Architecting Space Communication Networks

by

Marc Sanchez Net

Submitted to the Department of Aeronautics and Astronautics
in partial fulfillment of the requirements for the degree of

Master of Science in Aeronautics and Astronautics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2014

Signature redacted

Author ....
Department of Aeronautics and Astronautics
May 22, 2014

Signature redacted

Certified by ....
Prof. Edward F. Crawley
Professor of Aeronautics and Astronautics and Engineering Systems
Thesis Supervisor

Signature redacted

Accepted by ....
Prof. Paulo C. Lozano
Associate Professor of Aeronautics and Astronautics
Chair, Graduate Program Committee

Architecting Space Communication Networks

by

Marc Sanchez Net

Submitted to the Department of Aeronautics and Astronautics
on May 22, 2014, in partial fulfillment of the
requirements for the degree of
Master of Science in Aeronautics and Astronautics

Abstract

Reliable communication and navigation services are critical to robotic and human space missions. NASA currently provides them through three independent and uncoordinated network that consist of both Earth-based and space-based assets, all managed under the Space Navigation and Communication Program. Nevertheless, the ever increasing mission requirements and funding limitations motivates the need of revising the current network architectures in order to identify areas of potential performance and cost efficiency improvements.

The main objective of this thesis is to present a tool that helps decision-makers during the process of architecting a space communication network by (1) systematically enumerating and exploring the space of alternative network architectures, (2) identifying those with better performance and lower cost, and (3) providing traceability between the outputs of the tool and the architecting decisions. The tool is tailored to the high level design of near Earth space communication networks that support robotic and human activities in the Earth vicinity through a set of relay communication satellites and their supporting ground stations. The decisions available to the network architect (both technical and contractual) are presented and along with their couplings.

The tool is validated by comparing it to NASA's Space Network. The current operations of the system are analyzed and used as the baseline case for the validation process. Results demonstrate that the both performance model and spacecraft design algorithm are accurate to less than 10%, while the cost module produces estimates with a 15% error.

Finally, the utility of the tool is demonstrated through three case studies on the evolution of the Space Network. In particular, the impact of new radio-frequency and optical technology to increase the system capacity is analyzed based on the predicted demand for the 2020-2030 decade. Similarly, the savings of flying relay transponders in commercial satellites as hosted payloads are quantified and benchmarked with respect to NASA's current approach of procuring and operating the entire network. Lastly, the tool is used to compare the current Space Network bent-pipe architecture with a constellation of satellites that takes advantage of inter-satellite links to provide full coverage of low Earth orbits with only one ground station.

Thesis Supervisor: Prof. Edward F. Crawley
Title: Professor of Aeronautics and Astronautics and Engineering Systems

Acknowledgments

This thesis is the result of two amazing year of collaborative work in the System Architecture Lab. I want to first thank Dani, Bruce and Ed for all their support and valuable insights. Neither the work that I have done, nor the passion I have put into it would have been possible without your guidance, help and encouragement. I have really enjoyed every moment that we have shared, be it at MIT or somewhere else in our numerous trips.

I also want to acknowledge all my other friends in the lab - Peter, Morgan, Iñigo, Alessandro, Francisco, Alexander and John -, in 409 - Patricia, Paul, Narek, Sreeja, Sydney, and Ioana - and around campus - Ana, Hector, Alexandra, Roberto, Fernando, Ada, Carlos. You are a fundamental part of what it means for me to be at MIT and Boston, and I can hardly imagine going through these past two years without you. For those of you who are leaving us, all the best of luck in your future endeavors, you will be very much missed!

Thanks to Jim Schier, Bernie Seery, Antonios Seas, Greg Heckler and David Milliner. It has been a pleasure to work side by side with you as part of the SCaN program.

Romina, you have been the best source of support and affection that I could have ever wished for. These last eight months have been a mix of great moments, incredible experiences and unforgettable memories.

Finally, I would like to thank my family and in particular my mother. You have always encouraged me to be ambitious, curious and hardworking, and this has been key to the success of this thesis and all previous accomplishments I might have had.

THIS PAGE INTENTIONALLY LEFT BLANK

Contents

1 Introduction **15**

1.1 Background and motivation 15

1.1.1 History of NASA's space communication networks 15

1.1.2 Motivation 17

1.2 System Architecture 18

1.2.1 Tools for System Architecture 20

1.3 Generic problem statement 21

1.4 Architecting space communication networks 22

1.4.1 Network simulators 23

1.4.2 Point designs 24

1.4.3 Architecture studies 25

1.4.4 Tradespace exploration 26

1.5 Specific problem statement 27

1.6 Thesis overview 28

2 The Space Network Architecting Tool **29**

2.1 Introduction 29

2.2 Decisions to architect a space communication network 29

2.2.1 Network topology 31

2.2.2 Business model 32

2.2.3 Network technology 34

2.3 SNAT architectural decisions 35

2.4 Model overview 37

2.5 Model description 39
2.5.1 The VASSAR framework 40
2.5.2 Space and ground segment design 41
2.5.3 Network evaluator 49
2.5.4 Cost estimation 61
2.5.5 Search engine 67

3 Validation of SNAT **69**

3.1 Introduction 69
3.2 Validation strategy for the performance model 70
3.3 Validation of the performance model 70
3.3.1 Dataset description 70
3.3.2 Definition of network metrics 72
3.3.3 Analysis of TDRSS operational data 74
3.3.4 Validation of the scheduling algorithm 83
3.4 Validation of the spacecraft design algorithm 85
3.5 Validation of the cost model 88
3.6 Tradespace validation 88

4 Evolving NASA's Space Network **91**

4.1 Introduction 91
4.2 Network customer characterization 91
4.3 Case study 1: Valuation of new technology and hosted payloads . . . 94
4.3.1 Tradespace definition 94
4.3.2 Results: Infusion of new RF and optical technology 97
4.3.3 Results: Procurement vs. hosted payloads 100
4.4 Case study 2: Valuation of inter-satellite links 103
4.4.1 Tradespace definition 103
4.4.2 Results 105

5 Conclusions **111**

5.1 Thesis summary 111

5.2 Main contributions 113

5.2.1 Methodological and modeling contributions 113

5.2.2 Findings from the case studies 114

5.3 Future work 116

THIS PAGE INTENTIONALLY LEFT BLANK

List of Figures

2-1 Notional space of architectures 30

2-2 Model overview 38

2-3 Example of a tradespace from SNAT 39

2-4 Spacecraft design algorithm 44

2-5 Architecture satisfaction decomposition 62

2-6 Lifecycle cost breakdown 62

3-1 TDRSS daily data volume and service time 75

3-2 TDRSS number of granted contacts 75

3-3 Network level statistics 76

3-4 Metrics vs. service 77

3-5 Service popularity 78

3-6 Scheduled time per satellite 79

3-7 Data volume per satellite 79

3-8 Single Access antenna utilization 81

3-9 Scheduled contact data rate distribution 83

3-10 Evolution of the network capacity 89

4-1 Tradespace definition 96

4-2 Transponder selection 99

4-3 Transponder allocation 100

4-4 Evolution of a procured Space Network 101

4-5 Procurement vs. Hosted payloads 102

4-6 Combinations of procured and hosted payloads 102

4-7 Number of hosted payloads per satellite 104
4-8 Benefit-cost tradespace 106
4-9 Detail of the number of satellites and ground stations 108
4-10 Detail of inter-satellite links and ground station selection 108

List of Tables

2.1 Architectural decisions . . . 36

2.2 Correction factors for ISL and SGL data rates . . . 42

2.3 ΔV required to compensate drag for different orbits . . . 46

2.4 ΔV required for ADCS . . . 46

2.5 CERs coefficients for thermal, avionics, and structure subsystem . . . 49

2.6 Scheduling Algorithm Facts . . . 55

2.7 Scheduling Algorithm Heuristics . . . 59

2.8 Extract of launch vehicle database . . . 65

3.1 Satellite utilization . . . 78

3.2 SN nominal data rates . . . 82

3.3 Validation of $\%Ut$ for a typical day of operations . . . 85

3.4 Validation of DV_{total} for a typical day of operations . . . 85

3.5 Validation of $\%Ut$ for a high load scenario . . . 86

3.6 Validation of DV_{total} for a high load scenario . . . 86

3.7 Error on mass and power estimates for TDRS-J . . . 87

3.8 Error on mass fraction estimates for TDRS-J . . . 87

3.9 Error on cost estimates . . . 88

4.1 Detailed user scenario . . . 93

4.2 Case study architectural decisions . . . 97

4.3 Case study architectural decisions . . . 106

4.4 Pareto front architectures . . . 107

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 1

Introduction

1.1 Background and motivation

1.1.1 History of NASA's space communication networks

In 1956, the Space Studies Board of the National Academy of Sciences approved a plan by the Smithsonian Astrophysical Observatory to establish an optical tracking network to track the first American satellites [43]. In a few years, 12 optical ground stations were built around the world. The utility of these stations was limited due to the low degree of automation in the acquisition of targets. Microwave interferometric satellite tracking stations (Minitrack) were also developed in the 1950's. Minitrack was the primary Tracking, Telemetry and Command (TT&C) network for NASA during most of the late 1950's and early 1960's providing service to both Explorer 1 and Vanguard 1 missions.

In 1958 the National Aeronautics and Space Administration (NASA) was created in order to accelerate the pace of space exploration and start ambitious manned and unmanned programs. After Alan Shephard became the first American in space in 1960, the launch rate of unmanned and manned spacecraft started to grow, thus imposing tighter requirements on the Minitrack network. This caused the development of the higher performing Satellite Tracking and Data Acquisition Network (STADAN) in the early 1960's, which used 12-meter and 26-meter S-band antennas for TT&C.

Active tracking systems such as the Goddard Range and Range Rate (GRARR) were also developed as the passive interferometric systems were unable to track satellites in highly eccentric or high altitude orbits. Satellite Automatic Tracking Antennas (SATAN) were installed in order to enable data downlink for high data rate spacecraft. As NASA started to launch satellites into polar orbits, new ground stations such as Tananarive were added to support these new types of customers.

With the start of NASA's human spaceflight program (HSF) - Mercury, Gemini and Apollo programs - the Manned Space Flight Network (MSFN) was created. This new network complemented the already existing STADAN that provided service mainly to robotic missions. With the combination and expansion of the two, the use of the Minitrack network tampered off.

In 1971, after the end of Skylab, the STADAN and MSFN networks were consolidated into a single network, the Spaceflight Tracking and Data Network (STDN). It then became clear that using ground assets exclusively was not enough to meet the user requirements (especially those of manned spaceflight) due to line-of-sight constraints. The solution to this was to incorporate space assets to the network, namely the Space Network (SN), which includes a constellation of GEO satellites known as the Tracking and Data Relay Satellite System (TDRSS), and the supporting ground terminals in White Sands and Guam.

The first generation of TDRS was conceived in the early 70's to replace the MSFN. TDRSS' maiden launch occurred on April 4th 1983. Seven first generation TDRS satellites (TDRS-1 through TDRS-7) were launched between 1983 and 1995 into GEO orbits, although TDRS-2 never reached orbit as it was destroyed in the Challenger disaster. The second generation of TDRSS started with the launch of TDRS-8 in 2000, with two more satellites (TDRS-9 and TRDS-10) added in 2002. Finally, the third generation of TDRSS is currently being deployed, with TDRS-11 and TDRS-L launched in 2013 and 2014 respectively, TDRS-M scheduled for launch in 2015 and TDRS-N as a possible extension to the network by 2016.

On the other hand, the ground segment of STDN became the Near Earth Network (NEN). The current NEN consists of six NASA-operated and ten commercially operated ground stations featuring a broad range of antennas between 4 and 18 meters. These antennas are spread across the world, both in latitude and longitude, to provide contact opportunities to any LEO mission regardless of its inclination.

Finally, the last NASA owned and operated network is the Deep Space Network. It was created in the late 1950's with the goal of providing TT&C services to unmanned interplanetary missions. The DSN has been managed by the Jet Propulsion Laboratory (JPL) since its inception, and currently has three sites in Goldstone, Madrid, and Canberra with several 34-meter and one 70-m parabolic antenna. The three ground stations are approximately 120deg apart in longitude to provide full coverage, with the first two sites servicing the northern hemisphere and the last one operating in the southern one.

1.1.2 Motivation

Providing communication and navigation services is crucial to the success of any space mission. Since the start of the US space program, NASA has continuously developed and upgraded multiple networks of ground and space-based antennas that provide radio-frequency (RF) communications to spacecraft orbiting the Earth and in deep space. This networks have evolved to three independent set of assets, the Near-Earth Network (NEN), the Space Network (SN) and the Deep Space Network (DSN).

In 2006, the Space Communication and Navigation (SCaN) program was assigned management and systems engineering responsibilities for the SN, the NEN, and the DSN. The main rationale for this policy decision was to ensure that the architecture of the three formerly independent networks would evolve sinergistically to converge into a unified network that meets the needs of all user communities within the next decades. Since then, the SCaN program office has started multiple studies to explore architecture options for the SCaN network.

One of the first pieces of work within them has been the evolution of the SN and its TDRS satellites. Although the current system has been highly successful over the last thirty years, current funding limitations call for cheaper ways of maintaining and upgrading the network. As a reference, the SN operates approximately on a 10 year replenishment cycle, with three new satellites being launched each time. Based on the 3rd generation of TDRS, the cost of two satellites plus ground station upgrades is \$715M [1] approximately [2]. Additionally, NASA spends \$40M a year in operational costs, thus giving a total estimate of \$1.8B every 10 years.

Given that the SN is an expensive system, one should ask whether the current architecture is still the best alternative to meet the future demand and, at the same time, meet the expected budget limitations. Several technical and non-technical factors challenge this assertion. For instance, new RF and optical technology can be coupled with higher on-board processing capabilities to provide higher data rates with smaller transmit power and, therefore, simpler and cheaper satellite buses. Furthermore, communication payloads can be also be placed in commercial spacecraft as hosted payloads, thus reducing the cost of the system as neither the bus nor the launch vehicle have to be directly procured and paid for. These facts indicate that it is necessary to revise the architecture of a system like the SN and understand what changes can be made in order to optimize its performance and ensure its affordability within the SCaN program.

1.2 System Architecture

System Architecture as a discipline was conceived in the late 80's as a spin off of civil engineering. Since then, several other fields have embraced its methodologies and proven it successful. For instance, both aerospace and communication industries have applied it in the earliest phases of the design of complex systems [27].

The foundations of System Architecture lay in the principles of System Engineering, which considers a system as a *combination of interacting elements organized to achieve one or more stated purposes* [22]. Systems Architecture is a phase within the Systems Engineering practice in that it provides a framework to outline the high level design of a system and understand what decisions are important in order to meet the needs of those will eventually use it. In that sense, the architecture of the system

[1] Launch costs not included

can be thought of as the set of decisions that define its highest level design [44] and, in doing so, constrain the space of alternative designs and determine the majority of the system performance and cost.

Crawley defines a *system architecture* as "the embodiment of a concept: the allocation of physical/informational function to elements of form, and the definition of interfaces among them and with the surrounding context" [12]. This definition is based upon four main concepts: First, the *function* is what the system does to an external party in order to satisfy its needs. Second, the *form* is what the system is, i.e. its physical or informational representation; it is the sum of its elements and their structure. Third, the *concept* is the vision or mental model that helps humans understand how the system functions are performed with the available elements of form. And fourth, the *interfaces* are the connections between the system and its environment.

With these definitions and framework in mind, the architecture becomes the embodiment of the concept, the materialization of the mental model into a system that is implementable in the real world. Additionally, understanding the definition of a good architecture implies determining the sources of associated benefit and cost. Crawley argues that the benefit is delivered to an external party through the function of the system while the cost arises from its form. Being that the case, the value of the architecture can be measured as the benefit perceived by the external party given its cost.

In order to compare one architecture to another and determine which is "better", the system architect must specify which are the suitable metrics to measure benefit and cost. For the latter, life cycle cost (LCC) is generally used as a proxy for the economical investment required to design the system, manufacture it, operate it and finally dispose of it. However, other metrics are also used when not enough information is available to correctly price the different elements of the system throughout their life cycle (e.g. number of development projects, antenna aperture). On the other hand, the definition of metrics to measure performance is highly dependent on the problem at hand since they are, in general, tailored to particular functions that the system is performing in order to deliver value (e.g. miles per gallon for a car).

Finally, once the appropriate metrics have been defined, System Architecture advocates for systematically exploring the space of alternate architectures and understand what decisions are important (i.e. drive the system performance and cost), as well as perform trade-off and sensitivity analyses with respect to them and their associated design variables. With this process, System Architecture provides valuable information to the decision-maker by formally informing the process to transform a set of solution neutral requirements to a set of feasible conceptual designs.

1.2.1 Tools for System Architecture

At the highest level, System Architecture has benefited from computational tools for three main purposes [35]:

- To provide representation of different aspects and views of the system architecture (e.g. SySML [45]).
- To simulate the operational behavior of a system architecture using models (e.g. OPCAT [33])
- To assist decision-makers during the system architecting process (e.g. OPN [25])

Based on the notions introduced in section 1.2, architecting a system can be considered a decisions making process where the main goal is to maximize the delivered value by the system [44]. This can, therefore, be viewed as an optimization problem in which the architecting decisions are encoded as mathematical variables (continuous, discrete, logical) and the system value becomes the objective function. Variables can have lower and upper bounds so that all options in the architectural space are sensible, as well as constraints that capture relationships between them.

