and a Tundra orbit is similar to normal geostationary orbit. Under these assumptions the launch cost of a Tundra satellite is close to twice as high as a Molniya satellite [16].

As mentioned in section 4.2 the Molniya orbit will be at apogee two times each day. For the orbits used as examples above one of those will be over Europe as figure 9 show. The other appogee will then be above Alaska. Alaska and northern Canada is another area where satellite based broadband coverage is more or less non-existent. If a Molniya system is chosen it is possible to deliver satellite services to these areas as well. The cost of the satellites will probably be more or less the same, but the possible number of users increases dramatically.

With all this in mind it would seem that it is most cost efficient, and will provide the most profitable opportunities, to choose a constellation using Molniya orbits. The increased launch cost together with the larger space loss associated with a Tundra system will make it less expensive to design and build radiation shielded Molniya satellites. The possibility to get a second coverage area without addition of extra satellites makes it an even easier decision.

4.3.2 Constellation size

When it comes to deciding the number of satellites to use one of the most important factors is the antenna gain and complexity of the user terminals. If there are few users the cost of the space segment of the system should be minimized, and instead demand higher perfomance from the user equipment. Then a natural choice would be a constellation with three satellites and high gain tracking antennas at the user terminals. Should the system attract many users with high demands a constellation with six slightly more advanced satellites may be the best solution. That would allow the use of stationary non-tracking antennas at the user terminals with the same gain as a single axis tracking system get in a constellation of three satellites. Users needing even higher performance can then get that performance with antennas tracking in elevation only.

Between these two options lie a constellation with four or five satellites. The use of five satellites are probably not a good option as the performance increase from four satellites is limited, and not worth the extra cost. In a constellation of four satellites three user levels can be accomodated. Users that only need voice and low data rates can use fixed antennas with a fairly wide beam. A single axis tracking antenna can give users fairly high data rates, while advanced users employing fully tracking antennas can maximize the performance of the system.

One possible large user group for broadband communications in polar areas are ships. Ships will normally be moving, thus, tracking antennas are needed for broadband communications via satellite. However, this is a well known challenge with existing solutions available designed for communications with geostationary satellites that are not overly complex. Such a

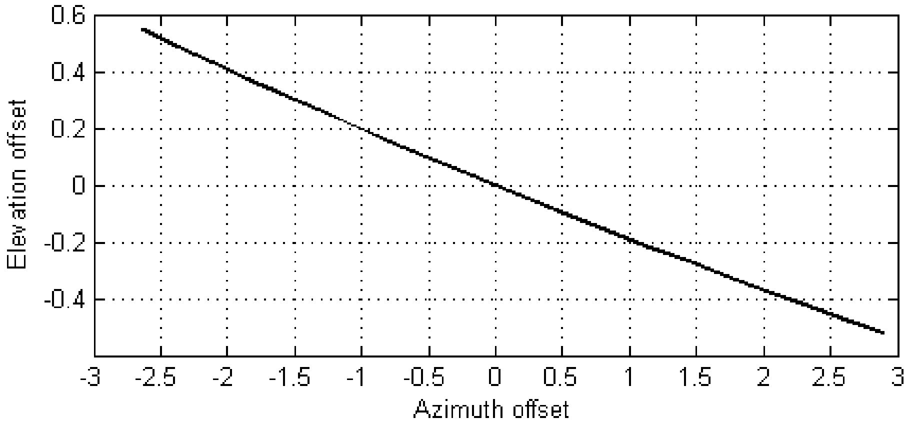


FIGURE 10: The offset in azimuth and elevation angle for eccentricities from 0.71 to 0.73 at the moment of handover for a user terminal positioned at 65° north and on the same longitude as the apogee

system is able to keep the antenna pointed in a specific direction making the antenna more or less fixed. Program tracking based on orbital parameters can be used and is fairly simple. Hence, the above discussion is also valid for communication to mobile ships.

4.3.3 Eccentricity

In the discussion above the eccentricity is used to adjust the position of the ground track intersection point. With four satellites in the constellation the correct eccentricity is 0.7199. In order to get such an excact eccentricity carefull orbit manoeuvers will be needed. That might require the use of excessive amounts of propellant. Thus, it is of interest to know how eccentricity errors will affect the handover situation.

Figure 10 show the offset in azimuth and elevation angle for eccentricities from 0.71 to 0.73 at the moment of handover for a user terminal positioned at 65° north and on the same longitude as the apogee. It is evident that large eccentricity errors can not be tolerated. How large the allowed error can be depends on the beamwidth of the user terminal antenna. If an angular distance between the incoming and outgoing satellites of 1° is acceptable for the user terminal antenna the eccentricity must lie between approximately 0.718 and 0.722. Should the user terminal antenna require a maximal angle between the satellites of 0.5° at handover the eccentricity must lie in the region from 0.719 to 0.721.

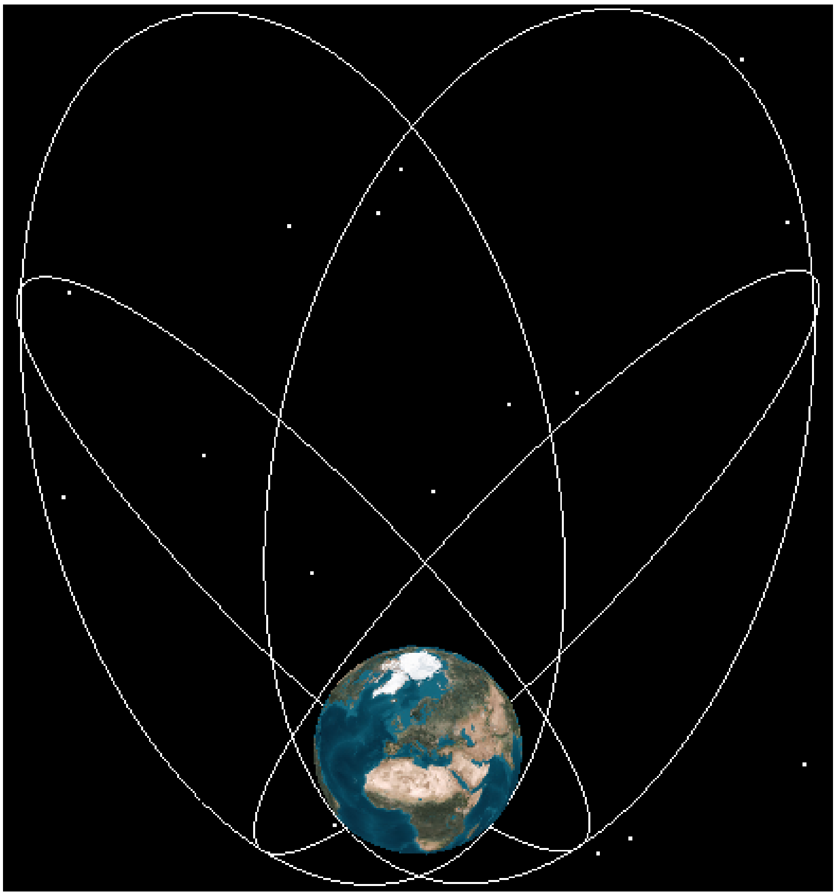


FIGURE 11: The four orbits of the chosen constellation design.

4.3.4 Conclusion

The conclusion is that the constallation option that seems to be the most apealing is a Molniya constellation with four satellites. It allows user flexibility whitout the need for many satellites. Another advantage it has over the option with three satellites is reliability. Should one satellite fail, continuous coverage is not lost for user systems that are tracking except at the moment of handover. In figure 11 the orbits of the four satellites is illustrated. As the figure show the satellites will have to be in individual orbital planes.

Chapter 5

The radio link

In a communications satellite the radio system is the payload. Thus, the spacecraft design is basically decided by the solutions chosen for the radio system. As a result it is imperative to reduce the size and complexity of the radio system without compromising the performance to much. This can be achieved through parametric analysis. The normal way to perform parametric analysis of a radio system is by setting up a link budget.

A link budget is a tabular method for evaluating the signal strenght. It is handy to use a spreadsheet when doing the calculations since the optimization process is iterative. When optimizing a radio system trade studies must be performed. Link budget calculations using spreadsheets facilitate this as the result of a change can be instantly observed.

5.1 Coverage area

When creating a link budget for a satellite based radio communications system it is advantageous to know which area the system should cover. That is because the coverage area has impact on the distance between satellite and user as well as possible antenna solutions. Also the coverage area has influence on the number of possible users which in turn affect the data rate the system should provide. For this system it has been preliminary specified that it should cover the area north of 65° northern latitude. Before proceeding it migth be usefull to understand why that area is suggested.

Firstly, for the system to be usefull at all it should cover areas around the north pole not already provided with satellite communications. As discussed in section 3.1 it is possible, but not desireable to communicate with satellites at elevation angles below about 10°. Thus, the area north of 70° north is either not covered of will get better performance with a Molniya based

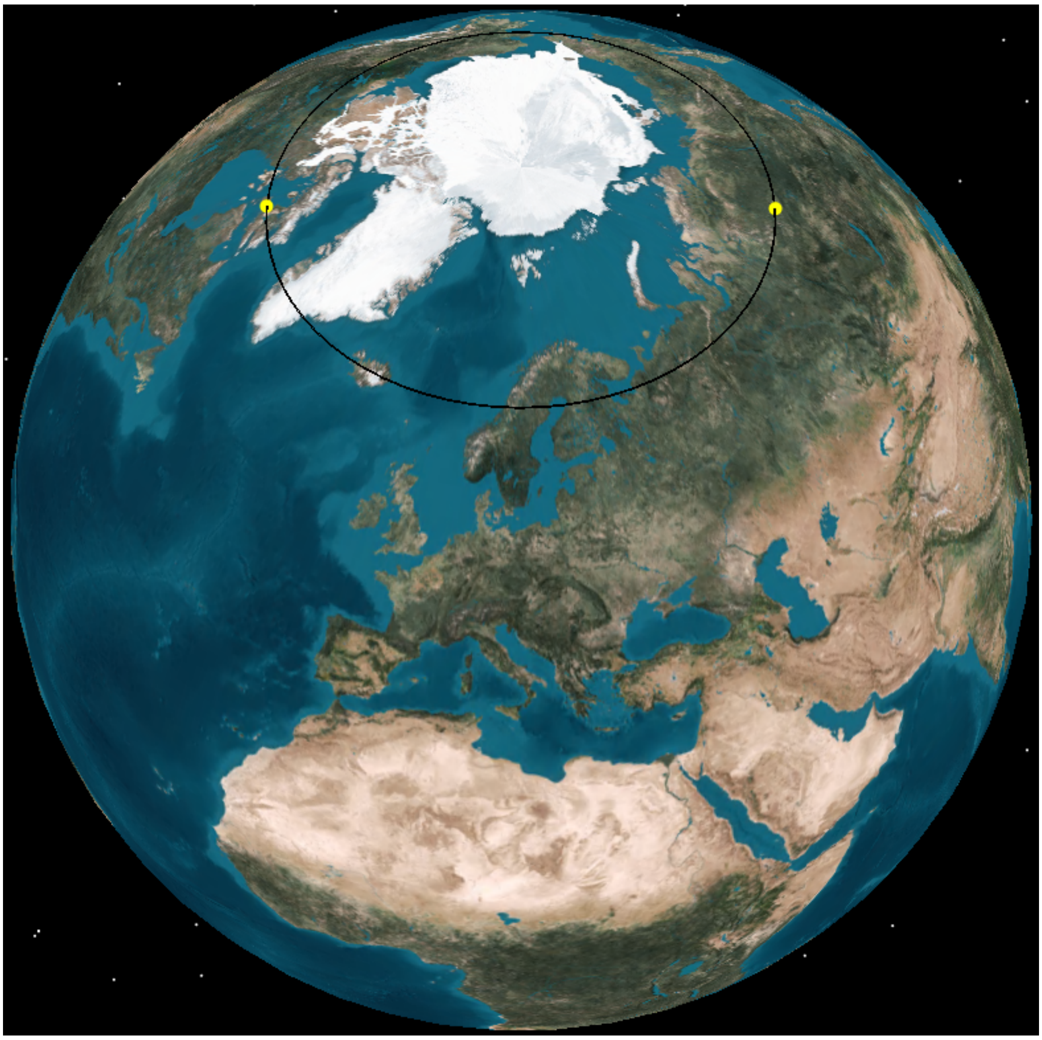


FIGURE 12: The view from the satellite above the ground track intersection point at 15° east. Inside the black circle is the suggested coverage area. The yellow dots show the two positions inside the coverage area that are farthest away from the satellite when the coverage area is divided between the two operational satellites.

system.

Many of the potential users are mobile, such as ships, and it would not be unnatural for them to move south of 70° north from time to time. Hence, if the coverage area is expanded down to 65° north the use of only one satellite system is facilitated for many users, and the costumer base is increased.

Figure 12 show the view towards earth from the satellite when above the ground track intersection point. The black circle is the 65th parallell.

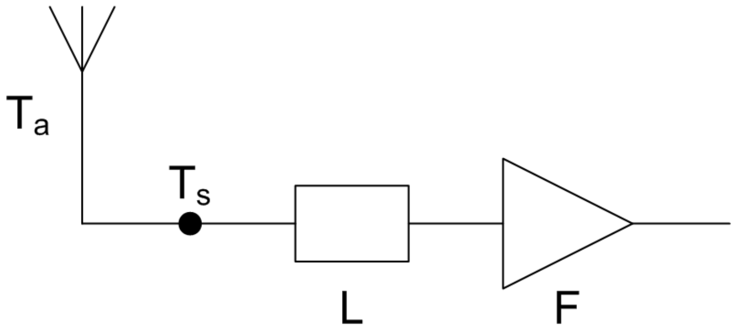


FIGURE 13: A simplified receiver with noisy elements.

From the figure it is evident that as long as the satellites are within the operational loop they see the whole coverage area. When using four satellites as concluded in chapter 4 two satellites will always see the whole coverage area. In order to maximize the elevation angle it is then natural to divide the coverage area in two equall parts. One part for the satellite over Europe and one for the satellite over Alaska. This coverage area is not carved out in stone, but is a good start for a baseline analysis.

5.2 Transmission theory

If the world had been ideal the received power in a radio system would have been given by equation 5. However, it is not that simple. In addition to all the possible losses discussed in section 2.3 there will be losses in wires, waveguides and connectors on both the transmitting and receiving side.

Noise is also present degrading the signal environment and reducing the chance of correct demodulation of a signal. In section 2.4 a few sources of noise is presented. Besides that noise the receiver generates its own noise. Figure 13 show a typical receiver system. At the 20 GHz downlink frequency the lossy line, L, will typically attenuate the signal 0.5 dB and the noise figure, F, of the amplifier will typically be in the area of 3 dB. The antenna noise, T_a, can be assumed to be about 100 K. Using these figures a typical system noise temperature, T_s, refered to the antenna output can be found:

$$\begin{aligned} T_s &= T_a + T_o(L-1) + T_o(F-1)L \\ &= 100\,\mathrm{K} + 290\,\mathrm{K}\cdot\left(10^{0.05}-1\right) + 290\,\mathrm{K}\cdot\left(10^{0.3}-1\right)\cdot 10^{0.05} \\ &= \underline{459\,\mathrm{K}} \end{aligned}$$

Thus, a typical system noise temperature refered to the antenna output at the 20 GHz downlink frequency is around 460 K, corresponding to 26.6 dBK. The system noise temperature for the uplink frequency will be about 200 K higher due to the more or less constant noise temperature of the earth [3].

A normal way to describe the signal quality of a radio link for digital transmission is its carrier to noise density ratio, C/N_o. The carrier power, C, is equall to the received power, P_r. As a result the carrier to noise density ratio can be written as:

$$\frac{C}{N_o} = \frac{P_t G_t G_r}{L_s L_a N_o} = \frac{P_t G_t G_r}{L_s L_a k T_s} \tag{16}$$

where N_o is the noise spectral density, P_t is the transmit power, G_t is the antenna gain at the transmitter, G_r is the receiver antenna gain, L_s is the free space loss as given in section 2.2, L_a is the attenuation caused by the atmosphere, k is Boltzmann's constant equal to −228.6 dBW/K/Hz and T_s is the clear sky system noise temperature of the receiver refered to the antenna output.

The values in a link budget can differ in several order of magnitude. In order to make the values more pleasent to work with it is customary to convert the values into dB. This has the added benefit of transforming the operations of multiplication and division into addition and subtraction:

$$\frac{C}{N_o} = P_t + G_t + G_r - L_s - L_a - T_s - 10\log(k) \tag{17}$$

Attenuation due to atmospheric phenomena, L_a, is discussed in section 2.3. Gaseous absorption is the largest single constant contributor. According to table 4 the attenuation is about 0.3 dB at 60° elevation for both the uplink and downlink frequencies. Allowing for lower elevation angles and some additional losses the atmospheric attenuation will here be set to 0.5 dB.

In order to calculate the free space loss the distance between the satellite and the ground station is needed. The link budget must be valid for all possible positions within the coverage area. Thus, the distance used when finding the free space loss for the link budget should be the largest distance possible to have between the satellite and a ground station inside the coverage area.

The largest distance occurs when the satellite is at apogee. Since one satellite only need to communicate with half the coverage area, as discussed in section 5.1, there will be two positions with the same maximal distance. This is due to symmetry. These positions will be at 65° north, and shiftet 90° in longitude away from satellite. The two yellow dots in figure 12 depict the two positions in question for a satellite with apogee and ground track intersection point at 15° east. With the satellite at apogee the distance to these two positions is about 40 700 km. Using equation 7 the maximum free space loss can then be found to be 210.6 dB at 20 GHz and 214.1 dB at 30 GHz.

5.3 User terminal antenna gain

The antenna gain is important and must be as high as possible in order to compensate to some extent for the massive free space loss. In general high gain is achived through large antennas with narrow beamwidths. However, the dimensions of the antennas can not be too large because that is inconvinient at the user terminals.

As discussed in chapter 4 it is necessary with fully tracking antennas at the ground stations in order to get the best performance. Such tracking kan either be achieved with a phased array antenna or a reflector antenna mounted on a steerable platform. The reflector antenna option is the conventional and most evolved technology. Phased array antennas are larger and more complex systems. Thus, steerable reflector antennas will be the basis for further calculations in this study with regard to the user terminals. However, phased array technology is progressing and it is perhaps the technology for the future.

For practical reasons the reflective dish should not be larger than about 1 m in diameter. Most installations and ships are able to accomodate such an antenna, but larger antennas also results in larger, heavier and more complex steerable platforms. That may be more difficult to accomodate.

The gain of an antenna can be found from equation 3. By substituting the effective area with the diameter of the antenna the expression becomes:

$$G = \eta \left(\frac{\pi d}{\lambda} \right)^2 \tag{18}$$

where η is the aperture efficiency, d is the antenna diameter and λ is the carrier wavelenght. The aperature efficiency for reflector antennas is typically between 55 and 70 % [4]. In calculations here a conservative efficiency value of 60 % will be assumed. With an antenna diameter of 1 m the gain then is 44.2 dB at 20 GHz and 47.7 dB at 30 GHz. The exact analytical expression for the 3 dB beamwidth depends on how the aperature is illuminated and is given by

$$\theta_{3\mathrm{dB}} = C \frac{\lambda}{d} \tag{19}$$

where λ and d are the same as above, and C is a constant that depends on the illumination function. With an aperture efficiency of 60 % C is about 74. By substituting in the frequency for the wavelength equation 19 can rewritten to

$$\theta_{3\text{dB}} \approx \frac{22.2}{f_{\text{GHz}} d} \tag{20}$$

where f_{GHz} is the frequency in GHz. At 20 and 30 GHz that corresponds to 1.11° and 0.74° respectively.