Therefore, this thesis is particularly interested in tools of the third type, i.e., tools that (1) can support the decision making process of the system architect and, at the same time, (2) explore the space of alternative designs to highlight those that are optimal. Simmons uses the term *architectural decision support tools* to refer to tools that can address the first part of the problem [39]. He identifies four desirable aspects that render them useful to the system architect:

- **Representational Aspect**: Methods to encode an architecture as a set of decision variables.
- **Structural Reasoning Aspect**: Methods to analyze and understand the structure of the problem and its decision variables.
- **Simulation Aspect**: Methods to list feasible architectures based on the constraints between decision variables and then evaluate them to obtain relevant metrics.
- **Viewing Aspect**: Methods to represent the output information of the tool in a way that is easily understandable by the decision-maker.

In turn, Selva introduces the concept of *System Architecting Tools* as tools that solve both problems (1) and (2) [44]. He classifies them into *decision support tools*, combinatorial optimization algorithms or search and constraint satisfaction algorithms. An exhaustive literature review of computational tools for system architecting and how they perform with respect to the previously mentioned desirable aspects can be found in his thesis.

1.3 Generic problem statement

Section 1.1.2 identified and motivated the need for re-evaluating the current SN architecture. More broadly, this can be interpreted as a need to develop tools that help decision-makers architect space communication networks, analyze them and understand the different trade-offs that arise when combining the architectural decisions.

In particular, for the tool to be useful during the system architecting process of space communication networks, it has to include the following properties: First, provide a flexible way to encode a space communication network architecture (i.e.

solve the *representational aspect*). Second, incorporate a flexible and scalable way to encode the decisions and design variables that characterize the network along with their inter-dependencies (i.e. solve the *structure reasoning aspect*). Third, provide a mechanism to generate feasible network architectures based on the architectural decisions, design variables and system constraints, and evaluate them with respect to multiple metrics (i.e. solve the *simulation aspect*). And fourth, implement a mechanism to trace the rationale(s) for the results of evaluating a network architecture (i.e. solve the *viewing aspect*). Additionally, the tool has to be able to handle and search through a large space of network designs and identify optimal alternatives.

1.4 Architecting space communication networks

Studies related to conceptual design, implementation and simulation of space communication networks are not scarce in the literature. At the highest level, they mainly follow four different approaches to tackle the problem:

- **Network simulators**: They model the different network architecture layers (physical layer, data link layer, network layer, and so on) in detail by specifying the protocols used in each of them. Then, they propagate the model in time to understand how data flows through the network and estimate the both point-to-point and end-to-end quality of service.

- **Point designs**: They specify the design of a set of communication assets and then use analytic expressions to quantitatively assess their performance and suitability with respect to the network requirements. The initial point for the study is generally a baseline architecture based on past designs or experience.

- **Architecture studies**: They propose multiple architectures for the network and then describe their desirability given the expected customer requirements. In most cases this process is conducted qualitatively, although in some cases the comparisons are partially quantified. The main difference with *point designs* is that they do not prescribe a baseline architecture but rather provide a generic exploratory view of the different feasible alternatives.

- **Tradespace exploration**: They explore a large space of network architectures (hundreds or thousands) and compute approximate metrics of desirability and cost. Then, these results are used to formally analyze the design space, understand trade-offs within the system's performance and cost, and identify architectures that are optimal.

Note that the main trade-off between the four approaches is model fidelity versus breadth in the design space. This trade-off arises from two facts: First, high fidelity models tend to be limited in the set of designs they can evaluate, i.e. they are tailored and optimized for a particular set of architectures but cannot be easily expanded to include other alternatives. Second, assuming that a higher fidelity model requires longer computational time to evaluate, then the number of alternatives that can be evaluated within a reasonable time decreases as the model fidelity increases.

The following sections provide a detailed explanation of each aforementioned category and highlight the model-fidelity vs. model-breath trade-off. They summarize relevant studies found in the literature and compare them with the desirable properties of a tool to architect space communication networks (see section 1.3).

1.4.1 Network simulators

Network simulators are widely used in the Telecommunication Industry. At the highest level, they are composed of two main parts: a discrete event simulator that replicates the system behavior through time, notifying the network model when its state needs to be updated. And a network model, that specifies the stack of protocols used by each of the network nodes and, in doing so, defines how information is transmitted through the network. Although different levels of abstraction for the network model are possible, the most common approach is to define a particular communications protocol for each layer of the OSI network architecture [47].

In the context of space communication networks, a third element is needed. Since the nodes of the network will be either space-based (communication satellites) or ground-based (ground stations in the Earth or other planets) an orbital propagator is required to determine their position over time. Then, this information can be used to infer line-of-sight (LOS) occultations. References [23], [6] and [5] present a tool that follows the previously described approach. They integrate two commercial pieces of software, STK [41] and QualNet [32], to obtain high precision simulations for the performance of the SN when it supports LEO spacecraft operations. Reference [46] introduces a similar tool, where the orbital propagator and line of sight analyses are based on custom developed modules, and the network simulator is implemented using NS-II [31].

Network simulators are highly accurate tools to solve the *simulation aspect.* Some of them also resolve the *viewing aspect* by providing graphical user interfaces (GUI) that facilitate understanding the results of the simulation and trace the rationale for some metrics. However, they neither address the *representational aspect* nor the *structure reasoning aspect* since they do not include flexible mechanisms to encode the decisions of the network architecture. As a result, they also do not provide mechanisms to search through a space of alternative network architectures. Therefore, they are of limited applicability and usefulness to system architects that aim at conducting fast trade-off analyses and understand the span of feasible designs.

1.4.2 Point designs

Point designs are typically structured as follows: First, they provide a generic vision for the need to architect a space network that supports reliable communications from different parts of the solar system (e.g. Earth commercial satellite system, lunar o Mars relay network, deep space communications). Second, they identify the high level needs of the network customers and provide quantitative requirements that can be used to assess the usefulness of a given architecture. Third, they propose a handful of alternative architectures (≤ 10), they evaluate their desirability based on both quantitative and qualitative metrics and they finally propose technical options to implement the system.

Reference [7] is a good example of a point design. It starts by indicating the need for affordable high data rate communications in the Moon vicinity to support the robotic and human exploration activities that were planned within NASA's Constellation Program. It then identifies six alternatives to create a lunar relay network and selects one of them for detailed analysis (satellites in inclined polar circular orbit at low altitudes). Next, it proposes a detailed design of the relay satellites by characterizing the communication payloads that they carry and communication technologies and protocols they utilize. Finally, it presents several use cases for the system and qualitatively analyzes the suitability of the network to address their needs.

Reference [9] is another example of a point design that studies architectures to integrate the current NASA networks to support Orion's exploration activities with end-to-end IP communication services. Similarly, [28] describes the evolution of Ka-band space communications for near Earth spacecraft assuming that the network architecture is based on the current design of the NEN. Finally, reference [3] is an example of a point designs for a commercial communication systems that provides broadband mobile services.

Point designs are mainly concerned with the *simulation aspect*, although they also devote some attention to the *representational aspect*. In particular, they identify the main characteristics of different alternative architectures without formally specifying the available decisions to the system architect. However, their analyses are tailored to the particularities of the baseline network architecture they study and cannot be applied to explore large spaces of alternative designs.

1.4.3 Architecture studies

Architecture studies provide a broader and more qualitative view of the network architecture than both network simulators and point designs. Their main goal is to frame the problem of architecting a space communications network based on the vision and needs for the system. For instance, [29] is the foundational document that identifies the need to perform architectural studies for a unified near Earth and deep space network. It describes the future needs that will have to be addressed by such a network and provides high level requirements for the different elements it will have to service. This work is further augmented by [8], [36] and [10], were specific requirements are specified and network protocols are suggested as appropriate.

Therefore, architecture studies generally focus on the *representational aspect* and *structure reasoning aspect* rather than on the *simulation* or *viewing aspect*. Their main concern is to understand the system at hand and what are the sensible alternatives given (1) the current and future technology trends and (2) the expected demands and requirements. Despite this fact, reference [21] is the only known study that couples the problem of defining architecture alternatives with that of evaluating them and providing recommendations based on quantitative findings. From the needs for NASA, it identifies four main architectural elements (ground-based Earth element, near-Earth relay element, lunar relay element and mars relay element) and proposes alternatives for each of them. Then it analyses their suitability with respect to both individual (particular to that element) and crosscutting (applicable to the entire network) requirements.

1.4.4 Tradespace exploration

Tradespace exploration is usually used to analyze large, complex and costly projects that need to satisfy the needs of several stakeholders with respect to multiple metrics. The framework has its roots in Multidisciplinary Design Optimization (MDO), a branch of Systems Engineering that provides a framework to formally search a space of alternative designs and identify its optimal solutions. Tradespace exploration has been applied in the design of multiple systems and industries, specially within the aeronautical and aerospace industry.

References [14], [15] and [38] explore the application of tradespace exploration during the design of LEO commercial communication networks. They encode the network architecture as a set of design variables that can take a discrete set of values, and then produce valid architectures by choosing one alternative for each of them. For instance, [15] identifies 5 design variables to characterize an architecture: Orbital altitude, minimum elevation angle, transmit power, antenna diameter and presence of inter-satellite links. Based on the range of values that each of them can take, a space of 600 alternative architectures is analyzed to compare their performance (measured as the number of communication channels available) and life cycle cost. Then, the authors identify the set of non-dominated architectures, i.e. those for which the system capacity cannot be increased without increasing the cost of the system.

On the other hand, reference [24] includes a case study on the design of broadband communication systems using MDO. Different constellation patterns in low Earth orbit, medium Earth orbit or geosynchronous orbit are possible, along with variations of the transmit power and antenna diameter of the satellite's transponders. The space of candidate architectures contains 42,400 alternatives which are evaluated through a combination of the GINA model [37] and parametric functions to size the spacecraft and compute its costs.

The surveyed tradespace exploration methods put emphasis on both the *representational* and *structure reasoning aspect*, and enable post-simulation analyses that provide useful insights to the decision-makers (*viewing aspect*). They also demonstrate the applicability of the framework in the design of commercial satellite system through simplified numerical models that capture the high level capabilities of the network. Additionally, they take advantage of some sort of search algorithm to enumerate large spaces of architectures and identify those that are non-dominated in a multi-objective optimization.

1.5 Specific problem statement

After reviewing the available literature on how to architect and design space communication networks, the following research objective can be formulated:

To identify the space network architecture(s) that better address the needs of future near Earth space missions **by**:

1. Characterizing the needs of future space missions with respect to communication services.

2. Identifying and characterizing the set of decisions that define a space communication network.

3. Exploring the space of network architectures defined by combinations of the decisions identified in 1.

Using a tradespace exploration tool that sizes the main elements of the network and evaluates its performance and cost.

1.6 Thesis overview

The rest of this thesis is structured as follows:

Chapter 2 starts by presenting the space of alternative architectures for near Earth space communication networks. It then formally defines the decisions that are available to the system architect and how they are modeled. Next, it provides a description of the tool developed in order to perform tradespace exploration studies in the context of space communications.

Chapter 3 is devoted to the validation of the tool. It first introduces the validation strategy against NASA's SN and summarizes the main findings of analyzing real SN operational data. Then, it benchmarks the outputs of the performance model with the SN operational data. Finally, it uses the design and cost of the 2^{nd} generation TDRS to validate both the spacecraft design algorithm and cost model.

Chapter 4 demonstrates the applicability of the tool by creating two case studies on how to evolve the SN. It initially characterizes the missions that are expected to use the system in the future and provides a high level description of the space of plausible network architectures. It then analyzes the evolutionary path for the SN from three perspectives: infusion of new RF and optical technology; choosing new contract modalities to build the infrastructure; and including inter-satellite links.

Finally, chapter 5 summarizes the contributions of the thesis, discusses the identified modeling limitations and proposes areas of improvement and future work.

Chapter 2

The Space Network Architecting Tool

2.1 Introduction

This chapter offers a detailed description of the Space Network Architecting Tool (SNAT), the newly developed computational tool that allows architecting space communication networks that provide service to missions in the Earth vicinity. The chapter is structured as follows: First, a discussion and analysis of the different decisions available to the network architect is presented. Each of them is formulated as a combinatorial problem that facilitates the process of enumerating the different architectures. This leads to the tool overview, where the different modules are introduced together with the inputs they require and the outputs they produce.

2.2 Decisions to architect a space communication network

Section 1.2.1 introduced the notion of formulating a system architecting problem as a decision making process. Therefore, in order to create a tool that assists this process it is necessary to specify what decisions are available to the network architect and

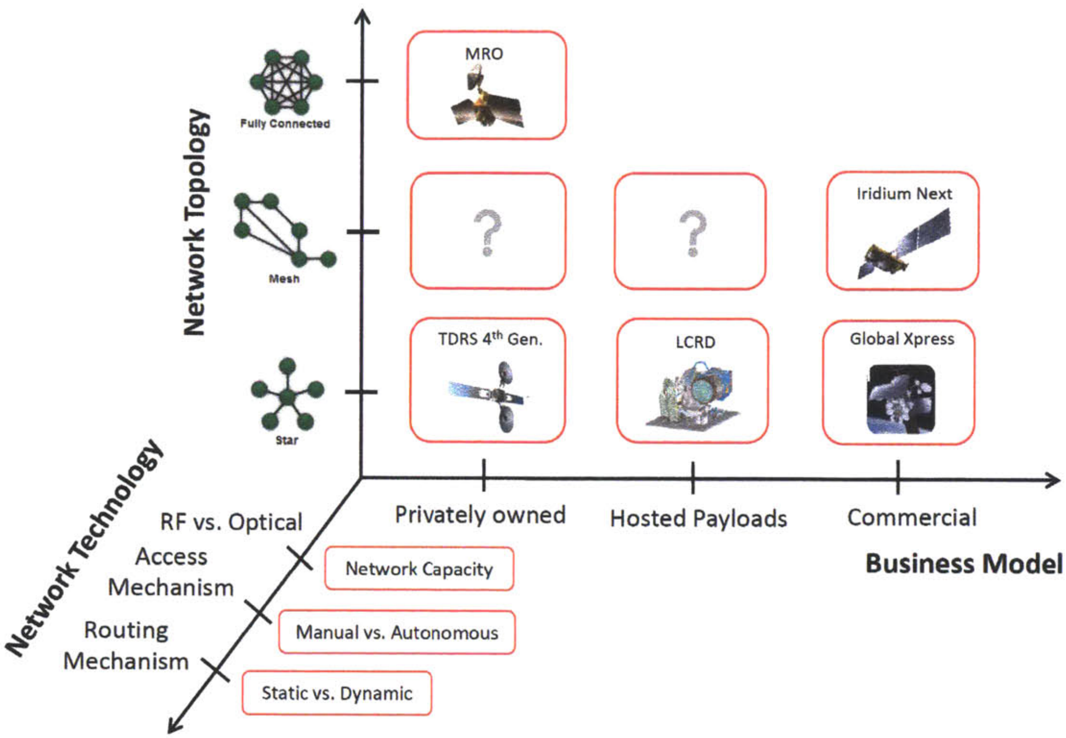


Figure 2-1: Notional space of architectures

how they are interrelated. Figure 2-1 presents a graphical representation of the main alternatives for the high level design of a space communication network. It structures them in three main axes, *network topology*, *network technology* and *business model* and indicates, when possible, systems already or under implementation that exemplify parts of the design space. For instance NASA's SN can be considered as a star topology network in that all communication channels to and from customer spacecraft are channelized through a TDRS satellite that downlinks the information directly to the supporting ground stations. Alternatively, a constellation of satellites like Iridium Next takes advantage of inter-satellite links send information from one satellite to another satellite thus creating a meshed network. Additionally, the infrastructure is owned by a commercial company that acts as a service provider to NASA and other government agencies.

On the other hand, the *network technology* axis is related to the communication technologies and protocols used to architect the network. Note that the three depicted elements have a one to one correspondence with the three first layers of the OSI model. The RF vs. optical decision can be more broadly represented as the choice of technologies to implement the physical layer of the network, while the access mechanisms is related to algorithms used to share the RF spectrum and avoid interference (data link layer). Finally, the choice of routing mechanisms is related to the need of ensuring reliable end-to-end communications within an environment with high delays and line-of-sight restrictions.

2.2.1 Network topology

The decisions regarding the network topology are related to physical configuration of the network assets. This translates to two coupled sub-problems: What is the accepted degree of inter-connectivity between network nodes? Where do we place communication assets and how do we configure them?

The first sub-problem leads to two architectural decisions than progressively determine the type of topology for the network:

- Fully connected vs. Meshed/Star: If the relay communication payloads are placed on-board the network customers then the network can be considered as a fully connected network. All nodes can talk to each other whenever there is line-of-sight, share information and route it from an origin to a destination. An example of this type of network in space is implemented by NASA to provide communications with the Mars surface. All scientific rovers carry a UHF transponder that connects to a relay transponder on-board the scientific orbiters around the planet, which in turn send the information back to Earth through a more capable X-band link. In contrast, communication payloads can also be placed on-board dedicated satellites that are specifically designed to provide communication services to other missions. This leads to either a meshed or star network and is the approach followed by most space networks nowadays (e.g. NASA's SN, ESA European Data Relay System (EDRS), or any of the commercial satellite communication providers).

- Meshed vs. Star: Assuming that relay communication payloads are placed on dedicated relay satellites, the next decision is related to using inter-satellite links. If that is the case, then the relay satellites can send information to one another thus creating a meshed network that minimizes points of failure. Without these inter-satellite links, all customers have to communicate to a relay satellite which is directly connected to the end terminal on the ground.