A steerable antenna platform cannot provide 100 % accurate pointing. Pointing error results in reduced gain. This gain reduction must be accounted for in the link budget, and can be found in dB using:

$$L_\theta = 12 \left(\frac{\delta}{\theta_{3\text{dB}}} \right)^2 \tag{21}$$

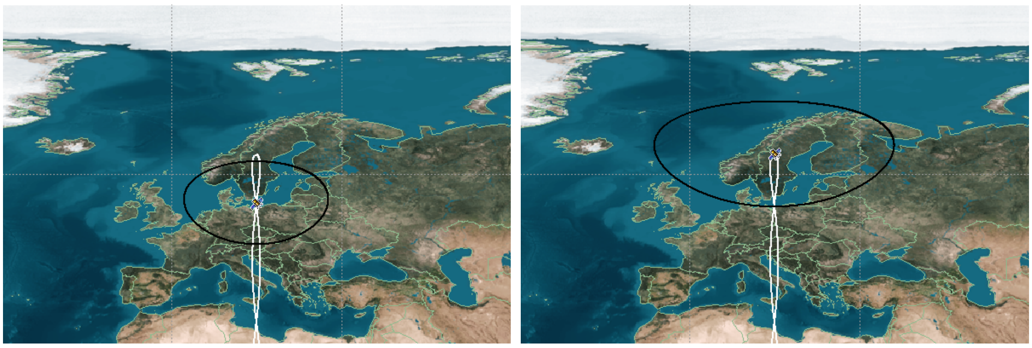
where $\theta_{3\text{dB}}$ is the 3 dB beamwidth and δ is the pointing offset from beam center [3]. Systems commercially available today can guarantee a pointing accuracy of ±0.2° [19]. The result is a gain reduction of 0.4 dB at 20 GHz and 0.9 dB at 30 GHz.

5.4 Satellite antenna gain

The antenna system on the satellite must provide coverage of the whole area of interest. There are several ways to achieve this, but they can all be divided into three general categories. The simplest solution is a single beam antenna covering the whole coverage area. The second option is a multiple beam antenna system where the available transmitt power is divided in a fixed manner between the beams. A third option is a multiple beam solution where the power is routed dynamically between the beams depending on where the traffic is comming from.

If the antenna system is directed towards the nadir direction of the satellite at all times the coverage area will move more than desired. This is illustrated in figure 14. The antenna used here have a 3 dB beamwidth of 3°, and the oval circle show the area that are within that angle. It is evident that the area covered by the antenna is no where near constant.

However, this can be achived by keeping the antenna system pointed towards one single spot on the earth. The result is shown in figure 15. It is apperent that the area within the 3 dB beamwidth is somewhat larger when the satellite is at apogee when compared to the ground track intersection point, but the area is generally the same. Hence, the antenna system must be controlled and adjusted to achive the necessary coverage while the satellite is operating within the loop. This is valid for all of the three types of antenna solutions.



(a) Ground track intersection point (b) Apogee

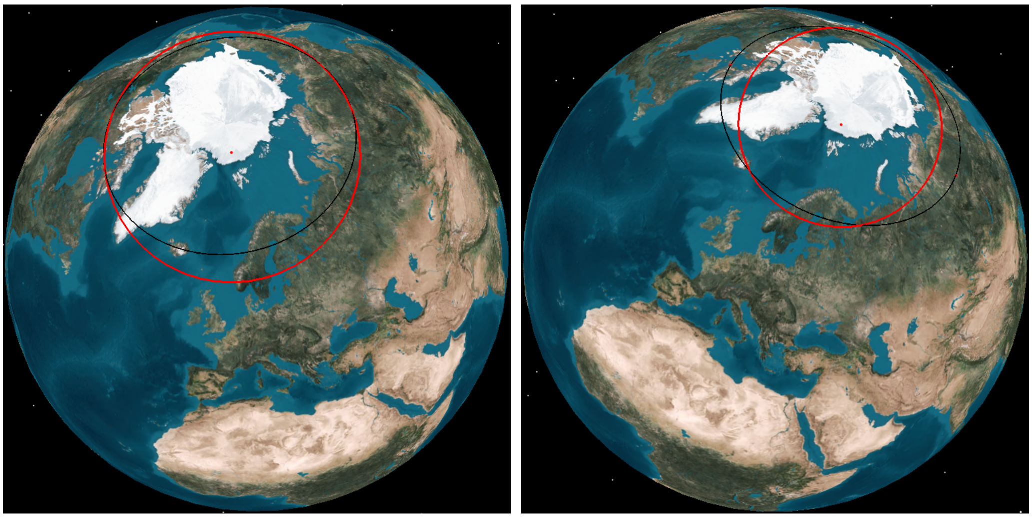
FIGURE 14: The direction of the antenna beam at two positions above earth with the antenna pointing in the nadir direction. The line depicts the 3 dB beamwidth.

(a) Ground track intersection point (b) Apogee

FIGURE 15: The direction of the antenna beam at two positions above earth with the antenna pointed towards a fixed point on earth. The line depicts the 3 dB beamwidth.

5.4.1 Single beam system

A simple parabolic reflector is assumed for the single beam solution. Calculations have shown that the antenna beamwidth that give the best performance is 7.7°. That corresponds to an antenna gain of about 27.4 dB. When the satellite is at apogee that beamwidth puts the whole coverage area within the red 3 dB contour as shown in figure 16a. It is not necessary to cover the complete coverage area, as discussed above, but it is the result when no advanced beamforming is used and the two extreme points are to be covered. With the satellite over the ground track intersection point the 3 dB contour does not enclose the necessary area, figure 16b illustrate this. However, the reduced distance of about 7 700 km between the satellite and the two extreme points



(a) Apogee (b) Ground track intersection point

FIGURE 16: The area covered by a single beam with a 7.7° beamwidth as seen from the satellite. Inside the black circle is the desired coverage area. The red circle show the 3 dB contour. Boresight of the antenna is the red dot positioned at 84° north and 15° east.

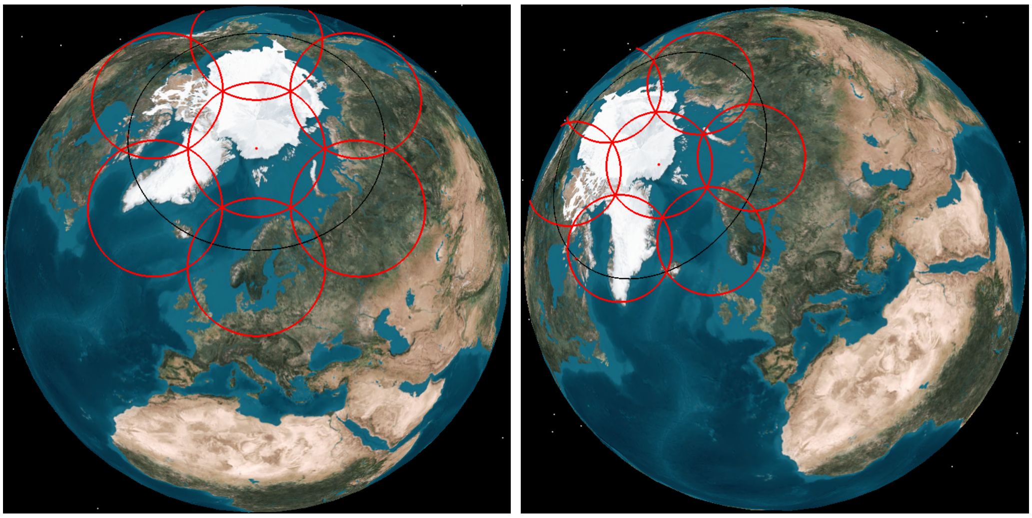
compensate for the reduced antenna gain experienced there. Calculations have shown that the signal strength actually is close to constant throughout the six hours a satellite is active.

With the use of equation 20 the diameter of the satellite antenna can be found. At 20 GHz an antenna beamwidth of 7.7° equalls a diameter of 14.4 cm. If the same antenna are used at 30 GHz it will have a beamwidth of 5.1°. As that is not a wide enough the antenne should be smaller for the uplink frequency. Thus, it is necessary to have separate antennas for downlink and uplink. This should not be problematic as the antennas will be very small with diameter of only 14.4 cm and 9.6 cm at 20 and 30 GHz respectively.

5.4.2 Multiple beam system

Parabolic reflectors is assumed for the multiple beams as well, but other antenna types can be used as well. The beams can either all be serving the same area or they can be directed to create a cell structure. If the beams are covering the same area it is more or less the same as having several single beam systems.

By directing the beams so they are arranged in a cell structure the an-



(a) Apogee (b) Ground track intersection point

FIGURE 17: The coverage area as seen from the satellite serviced by seven beams with a beamwidth of 4.0°. Again the black circle show the boundry of the coverage area, while the red circles are the 3 dB contour of the beams. Boresight of the centerbeam is fixed towards the red dot positioned at 84° north and 15° east.

tenna beamwidth can be reduced and higher antenna gain achieved without reducing the total area covered by the satellite. With the appropriate cell design it is also possible to reuse frequencies in a similar way as in ground based mobile communications systems. In principle there are no limits on how small the beams, or cells, can be. However, as equation 20 states the beamwidth is inverse proportional to the diameter of the antenna. In other words, for the beamwidth to be halved the dimensions of the antenna must be doubled. Thus, dividing a large coverage area into small cells may require many large antennas, and many large antennas can be difficult and expensive to put on a satellite.

One possibility is a seven beam solution with one beam in the center and six beams spaced 60° apart around it. Figure 17 show the layout when the beams have a beamwidth of 4.0°. A close look at figure 17 reveal that except for the centerbeam the beams are not pointed towards the same area at the two satellite positions. That is caused by the rotation of the earth. During the six hours a satellite is operational earth will rotate 90°. If that rotation is countered by the antenna system at the satellite the number of beams or their beamwidth can be reduced to some extend while still covering the area discussed in section 5.1. With 4.0° beams only three or four beams are needed.

However, the possibility to be able to cover the whole coverage area with one satellite can be usefull. It provides redundancy in the system as it is possible to provide continous coverage with only two satellites. This is a great advantage in the buildup of the constellation as it facilitate almost full performance with only half of the satellites operational. These advantages come in addition to the reduced complexity of the satellites as they do not need to counter the earth rotation.

As it can be seen from figure 17b a small patch of the desired coverage area is outside all of the 3 dB areas when the satellite is above the ground track intersection point. However, as is the case for the single beam solution discussed above the reduced free space loss compared to at apogee compensate for the reduced antenna gain.

Figure 17 also indicate that the suggested multibeam solution provide coverage to large areas outside of the stipulated coverage area. This is caused by the circular properties of the beams. The increase in possible users from these areas can be large, but they do not all get continous coverage. Then it may be more profitable to use advanced beamforming and get higher gain inside the given coverage area.

The suggested beamwidth of 4.0° will provide a maximum antenna gain of 33.1 dB. With help from equation 20 the corresponding antenna diameters can be found to be 27.8 cm and 18.5 cm at 20 and 30 GHz respectively. This is a manageable size, and it should be possible to design a satellite with seven antenna feeds for each of the two reflectors. Any significant reduction in beamwidth could on the other hand be more problematic as the antennas will then become larger, and more feed horns will be needed to cover the whole coverage area. That may also trigger the need for extra reflectors.

Instead of reflector antennas with multiple feeds, horn antennas can in theory be used. For each beam a horn with about the same dimensions as the proposed reflector antennas is then required. The array of horn antennas will then be significacantly larger than the corresponding reflector antenna. Exactly how large the horn array will be depends on the type of horns used and how they are positioned in releation to each other, but a diameter of at least 1 m at 20 GHz and 65 cm at 30 GHz is a realistic estimate. The depth of the horns will have to be quite large in order to get the desired beamwidth, probably more than a meter [20]. As a result a horn array will occupy a very large volume and have a significant mass. Thus, an array of horn antennas is not viable solution for this system and can easily be rejected.

5.4.3 Active phased array antenna system

In order to achive higher antenna gain than discussed in the previous section while still providing coverage to the whole coverage area an active phased array antenna will probably have to be used. A phased array antenna is a linear or two-dimensional array of radiators that have individually controlled phase and amplitude. When combined and adjusted correctly it is then possible to shape the antenna beam into almost any form. If electronically controlled the direction of the beam can be changed and adjusted even after the satellite is launched [20].

A classic narrow beamed system will require the satellite to generate a significant number of beams. How many depends on the desired beamwidth or required gain, but active phased arrays are better suited for this task than conventional reflector systems [21]. In addition to beeing dynamically adjustable it is possible to generate many beams from the same aperture. The flexibility offered by active phased antennas provide the possibility to direct the resources of a communications satellite to the areas where there is traffic at any instant [22].

Since the active phased array allow full controll of the beam direction a virtual multibeam system can be created. Instead of having a large number of beams with somewhat fixed direction, a low number of moveable beams can be used. The most basic solution is one single very narrow beam that are targeted at individual users in time slots. The term virtual multibeam is used as there is only one beam, but due to the high gain and individual targeting it can be perceived as a multibeam system. Another term that can be used to describe this is beam hopping.

The hopping beam does not have to follow a fixed pattern. Thus, it is possible to dynamically allocate time slots to users, regardless of their position, dependant on traffical needs. This flexibility will allow for maximum utilization of the satellite resources, and should be exploited in this communications system.

The phased arrays can be mounted on deployable panels, similar to solar cells, paving the way for very large apertures. Larger apertures can provide more beams that are more accurately shaped and smaller. Another advantge with such a deployed solution is related to heat dissipation. RF amplifiers generate a large amount of heat that can be challenging to deal with. In an active phased antenna each radiating element has a separate power amplifier. When these amplifiers are placed on the panel they can radiate their excessive heat directly into space, thus, greatly reducing the thermal dissipation into the spacecraft bus [23].

With many hundreds and perhaps thousands of RF amplifiers each of them will of course have to be small and relative low-power. This facilitate the use of fairly lightweight solid state amplifiers instead of the usual large and massive traveling wave tubes. The large number of elements also provide redundancy. Should one element fail it will have little impact on the total performance of the array. As a result the antenna system will degrade gracefully [23].

As mentioned above the beam from a phased array antenna can be shaped into almost any form. For easier comparison with the two other antenna options discussed in the previous sections a circular beam will be assumed here. A circular beam can be created by using a circulary shaped array. Thus, the beam control capability will only be used to steer and move the beam or beams inside the coverage area and adjustment of the beamwidth. How narrow the beamwidth can be is determined by the size of the array. With a large number of radiators the directivety of a phased array closely approximates that of a continuous antenna [20].

The maximum gain from a continuous antenna is given by equation 3. This expression is, as mentioned above, also valid for a phased array antenna. However, as the beam is moved around the projected area of the array changes. This change is proportional with cosine of the angle between the normal to the array and the direction of the beam. As a result the gain of a phased array antenna can be written as:

$$G = \frac{4\pi A}{\lambda^2} \eta \cos(\theta_0) \tag{22}$$

where A is the array area, η is the illumination efficiency and θ_0 is the angle between the array normal and beamcenter [20]. Equation 22 can be converted into dB. The expression for the maximum antenna gain is then

$$G = 21.5 + 20\log(f_{\mathrm{GHz}}) + 10\log(A) + 10\log(\eta) + 10\log(\cos\theta_0) \tag{23}$$

where f_{GHz} is the frequency in GHz.

As the direction of the beams from an active phased array can be controlled electronically, there is no need for mechanical adjustments of the antenna array's pointing direction as discussed above. This simplifies implementation of the antenna system. The angle θ_0 is then the angle between the nadir direction and beamcenter. Given the coverage area discussed in section 5.1 the largest possible value of θ_0 is a beam pointing towards one of the exreme points when the satellite is at the ground track intersection point. θ_0 is then about 7.5°. The reduced antenna gain in dB due to this angle is according to equation 23 equall to $10\log(\cos\theta_0)$. With a maximum angle of 7.5° the reduction in gain is then less than $-0.04\,\mathrm{dB}$. Such a low value can at this stage be ignored without introducing any significant errors.

With the frequencies allready more or less fixed the array area is the parameter that can most easily be adjusted in order to achive higher antenna gain. Larger array area will, as always when it comes to antennas, increase the maximum antenna gain. If it is assumed a circular array is used to create a circular beam the area and in turn the antenna gain can be expressed using the diameter, d, of the array:

$$G \approx 20.4 + 20\log(f_{\mathrm{GHz}}) + 20\log(d) + 10\log(\eta) \tag{24}$$

Substituting the area with the diameter causes the constant part of the expression to be reduced to 20.4. Here the effect of the angle to the array normal is ignored due to its negligible value.

The largest gain is therefore achived using the largest arrays possible to fit inside the fairing of the launcher that are eventually chosen. However, it must not be forgotten that higher gain give a narrower beamwidth. If the required beamwidth is too small the necessary beam steering accuracy may be difficult and very costly to achieve. The array can of course then be divided into several wider beams, but if this question is not adressed early enough there may not be enough power available to deliver the desired performance.

Realistic values for the aperture efficiency, η, of an active phased array antenna lies in the region between 0.7 and 0.8 [24]. Thus, an aperture efficiency of 75 % will be used in calculations here. The 3 dB beamwidth can then be found using equation 19 with $C = 1.27$ when calculating in radians [20]. In degrees that corresponds to 72.8. As a result the 3 dB beamwidth is given by

$$\theta_{3\mathrm{dB}} = 72.8\frac{\lambda}{d} = \frac{21.8}{f_{\mathrm{GHz}}d} \tag{25}$$

where f_{GHz} and d are, as above, the frequency in GHz and the diameter of the array.

A satellite with an antenna array having a diameter of 2 m can be accomodated on most launchers designed for putting satellites into orbit. If folded correctly probably even two such arrays. The antenna gain from a 2 m array will be around 51.2 dB at 20 GHz and 54.7 dB at 30 GHz. Equation 25 give beamwidths of 0.55° and 0.36°. If a beam only handles one user at a time it should not be too difficult for the array to provide a pointing accuracy that ensures the user is within the 3 dB contour.

The choice of a 2 m antenna array is somewhat arbitrary except that it gives high gain, not too narrow beam and should fit on most launch systems. It can be used for comparison between the three options presented here, and if so warranted is a good basis for further studies.

5.5 Link budget

In the privious sections the different parts of equation 16 have been discussed and assumptions made about their values. It is now possible to set up link budgets for the options given in section 5.4. These can then be compared and the best solution can be chosen.

5.5.1 Bit rate

When comparing options for a radio link it is advantageous to set some kind of common performance level all options should reach. This can either be fixed available power or a minimum bit rate demanded. In the linkbudgets given in table 14 and 15 a fixed throughput to the coverage area of a satellite is used as benchmark. The requirement used as a benchmark is that 250 simultaneous users should have available a bit rate of 1 Mbps in both directions. There is no specific reason for this choice, only the need for some values to use as a baseline.

Applied to the single beam solution the downlink must be able to provide a total bit rate of 250 Mbps. On the uplink each user must have enough available power to provide a bit rate of 1 Mbps to the satellite. For the multiple beam antenna system it is assumed that only five beams are needed to cover the coverage area discussed in section 5.1 at any instant. As a result each beam must have the ability to provide a total bit rate of 50 Mbps on the downlink. The uplink have the same requirements as the single beam system. These requirements is based on all user terminals beeing continuously covered by the satellite facilitating continuous transmissions.

For the active phased array option it is assumed that one beam is used in each direction. This beam is then moved between the users. As a result transmissions can not be continuous, but must happen in time slots. The instantaneous bit rate between a user and the satellite will then have to be a great deal larger than the average bit rate. This will have little effect on the downlink requirements, but the uplink will require an instant bit rate of 250 Mbps in order to achive the average bit rate of 1 Mbps.

5.5.2 Bit energy to noise density ratio

When a digital signal is demodulated bit errors can occur. The probability of bit errors, also known as the bit error rate (BER), is dependant on the bit energy to noise density ratio, $\frac{E_b}{N_o}$. A high ratio give a low bit error rate, while a low ratio results in a high bit error rate. In radio communications systems it is common to specify that the bit error rate is not to exceed 10^{-5}. That is equall to only one erroneously detected bit for every 100 000 bit received.

The bit energy to noise density ratio needed to meet this requirement depends on the type of modulation and coding scheme used. Error free communications is not possible when the bit energy to noise density fall bellow -1.6 dB. This is known as the Shannon limit. In practice it is not possible to reach the Shannon limit as the bandwidth and coding complexity increases whitout bound [3].

Binary phase shift keying (BPSK) and quadriphased phase shift keying (QPSK) is the two most basic and common modulation techniques used in satellite systems. Uncoded they both need a bit energy to noise density ratio of at least 9.6 dB to meet the BER requirement [3]. The largest difference between BPSK and QPSK is the spectrum efficiency. With BPSK it is possible to send one bit per symbol, while QPSK allows for two bits per symbol. Higher order modulation systems that sends more bits per symbol require a larger bit energy to noise density ratio or have inferior spectrum utilization [25].