Once the topology of the network has been specified, the next step is to define the positions of both the space and ground segment. For the space part, and recalling that SNAT is intended to help design near Earth networks, this is equivalent to the constellation design problem, i.e. selecting the optimal orbital positions to place the relay assets. In turn, the design of the ground segment can be considered as a selection problem in which the goal is to pick the optimal subset of locations to place ground stations in order to maximize support to the relay satellites.

Finally, the last decision related to the network topology is tied to how communication payloads are allocated into relay satellites. For instance, if three payloads operating at different frequency bands are to be launched, is it better to put them all into a single spacecraft or should they be separated so that there is no interference between them. If that is the case, then each orbital position will not contain one but several satellites flying in formation.

2.2.2 Business model

The business model decision intends to capture the trades on how to financially sustain the provision of communication and navigation services for space missions. From a historical perspective, the default business model has been direct procurement of all network assets by the entity deploying the infrastructure. For instance, over the last 30 years NASA has bought and then privately operated all the assets of the Space Network - albeit through subcontractors in some cases. This fact has forced the agency to pay for the whole life cycle cost of all network assets.

Nevertheless, other alternatives are now becoming available. The most well known option is probably "hosted payloads", an approach that allows companies to put secondary payloads on-board satellites that they do not own in exchange for an economic compensation. The owner of the satellite commits itself to provide satellite resources (mass, power, volume from the bus) to that payload and launch it into orbit. As a result, the hosted payload does not need to incur in the cost of designing and manufacturing the satellite bus, nor does it have to directly procure the entire launch vehicle. However, it is still responsible for part of the integration and testing cost, as well as the operations cost of the hosted payload.

Despite the attractiveness of hosted payloads from a financial perspective, there are several challenges that hinder their suitability for architecting a space network. For instance, concerns have been raised on how unexpected failures in the host spacecraft would affect the hosted payload and the level of service it is providing. This is particularly critical if the hosted payload has to provide contingency communications to other spacecraft, thus requiring a high degree of availability and flexibility. On the other hand, since only a handful of missions have used them in the past, there is little expertise and quantification of the programmatic burden incurred when flying a hosted payload. This is also a problem when negotiating the legal and contractual agreements between the payload and satellite owner.

Finally, the third alternative to obtain communication and navigation services in the Earth vicinity is to take advantage of commercial providers that own a constellation of relay satellites. With this alternative an agency like NASA would not have to own any network asset, and instead would have to pay a fixed fee to use the transponders of the commercial operators. Two main drawbacks for these options can be envisioned: First, the commercial providers typically size their network to provide low (kbps) or moderate data rate links (tens of Mbps) to numerous customers. In contrast, a network that supports space missions is more prone to use a limited amount of links that can be configured to provide very high data rates (up to Gbps). Second, the type of information to send over the network might impose high confidentiality and integrity requirements that cannot be guaranteed by a third-party owned network.

2.2.3 Network technology

The network technology decisions are related to the configuration of the network and the choice of protocols used for providing the communication services. In that sense, the first decision to make is what frequency bands will be supported in the system. This choice has huge implications in the amount of information that can be transmitted since there are tight bandwidth allocations and limitations imposed by national and international regulatory organizations (e.g. Federal Communications Commission, National Telecommunications and Information Administration, International Telecommunication Union).

Once the frequency bands have been selected, the next decision is related to the available transponder technology. In particular, traditional systems rely on bent-pipe technology that can only process the RF signal at the analogical level to provide signal filtering, mixing and amplification. Alternatively, newer transponders are able to fully process incoming analog signals by demodulating them, interpreting the digital information at either the frame or the packet level, and finally re-modulate. Based on these differences, three main types of networks can be envisioned:

- Bent-pipe: The relays only process signals at the analogical level. When a transmission is received it is immediately re-transmitted to the next node with which continuous connection is already available.

- Circuit-switched: The relays are able to demodulate the signal and process digital information up to the data link layer. This option allows increasing the link performance since bit and frame error correction techniques can be utilized in the intermediate nodes. However, circuit-switched network do not have buffers to store information and, therefore, immediately re-transmit similar to how a bent-pipe transponder would do it.

- Store-and-forward: The relays can do full processing of the incoming signals, convert them into a bit stream and store the information locally if it is determined that the next hop is not available. They have to implement both routing and transport mechanisms to ensure that packets reach their end destination reliably. Additionally, they can also implement a *bundle layer* that offers reliable communications in high delay environments like space networks. If that is the case, then a Delay Tolerant Network is achieved.

Finally, the last two decisions are related to options regarding the access, routing and transport mechanisms. Since all communication channels within the network utilize the same spectrum bandwidth, access mechanisms are required in order to avoid interference between different customers. The classical approach to this problem in the space network community has been to use either scheduled systems or a combination of time, frequency or code division multiplexing (TDMA, FDMA or CDMA). Similarly, if the network under consideration is store and forward then it is necessary to select the appropriate routing and transport mechanisms to support its operation.

2.3 SNAT architectural decisions

Section 2.2 presented an overview and discussion of all high level decisions that are needed in order to architect a space communication network. Although SNAT aims at capturing the entire space of alternative designs, data and modeling limitations restrict its applicability by making the following assumptions:

- All communication payloads are placed in relay spacecraft effectively eliminating fully connected networks.
- The access media scheme will be based on a scheduled system. This assumption comes from the fact that all networks that provide service to space missions (SN, NEN, DSN, EDRS, and so on) utilize this type of access mechanism.
- The routing and transport mechanisms will be assumed to be ideal. In other words, if a customer connects to the network, then the information that is being sent will reach its destination seamlessly and reliably.

As a result, the set of available decisions to architect the system and their corresponding values are presented in table 2.1.

Decision	Range of values
Orbit selection	GEO, MEO or LEO
Constellation design	Number of planes: N_p Number of sats per plane: N_{sp}
Inter-Satellite Link payload allocation	Yes or no for each constellation of satellites
Transponder-to-spacecraft allocation (disaggregation)	All the possible partitions of N transponders into $1 \leq N_{sat} \leq N$ satellites
Ground station	Subset of White Sands, Guam and a new site
Contract modality	100% procurement, hosted payloads, or 100% commercial
Transponder selection	Any parabolic antenna supporting S-, X-, Ku-, Ka-band. Optical telescopes (1550nm)
Transponder technology selection	bent-pipe, circuit-switched, store-and-forward

Table 2.1: Architectural decisions

The orbit selection and constellation design are two coupled decisions that dictate the orbital positions where network assets will be placed. For instance, a particular architecture can be based on *GEO-1-3* constellation augmented by a *MEO-1-1* constellation. This means that three clusters of satellites will fly in geosynchronous orbit with a longitudinal separation of 120 deg. Additionally, another constellation of only cluster will fly in medium Earth orbit. For the geosynchronous case, the inclination of the orbit is assumed to be 0, while for MEO and LEO it becomes a variable that the user can specify.

Once the number of constellations and their shape has been defined, the next step is to select whether each of them will have an inter-satellite link (ISL). Each constellation is completely independent in that regard, thus enabling networks where the satellites in the *GEO-1-3* constellation carry ISLs, while the *MEO-1-1* satellite does not.

Similarly, the set of ground stations to support the space segment of the network has to be selected. This can be done from a sub-set of predefined ground stations. If the selected one already exists, this will be reflected in the system cost by setting its construction cost to 0. Alternatively, if the site has to be built, then both the construction and operation cost will be accounted in the life cycle cost estimate.

Next, the contract modality for the network is chosen. Similar to the ISL decision, each constellation can have a different contract modality. As a result, the *GEO-1-3* can be 100% procured while the augmentation transponder flying at *MEO-1-1* can be a hosted payload. The cost of the two constellations will be assessed independently, thus quantifying the cost of flying the augmentation system as hosted in a commercial satellite.

The next two decisions capture the alternatives in transponder selection and how to place it on orbit. The first choice is related to the supported transponders, what band they utilize and what are their nominal data rates. This decision is done at the constellation level, thus enabling networks where the *GEO-1-3* constellation is supporting RF communications, while the augmentation system *MEO-1-1* is based on an optical telescope.

Finally, the last decision is related to the transponder technology. This is done at the network level, so all the constellations will have the same alternative: bent-pipe, circuit-switched or store-and-forward.

2.4 Model overview

Figure 2-2 presents an overview of the high level structure of SNAT. Two main types of inputs are required in order to explore the space of architectures defined by the decisions presented in the previous section:

- *Tradespace definition*: It defines the subset of decisions that are selected to architect the network, along with their allowable values. For instance, in figure 2-2 only two decisions have been selected, transponder selection and antenna allocation. Four antennas have been defined, a TDRS-like antenna with a 5 meter dish and supporting communications at S, Ku and Ka-band. A fast and slow RF antenna, both with 5 meter dishes but one supporting high data rate communications at Ku and Ka-band, and the other one only supporting S-band.

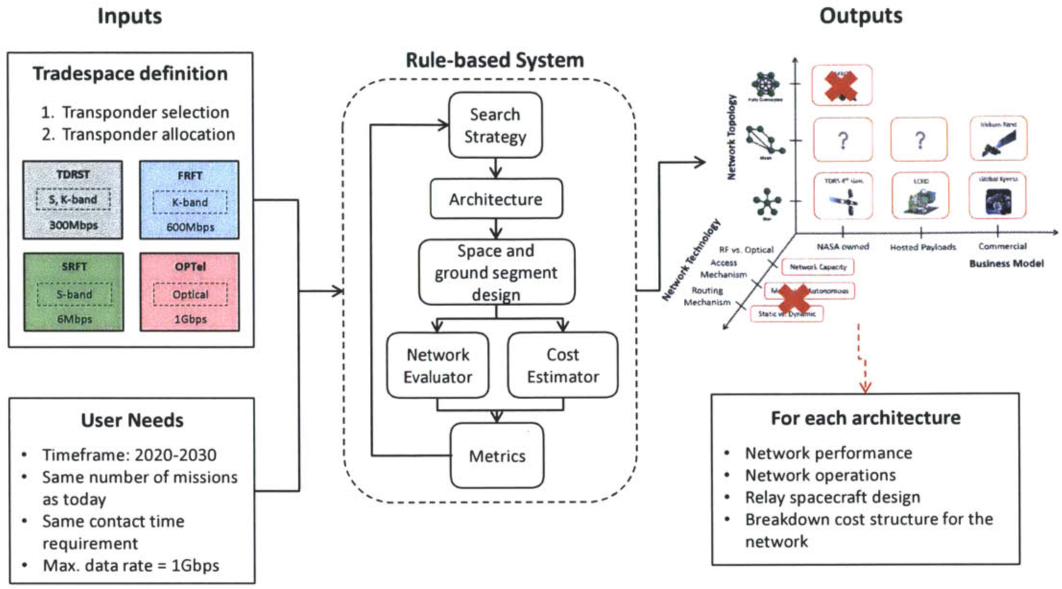


Figure 2-2: Model overview

And an optical telescope that can provide up to 1Gbps.

- *User needs*: It defines the set of missions that will be customers of the network along with their expected concept of operations (number of contacts per day, nominal contact data rate, etc.).

These inputs are run through the core of the model, a rule-based expert system that implements four main parts: A search strategy, i.e. an algorithm that iteratively looks for the optimal subset of architectures with respect to some objective metrics and identifies the best options for each of the architectural decisions. A space and ground segment design algorithm that sizes the different elements of the network (relay satellites, ground stations and their respective antennas). A network evaluator that simulates the operations of the system for a typical day of operations by producing a plausible schedule for the different network assets. And finally, a cost estimator module that, given the sized elements, costs the different parts of the space and ground segment and provides an estimate of the total life cycle cost.

The output of the model is a tradespace of architectures that have been evaluated to obtain both a normalized metric for their benefit and an estimate for their life

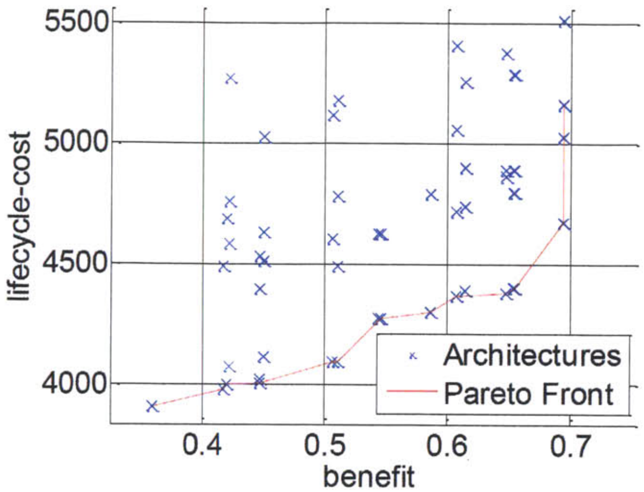


Figure 2-3: Example of a tradespace from SNAT

cycle cost. Figure 2-3 presents an example of the output of the tool in the form of a notional tradespace in the benefit-cost space, and highlights the resulting Pareto front. Each point in the plot represents one particular evaluated architecture, with sized communication payloads, relay satellites and their supporting ground stations.

2.5 Model description

This section presents a detailed description of the SNAT tool and the different modules that compose it. Most of the information herein presented is extracted from references [35] and [34]. Before delving into SNAT's modeling and implementation details, this section presents VASSAR, a methodology for Value Assessment in System Architecting using Rules, developed by Selva in reference [44]. SNAT is based on this framework in that it has been implemented using a rule-based expert system. Once VASSAR has been introduced, this section provides a detailed description of the modules used to size the space segment, ground segment, and network capacity. It also provides an overview of the cost module and search engine since both of them have been partially adapted from Selva's previous work.

2.5.1 The VASSAR framework

VASSAR is a methodology developed by Selva in order to assess the value of an architecture using expert knowledge. At the highest level, VASSAR accomplishes this goal through a three step process: First, the system architecture is decomposed from its architecture to the component level in order to understand its level of performance or capabilities. Second, these capabilities are combined and matched against the requirements from the stakeholders that will eventually benefit from the system. This matching process allows quantifying the satisfaction of each of the customer requirements. Finally, these satisfactions are aggregated into a single metric that summarizes the value of that particular architecture.

SNAT's high level structure is based in this same three step process. Given a network architecture, the different space and ground assets are appropriately placed with their respective communication transponders. Then, the network capabilities are computed using a rule-based scheduling system that grants contacts to the customers based on their priorities and compatibility with the system (e.g. frequency compatibility, line of sight considerations). The analysis of this schedule provides the metrics that define the customer quality of service and, in doing so, quantifies his satisfaction with respect to his stated requirements. Finally, the satisfaction of each customer is aggregated into a single value of network benefit using a weighted average approach.

Additionally, SNAT also benefits from Selva's work in the spacecraft design algorithm and cost estimation module. Both of them were included in his tool to architect Earth Observation Satellite Systems [44] and have been modified in order to tailor them to the specifics of space communication networks. Similarly, the genetic algorithm used to explore the space of alternative designs was first developed by Selva and has been adapted to the fit within the architectural decisions presented in section 2.3.

2.5.2 Space and ground segment design

The first step in order to evaluate a space communication network is to size the relay satellite communication payloads and antennas so that they can provide the data rates desired by the customers. This is accomplished through a two step process: First, the antennas of the relays are sized based on the network nominal data rates and available frequency and bandwidth allocations. Second, the mass of these antennas and accompanying electronics is estimated an inputted into a spacecraft design algorithm that sizes the bus of the satellite accordingly.

2.5.2.1 Antenna design module

SNAT differentiates between three different types of communication links: Relay-to-user links (RUL) are used to communicate relay satellites with the network customers. They have a fixed nominal data rate that is an input to the tool and has to be consistent with the expected technological limitations. On the other hand, the inter-satellite links (ISL) are links between relay satellites. Finally, space-to-ground links (SGL) are used to downlink the information from the relay spacecraft to the ground.

The main objective of the antenna design module is to size the ISL and SGL communication payloads and antennas based on the RULs that the spacecraft has to support. This is accomplished through an iterative process that progressively increases the frequency band of the ISLs and SGLs so that they can operate at the required data rate within the existing bandwidth limitations. In particular, the steps to size an ISL or SGL antenna are as follows:

1. Find the bands were transmission is allowed (typically S, X, Ku and Ka-band) and order them according to increasing bandwidth. Assign the optical band as the last option since it has no bandwidth restrictions.

2. Compute the nominal data rate R_b that the link has to support as a fraction of the total data rate provided to the customer (equation 2.1). The values of the correction factor α are presented in table 2.2. For bent-pipe and circuit

Transponder technology	ISL	SGL
Bent-pipe	0.8	1
Circuit switched	0.8	1
Store and forward	0.6	0.75

Table 2.2: Correction factors for ISL and SGL data rates

switched transponders, α can be viewed as a simultaneity factor, that is, the probability of having all the RULs at nominal capacity at the same time. For store and forward technology, α is a combination of the simultaneity factor and the multiplexing gain obtained when packetizing the information prior to its transmission.

$$R_b = \alpha \cdot \sum_{\forall RUL} R_{b_i} \tag{2.1}$$

3. Compute the transmit frequency f_c and bandwidth BW for the first band in the list created in (1). If $BW = 0$, then there is no bandwidth allocation for that particular type of link (e.g. NASA does not have frequency allocations at X-band for ISLs).

4. Compute the number of modulation levels M required to provide R_b with BW, assuming that a phase-shift keying (M-PSK) modulation is used.

5. If $M \leq 8$ then select this band for the link. Use the link budget equation to size the antenna of the relay satellite assuming that an ISL has identical antennas at both ends of the link. For SGLs, assume a 12 or 18 meter dish antenna depending on the architect inputs.