Coding techniques is a wast field in itself, and there are countless ways to code a signal in order to make it more robust against detection errors. In general coding a signal means inserting extra bits in the bit stream. These extra bits help the decoder in the receiver find and correct errors. The most basic type of coding is the block code. When a block code is used the data bits are not altered but divided into blocks. At the start of each block a certain number of parity check bits are inserted. This technique imply the use of buffer with the same size as a message block. The buffer can be avoided if convolutional codes are used. In a convolutional code the data bits are fed into a shift register that outputs a coded bit stream with a higher rate than the orignal bit stream. A convolutional code changes the values of the original data bits and fuse them togheter with the extra bits. At the receiver the original message is regained simply by feeding the received bit stream through a matching shift register. In addition to block and convolutional codes there is a third more advanced class of codes called compounded codes [25].

A thorough and extensive examination of possible modulation and coding schemes will not be performed at this point. What is needed here is a baseline suitable for comparison of the three satellite antenna options. QPSK have excellent spectrum efficiency, and a bit error rate performance equall to that of BPSK which is good. If a basic half rate convolutional code is implemented together with viterbi decoding it is possible to reduce the bit energy to noise density ratio down to 4.4 dB without increasing the bit error rate [3]. Thus, when comparing the three satellite antenna options, QPSK modulation with viterbi decoded half rate convolutional code is used.

5.5.3 Rain margin

In section 2.3.2 the attenuation due to rain was discussed, and table 5 presented the attenuation that can be expected at different time percentages. Assuming a link availability of 99.9 % is necessary a margin of about 3.7 dB and 7.8 dB must be taken into account due to rain attenuation at 20 and 30 GHz respectively. These values are valid for the northern coastal areas in Norway at a 60° elevation angle. At the most remote locations inside the coverage area, as discussed in section 5.1, the elevation angle can be as low as 40°. However, in these areas the rainfall rate is just above 10 mm/h. This low rainfall rate reduces the attenuation far more then the lower elevtion angle increases it. Thus, the above mentioned attenuation values are conservative worst case assumptions.

In addition to attenuation rain also increases the system noise temperature. The noise caused by rain is given some attention in section 2.4. There it is found that only the downlink is affected. Table 8 indicate that the user terminals must be able to handle an increase in noise temperature of close to 160 K from the clear sky conditions discussed in section 5.2. Such an increase in noise corresponds to a 1.3 dB reduction of the carrier to noise density ratio that comes in addition to the attenuation caused by the rain.

The rain margin needed in order to provide a link availability of 99.9 % is then 5.0 dB on the downlink, and 7.8 dB on the uplink. These are conservative figures that take into account both the attenuation and increased radio noise, where applicable, caused by rain.

5.5.4 Evaluation

Satellite parameters for all the three antenna options is summarized in table 12. For the single and multiple beam options the only difference between the uplink and downlink antenna is the size. The phased array option has fixed size and the differences between the uplink and downlink is due to the different frequencies used. In table 13 the parameters for the user terminals is listed. As the same antenna is used for both uplink and downlink the variation in beamwidth and gain is due to the difference in frequency on the uplink and downlink.

The satellites does not have to use the same solution for both the uplink and the downlink. Thus, the uplink and downlink should be evaluated separately.

TABLE 12: Satellite parameters.

	SINGLE BEAM	MULTIPLE BEAM	PHASED ARRAY
DOWNLINK ANTENNA			
Beamwidth	7.7°	4.0°	0.55°
Efficiency	0.6	0.6	0.7
Diameter	14.4 cm	27.8 cm	2.0 m
Gain	27.4 dB	33.1 dB	51.2 dB
UPLINK ANTENNA			
Beamwidth	7.7°	4.0°	0.36°
Efficiency	0.6	0.6	0.7
Diameter	9.6 cm	18.5 cm	2.0 m
Gain	27.4 dB	33.1 dB	54.7 dB
RECEIVER			
System noise temperature		660 K = 28.2 dBK	

TABLE 13: User terminal parameters.

	DOWNLINK	UPLINK
ANTENNA		
Diameter	1.0 m	1.0 m
Beamwidth	1.11°	0.74°
Efficiency	0.6	0.6
Gain	44.2 dB	47.7 dB
RECEIVER		
System noise temperature	460 K = 26.6 dBK	

5.5.4.1 Downlink

Table 14 show the link budgets for the downlink at 20 GHz. The only difference between the three options lie in the satellite antenna gain, which in turn affects the needed output power. Calculated in dB the difference in output power is equal to the difference in antenna gain. For the values given for the multiple beam solution this does not match up, but that is only for one beam. When all the five beams assumed necessary to provide the 250 Mbps is taken into account the needed power is five times as high. That corresponds to the "missing" 7 dB in power difference.

For the single beam solution to be able to provide the bit rate set as benchmark a output power of 36.1 dBW is needed. In linear value that is equal to 2 715 W. Several geostationary broadcast and communications satellites providing more than 15 kW have been launched, but the power is

Table 14: Link budgets for the downlink at 20 GHz.

	Single beam [dB]	Multiple beam [dB]	Phased array [dB]
Satellite			
Output power	34.3	21.6	10.5
Antenna gain	27.4	33.1	51.2
Edge of beam loss	3.0	3.0	3.0
User terminal			
Antenna gain	44.2	44.2	44.2
Pointing error loss	0.4	0.4	0.4
System noise temperature	26.6	26.6	26.6
Other parameters			
Free space loss	210.6	210.6	210.6
Atmospheric loss	0.5	0.5	0.5
Boltzmann's constant	−228.6	−228.6	−228.6
Result			
$\frac{C}{N_o}$	93.4	86.4	93.4
Margin[1]	5.0	5.0	5.0
$\frac{E_b}{N_o}$	4.4	4.4	4.4
Maximum bitrate	84.0 250 Mbps	77.0 50 Mbps	84.0 251 Mbps

[1]Margin necessary to handle attenuation and increased noise caused by rain.

then distributed over several beams and frequencies. Using a single beam to channel an output power of almost than 3 kW will probably leave little room for a possible increase in throughput capacity, and the single beam can by all accounts be deleted from the list of useable alternatives.

Each beam in the multiple beam option must have a transmit power of at least 21.6 dBW, equal to 145 W, in order to provide the required bit rate. When taking into acount all five beams that will be active a total output power of 725 W is needed in the downlink. The total power demand is within reason and a communications system using satellites with this performance should be realisable. However, the power required to each beam is probably slightly above what available travelling wave tube amplifiers can provide today at this frequency. This might change in the future, but any significant increase in throughput above the baseline will certainly have a pronounced effect on the size and complexity of the satellites limiting the potential of this solution.

One aspect that must be addressed when multiple beams are used is routing. Since the users will be distributed among the beams a routing protocol must be used for efficient utilization of satellite resources. This imply a certain degree of onboard processing, but the extend of it is somewhat coupled with the choice of antenna system for the uplink.

With the active phased array as the downlink antenna a bit rate of 251 Mbps can be reached with only 10.5 dBW of output power, corresponding to 11.3 W. This low power value leaves a huge potential for increased throughput above the benchmark used for comparison here. The idea with the phased array is that each user is targeted individually by the satellite antenna beam when transmitted to. This puts high a high demand on the pointing accuracy of the array, and demands knowledge about the approximate position of all possible receivers. Some transmission time will be unusable as the direction of the antenna beam is changed between users. However, this should not degrade the system significantly. Another disadvantage with the phased array antenna is caused by the time slotted transmission. The inability to continuously transmit to all users can in some cases make it impossible for the satellite to simply retransmit a received signal as traditional bent pipe communications satellites does. Thus, onboard processing with a buffer is needed onboard the satellite for storage of received data packages not yet retransmitted. The size and complexity is largely dependant on how the uplink system is designed.

The active phased array will control the beamdirection electronically, thus, mechanical antenna steering systems will not be needed as is the case for the multibeam solution. Moving parts that are critical for the function of a satellite should be avoided as far as possible. Hence, the phased array have an advantage over the multibeam antenna in addition to the lower power requirements. However, the design of an accurate beamsteering system for the phased array will be challenging. Never the less even though it is the most complex and difficult system to implement the inherent posibilities of the active phased array makes it the most interesting option. It may be more expensive to develop and implement, but the increased flexibility should make it the best choice regardless of which uplink system is used.

5.5.4.2 Uplink

The three link budgets for the 30 GHz uplink is shown in table 15. As the case was for the downlink budgets the difference in output power is mainly due to the difference in antenna gain. One change here is that the phased array option must have a higher instantaneous bit rate because of the time slotted transmissions. As a result the multiple beam actually requires the

TABLE 15: Link budgets for the uplink at 30 GHz.

	SINGLE BEAM [dB]	MULTIPLE BEAM [dB]	PHASED ARRAY [dB]
USER TERMINAL			
Output power	15.4	9.5	13.6
Antenna gain	47.7	47.7	47.7
Pointing error loss	0.9	0.9	−0.9
SATELLITE			
Antenna gain	27.4	33.1	54.7
Edge of beam loss	3.0	3.0	−3.0
System noise temperature	28.2	28.0	−28.0
OTHER PARAMETERS			
Free space loss	214.1	214.1	−214.1
Atmospheric loss	0.5	0.5	−0.5
Boltzmann's constant	−228.6	−228.6	228.6
RESULT			
$\frac{C}{N_o}$	72.4	72.2	96.2
Margin[1]	7.8	7.8	−7.8
$\frac{E_b}{N_o}$	4.4	4.4	−4.4
Maximum bitrate	60.2 1 Mbps	60.0 1 Mbps	84.0 250 Mbps

[1]Margin necessary to handle attenuation caused by rain.

least amount of output power from the user terminals.

The output power of 15.4 dBW required for the single beam option is equal to 35 W. User terminals with output power up to 100 W at 30 GHz should be possible to produce. Thus, the single beam uplink solution is feasible even with increased bit rate requirements. The possible increase in bit rate is proportional with the increase in output power. As a result the maximum bit rate with 100 W transmitters is 3.0 Mbps.

A multibeam solution at the satellite will require a output power from the user terminals of 9.5 dBW, corresponding to 9 W. If the output power of the terminal is increased to 100 W an uplink bit rate of about 11.1 Mbps can possibly be achieved for each user. Thus, the multiple beam option is much more efficient and has more potential than the single beam option without complicating the total system design in any marked extent.