6. If $M > 8$ then jump to the next frequency band and go back to (3). If the current selected band is optical, then use it in conjunction with a Pulse Position Modulation with 16 levels (16-PPM) to size the diameter of the optical telescope.

Inherent to the previously presented algorithm are both technical choices and regulatory limitations. For instance, the choice of PSK modulations comes from the fact that the current Space Network uses both BPSK and QPSK [17]. SNAT assumes that in the future at least 8-PSK will be available. Similarly, the choice of 16-PPM as a baseline optical modulation is grounded on the optical technology that NASA is currently developing and demonstrating [20]. However, SNAT can easily integrate other technologies for link budget calculations should data their performance be available. Finally, the frequency and bandwidth allocations are based on the current NASA allocation for the SN [30] although other regulations can be easily encoded.

2.5.2.2 Spacecraft design algorithm

The spacecraft design module is an iterative module that provides a subsystem-level design of the spacecraft bus including a very rough configuration of the spacecraft from the payload requirements. It was originally developed by Selva in [44] and later adapted by him to the specifics of SNAT. The discussion herein presented was written by Selva as part of the co-authored reference [35] and is included here in order to provide a holistic view of the tool.

Figure 2-4 presents the high level structure of the iterative algorithm. Based on the communication payloads requirements, an initial guess for the spacecraft mass, power and dimensions is obtained. Then the different bus subsystems are progressively sized and compared to the initial guess until the algorithm has converged. The following subsections present the equations used to size the different bus subsystem. Most of them where extracted from [26].

2.5.2.2.1 power subsystem design The Electrical Power Subsystem (EPS) is designed based on a very basic power budget. The mass of the EPS is given by:

$$m_{EPS} = m_{SA} + m_{batt} + m_{other} \tag{2.2}$$

m_{SA} is the mass of the solar array, m_{batt} is the mass of the batteries, and m_{other} is the mass of the other electrical components (e.g. cables, regulators). The solar array is designed to provide enough power at end-of-life, assuming a certain yearly

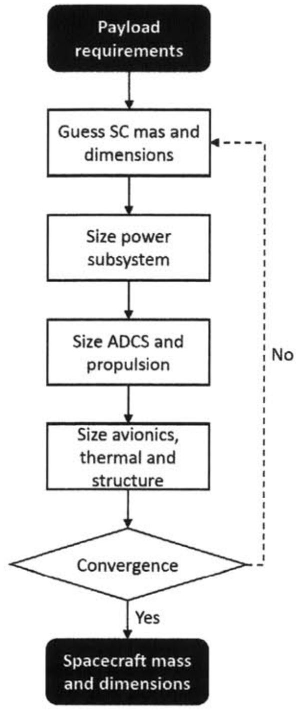


Figure 2-4: Spacecraft design algorithm

degradation $\eta(\%/yr)$. Its mass is calculated assuming a given specific power density $\rho_p(W/kg)$:

$$
\begin{aligned}
P_{SA}(W) &= \frac{P_e\frac{T_e}{X_e} + P_d\frac{T_d}{X_d}}{T_d} \\
W_{BOL}(\frac{W}{m^2}) &= W_0 I_d \cos\theta \\
A_{SA}(m^2) &= \frac{P_{SA}}{W_{BOL}(1-\eta)^t} \\
m_{SA}(kg) &= \frac{W_{BOL}A_{SA}}{\rho_p}
\end{aligned}
\tag{2.3}
$$

$T_e(s)$ is the average eclipse time per orbit, which is calculated from geometrical considerations, $T_d(s) = T - T_e$, $T(s)$ is the orbital period, $P_e(W)$ are the power requirements during eclipse, $P_d(W)$ are the power requirements during daylight, X_d and X_e are the energetic efficiencies between the solar array and the power bus (through the batteries in case of eclipse), $W_0(W/m^2)$ is the power density given by the solar array technology, I_d is an efficiency, θ is the Sun angle, $W_{BOL}(\frac{W}{m^2})$ is the power density of the solar array at BOL, and $A_{SA}(m^2)$ is the solar array area.

The mass of the batteries is calculated from its capacity assuming a certain specific energy $\rho_e(Wh/kg)$:

$$C_r(Wh) = \frac{P_e T_e}{3600 \cdot DOD \cdot n}$$
$$m_{batt}(kg) = \frac{C_r}{\rho_e} \tag{2.4}$$

DOD is the depth of discharge (which depends on the orbital parameters) and n is the efficiency from the batteries to the load. In particular, the DOD is assumed to be 0.8 for GEO, 0.6 for dawn-dusk SSO, and 0.4 for all other orbits.

The mass of the rest of components (regulators, converters, and wiring) is estimated as a function of the power at beginning of life P_{BOL} and the spacecraft dry mass m_{dry} as suggested in [11]:

$$m_{other} = \alpha P_{BOL} + \beta m_{dry} \tag{2.5}$$

α has a component of regulated power and a component of converted power, $P_{BOL} = W_{BOL} A_{SA}$ is the power available at BOL, β accounts for the EPS wiring, and m_{dry} is the spacecraft dry mass.

2.5.2.2.2 Delta-V and propellant mass budgets The design of the ADCS and propulsion subsystems is based on a rough ΔV budget of the spacecraft, which consists of four components: injection, drag compensation, ADCS, and de-orbiting:

$$\Delta V = \Delta V_{inj} + \Delta V_{drag} + \Delta V_{ADCS} + \Delta V_{deorbit} \tag{2.6}$$

The ΔV_{inj} is computed assuming that the spacecraft is injected into a transfer orbit that has the perigee at 150km and the apogee at the final orbit altitude:

Orbit	$\Delta V_{drag}(m/s/yr)$
LEO($h < 500km$)	12
LEO($500km < h < 600km$)	5
LEO($600km < h < 1000km$)	2
MEO	0
GEO	0
HEO	0

Table 2.3: ΔV required to compensate drag for different orbits

ADCS configuration	$\Delta V_{ADCS}(m/s/yr)$
Three-axis	20
Spinner	0
Gravity gradient	0

Table 2.4: ΔV required for ADCS

$$V(r_p, r_a, r) = \sqrt{2\mu(\frac{1}{r} - \frac{1}{r_p + r_a})}$$
$$\Delta V(r_{p1}, r_{a1}, r_{p2}, r_{a2}, r) =$$
$$= |V(r_{p2}, r_{a2}, r) - V(r_{p1}, r_{a1}, r)|$$
$$\Delta V_{inj} = \Delta V(R_E + 150km, r, r, r, r) \tag{2.7}$$

$\Delta V(r_{p1}, r_{a1}, r_{p2}, r_{a2}, r)$ is the ΔV necessary to perform a change of orbit semimajor axis from orbit (r_{p1}, r_{a1}) to orbit (r_{p2}, r_{a2}) when the spacecraft is at distance r from the Earth.

The ΔV_{drag} necessary to compensate drag is strongly dependent on orbit altitude. The values shown in table 2.3 were taken from [40] and used as first order approximations for SNAT. Similarly, the ΔV_{ADCS} required for ADCS depends on the type of ADCS system that the satellite is using (see table 2.4). These values were adapted from [40] and also used in SNAT as first order approximations.

Finally, the $\Delta V_{deorbit}$ is computed assuming that LEO spacecraft are de-orbited using atmospheric drag, and all other spacecraft are de-orbited using solar radiation pressure. For drag-based de-orbiting, the $\Delta V_{deorbit}$ is computed based on a change of semi-major axis from the current circular orbit to an elliptical orbit that has the perigee at 0km and the apogee at the orbit altitude:

$$\Delta V_{deorbit,drag} = \Delta V(r, r, R_E, r, r) \tag{2.8}$$

For solar radiation pressure-based de-orbiting, the $\Delta V_{deorbit}$ is computed based on a change of semi-major axis from the current circular orbit to an elliptical orbit that has the same perigee and a slightly higher apogee:

$$\begin{aligned} \Delta V_{deorbit,SRP} &= \Delta V(r, r, r, r + \Delta h, r) \\ \Delta h &= 200km + 35km + \frac{1000 C_R A}{m_{dry}} km \end{aligned} \tag{2.9}$$

The 200km are due to the geosynchronous belt restricted zone, the 35km are to allow for gravitational perturbations, and the remaining margin depends on the magnitude of the effect of solar radiation pressure on the spacecraft (the larger the effect, the larger the margin); C_R is the solar radiation pressure coefficient, and A is the surface area of the spacecraft.

Once the ΔV has been calculated, it is possible to compute the propellant mass required to satisfy this ΔV budget. The tool assumes that ΔV_{inj} is performed by the apogee kick motor (AKM), while the other ΔV are performed by the ADCS subsystem. For each of these propulsion systems, the propellant mass can be computed using the rocket equation:

$$\Delta V_j = g I_{sp,j} \log \frac{m_i}{m_f} \tag{2.10}$$

I_{sp} is the propelllant specific impulse, which can be different for the AKM and the ADCS subsystem, m_i is the initial mass with propellant and m_f is the final mass without the propellant.

2.5.2.2.3 Attitude Determination and Control and Propulsion Subsystem The mass of the ADCS is mostly given by the mass of the sensors and the mass of the actuators. The mass of the sensors is driven by the attitude knowledge accuracy requirement acc (equation 2.11), while the actuators are sized to satisfy by the momentum storage h required (equation 2.12):

$$m_{sen} = 10acc^{-0.316} \tag{2.11}$$

$$m_{act} = 1.5h^{0.6} \tag{2.12}$$

Note that acc can vary depending on the architecture, as the pointing requirements of a high gain antenna, or an optical payload, are very different from those of a low gain antenna. On the other hand, the momentum storage h is assumed to be sized to counter the different disturbance torques produced by atmospheric drag, gravity gradient, solar radiation pressure and the Earth's magnetic field. Expressions for these disturbance torques were taken from [40].

In addition to sensors and actuators, the ADCS has additional mass that can be estimated as a fix fraction of the spacecraft dry mass. This results in the following total mass estimate:

$$m_{ADCS} = 3m_{sen} + 4m_{act} + 0.01m_{dry} \tag{2.13}$$

In turn, the mass of the propulsion subsystem is based on the AKM propellant mass estimate, assuming a certain corrective mass fraction:

$$m_{AKM} = \frac{(1-\mu)}{\mu} m_{prop,inj} \tag{2.14}$$

2.5.2.2.4 Thermal, avionics, and structure subsystems The thermal, avionics, and structure subystems are designed using simple parametrics of the form

Subsystem	k
Thermal	0.0607
Avionics	0.0983
Structure	0.5462

Table 2.5: CERs coefficients for thermal, avionics, and structure subsystem

$m_{subsystem} = k \cdot m_{payload}$. The constants k that are used for each subsystem are summarized in table 2.5. Furthermore, the mass of the launch adapter is computed as $m_{LA} = 0.01 \cdot m_{dry}$ and added to the mass of the spacecraft.

2.5.2.2.5 Algorithm convergence criteria After the first iteration, the dry and wet mass of the spacecraft are updated. Spacecraft dimensions are estimated assuming a perfect cube of 100 kg/m^3. The mass and dimensions of the solar panels are taken into account to update the inertial properties of the spacecraft, as illustrated in equation 2.15:

$$
\begin{aligned}
L_A &= 1.5s + 0.5\sqrt{\frac{A_a}{2}} \\
I_z &= 0.01 m_{dry} \\
I_x = I_y &= I_z + L_a^2 M_a
\end{aligned}
\tag{2.15}
$$

Finally, the convergence criterion used in SNAT is described in equation 2.16, where the subscript i indicates the iteration number.

$$
|m_{dry}^{(i+1)} - m_{dry}^{(i)}| < 10kg \tag{2.16}
$$

2.5.3 Network evaluator

The performance model or network evaluator intends to capture the ability of the network to satisfy a set of stakeholders given their relative importance and expected missions. In particular, it has been designed to satisfy two purposes: First, properly simulate the operations of the network over a typical day of operations; and second, effectively capture how value is delivered to each stakeholder by understanding how the performance of the network compares to the customer requirements.

2.5.3.1 Modeling the network topology

The first step to model a space network is to understand the movement of both relay satellites and end users over a representative time period. This information is needed to define the contact opportunities between them and therefore assess when a particular path between an origin and a destination is available.

The performance model uses STK [41] in order to simulate the movement of both the space and ground segment. For each constellation, the following parameters can be inputted:

1) Constellation design: number of planes, number of satellites per plane.

2) Orbit design: altitude, eccentricity, inclination, argument of perigee.

3) STK database: a list of satellite identifiers on the STK database ($SatId_1$, ..., $SatId_N$).

If (3) is specified then STK creates a constellation of satellites by directly importing their orbital information from the database. Otherwise, STK combines (1) and (2) to create a simplified constellation of equally spaced spacecraft both in latitude and longitude. The same process is followed in order to add the set of desired ground stations. In this case, one can input their Earth coordinates or their STK database identifier.

Once all the nodes of the network have been put in place, the next step is to define the field of view (FOV) of their antennas. This is done as a post-processing step in which the elevation angle (with respect to the nadir direction) for a contact is used to trim it accordingly. Furthermore, choosing between the nadir and zenith orientation is generally not trivial and depends on the relative position of the two communicating nodes. The model solves that problem by letting STK compute the four possible combinations (two for the transmitting antenna and two more for the receiving antenna) and automatically selecting the best one.

The output of the STK simulation is a set of reports that contain all the visibility windows between different customers of the network and the relay satellites. These visibility windows only take into consideration line-of-sight and field-of-view limitations. Additionally, STK is also used to estimate the maximum coverage gap between a relay satellite and its supporting ground station. This information is then used to size the On-board and Data Handling (OBDH) subsystem for relay satellites that contain store-and-forward technology.

On the other hand, the STK simulation is used to compute *paths* between relay satellites and their supporting ground stations, with and without inter-satellite links as intermediate hops. These *paths* are used as a proxy to differentiate the operation of the network when inter-satellite links are included and when store-and-forward technology is present. In particular, three possible combinations are possible:

1. The network uses bent-pipe or circuit-switched technology and does not have inter-satellite links: A contact opportunity between a customer and a relay spacecraft is valid only if that relay spacecraft is in view and connected to a ground station.

2. The network uses bent-pipe or circuit-switched technology and also has inter-satellite links. A contact opportunity between a customer and a relay spacecraft is valid only if that relay spacecraft is in view to either a ground station or another relay spacecraft that can redirect the information to the ground.

3. The network uses store-and-forward technology: A contact opportunity between a customer and a relay spacecraft is always valid. If the relay spacecraft is not in view of a ground station or relay satellite, it will store the information locally until that happens.

Based on this three cases, the *paths* are used to trim the contact opportunities (from now on *visibility windows*) accordingly. For instance, if the network operates under the conditions specified in option (1), then all relay satellites must be in view with a ground station to be able to grant contacts from the network customers.

2.5.3.2 Modeling the network schedule

Once the contact opportunities between the different customers and assets of the network have been computed, the next step is to utilize them during the scheduling process. This will result in a proxy for the operational performance of the network, i.e. what relays satellites are used to provide service to what satellites and how much information is sent over each contact.

2.5.3.2.1 Assumptions The first and most important assumption for the proposed rule-based scheduling algorithm is that of non-optimality. This assumption is grounded in two facts: First, the scheduling process is done following a greedy approach, i.e. contacts are scheduled whenever it is possible regardless of future states of the network. Second, when more than one network resource can serve a contact, deciding which one to chose is done based on a set of heuristic rules.

On the other hand, it is also important to note that the proposed scheduling algorithm only addresses the problem of the access network, that is, the allocation of resources between network customers and network nodes. This simplification is done in order to facilitate the scheduling process and decouple it from the intricacies of the routing algorithm implemented in the backbone network.

Finally, the third assumption for the scheduling algorithm is related to its inputs. It is assumed that all the required information will be available at the start of the process and no numeric computations (e.g. propagating satellites) will have to be done inline with the scheduling process. This fact is further explored in the following subsection.

2.5.3.2.2 High-level Structure of the Scheduling Algorithm

As previously mentioned, the inputs of the scheduling algorithm come from (1) the set of users that want to connect with the network and their relative priorities and (2) the *visibility windows* between them and the relay satellites. With these available, the scheduling algorithm is ready to start allocating network resources to customers.

The first step in the scheduling process is to decide the order with which requested contacts by the customers (from now on *required contact*) will be served. In other words, contacts are prioritized according to their importance.

Next, the scheduling algorithm compares and assigns the *visibility windows* to the *required contacts* until there are no more contacts to schedule or there are no more free resources on the network. In order for a *visibility window* to be assigned to a *required contact*, the following criteria are used:

1. Frequency band compatibility: A visibility window is suitable for scheduling a contact if it uses the same frequency band (or optical wavelength) than the customer mission.

2. Data rate: A visibility window is suitable for scheduling a contact if the nominal data rate it can support is higher than the requested data rate by the customer mission.

3. Duration: A visibility window is suitable for scheduling a contact if its duration is longer than that of the contact, or if it can be concatenated with other visibility windows to achieve an overall duration longer than that of the contact.

4. Network status: A visibility window is suitable for scheduling a contact only if its state (starting and ending times) has been updated based on the past scheduled contacts.

Note that (1) and (2) are not tied to the physical geometry of the system but rather to the communication payloads that are carried by the customers and relay satellites. Alternatively, (3) depends primarily on the satellites' motion and captures the idea that the system performs handovers as the customer loses support from a relay and then connects to another one. The specific heuristics used for selecting the best visibility window and performing handovers is explained later in this section. Lastly, (4) takes into account the evolution of the network resources as contacts are progressively scheduled. In particular, once a contact is allocated it is mandatory to update the status of the visibility windows so that they are accordingly trimmed or eliminated.

2.5.3.2.3 Data Structures Based on the description from the previous subsection, four main data structures are required in order to capture the state of the network and perform the scheduling process: *visibility-window*, *path*, *required-contact* and *customer*. In a rule-based expert system these data structures come in the form of *facts*, i.e. a container of any sort of information. As an example, a fact *car* can contain multiple car properties (slots) such as color, mass, power or make.

Table 2.6 lists the four main facts of the scheduling algorithm along with their slots. It can be seen that the *visibility-windows* and *paths* have a similar structure: an origin node (a customer for the first and a satellite for the second), a destination node (a satellite vs. a ground-station), a starting time and an ending time. Note also that they have the slot "orbital-position" that is used to pre-compute these facts. The *visibility-windows* also include information regarding the band and nominal-data-rate so that it is easily accessible when determining their suitability to service a particular contact. Similarly, a *path* has a slot "ISL" that is used as proxy for multi-hop or single-hop contact between a relay and a ground station. As previously stated, if the architecture under evaluation does not support ISLs, then the contact opportunities will be trimmed according to paths computed without ISLs.

On the other hand, the facts *required-contact* contain two different types of data: information regarding the characterization of the contact such as the user requesting it, the service, expected duration, priority, frequency band and nominal data rates. On the other hand, it also contains information on how well the contact has been served by the network.

Finally, the *customers* facts keep track of the temporal evolution of the customer satellites during the scheduling process. All satellites start with their slot "time" at 0

Fact Name	Fact Slots
Visibility-Window	id, customer, satellite, orbital-position, tStart, tEnd, band, data-rate
Path	id, satellite, orbital-position, ground-station, ISL, tStart, tEnd
Required-Contact	id, user, service, priority, duration, band, nominal-data-rate, satisfied, tStart, tEnd, data-volume
Customer	id, service, time

Table 2.6: Scheduling Algorithm Facts

and progressively increase its value as contacts get scheduled. If at a particular instant in time the network cannot provide access to the customer, then the "time" slot is increased $n \cdot \Delta t, n > 0$ until the access can be granted or the scheduling algorithm reaches its simulation end. Note also that this process is repeated for each customer and service, i.e. a satellite will be able to independently schedule contacts used for telemetry and science data return.

2.5.3.2.4 Prioritization As previously stated, in order to build a schedule that maximizes the system benefit it is important to prioritize the customers and the services they require. In this case, this is done through the following heuristic:

$$Priority = M \cdot mission\ class + service\ priority \tag{2.17}$$

The *mission class* is used to capture the idea that some missions will always have higher priority due to specific circumstances. For instance, a human space flight mission will always have more priority than any robotic mission since astronaut lives are at stake. Alternatively, *service priority* captures the importance of the mission and service in the stakeholder-objective-mission-service decomposition used to compute the benefit of the system. This metric is included so that the computed scheduled is maximizing the benefit score of the network. Finally, M is a control parameter that captures the relative importance between *mission class* and *service priority*. It is typically set to high value (e.g., 10000) to ensure that *mission class* is always the most important factor.

2.5.3.2.5 Visibility window selection and concatenation Let's assume that a *required contact* has been selected as the next one to be served based on the prioritization. When tackling the problem of scheduling it based on the available *visibility windows*, three questions arise:

1. Are there any restrictions that limit the time at which the contact should be scheduled?
2. If there is more than one visibility window that can completely overlap a contact, which one has to be chosen?
3. If there is more than one visibility window that can partially overlap a contact and none that can fully serve it, which one has to be chosen for concatenation?

The answer to the first question is, in most cases, yes. Missions tend to prefer scheduling contacts when they have significant amounts of information to dump, or when their batteries are almost fully charged. Even operational constraints might arise, with contacts being scheduled preferably during the normal working hours of the staff at the mission control center. As a result, most schedulers allow customers to provide their desired contact times along with a flexibility window over which the event should be preferably scheduled [1].

Despite the importance of these factors in real-life operations of satellite communication networks, they are currently disregarded in the current scheduling algorithm. The rationale for this is as twofold: On the one hand, the goal of the scheduler is to assess the overall capacity of network and capture the effect having extra satellites and communication payloads in orbit. The authors realize that having extra scheduling constraints reduces the flexibility of allocating the network resources, thus potentially reducing the overall data flows through the network. However, they also assume this fact to be a second order effect compared to extra amount of resources that become available if more relay satellites are put into orbit. On the other hand, the scheduling process for the tool is based on a demand forecast beyond the 2020 time frame. This implies high levels of uncertainty both on the number and types of missions that will be flying, as well as on the operational aspects of the services they will demand. Therefore, taking into account scheduling constraints will be of limited usefulness since they will typically be unknown and difficult to predict from existing missions.

Alternatively, answering questions (2) and (3) is a complex problem that requires taking into account the duration of the *visibility windows* and how they can be concatenated in order to provide long periods of mission support. In particular, the following distinction is made: A *visibility window* is said to *completely overlap* with a contact if its duration d_{wnd} satisfies equation 2.18, where d_{con} is the contact duration. If this inequality does not hold, then the visibility window is said to *partially overlap* with the contact and requires concatenating multiple visibility windows until complete overlap is achieved. This happens when equation 2.19 becomes true, where d'_{wnd_i} represents the duration of the i-th concatenated *visibility window*, trimmed accordingly with of the previous selected windows $d'_{wnd_{j-1}}$.

$$d_{wnd} >= d_{con} \tag{2.18}$$

$$\sum_{i=1}^{n} d'_{wnd_i} >= d_{con} \tag{2.19}$$

Distinguishing between *partial* and *complete overlap* is important because it provides the condition to identify when a *required contact* can be immediately fully satisfied served versus the problem of using handovers to provide continuous support. As it will now be explained, these two situations entail different heuristics.

2.5.3.2.6 Partial vs. complete overlap heuristics Having identified the difference between *partial* and *complete overlap*, the next step is to identify the temporal conditions that characterize both situations.

Let t_0 be the time at which a customer wants to schedule a contact, d the duration of the contact, and t_{s_i} and t_{e_i} the starting and ending times of the i-th visibility window. Then, a *partial overlap* exists if equation 2.20 holds, while *complete overlap* requires equations 2.20 and 2.21 to be true at the same time.

$$t_{s_i} \leq t_0 < t_{e_i} \tag{2.20}$$

$$d \leq t_{e_i} - t_0 \tag{2.21}$$

Note also that equation 2.20 is a necessary but not sufficient condition for *partial overlap*. That being the case, two strategies can be used to identify a *partial overlap*: Look for all the visibility windows that satisfy condition 2.20 and dissatisfy condition 2.21; or eliminate all visibility windows that satisfy conditions 2.20 and 2.21 at the same time.

Once the conditions for *complete* and *partial overlap* are understood, the final and most important step is to define the heuristics that will chose the best visibility window among the set that satisfy equations 2.20 and 2.21 (dissatisfy in the latter case). For *complete* overlap, the following heuristic will be used: *The best visibility window is the one that minimizes the non-overlapping time with the contact.* This is mathematically expressed in equation 2.22, where $(t_0 - t_{s_i})$ expresses the non-overlapping time between the start of the visibility window and the start of the contact, and $(t_{e_i} - (t_0 + d))$ accounts for the non-overlapping time between the end of the visibility window and the end of the contact.

$$\begin{array}{c}(t_0 - t_{s_i}) + (t_{e_i} - (t_0 + d)) < \\ \left(t_0 - t_{s_j}\right) + \left(t_{e_j} - (t_0 + d)\right)\end{array} \qquad \forall j, i \neq j \tag{2.22}$$

Simplifying equation 2.22 is a straightforward process that leads to equation 2.23, which simply states that the best visibility window is the shortest one. Recall, however, that equation 2.21 must also be true for *complete overlap*, thus indicating that the selected window will be the one with duration closest to the contact duration.

Table 2.7: Scheduling Algorithm Heuristics

Heuristic	Conditions	Outputs
Prioritization	$\exists$ *req-con* $\mid$ priority = *null*	priority = M $\cdot$ *mission class* + *service priority*
Complete overlap	$\exists$ *vis-wnd* $i \mid t_{s_j} \leq t_0 < t_{e_j}$ AND $d \leq t_{e_j} - t_0$ $\nexists$ *vis-wnd* $j \mid t_{s_i} \leq t_0 < t_{e_i}$ AND $d \leq t_{e_i} - t_0$ AND $t_{e_j} - t_{s_j} < t_{e_i} - t_{s_i}$ $i \neq j$	Select i-th vis-wnd
Partial overlap	$\nexists$ *vis-wnd* $k \mid t_{s_k} \leq t_0 < t_{e_k}$ AND $d \leq t_{e_k} - t_0$ $\exists$ *vis-wnd* $i \mid t_{s_j} \leq t_0 < t_{e_j}$ $\nexists$ *vis-wnd* $j \mid t_{s_i} \leq t_0 < t_{e_i}$ AND $t_{s_i} - t_{s_j} < t_{e_j} - t_{e_i}$ $i \neq j$	Select i-th vis-wnd

$$t_{e_i} - t_{s_i} < t_{e_j} - t_{s_j} \forall j, i \neq j \tag{2.23}$$

The other heuristic that needs to be explored is the one used to select the visibility window to concatenate when the contact is under a *partial overlap* situation. In this case, the following heuristic will be used: *The best visibility window is the one with longer duration and shorter non-overlapping time between the start of the contact and the start of the visibility window.* This is mathematically expressed in equation 2.24, where $((t_0 + d) - t_{e_i})$ is the time between the ending of the contact and the ending of the visibility window, and $(t_0 - t_{s_i})$ is the non-overlapping time between the start of the contact and the start of the visibility window. As previously demonstrated with equations 2.22 and 2.23, equation 2.24 can be simplified to more compact and manageable form - equation 2.25.

$$\begin{matrix} ((t_0 + d) - t_{e_i}) + (t_0 - t_{s_i}) < \\ ((t_0 + d) - t_{e_j}) + (t_0 - t_{s_j}) \end{matrix} \quad \forall j, i \neq j \tag{2.24}$$

$$t_{s_j} - t_{s_i} < t_{e_i} - t_{e_j} \forall j, i \neq j \tag{2.25}$$

Table 2.7 summarizes all the heuristics presented in this section and formulates them in the form of *if-then* statements that can be directly input into the rule-based scheduling algorithm.

2.5.3.3 Computing the network benefit

Once the network schedule has been obtained, it can be analyzed in order to understand the service provided to the customers with respect to their communication requirements. In particular, two metrics are used, number of granted contacts and latency.

The number of granted contacts indicates how many times the customer mission connects with a relay satellite to download and/or upload information. Since all contacts for a particular service and mission have the same nominal data rate and duration, computing the number of granted contacts is effectively equivalent to computing the fraction of the data volume returned to Earth (scaled by a factor).

$$\% \text{ returned DV} = \frac{DV_{ret}}{DV_{des}} = \frac{R_b \cdot T_c \cdot \#GC}{R_b \cdot T_c \cdot \#DC} = \frac{\#GC}{\#DC} \tag{2.26}$$

where DV_{ret} is the returned data volume, DV_{tot} is the total data volume to return, R_b is the contact nominal data rate, T_c is the contact duration, $\#GC$ is the number of granted contacts and $\#DC$ is the number of desired contacts.

On the other hand, the maximum latency computes the maximum amount of time that a mission will have to hold its collected data in the on-board memory before being able to download it. This metric is important in some applications where the scientific information from the satellites is used to obtain periodic data products (e.g., images from weather satellites might be needed every three hours to update weather predictions and issue alerts if necessary). Similarly, this metric can be important for mission controllers that do not want satellite operating without supervision for long periods of time. As the worst case estimate, this maximum latency is computed as the maximum time between consecutive contacts.

Once these three metrics have been computed, the satisfaction of a particular service can be directly assessed by comparing their values to a set of step functions that capture each mission preferences. These step functions transform the metrics to a normalized dimensionless measure of benefit for that particular type of service. For instance, a human space flight mission might heavily value not dropping contacts because continuous communication with the astronauts is essential to the mission safety. Therefore, if a contact is lost then its satisfaction might be 20% (in a normalized 0 to 100% scale) or even almost null. Alternatively, a small astrophysics mission might not consider critical to have all its desired contacts scheduled since the scientific data it gathers is not time sensitive. Thus, if a contact is lost then only a 5% of the satisfaction is lost.

Based on the metric normalized benefit, the satisfaction of a particular service can be computed using a weighted sum approach. Similarly, the satisfaction of the customer can be assessed by aggregating the satisfaction of the difference services it requests. This process can be repeated for a set of customers that belong to a same class or stakeholder, and finally, the satisfaction of all the network stakeholders can be averaged into a single figure of benefit for the entire architecture.

The result of this process is illustrated in figure 2-5, where, the satisfaction of the entire architecture is 0.87242. Three main stakeholders for the network have been identified, namely NASA, NOAA and the USGS. Within NASA, four different communities are expected to request services from the network, Earth Observation missions, Astrophysics and Heliophysics missions, and Human Space Flight missions. Each of them has a normalized satisfaction that comes from aggregating the satisfaction of all the missions under this community.

2.5.4 Cost estimation

SNAT's cost estimation module was developed by Selva as an adaptation from his prior work in Earth Observation systems [44]. It was tailored to SNAT so that it can provide an estimate of the life cycle cost of a network architecture (i.e., a set of constellations and ground stations) that is good enough for relative comparisons between architectures. This includes, in particular, differentiating between different contract modalities (procurement vs. hosted payloads vs. 100% commercial).

The life cycle cost estimate consists of several parts: transponder cost (antenna

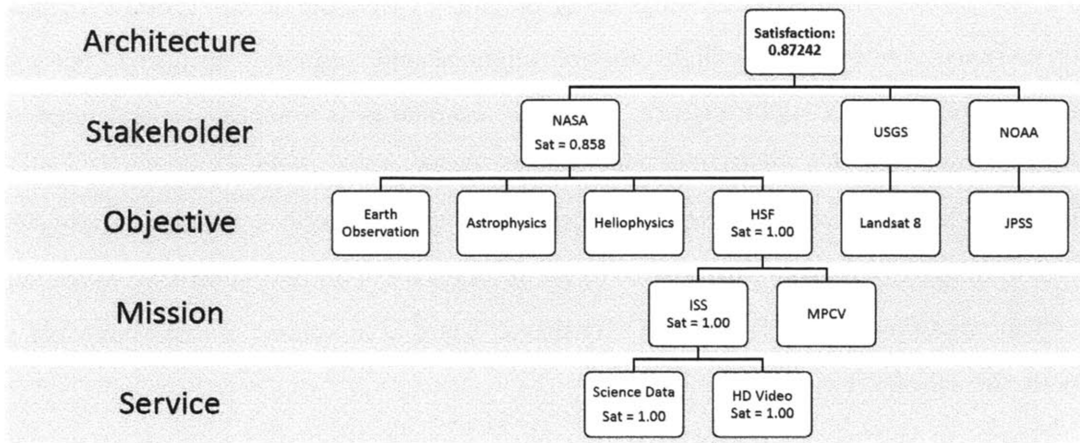


Figure 2-5: Architecture satisfaction decomposition

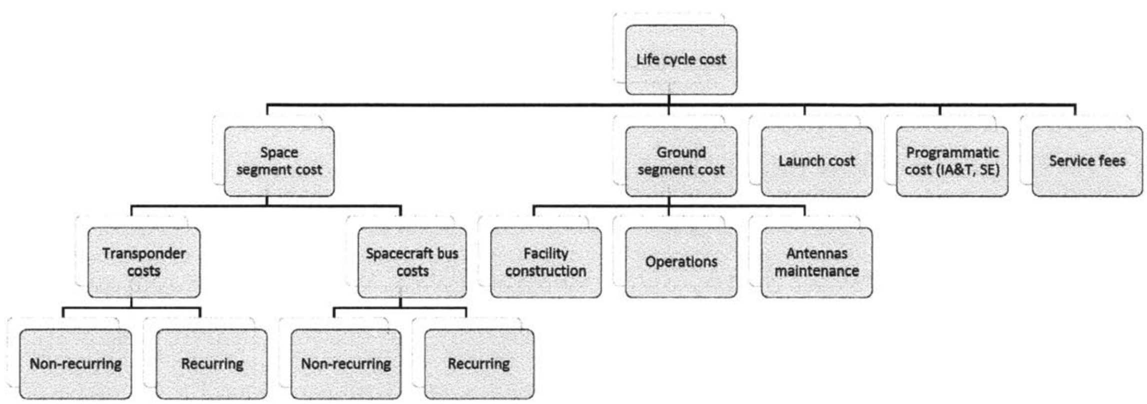


Figure 2-6: Lifecycle cost breakdown

plus electronics), bus cost, launch cost, IA&T cost, operations cost, and program overhead. Some of these are further divided into non-recurring and recurring costs, as illustrated in figure 2-6.