Due to the high instantaneous bit rate required for the active phased array antenna a slightly higher output power is needed from the user terminal if that option i chosen. The 11.9 dBW that give the necessary bit rate of 250 Mbps corresponds to 15.4 W. Increasing the output power to 100 W will allow for an instantaneous bit rate of 1.6 Gbps. Under the given conditions with 250 simultaneous users that is equal to an average bit rate of 6.5 Mbps. This average rate is lower than that provided by the multibeam solution, but the time slot system of the phased array system have an advantage. By allocating time slots dynamically it is possible to assign extra slots to users with much data to transmit. In the extreme case all slots can be given to one user allowing that user the full capacity of the satellite. Such flexibility can be very valuable.

Another aspect is that the users are not going to be uniformly distributed around the coverage area. They will to some extend cluster in some areas. For example along shipping routes and in areas with oil and gas operations. With an unflexibile system such as the multibeam system the result may then be that, even though the total throughput capacity of the satellite is high enough, one of the beams are swamped with traffic while others have available capacity. This is not a problem for the phased array as the beam are constantly moved according to the traffical needs. In other words it is possible to utilize the full capacity of the satellite regardless of the position of the users.

From this it can be concluded that the option with the active phased array antenna should be chosen for the uplink as well. This has the added benefit of beeing a good match with the downlink system. Since both systems then uses a time divison multiple access (TDMA) technique the size and complexity of the buffer and onboard processing can be somewhat reduced. As data packages are received in the satellite one by one they can be retransmittet more or less instantly without extensive use of interim storage.

Chapter 6

System overview

In the previous chapters several important parts of a satellite communications system have been discussed with the aim of finding the right way to provide polar areas with broadband communications.

Chapter 2 looked at propagation issues for a range of frequencies, and it was concluded that downlink at 20 GHz and uplink at 30 GHz will give the best result. At these frequencies the attenuation in the atmosphere is manageable, and the short wavelength allow for small antennas with fairly high gain.

In chapters 3 and 4 satellite orbit and constellation design was addressed. A higly elliptical orbit with a nearly twelve hour orbit period and an inclination of $63.4°$ was found to be best suited. One of the biggest challenges when communicating with a non-geostationary satellite constellation is the handover problem. Here this is suggested solved by adjusting the eccentricity of the orbit so the satellite ground track forms a loop. Handover can then be performed when the satellites are above the exit and entrance to this loop even if highly directive antennas are employed at the user terminals. This point has been refered to as the ground track intersection point. In a constellation of four satellites each satellite will spend six hours in the operational loop.

6.1 Radio system

With the frequency and orbit decided it is possible to do initial studies and assessments of the radio system and its performance. This was done in chapter 5. The result from the discussion in section 5.5.4 is that active phased array antenna systems should be employed on the satellites, both on the uplink and the downlink. Each transmission direction will use one beam. Users

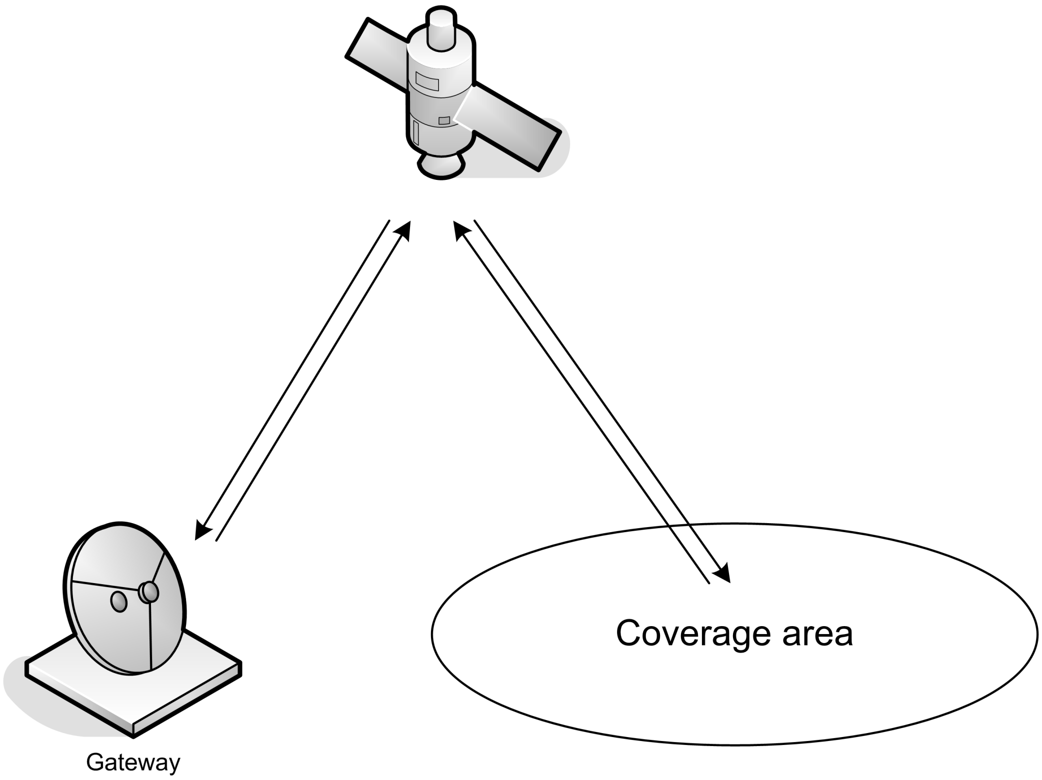


FIGURE 18: The system setup with the necessary gateway.

are individually targeted by the beams in assigned time slots. The allocation of the time slots are done dynamically which makes exploitation of the full capacity of the satellite possible, regardless of the users needs and positions.

In addition to communications between users of the system the satellites must provide communication with the rest of the world through a gateway. Transmission between the gateway and the satellite should happen via a dedicated beam. Thus, the satellites empolyed in this communications system must have four beams. The two user beams, downlink at 20 GHz and uplink at 30 GHz, that are moved between the users as needed, and two gateway beams, also downlink at 20 GHz and uplink at 30 GHz, that are fixed towards the gateway station. Placement of the gateway station is not limited to the coverage area. Figure 18 show a simplified version the system setup. The four beams can in theory be generated by the same array, but because of the large frequency difference between uplink and downlink it may be more practical to use two apertures. One for downlink beams and one for uplink beams. Further studies must be performed in order to verify this assumption.

The capacity of the satellite must be balanced. There is no need for more uplink capacity than there is downlink capacity. The gateway downlink will only be used when user terminals are transfering data out of the system through the gateway. As a result it does not need a higher capacity than the user uplink. For the user downlink the highest possible traffic is the combined traffic from both the gateway and the user uplinks. It is not very likely that this will happen over long periods of time and it can probably be handled with buffering, but the user downlink should probaly have somewhat higher capacity than the user uplink. Of course only as long as it doable at a reasonable cost. In theory the gateway uplink can have the same capacity as the user downlink, but that will increase the risk of congestion in the user downlink. A somewhat lower capacity will then reduce the necessary buffer capability.

As mentioned in the uplink discussion in section 5.5.4 it should be possible to produce user terminals with 100 W of output power that will allow for a instantaneous bit rate as high as 1.6 Gbps. Then it is assumed a QPSK modulated signal with a half rate convolutional coding is used. This modulation and coding scheme has a spectral efficiency of 1 [3]. In plain terms that means that 1 bit of data require 1 Hz of bandwidth. Thus, the available bandwidth must be equal to the instantaneous bit rate. Then it is necessary for the system to use a bandwidth of 1.6 GHz both around 20 and 30 GHz. This is an extremely large bandwidth and allocation of such a large bandwidth will probably not be possible. Suitable components able to handle the wide bandwidth will also be difficult to find. As a result the bandwidth usage have to be reduced from the maximum, either by reducing the bit rate or by employing a different modulation technique with higher spectral efficiency.

Here the modulation and coding scheme will be kept and the instantaneous bit rate reduced to around 1 Gbps. The resulting bandwidth requirement of 1 GHz may stil cause difficulties when it comes to allocation and component selection and design, but they should be easier to solve. A revised link budget for the uplink where the output power is set to 18.3 dBW, equal to 67 W, is shown in table 16. The maximum bit rate is the maximum instantaneous bit rate possible with the given user terminal output power. The average bit rate experienced by a user will depend on the total number of users and the amount of data they want to transmit.

Some transmission time will be lost each time the beam is moved. The beam may move from one user to another between two time slots. Thus, some guardtime must be inserted between the slots to allow transients in filters and phase shifters to die out. How much guardtime that is necessary is, however, difficult to say for certain at this point, and only assumptions can be made. If the guard time is set to about 8 % of a time slot the satellite can receive approximately 1 Gbps.

Table 16: Revised link budget for the uplinks at 30 GHz.

	Value	Unit
Ground station		
Output power	18.3	dBW
Antenna gain	47.7	dB
Pointing error loss	0.9	dB
Satellite		
Antenna gain	54.7	dB
Edge of beam loss	3.0	dB
System noise temperature	28.2	dBK
Other parameters		
Free space loss	214.1	dB
Atmospheric loss	0.5	dB
Boltzmann's constant	−228.6	dBW/K/Hz
Result		
$\frac{C}{N_o}$	102.6	dBHz
Margin[1]	7.8	dB
$\frac{E_b}{N_o}$	4.4	dB
Maximum bitrate	90.4	dBHz
	1.089	Gbps

[1]Margin necessary to handle attenuation caused by rain.

The uplink from the gateway can in theory have a higher bit rate than the user uplink. However, that would require the gateway uplink to use more bandwidth or a different modulation technique than the user uplink. The resulting difference in the incoming transmissions will complicate the design of the radio system. This complication is not present if the same time slotted transmission structure is used in both the gateway and user uplink. Thus, the link budget given in table 16 can also be used for the gateway uplink. The gateway then have the possibility to transmit up to about 1 Gbps.

Since the gateway downlink rate never will exceed the user uplink rate a bit rate of 1 Gbps is enough for the gateway downlink. As mentioned above it would be advantageous if the user downlink has a higher capacity, but as the case for the uplink system it may complicate the design if the two downlinks have different bit rate or bandwidth. The capacity of the gateway downlink could be increased, but that increased capacity will never be utilized. In other words a cost increase without benefit. Probably the best solution is then to set the bit rate of the user downlink to 1 Gbps as well. This has

TABLE 17: Revised link budget for the downlinks at 20 GHz.