2.5.4.1 Payload Cost

Payload cost is only incurred when the contract modality is procurement or hosted payloads. If a 100% commercial approach is taken, payload cost is set to zero, as it is included in the service fee charged by the commercial service provider. On the other hand, payload cost is the sum of a non-recurring cost and a recurring cost. When these values are not provided by the user, they are estimated using CERs that utilize payload mass m and number of communication channels n as independent variables. The CERs are taken from the USCM8 model [4] and are provided below. All values are in FY2010$k.

$$C_{payl,NR} = 339m + 5127n \tag{2.27}$$

$$C_{payl,R} = 189m \tag{2.28}$$

$C_{payl,NR}$ is the development cost, including the cost of fabricating a qualification unit, and $C_{payl,R}$ is the cost of fabricating the first flight unit. The standard error of the estimate (SEE) of equation 2.27 inside the domain $160 - 395kg$ and 2-32 channels is 40%. In turn, the SEE of equation 2.28 inside the domain $38 - 928kg$ is 28%. All SEEs are corrected for the number of degrees of freedom. Total cost for development and fabrication of N identical payloads is thus given by equation 2.29 where $b < 1$ is chosen to model a cumulative average learning curve of 95%.

$$C_{payl} = C_{payl,NR} + N^b C_{payl,R} \tag{2.29}$$

2.5.4.2 Bus Cost

Bus cost is only incurred when the contract modality is procurement. If a 100% commercial approach or hosted payloads approach are taken, bus cost is set to zero and the respective cost is included in the service fee charged by the commercial service provider.

Similar to the costing of payloads, bus cost is the sum of a non-recurring cost and a recurring cost. When these values are not provided by the user, they are estimated using CERs that utilize subsystem mass as the independent variable. The CERs are taken from the USCM8 model [4] and are provided below. All values are in FY2010$k.

$$C_{bus,NR} = 110.2m_{dry} \tag{2.30}$$

$$C_{bus,R} = 289.5m_{dry}^{0.716} \tag{2.31}$$

m_{dry} is the satellite dry mass, $C_{bus,NR}$ is the development cost, including the cost of fabricating a qualification unit, and $C_{payl,R}$ is the cost of fabricating the first flight unit. The standard error of the estimate (SEE) of equation 2.30 inside the domain $114 - 5127kg$ is 47%. The SEE of equation 2.31 inside the domain $288 - 7398kg$ is 21%.

Finally, the total cost for development and fabrication of N identical buses is computed as illustrated in equation 2.29, with a learning factor of 95%. Note that the computation of bus cost depends on the dry mass of the spacecraft. This can be provided by the user, or it can be estimated by the spacecraft design algorithm.

2.5.4.3 Launch Cost

Launch cost is only incurred when the contract modality is procurement. If a 100% commercial approach or hosted payloads approach are taken, launch cost is set to zero, as it is included in the service fee charged by the commercial service provider. Launch cost is given by the sum of the costs of launching all constellations in the architecture. Computing the launch cost for a constellation is based on the assumption that, given the number of planes P and satellites per plane S, $P < N_L < PS$ launches are necessary to launch the constellation. In other words, at least one launch per plane and at most one launch per satellite are required to fly the entire fleet of satellites. The exact value of N_L is obtained by taking into account both launcher performance and orbital considerations.

In particular, SNAT contains a database of launchers that are available for each constellation. An extract of this database is shown in table 2.8. Note that the data concerning performance is provided in terms of the 3 coefficients of a quadratic function of the orbit altitude. In other words, if the entry of the table for a certain

	Atlas-V	Delta-7920
Payload GTO	$[10^4,0,0]$	$[5 \cdot 10^2,0,0]$
Payload LEO polar	$[15 \cdot 10^3,-4 \cdot 10^{-2},0]$	$[4 \cdot 10^3,-1 \cdot 10^{-2},7 \cdot 10^{-5}]$
Diameter (m)	4.8	2.7
Height (m)	10.0	7.53
Cost (FY2010$M)	100	65

Table 2.8: Extract of launch vehicle database

orbit type (e.g., LEO polar) gives the coefficients $[a, b, c]$, then the performance at altitude h can be computed as:

$$perf(h) = a + bh + ch^2 \tag{2.32}$$

Given these data, N_L is determined as follows:

$$N_L = \max\{N_{L,mass}, N_{L,vol}, N_{L,dim}\} \tag{2.33}$$

where $N_{L,mass}$ is the minimum number of launches required given the total spacecraft mass and the performance of the launch vehicle to the desired orbit; $N_{L,vol}$ is the minimum number of launches required given the total spacecraft volume and the volume of the launch vehicle; $N_{L,dim}$ is the minimum number of launches required given the sum of the maximum dimension of all spacecraft and the height of the launch vehicle:

$$N_{L,mass} = \left\lceil \frac{\sum_{i=1}^{N_{S/C}} m_{wet,i}}{perf(lv, orbit)} \right\rceil \tag{2.34}$$

$$N_{L,vol} = \left\lceil \frac{\sum_{i=1}^{N_{S/C}} vol_i}{vol_{lv}} \right\rceil \tag{2.35}$$

$$N_{L,diam} = \left\lceil \frac{\sum_{i=1}^{N_{S/C}} dmax_i}{h_{lv}} \right\rceil \tag{2.36}$$

In equations 2.34, 2.35 and 2.36 $N_{S/C}$ is the number of spacecraft in the constellation, $m_{wet,i}$, vol_i, and $dmax_i$ are the wet mass, volume, and maximum dimension respectively of spacecraft i. In turn, $perf(lv, orbit)$, vol_{lv}, and h_{lv} are the launcher performance at the desired orbit, and its maximum fairing volume and height. Once the number of launches has been computed, launch cost is simply given by the product of number of launches and individual launch cost.

2.5.4.4 Ground segment cost

Ground segment cost is the sum of a non-recurring cost and a recurring cost. When these values are not provided by the user, they are estimated using CERs that utilize location of the facility, the number of spacecraft, and spacecraft lifetime as independent variables. The CERs are taken from [4] produce cost estimates in FY2010$k.

$$C_{ground,R} = C_{ground,NR} + C_{ground,R}t(yr) = \\ = F(loc)6,471(\frac{\$}{m^2})A(m^2) + 0.5\frac{\$M}{S/C/yr}N_{S/C}t(yr) \quad (2.37)$$

In equation 2.37 $F(loc)$ is an adjustment factor that takes into account differences in construction cost in different locations, A is the floor area of the facility in m^2, and $t(yr)$ is the lifetime in years.

2.5.4.5 Service fees

Service fees are only applicable for hosted payloads and 100% commercial approaches. For 100% hosted payloads, the tool uses two different pricing models taken from [13], namely "pay per mass" and "resource pricing." The pay per mass pricing model computes the final cost estimate as the sum of the cost of systems engineering, integration and testing, a fraction of the total bus and payload cost, the insurance cost, operator cost, and a certain profit margin.

Alternatively, the resource pricing model takes into account everything in the pay per mass mode, plus a marginal bus and launch cost. These delta costs are based on the fraction of resources (mass, power) of the host spacecraft that are used by the hosted payload. Lastly, 100% commercial services costs are computed at a fixed price per MB of data transmitted.

2.5.5 Search engine

The main goal of the search engine is to provide a mechanism that systematically explores the space of alternative network designs given the decisions available to the system architect. Three main approaches have been implemented.

First, a *test mode* that allows the system architect to manually input a particular architecture and rapidly obtain its benefit and score cost. This strategy is useful in two cases: The architect wants to evaluate a particular design that is dictated from external sources (e.g. use SNAT to evaluate the current TDRSS architecture with a set of future customers). The architect has performed conducted an optimization over a given design space and wants to benchmark its outcome to a reference architecture that was not part of the original space.

The second search engine is based on a *full factorial enumeration.* This approach uses rules adapted from [44] to list all the possible architectures and then evaluate them sequentially. This has the advantage of exploring the entire space of alternatives and thus always identify the optimal solution (or Pareto front for a multi-objective optimization). However, it can only search spaces of less than 2000 architectures due to limited computational resources.

Finally, the third approach uses a heuristic optimizer in the form of a mutli-objective genetic algorithm. In particular, the non-dominated sorting algorithm is based on work by Selva in [44] and follows the prescriptions of non dominated sorting algorithm II (NSGA-II [16]). In contrast, the mutation and crossover operators, as well as the random architecture generator, have been adapted and customized to the particularities of the architectural decisions included in SNAT.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 3

Validation of SNAT

3.1 Introduction

Chapter 2 presented SNAT, a computational tool that can be used to conduct tradespace exploration studies in the context of space communication networks. This chapter describes the validation of SNAT, i.e. the process of benchmarking its results with those of an actual system in order to assess (1) whether the underlying models produce reasonable results and (2) what are the effects and limitations from the implemented modeling assumptions and simplifications.

SNAT can be basically decomposed in three main parts, the performance model, the spacecraft design algorithm and the cost model. The performance model has been specifically built around the problem of understanding the performance and capacity of a space communication network. In contrast, both the spacecraft design algorithm and cost model are based on parametric relationships extracted from multiple references (e.g. [26]) that, if necessary, have been tailored to communication satellites. Therefore, the majority of this chapter is devoted to the validation of the performance model.

In particular, SNAT will be validated against NASA's Space Network. The scheduling algorithm will be compared to operational data for the TDRS satellites in order to see if similar schedules can be produced. At the same time, the spacecraft design algorithm will be validated by comparing the design of an actual TDRS satellite to the spacecraft configuration produced by the tool. Finally, the different cost estimates from the cost model will be benchmarked against publicly available data on the cost of the Space Network.

3.2 Validation strategy for the performance model

The proposed validation strategy for the performance model can be divided three main parts: First, gather and analyze operational data from the SN; second, perform both qualitative and quantitative comparisons between the outputs of SNAT and the SN data; and third, produce basic intuitive tradespaces to visualize the output of tool. These parts can be decomposed into the following tasks:

1. Gather data on the operations of the SN (see section 3.3.1).

2. Define aggregate metrics that capture the performance of the overall SN scheduling system (see section 3.3.2).

3. Examine the validation data in order to characterize the SN and identify generic trends and typical days of operation. Define representative mission scenarios that can be used as inputs to the tool (see section 3.3.3).

4. Benchmark the results from the rule-based scheduler with the real SN operational data based on the metrics identified in (2) (see section 3.3.4).

5. Run a basic case study to demonstrate the ability of the tool to produce and analyze tradespaces, and ensure that its outputs are sensible and consistent with engineering intuition (see section 3.6).

3.3 Validation of the performance model

3.3.1 Dataset description

The Network Control Center Data System (NCCDS), located in White Sands NM, is the operations control facility for the SN. It provides the following functional capabilities:

- Service Planning: Generates a conflict-free schedule based on customer requests. The services that can be scheduled are real-time data transport (digital and analog/video), tracking and spacecraft simulations.
- Service Control: Dissemination of the service schedule and real-time update if needed. Dissemination of the SN performance data to customers.
- Service Assurance: Monitor the ongoing SN service as scheduled by the service planning function.
- Service Accounting: The Service Accounting Segment Replacement (SASR) database collects stores and reports usage data on SN services.

The NCCDS scheduling capabilities are managed by means of a Structured Query Language (SQL) database that acts as a pool for all the requests, tentative schedules and final implementable schedules for the SN. This SQL database is tailored so that it easily interfaces with the SN ground segment, the NASA Integrated Network Services (NISN) and the Space Network Access System (SNAS).

The validation of SNAT is based on an extract of the NCCDS SQL database that contains 164 tables with a total of 631.928 entries. They accounted for 12 days of SN scheduling data (from 05/08/2012 to 05/20/2012). This data was partially incomplete due to ITAR restrictions (e.g. no users from the Department of Defense were included, the names of missions using the SN were not disclosed).

Furthermore, the data in the NCCDS database uses a codification scheme and a set of acronyms that are specific to the SN. A comprehensive list for them can be found in [17], but the following subset is herein presented to facilitate the comprehension of the analysis of the SN operational data:

- S-band Single Access Forward service (SSAF): S-band communication link from TDRS to the customer platform for commanding.
- S-band Single Access Return service (SSAR): S-band communication link from the customer platform to TDRS for telemetry and science data.

- Ku-band Single Access Forward service (KuSAF): S-band communication link from TDRS to the customer platform for commanding.

- Ku-band Single Access Return service (KuSAR): S-band communication link from the customer platform to TDRS for telemetry and science data.

- Multiple Access Return (MAF): S-band communication link from TDRS first generation to the customer platform for commanding. It uses CDMA and beamforming to support up to 1 simultaneous channel.

- Multiple Access Return (MAR): S-band communication link from TDRS first generation to the customer platform for telemetry. It uses CDMA and beamforming to support up to 5 simultaneous channels.

- Multiple Access Return (SMAF): S-band communication link from TDRS second generation to the customer platform for commanding. It uses CDMA and beamforming to support up to 2 simultaneous channels.

- Multiple Access Return (SMAR): S-band communication link from TDRS second generation to the customer platform for telemetry. It uses CDMA and beamforming to support up to 5 simultaneous channels.

3.3.2 Definition of network metrics

The goal of this section is to define a set of metrics that capture the load of the network as a whole, rather than understanding the amount of information that flows through each of the different satellites. The rationale for this approach follows from the initial intent of having a scheduling algorithm as part of this tool, i.e. understanding the *global network capacity* rather than trying to mimic the exact operations of the SN. Two metrics were identified as useful: Total data volume DV_{total} and percentage utilization $\%Ut$.

The total data volume is used as a summary of how many contacts can be scheduled through the network. The available SN dataset contains information on the nominal data rate at which the scheduled contacts run, the same information that can be used as an input for the model scenario. Therefore, the data volume can be simply be computed as

$$DV_{total} = \sum_{i=1}^{N} \sum_{j=1}^{M_i} DV_{ij}, \ DV_{ij} = R_{b_{ij}} \cdot T_{c_{ij}} \tag{3.1}$$

where N is the number customers for the network, M_i is the number of contacts the i-th customer wants to schedule, and $R_{b_{ij}}$ and $T_{c_{ij}}$ are the data rate and contact duration for the j-th contact of the i-th user. It is immediate to see the if the rule-based scheduler drops contacts due to lack of network resources, the total data volume DV_{total} metric will be penalized.

Alternatively, the percentage of utilization ($\%Ut$) measures how much time a payload, antenna or satellite has been active and serving a customer. The percentage of utilization is defined as follows:

$$\%Ut_{pay_i} = \frac{\cup_{p=1}^{P} C_{i_p}}{86400} \tag{3.2}$$

$$\%Ut_{ant_j} = \frac{\cup_{a=1}^{A} C_{j_a}}{86400} \tag{3.3}$$

$$\%Ut_{sat_k} = \frac{\cup_{s=1}^{S} C_{k_s}}{86400} \tag{3.4}$$

where the operator $\cup$ indicates that all contacts served through that element are placed in a common time axis and consolidated into single contacts when more than one overlap. Therefore, $\cup_{p=1}^{N} C_{i_p}$ indicates that all contacts C_{i_p} routed through the i-th payload of a particular satellite at the same time are consolidated.

Note that for validation purposes only $\%Ut_{ant}$ is really meaningful since the antennas are the effective network resource that the scheduler allocates. In other words, if an antenna is pointed to a customer it will not be able to serve another mission, therefore decreasing the pool of available resources. In contrast, if a contact is served through a particular satellite that does not imply that this same satellite cannot be serving another contact at the same time.

3.3.3 Analysis of TDRSS operational data

Once the two metrics for validation have been identified, the next step is to analyze the SN operational data to understand its current level of performance. This is done at the network, service, satellite, antenna and payload level metrics.

3.3.3.1 Network-level characterization

Three network level metrics were analyzed: Total daily data volume (Tbit), total service time (h), and number of granted contacts. TDRS communication payloads allow a contact to be scheduled using multiple frequency bands at the same time. When that happens, it is considered a single granted contact, but each band is accounted as independent service time through the same antenna.

Figure 3-1 and figure 3-2 present the collected summary metrics for the SN as a whole. Results indicate that the network serves an average of 835 contacts per day that account for a total scheduled time of 407 hours and 50 minutes, and a total data volume of 34.3 Tbits.

Results also indicate that the SN operates under great variability from one day to the next one. In particular, figure 3-1 shows how the network can currently support days with 50 Tbits and more than 500 hours of scheduled time, while other times it relays half the same amount of data and schedules half of the contact hours. Additionally, it is also interesting to note the limited correlation between the data volume, service time and number of scheduled contacts. A reduced number of scheduled contacts does not imply, in general, less scheduled time nor does it imply higher or lower total daily data volumes.