	VALUE	UNIT
SATELLITE		
Output power	16.9	dBW
Antenna gain	51.2	dB
Edge of beam loss	3.0	dB
GROUND STATION		
Antenna gain	44.2	dB
Pointing error loss	0.4	dB
System noise temperature	26.6	dBK
OTHER PARAMETERS		
Free space loss	210.6	dB
Atmospheric loss	0.5	dB
Boltzmann's constant	−228.6	dBW/K/Hz
RESULT		
$\frac{C}{N_o}$	99.8	dBHz
Margin[1]	5.0	dB
$\frac{E_b}{N_o}$	4.4	dB
Maximum bitrate	90.4	dBHz
	1.088	Gbps

[1]Margin necessary to handle attenuation and increased noise caused by rain.

the added benefit of allowing a uniform signal structure through the whole communications system.

The downlink budget presented in table 14 should then be modified to the desired bit rate. In table 17 these revisions are done. Assuming a guardtime of about 8 % of a time slot an instantaneous bit rate of 1.088 Gbps will give the desired capacity of 1 Gbps, the same as for the uplinks. An output power of 16.9 dBW, or 49 W, is needed from the satellite for that to be possible.

6.2 Challenging issues

There are issues with the suggested system that may require extra carefull planning and innovative design solutions. Some of them have been discussed and addressed to a certain extent in previous chapters while others have not.

The handover problem was given some consideration in chapter 3 in terms of orbit design. As a result the suggested orbit and constellation design minimizes the handover issues. However, only the question of when and where handover should take place is answered. The technical issues when it comes to moving traffic from one satellite to another on a regular basis have not been adressed properly. A simple solution is to turn the incomming satellite on and the outgoing satellite off at the same time when their paths cross. Unfortunately that may not work as seemless as desired if for eksample the message buffer in the outgoing satellite is not empty. The simple on/off switching will then stop that buffered data from reaching its recipients, which of course is an undesireable situation. Hence, a more advanced handover system is necessary to ensure seamless handover between satellites.

Another issue that is just briefly mentioned in chapter 5.5.4 is the beam-steering system for the active phased array antennas. The necessary accuracy is probably not too difficult to achieve if the target is fixed and stationary, but that is not the case here. During transmission in a timeslot the antenna beams can probably be kept stationary given that the timeslots are not to long. However, between the timeslots the user beams can be required to adjust the pointing direction with several degrees. The gateway beams will only have to move slightly and in a predictable manner. Correct phase and amplitude for each array element can then be calculated in advance. For the user beams it may be necessary to do new calculations for each time slot. Depending on the number of elements in the arrays that may be a large task requiring substantial computational power.

In order to do these calculations the satellite must know when and where to direct the beams. Knowledge about the position of all active user terminals is probably necessary. Downlink transmisson will require little coordination and planning if it is assumed that all user terminals are in receiving mode unless they are transmitting. The satellite can then simply retransmit all received messages as long as the position of the recipient is known. Uplink transmission from the user terminals will on the other hand require much coordination and planning. Since the satellite uplink beam only covers a small fraction of the coverage area at a time a user teminal can not simply transmitt data when that suits it. For a transmission to be successfull the user terminal must be assigned a time slot and the satellite antenna beam must be targeted at that user. However, before the satellite can assign a time slot to an user the satellite must be notified of that users desire to transmit data. The challenges concerning access controll are many and it is a problematic issue that will require an innovativve and clever solution.

Chapter 7

Conclusions and future work

In this report a satellite system providing broadband communications to polar areas have been discussed. Various important parameters such as frequency selection, orbit and constellation design have been addressed along with a more detailed look at the communications system.

Four different frequency configurations ranging from 20 to 70 GHz was evaluated with focus on their propagation properties. Rain effects was found to dominate all other propagation phenomena. Attenuation and increased radio noise caused by rain is frequency dependant, with higher frequency resulting in raised attenuation and noise. Thus, 20 GHz was chosen as the downlink frequency and 30 GHz as the uplink frequency. This configuration is actually distinctively better than the three other alternatives even if rain is ignored.

Calculations and simulations using STK have proved that the satellite system should utilize satellites in Molniya orbits. That give the best coverage of the desired coverage area, and with the right eccentricity it drastically simplifies the handover problem. It was found that the best constellation uses four satellites in Molniya orbits with an eccentricity of 0.72. The ground track of the satellites then form a loop where they spend six hours in operating mode. With four symetrically positioned orbital planes handover can then be done when one satellite exits the loop, and a new enters.

Link budget calculations has shown that the most capable system is achieved when active phased array antennas with a diameter of 2 m are used on the satellite. The high gain antenna beams can then be electronically controlled and individually targeted at users. A downlink beam then requires an output power of less than 50 W in order to be able to provide an instantaneous bit rate of about 1 Gbps. On the uplink the same bit rate requires an output power from the user terminals of approximately 67 W, assuming 1 m reflector antennas is used. It is suggested that the hopping beams should not follow fixed patterns, but be dynamically allocated to users who wish to transmit data.

The findings presented here is only the very beginning of the long process it is to realize a satellite system able to provide broadband coverage to polar areas. Huge amount of work remain as spcacraft design, gateway and user terminal composition, access controll and other important tasks require attention. However, this study has shown that it is definetively possible to provide polar areas with broadband communications using a non-geostationary satellite system.

7.1 Future work

Before this system can be realized and provide broadband access to people in polar areas a multitude of questions must be answered. Some of them have been adressed here to some degree while others can be answered with basis in this work. Other issues again are not even thought of yet, but they all must be looked into and solved before this communications system can be realized. The following sections give a brief summary of some issues that must be addressed. To some of the issues possible solutions is offered, but they are only ideas at an early stage and not yet thourougly worked over. This brief account is not complete as more problems and challenges will arise as the work progresses.

7.1.1 Spacecraft design

Design of the spacecraft and its composition is a very important and comprehensive task. It includes the spacecraft bus with all its subsystems and the communications system, with the antenna structures, as the payload. The job of the spacecraft bus is to support the payload and make sure it can function as intended. Before any major work can be done on the spacecraft bus at least a preliminary design of the payload must be available. Without that it is not possible to know what is required by the various subsystems.

The attitude determination and control system design depends on the pointing requirements and termal control depends on the type of amplifiers and their configuration. In order to make decision about the power system with its solar cells and battery bank some knowledge about the power consumption levels is necessary. These are only a few examples as all subsystems interact with the payload to some extend. The design of the subsystems also depends on each other and influences the payload design, making the design of the whole satellite a complex and demanding process.

It should be given some consideration if the satellite design should start from scratch or use a standardized spacecraft bus and only adapt it to the mission. The second option can potentially reduce the cost significantly, while a fully costumized satellite may give better performance. However, it does not necessarily have end up that way, as both risk resulting in poor performance at a high cost.

Satellite antenna design have been addressed in this study, but mainly on the general technology level. The solutions suggested as the best here is not the one that has conventionally been used for this application. Research should therefore be done to verify that the desired performance is achievable, and of course how this can be done.

7.1.2 Access control

The issues mentioned in section 6.2 must of course be given attention. Slot assignment and access control is there mentioned last, but that is not because it is the least important. It is crucial that this problem is solved in a good and effective way. Luckily there are numerous approaches to this problem, but if the wrong solution is chosen the whole system will function suboptimal.

One option is to use a separate control system with its own frequency on both uplink and downlink. User terminals can then use the uplink control frequency to ask for time slots on the regular frequency. After receiving a slot request the satellite send back slot assignments on the downlik control frequency. These control messages can be very small so there is not necessary to design the control system for high bit rates. As a result there is no need for a high gain antenna allowing the control system to use widebeam antennas that covers the whole coverage area. When a user terminal wishes to transmit data it then sends a slot request on the control frequency, using for example slotted aloha or a similar accsess technique, and receives back a suitable transmisson schedule for the regular frequency. This control system should also be used to update the satellite with the positions of user terminals. The drawback with this option is the need for a separate radio system, possibly with its own antennas.

Another option is to give all users a fixed transmission slot. This slot they then use to update their positions and request additional transmission slots. Some data transmission can of course also be done in this fixed slot. With a large number of users this is not a good solution as the amount of free slots then is reduced. In addition it is probable that many of the fixed slots will more or less be unused. Also new users will have difficulties accessing the system.

A combination of the two solutions is an alternative. With the active phased array it is possible to control the width of the beam. Some time slots where the beam is adjusted to cover the complete coverage area can then be reserved for control messages. User access to these slots can again be done through slotted aloha or a similar access technique. This option reduces the total transmission capacity of the satellite a little, but it may not require a separate radio system to control user access and their positions.

In addition to the options for access control outlined here there may be others, and a study that looks at all the possibilities should be performed. Access control is, as mentioned above, important in order to maximize the throughput of a communications system, and the success of a satellite system can in the end be decided by that.

7.1.3 Gateway and user terminals

Another topic for future research and study is the composition and design of the gateway station and the user terminals. The possibility of using allready existing user terminal systems should also be looked at. If that is possible it will not be necessary to develop these systems from scratch, and that will save time and money. It will pobably also secure the system easier access to the communications market as potential user may allready be familiar with and even have the necessary equipment. The user terminals must comply with the transmission protocol and the access control scheme discussed in the privious section. If an existing user terminal design can be used with only small modifications, either on the user terminal or the transmission and access control scheme, it will by all accounts be worth exploring.

The satellite system will be best served with two gateways for regular service. One for each operational satellite. In other words one gateway located in northern Europe and one in the northern part of North America. The exact location should be decided based on the final orbit positions and infrastructural needs. An operation control centre for the satellites is also needed. If positioned correct it is enough with one facility, but a backup control centre should be available. Thus, it can be argued that two operation control centers should be colocated with the two gateways. If these facilities can be placed in conjunction with allready existing satellite control stations further synergies might be achieved, and that justifies that at least some attention is given to the issue.

7.1.4 Handover

Handover is probably the least critical of the issuses mentioned in section 6.2, both in time and importance, but for some users continuous communications is important and a connection loss, even if only for a few seconds, can have catastrophic consequences. One possible and simple solution is mentioned in section 6.2, but as it is stated there it is not the ideal one. The task of finding good handover solutions should therefore get attention in the future.