Figure 3-3 presents a summary table with basic statistics for each of the three network level metrics. The table is color coded based on mean and standard deviation for each metrics, that is, red colored days indicate that the metric is at least one standard deviation above the mean, and green colored days are at least one standard deviation below the metric. Visual inspection indicates that both the 05/08/2012 and 05/09/2012 were high load operation days, whereas 05/14/2012 or 05/17/2012 can be

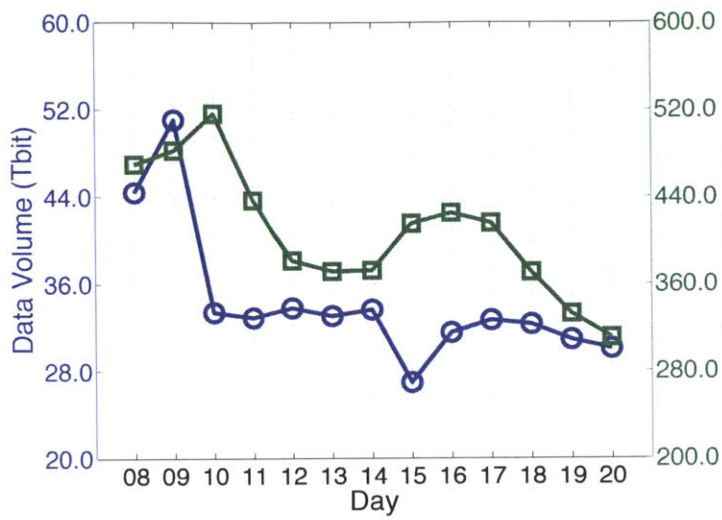


Figure 3-1: TDRSS daily data volume and service time

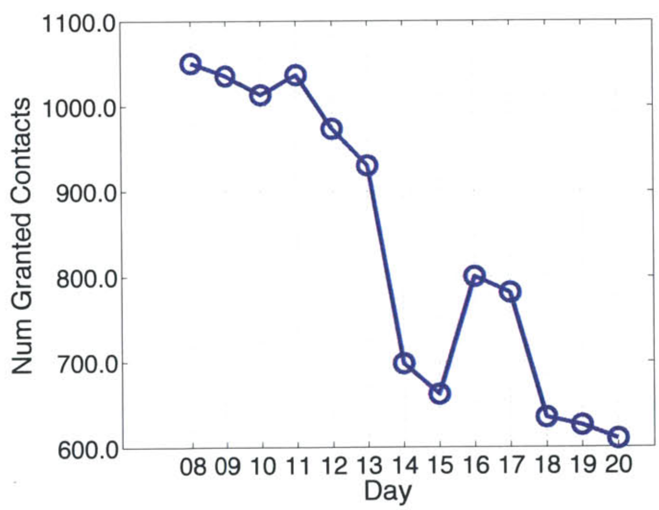


Figure 3-2: TDRSS number of granted contacts

Day	Data Volume (Gbit)	Scheduled Time (s)	Scheduled Contacts	% utilization
5/8/2012	44398.00	1691333	1050	48.79
5/9/2012	51023.00	1736937	1035	54.87
5/10/2012	33367.00	1858700	1013	60.78
5/11/2012	32883.00	1572530	1036	45.51
5/12/2012	33792.00	1374366	973	43.02
5/13/2012	33046.00	1337658	930	38.66
5/14/2012	33610.00	1340395	698	46.82
5/15/2012	26950.00	1495720	662	46.96
5/16/2012	31510.00	1530990	800	43.53
5/17/2012	32673.30	1497578	781	45.23
5/18/2012	32279.80	1334937	635	40.61
5/19/2012	30871.90	1197064	626	38.74
5/20/2012	30045.30	1119015	610	32.08
Mean	34342.25	1468247.92	834.54	45.05
Std Dev	6109.13	205000.93	169.51	7.02

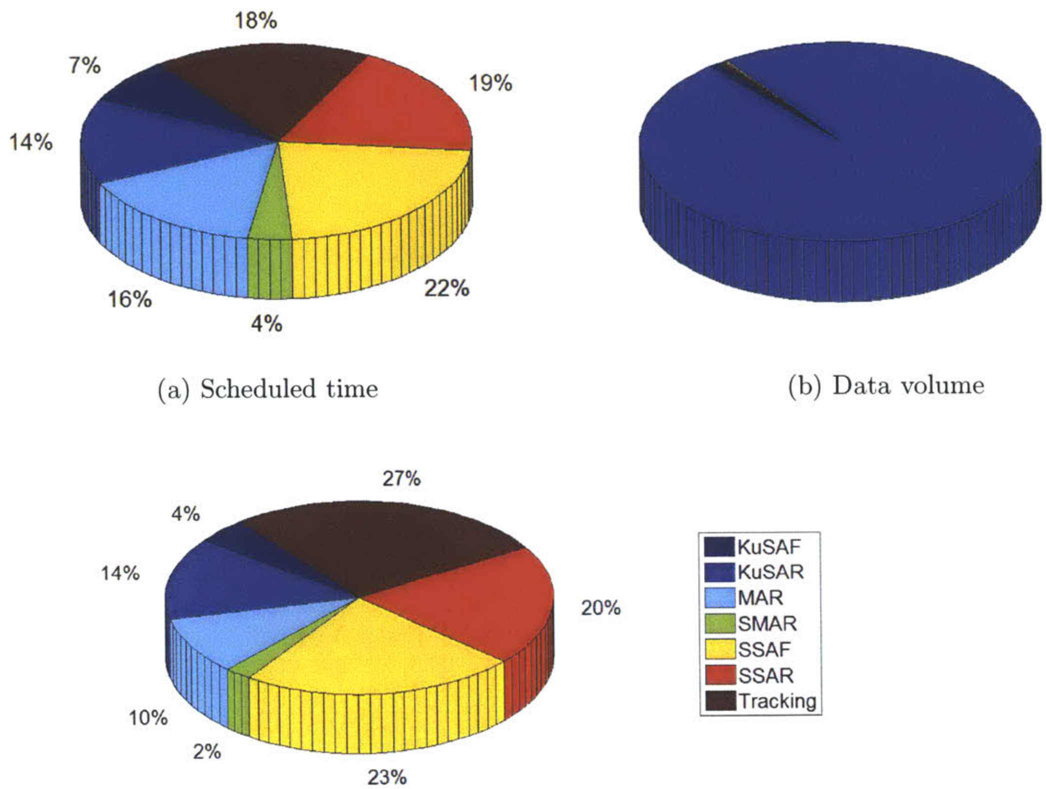
Figure 3-3: Network level statistics

considered a typical day of operations for the SN. Therefore, for validation purposes day 05/14/2012 will be used a high load operations day, while 05/14/2012 will be used as a typical day of operations. For both days, the customers of the network will be carefully characterized and then used as an input to SNAT for benchmarking purposes.

3.3.3.2 Service-level characterization

Figure 3-4a, figure 3-4b and figure 3-4c present the three network metrics from section 3.3.3.1 categorized based on the type of requested service. Results indicate that the SSAF, SSAR and tracking services account for the majority of scheduled contacts and number of requests (Ku-band services represent only 20% of the total scheduled time while S-band SA and MA services account for a 61%). Nevertheless, they represent a minimal fraction of the daily data volume (less than 2%). This fact indicates that although the Ku-band services dominate the utilization of the network in terms of data transfer, the utilization of the antennas is mainly due to S-band services.

Figure 3-5 plots the percentage of users that demand each type of service. S-band single access services are currently the most popular along with tracking services. This is consistent with SSAF, SSAR and tracking being the services that have more requests and more scheduled time. In contrast, MA services are also popular (40%



(a) Scheduled time

(b) Data volume

(c) Number of requests

Figure 3-4: Metrics vs. service

of users request them) but only account for a 12% of the total requests. This fact hints that most missions utilize MA services only for contingency operations or special events while they usually rely on the SSAF and SSAR for their routine contacts.

Finally, no Ka-band services were observed during any of the 20 operational days. This can be due to either this service being heavily underutilized, or to the classified nature of the users that request it. Similarly, there no instances of MAF or SMAF services were identified for any of the available days.

3.3.3.3 Satellite-level characterization

The satellite-level characterization can be used to understand the differences in the scheduling load for the different satellites of the SN. Two main pieces of information

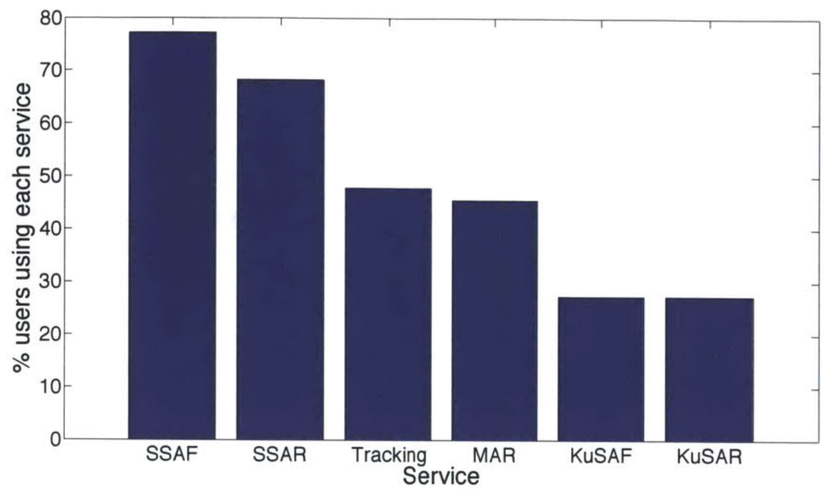


Figure 3-5: Service popularity

	Scheduled Time (h)		Data Volume (GBit)	
Satellite	Mean	Std. dev.	Mean	Std. dev.
TDRS C	114.1	8.5	5303	669
TDRS E	131.3	7.6	11231	1591
TDRS G	51.6	24.3	6116	2038
TDRS I	49.8	11.8	7541	3723
TDRS J	61.1	26.2	4152	1964

Table 3.1: Satellite utilization

were considered: the total scheduled time per satellite (figure 3-6), and the total daily data volume that each of them relays to either White Sands or Guam (figure 3-7). Additionally, table 3.1 also summarizes the utilization of the different TDRS satellites.

Results indicate that there are major differences in how satellites are used during real operations. The Atlantic Ocean satellites (TDRS C and I) consistently relay more information than the others. Special attention must be drawn to TDRS C, which on average relays twice as much data as most of the other satellites.

Similarly, the first generation satellites typically get more scheduled time per day than the second generation satellites. Two possible explanations for this effect have

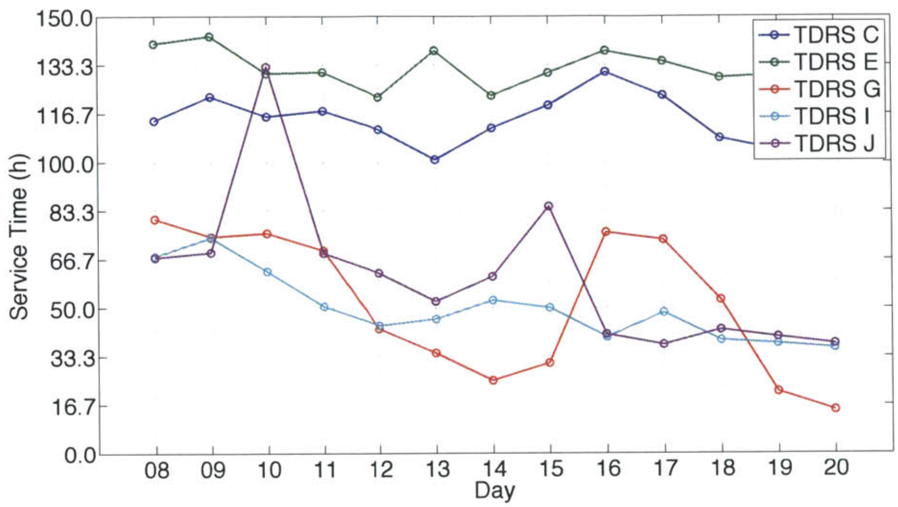


Figure 3-6: Scheduled time per satellite

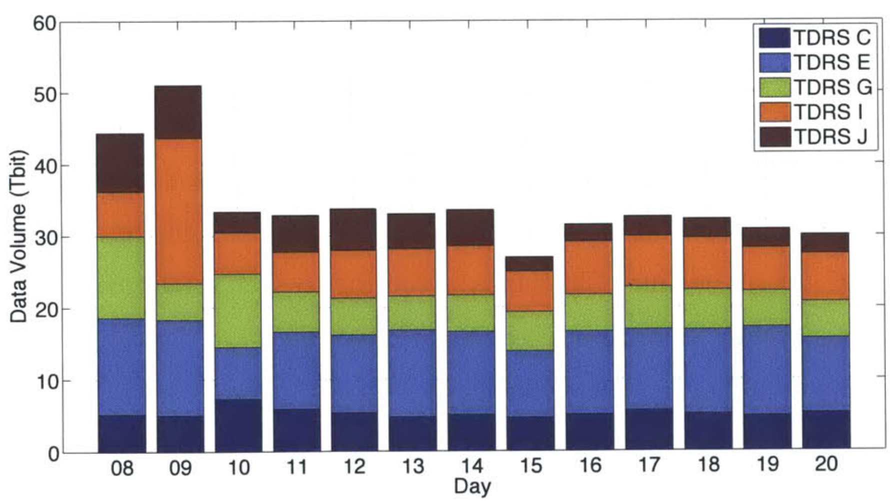


Figure 3-7: Data volume per satellite

been theorized (and not contrasted since the dataset does not capture the decisions made during the scheduling process):

- The SN prefers to concentrate the maximum possible load to its oldest satellites in order to extend the life of the new ones as much as possible.

- When a contact is requested, the MOC specifies the beginning and end times of the service as well as the specific desired satellite and antenna. This strategy encourages MOCs to use the same scheduling patterns, i.e. if they have scheduled services through a particular satellite they should continue to do so as long as it meets their constraints. Since the first generation satellites have been working for longer periods of time, more customers are utilizing them for heritage reasons.

3.3.3.4 Antenna-level characterization

The antenna-level metrics are mainly focused at capturing the percentage of time that the antennas are being utilized. Figure 3-8 plots this percentage for the two single access antennas of the TDRS satellites over the twelve days of operation. This percentage is computed as the fraction of time that the antenna is committed and pointing to a user over the total available time in a day.

The analysis indicates that the average utilization of a TDRS single access antennas is 45%. NASA internal studies have assessed that if the system is operated at full capacity this percentage increases to maximum of 70 to 75% due to scheduling and customer orbital constraints. Therefore, according to the current available dataset the SN is, as of today, running at medium load.

Nevertheless, SN experts have also stated that support to the DoD and other sensitive users not present in the dataset can increase the network utilization by 30%. Although there is no indication onto whether this percentage relates to data volume or antenna utilization, it is expected that the real TDRS antenna utilizations are significantly bigger than those shown here.

Finally, a similar analysis can be conducted for the MA services. Although there is

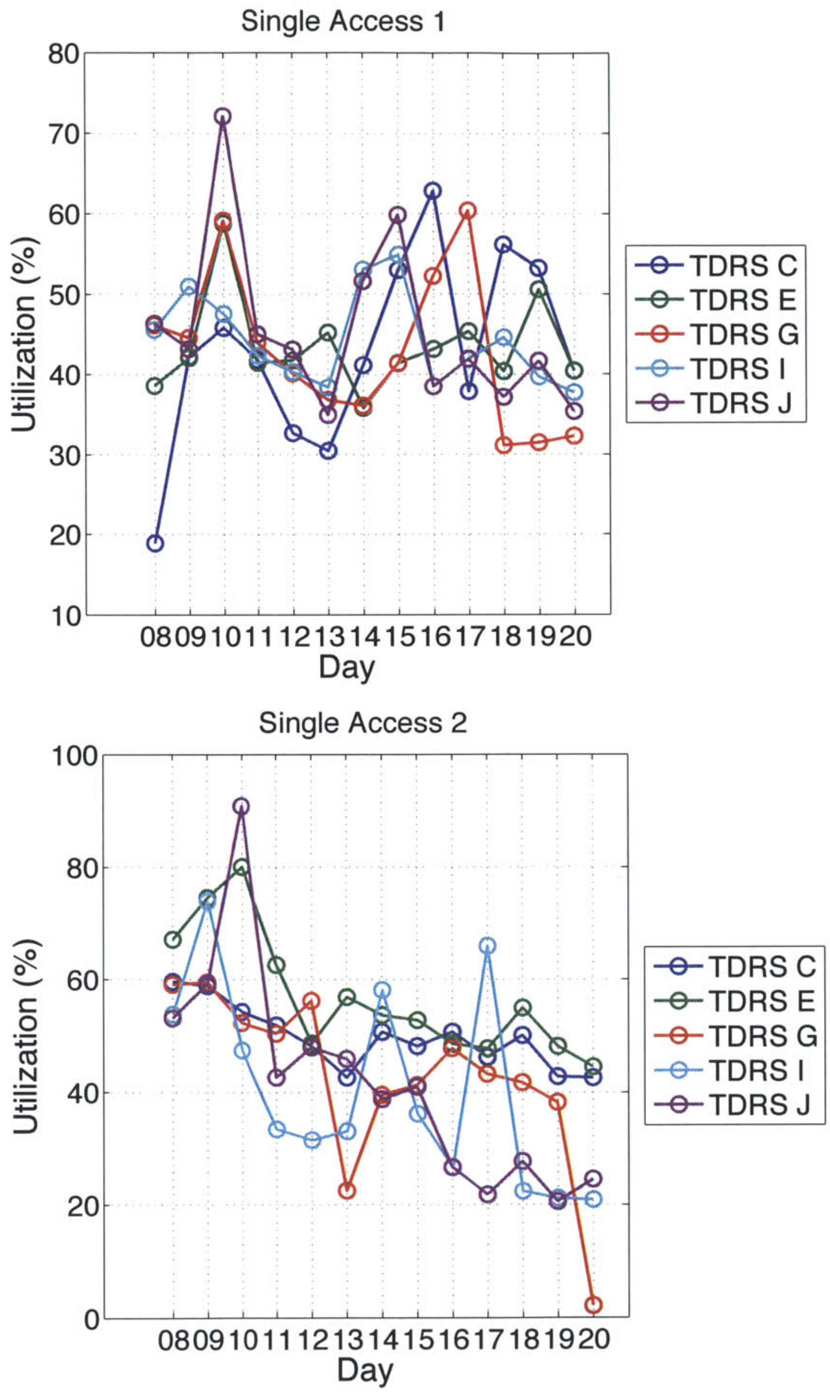


Figure 3-8: Single Access antenna utilization

		TDRS A-G		TDRS H-J	
SA	SSAF	7 Mbps	68.45 dB-bps	7 Mbps	68.45 dB-bps
	SSAR	6 Mbps	67.78 dB-bps	6 Mbps	67.78 dB-bps
	KuSAF	25 Mbps	73.98 dB-bps	25 Mbps	73.98 dB-bps
	KuSAR	300 Mbps	84.77 dB-bps	300 Mbps	84.77 dB-bps
	KaSAF	-	-	25 Mbps	73.98 dB-bps
	KaSAR	-	-	300 Mbps	84.77 dB-bps
MA	(S)MAF	1 @ 300 kbps	54.77 dB-bps	2 @ 300 kbps	54.77 dB-bps
	(S)MAR	5 @ 300 kbps	54.77 dB-bps	5 @ 3 Mbps	64.77 dB-bps

Table 3.2: SN nominal data rates

only one MA antenna per satellite, it is in fact capable of supporting five independent simultaneous users for return services and only one for forward services. As previously stated, no instances of SMAF or MAF services were encountered on the dataset. Therefore, the analysis for MA services was solely conducted based on return services. Its results indicate that, on average, the utilization of an MA beam is as low as 13% and although particular beams can be in operation for more than 50% of the time, an overall MA antenna (the five beams) hardly ever operates more than 25% of the time.