In relation to the handover problem the issues regarding the required accuracy of the satellite orbit's eccentricity should be looked into. It would be of great interest to know if the required eccentricity accuracy is difficult to aquire and maintain, and the effect this has on the propellant budget.

7.1.5 Antenna beamsteering

In this study assumptions about the active phased array antenna system have been done. Among those are the possible accuracy and the ability to quickly change the pointing direction. These assumptions needs to be verified, at least to some extent, before work on related matters can continue. There is no indication that the assumptions are erroneous, but if they at a later stage turn out to be very wrong huge amount of work may have been wasted. For example, if the assumed pointing accuracy is not possible to guarantee the beam must be wider. That will reduce the antenna gain resulting in a weaker recived signal and lower possible bit rate. A design of the radio system based on the original bit rate may then be invalid and require a complete redesign, which should be avoided.

7.1.6 Bandwidth

In section 6.1 it is breifly mentioned that it can possibly be difficult to aquire the rights to use the necessary frequency resources. Before any detailed radio design is started the bandwidth available to the system should be known with at least some degree of certainty. Preliminary design studies of the radio system can also end up concluding that a bandwidth of 1 GHz will complicate the design to much and give a suboptimal result. It might also be possible to experience selective fading over such a large bandwidth.

If the bandwidth must be reduced from the one stipulated in this study the bit rate is reduced proportionally. This reduction can possibly be countered if a higher order modulation technique with higher spectral efficiency is used instead of QPSK. As discussed in section 5.5.2 that will increase the necessary bit energy to noise density ratio, but that will probably be available since the system then are bandwidth limited.

Even if the bandwidth of a little more than 1 GHz is available at both 20 and 30 GHz the possibility of using a different modulation and coding scheme than suggested here should be considered. Not because it is believed to be a poor choice, but because it might exist a better solution. As it is stated in 5.5.2 modulation and coding is a vast field, and this study barely touches it. A closer look can therefore be beneficial.

7.1.7 Propagation

In this study all propagation effects have been estimated. Estimations is of course vastly inferior to actuall measurements, but there exist very little relevant data for the polar areas. There has been performed large amounts of measurements at Svalbard, but more or less only towards geostationary satellites. Due to the low elevation angles these results are difficult to use directly when planning a Molniya based system. Thus, it would be advantageous if measurements with higher elevation angles can be done and data collected, allowing for a better understanding of the propagation conditions in polar areas. That would pave the way for more accurate prediction and estimation methods which in turn results in more exact link budgets and removes some uncertainties.

7.1.8 Users and economy

The number of possible users and their communications needs are unknown. That is not a good situation to be in when designing a communications system. Due to the lack of information on this subject the results in this study is found independent of the users and their cravings.

This is not a problem while discussing the total capacity of the system, but in order to know what that translates into for the individual costumer some grasp of at least the number of users are needed. Some might say that the information is not necessary until the time comes for selling the services, but those people are wrong. If the number of users and their needs are known to some extent early on it is possible to dimension the system and its equipment accordingly, but maybe most importantly it can give valueable insigth into how the system can be financed.

In order to achieve success in the commercial world of today a good finacial plan is paramount. No investors will open their wallet before they understand how the satellite system will bring in revenue, and if that revenue is large enough to pay the bills. It does not matter how smart and usefull a product is, if it is not profitable it will not live long and that will scare investors away.

A financial plan rely on cost estimates. Therefore it is necessary to estimate the costs of the satellite system from design and deployment to daily operations, including launch costs and insurance. Replacement of satellites at their end of life should also be contemplated even at this stage. These cost estimates can then be used to estimate the revenue required for long term profitable operation of the satellite system. If this information is coupled with estimates of the number of users and their needs a financial structure and plan can be created. Based on this plan it is possible to decide if a satellite based system for broadband communications to polar areas has the right to live.

7.1.9 Intersatellite links

It may possible to use intersatellite links to improve the throughput and reduce transmission delays. Such links is not only limited to satellites that are part of this system. Links to geostationary satellites and low earth orbit constellations is also possible. A study that looks at this topic to see if there can be gained any substantial benfits from various types of intersatellite links would be interesting.

A related issue that deserve some attention is the posibility of single hop communications between the two coverage areas. For example between northern Norway and Alaska. That is not possible with a geostationary satellite. Thus, possibly another service this satellite system can provide that is not available today. If this is a service that adds value to the satellite system should be investigated.

7.1.10 Additional payloads

The downlink budget in section 6.1 indicate that the total output power for the two downlinks is about 100 W. There will of course be additional power consumption in the rest of the radio system and other satellite subsystems. How much power the whole satellite will require can only be loosely estimated at this point. The best estimate is somewhere between 250 and 350 W. As a rule of thumb a solar array produce about $100\,\mathrm{W/m^2}$ [3]. Hence, a solar array of around $3.5\,\mathrm{m}^2$ will cover the satellites needs.

For a communications satellite that is not very much. These low requirements can be utilized to design a small and fairly inexpensive satellite. However, the total size of the satellite will be affected by the large antenna arrays suggested. With a larger size comes more surface area and the possibility for more solar cell area. The result can be more available power than needed. Therefore the possibility of adding additional payloads to the satellites should be investigated.

One posibility is broadcasting of television. Satellite based broadcasting systems available today all use the geostationary orbit. These satellites are not available far north as discussed in chapter 3. Also they tend to focus their beams on land where the bulk of the television viewers are. Thus, there are large areas inside of the suggested coverage area where television is not available today. A broadcasting payload on the satellite will probably need more than just the surplus power, but the design of a larger and more complex satellite may be warranted by the increased revenue.

A navigation payload is also a possibility. Most likely then as part of the coming European Galileo system as it is a civilian system. The satellites in both the developing Galileo system and the fully functional American NAVSTAR system (GPS) have an inclination of about 55-56°. North of 56° latitude all the satellites will then be south of a user, and when moving north the elevation angle will be reduced more and more. The result is lower position accuracy, especially in the vertical direction. A navigation payload on a Molniya based satellite system would then improve the accuracy of satellite aided navigation in polar areas. In a simplified version it might be possible to implement this additonal payload without enlarging the satellite to any great extend.

Another possible option is to equip the satellites with one or more C-band, 4/6 GHz, transponders. That will allow the use of more or less standard VSAT terminals. Such terminals are in use world wide in connection with the Intelsat satellite system. VSAT terminals are usually equipped with 1 m reflector antennas. At C-band that results in a fairly wide beam which reduces the accuracy required by tracking systems if used with the satellites suggested in this study. Because of the widespread use and availability of VSAT systems a C-band payload can possibly function as a fast and smooth transition to the proposed satellite system for many users, and also increase the number of potential costumers.

These are just three examples of possible additional payloads, but there probably are more. Extra payloads will complicate the satellite design, but it can also increase the possible revenue and make it easier to find funding as more needs are satisfied. However, it is important to not overreach as that can topple the whole project. A balanced solution must be found.

References

[1] ITU-R, *Radio Regulations*, International Telecommunication Union, 2001.

[2] D. M. Pozar, *Microwave and RF Wireless Systems*, John Wiley & Sons, 2001.

[3] W. J. Larson and J. R. Wertz, editors, *Space Mission Analysis and Design*, Microcosm Press, 3rd edition, 1999.

[4] T. Pratt, C. Bostian, and J. Allnutt, *Satellite Communications*, John Wiley & Sons, 2nd edition, 2003.

[5] G. Stette, *Radiokommunikasjon*, Dept. of electronics and telecommunications, Norwegian University of Science and Technology, 2004.

[6] J. Allnutt, *Satellite-to-ground radiowave propagation*, Peter Peregrinus Ltd, 1989.

[7] L. J. Ippolito, *Radiowave propagation in satellite communications*, Van Nostrand Reinhold Company Inc, 1986.

[8] ITU-R Recommendation P.676-7, *Attenuation by atmospheric gases*, International Telecommunication Union, 2007.

[9] ITU-R Recommendation P.839-3, *Rain height model for prediction methods*, International Telecommunication Union, 2001.

[10] ITU-R Recommendation P.837-5, *Characteristics of precipitation for propagation modeling*, International Telecommunication Union, 2007.

[11] ITU-R Recommendation P.618-9, *Propagation data and prediction methods required for the design of Earth-space telecommunicaton systems*, International Telecommunication Union, 2007.

[12] O. Gutteberg, *Effects of atmosphere on earth-space radio propagation*, pp. 63–70, Telektronikk, Televerkets forskningsavdeling, 1992.

[13] L. E. Bråten, T. Tjelta, and D. Larsen, *Excess attenuation caused by sleet on coastal terrestrial radio links in Norway*, Telenor R&D, 2003.

[14] ITU-R Recommendation P.840-3, *Attenuation due to clouds and fog*, International Telecommunication Union, 1999.

[15] E. Salonen and S. Upphala, *New prediction method of cloud attenuation*, pp. 1106–1108, Electronics Letters, June 1991.

[16] J. R. Wertz, *Mission Geometry; Orbit and Constellation Design and Management*, Microcosm Press, 2001.

[17] A. K. Maini and V. Agrawal, *Satellite Technology*, John Wiley & Sons, 2007.

[18] P. Berlin, *Satellite Platform Design*, Dept. of Space Science, Universities of Luleå and Umeå, 4th edition, 2005.

[19] Orbit Technology Group, *http://www.orbit-tracking.com*, Last visited: 11.04.08.

[20] R. C. Johnson and H. Jasik, editors, *Antenna engineering handbook*, McGraw-Hill Book Company, 2nd edition, 1984.

[21] A. I. Zaghloul, Y. Hwang, R. M. Sorbello, and F. T. Assal, *Advances in multibeam communications satellite antennas*, pp. 1214–1232, Proceedings of the IEEE, IEEE, July 1990.

[22] E. Lier, D. Purdy, and K. Maalouf, *Study of deployed and modular active phased-array multibeam satellite antenna*, pp. 34–45, IEEE Antennas and Propagation Magazine, IEEE, October 2003.

[23] A. Jacomb-Hood and E. Lier, *Multibeam active phased arrays for communications satellites*, pp. 40–47, IEEE Microwaves Magazine, IEEE, December 2000.

[24] J. A. Aas, *(private communication)*, Dept. of electronics and telecommunications, Norwegian University of Science and Technology, 2008.

[25] S. Haykin, *Communications systems*, John Wiley & Sons, 4th edition, 2001.