3.3.3.5 Payload-level characterization

The payload-level characterization is not centered around the percentage of utilization like the previous analyses, but rather focuses on the distribution of data rates the customers require from the network. Its main goal is to assess whether the nominal data rates of the SN are currently sufficient or whether a vast majority of users are already using the links maximum offered capacity. Table 3.2 presents the nominal data rates for the different SN services.

Figure 3-9 plots six histograms, one per SN service type, for the data rates that all the SN customers utilize in their contacts. The red dashed vertical line indicates the nominal data rate for that particular service as specified in table 3.2, thus setting a reference against which customer data rates can be compared. Note that in most

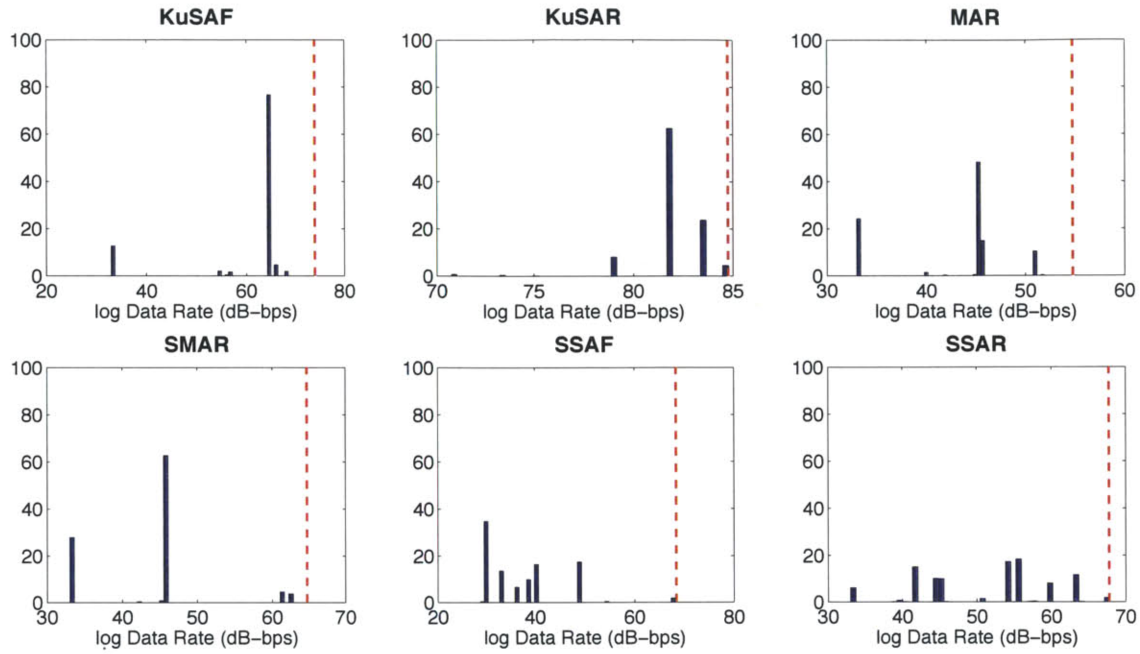


Figure 3-9: Scheduled contact data rate distribution

cases customers requests are at least one order of magnitude below the maximum achievable data rate. This assertion is especially true for the SSAF and SSAR services. In contrast, the KuSAR services tend to be more heavily utilized.

On the other hand, MAR and SMAR users tend to use data rates on the order of tens and hundreds of bps. This is consistent with the data rates offered by MAR (first generation TDRS) while very few users take advantage of the increased data rates available with SMAR.

3.3.4 Validation of the scheduling algorithm

As previously mentioned, the goal of the heuristic scheduler is not to replicate the exact SN schedule but rather assess the capacity of the network by cleverly allocating its resources. Therefore, the main objective of the scheduling algorithm validation is to understand whether, given a predefined set of customers, the model can predict the network load similar to that of the SN. Note that this is a crucial part of tool validation since all the performance metrics are computed from the network's schedule.

The validation of the scheduling algorithm is based on the *typical day of operations* and *high load day of operations* identified in section 3.3.3.1. Additionally, the benchmarking metrics presented in 3.3.2 are used for comparison purposes. Tables 3.3, 3.5 and 3.4, 3.6 present the relative error for $\%Ut$ and DV_{total} respectively, both in the case of a typical day of operations and high load scenario. These errors are computed by comparing the two aggregate metrics for the real SN schedule vs. the rule-based schedule. Results indicate that the rule-based scheduling algorithm is able to approximate the SN load within less than a 10% error in all cases except for the satellite percentage of utilization, and S-band total data volume and percentage of antenna utilization.

The mismatch between the SN's percentage of satellite utilization $\%Ut_{sat}$ and what is predicted from the model comes from the fact that the rule-based scheduler does not take into account whether antennas of the same satellite should or should not be activated simultaneously. If one of the antennas of a satellite is relaying information, then that satellite is considered to be active. Therefore, the rule-based scheduling algorithm could potentially reduce the mean $\%Ut_{sat}$ by including a rule that encourages contacts to be scheduled through the same satellite using the two available SA antennas at the same time. However, since no such heuristic is included, there is no guarantee that the heuristic scheduler will obtain a $\%Ut_{sat}$ similar to that of the SN. Furthermore, the SN schedule is produced as an iterative process where each mission makes requests based on their specific operational requirements (e.g. contacts are scheduled during working hours in the mission control center). These considerations are also not captured by the tool schedule.

On the other hand, the differences in $\%Ut_{pay}$ and DV_{total} for the S-band payload are due to the TDRS ability to support multi-frequency contacts. In particular, missions typically schedule a Ku-band event overlapped with an S-band event so that science data, telemetry and commands are transmitted simultaneously. This feature is currently not implemented in the rule-based scheduler. As a result when modeling this type of contacts it is assumed that only the Ku-band link is active since it drives the amount of data being sent over the contact. This modeling inaccuracy results in

	TDRSS $\%Ut$	**Model** $\%Ut$	**Absolute Error**	**Relative Error**
Satellite	76.75	62.98	13.77	17.94%
Antenna	46.82	43.46	3.36	7.19%
S-band	39.69	17.72	21.97	55.35%
Ku-band	21.75	20.43	1.32	6.08%
Ka-band	0.00	0.00	0.00	0.00%

Table 3.3: Validation of $\%Ut$ for a typical day of operations

	TDRSS DV_{total}	**Model** DV_{total}	**Absolute Error**	**Relative Error**
Satellite	33610	35613	2003	5.96%
Antenna	33575	35614	2038.93	6.07%
S-band	290	121	169	58.30%
Ku-band	33052	35492	2440	7.38%
Ka-band	0.00	0.00	0.00	0.00%

Table 3.4: Validation of DV_{total} for a typical day of operations

big disparities for the S-band payload both in terms of $\%Ut$ and DV_{total}.

That said, it is important to note that the results in tables 3.3, 3.5 and 3.4, 3.6 do validate the scheduler. It can be seen that both the antenna total data volume and the percentage of utilization of the antennas is accurate to less than 10%. This indicates that despite the fact that the rule-based scheduler makes several simplifications and cannot mimic all the factors included in an operational schedule, it is able to grant the same amount of contacts and transmit the same amount of information through the network. Therefore, it is correctly assessing the network capacity.

3.4 Validation of the spacecraft design algorithm

The spacecraft design algorithm provides an estimate of the different satellite subsystems given the communication payloads that it carries. Therefore, its validation must ensure that it does not oversize any them and, if that is the case, estimate the

	TDRSS %Ut	Model %Ut	**Absolute Error**	**Relative Error**
Satellite	82.88	65.94	16.94	20.44%
Antenna	48.79	47.14	1.65	3.38%
S-band	42.73	19.34	23.39	54.74%
Ku-band	29.75	27.80	1.95	6.55%
Ka-band	0.00	0.00	0.00	0.00%

Table 3.5: Validation of %Ut for a high load scenario

	TDRSS DV_{total}	**Model** DV_{total}	**Absolute Error**	**Relative Error**
Satellite	44398	41087	3311	7.46%
Antenna	44386	41087	3299	7.43%
S-band	6350	3603	2747	43.26%
Ku-band	35801	37483	1681	4.70%
Ka-band	0.00	0.00	0.00	0.00%

Table 3.6: Validation of DV_{total} for a high load scenario

	TDRS-J	Model	Relative Error
Launch Mass (kg)	3196	3091	-3%
BOL mass (kg)	1786	1905	7%
EOL power (W)	2300	2170	-6%

Table 3.7: Error on mass and power estimates for TDRS-J

Subsystem	TDRS-J Mass Fraction	Model Mass Fraction	Relative Error
Payload	36%	35%	-1%
Power	25%	25%	0%
Structure and thermal	22%	24%	5%
ADCS	8%	7%	-1%
Propulsion	6%	5%	-13%
TT&C	3%	4%	8%

Table 3.8: Error on mass fraction estimates for TDRS-J

incurred error. The results herein presented were compiled by Selva and are included in this thesis in order to provide a more comprehensive view of the validation process.

The validation of the spacecraft design algorithm is grounded on the second generation of TDRS, specifically the TDRS-J spacecraft. Information on its launch mass, beginning of life mass and end of life power was gather through publicly available data. Additionally, the subsystem mass ratios were obtained through NASA officials.

Based on the communication payloads that TDRS-J is carrying, the spacecraft design algorithm was used to size the satellite and estimate the mass of each subsystem. Then, these were compared with the real values. Tables 3.7 and 3.8 present the results of the validation of the spacecraft design algorithm. It can be seen that it is accurate to approximately 10% for both the mass fractions and the mass and power estimates.

	TDRSS	Model	Relative Error
LCC TDRS 1st gen over 10 years	3.5 FY00$M	3.0 FY00$M	-14%
Bus and launch cost for TDRS 2nd gen.	544 FY00$M	566 FY00$M	4%
Annual ground segment operations cost	97 FY10$M	83 FY10$M	-14%

Table 3.9: Error on cost estimates

3.5 Validation of the cost model

The validation of the cost model is particularly challenging due to the lack of reliable public information. Similar to the spacecraft design algorithm, the results herein presented were computed by Selva and are included to provide a holistic view of the validation process.

Table 3.9 provides the only three data points that could be gathered regarding the cost of the SN. It can be seen that the tool provides an estimate that is accurate to 15% approximately, although a more thorough validation would be advisable should more data become available.

3.6 Tradespace validation

The goal of the tradespace validation is to run a simplistic case study that demonstrates the ability of the tool to produce tradespaces and link its outcomes to the architectural decisions. To that end, this validation is done by analyzing the results of increasing the capacity of a space network by increasing the number of launched satellites and varying the number of single access antennas that they carry.

Intuitively, as more satellites and single access antennas are deployed in the network more customers can be served. However, the system becomes increasingly expensive. Therefore, there is an optimal network capacity for which all customers are fully satisfied and adding extra nodes on the network only increases its cost without

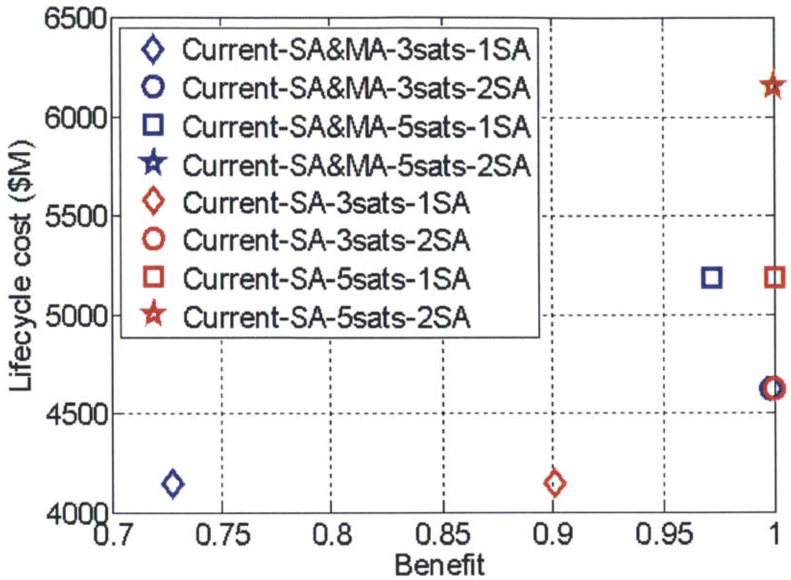


Figure 3-10: Evolution of the network capacity

any added benefit.

The validation case study tested the following network architectures: 3 satellites with 1 single access (SA) antenna per satellite, 3 satellites with 2 SA antennas per satellite, 5 satellites with 1 SA antenna per satellite, and 5 satellites with 2 SA antennas per satellite (red markers on figure 3-10). The same four cases were also run with a different user set where all low data rate SA users are switched to the multiple access (MA) services (blue markers).

The results of this test show that:

- With the user set extracted from the SN operational data 5 satellites and 2 antennas per satellite are needed to provide 100% satisfaction (red square marker). This is consistent with the current status of the TDRSS constellation.

- 3 satellites with 2 SA antennas (red circular marker) per satellite yield 99% satisfaction and reduce the cost of the system by approximately $1B.

- With the low data rate users switched to MA, 3 satellites with 2 SA antennas per satellite (blue circular marker) suffice to provide 100% satisfaction, thus saving around \$1B in life cycle cost with respect to the original case.

- 3 satellites with 1 SA antenna per satellite yield 70% satisfaction when users are serviced through the SAs. However, the overall satisfaction level increases to 90% if low data rate users are switched to MA.

- Deploying 5 satellites, either with 1 or 2 SA does not provide any extra satisfaction and increases the cost of the system by nearly \$2B.

Therefore, the qualitative validation case matches the initial expectations. Increasing the capacity of network improves the quality of service provided to the customers by granting more contacts at the desired instants in time. However, it also translates to higher deployment and operational costs.

Chapter 4

Evolving NASA's Space Network

4.1 Introduction

Now that SNAT has been described and validated, it is time to demonstrate its usefulness during the architecting phase of a space communications network by building two case studies on the evolution of the SN. To that end, this chapter is structured in two main parts: First, the expected customer base for the 2020-2030 time frame is discussed and turned into a user scenario that can be inputted to SNAT. Then, two case studies are presented, the first one focusing on new technology and hosted payload opportunities, and the second one built around the valuation of inter-satellite links.

4.2 Network customer characterization

The network customer characterization was conducted through three complementary approaches: First, a literature review of future mission concepts (e.g. NRC Astronomy and Astrophysics Decadal Survey [18]). Second, interviews with experts from different scientific communities (Earth observation, astrophysics, heliophysics and human space flight - HSF) were conducted in order to better understand their predicted needs. Third, the dataset with SN operational data was analyzed with respect to the current users in order to identify typical utilization patterns.

The result of this exercise for the 2020-2030 time frame are presented in table 4.1, where the different types of services provided are: Science data return[1] (SD), tracking, telmetry and command (TT&C), and high definition video (VD). This table specifies the primary services that each type of mission will require, although all of them would use the network for contingency communications should they be needed.

The resulting scenario consists of a set of 15 missions from NASA, NOAA and USGS. The NASA Earth observation community has eight missions using the network, two of them being major drivers (e.g. DESDYNI class missions [42]), two more being medium sized missions and four more being small-sats. A mission is considered to be a driver for the network if its required data rate at least doubles the capabilities of the current SN and the total desired data volume is greater than 10Tbit/day. On the other hand, the NASA astrophysics and heliophysics community are represented through three missions, one of which is considered to be a driver and the other three are medium sized missions. These medium sized missions typically return between 1 and 10Tbit/day approximately.

Similarly, the human spaceflight stakeholder is considered to have two missions in orbit, one similar to the International Space Station (ISS) and another one representative of a Multi-Purpose Crew Vehicle (MPCV) capsule. The ISS-like is a driver both in terms of required scheduled time (continuous contact throughout the day is required) and data volume. In turn, the MPCV-like missions only requires support for a long continuous periods of time (5 hours approximately). Finally, the NOAA and USGS mission are also included in the scenario, but are expected to be opportunistic users of the system. In this case, it is assumed that they will use the network for contingency situations and therefore will schedule a maximum of two 5 to 10 minute contacts.

Once the set of customers has been identified, the next step is to define their operational requirements with respect to the network. Since mission concept documents for future missions rarely define these parameters, it is assumed that in most cases these will similar to the concept of operations of current missions. In particular, the

[1] For HSF missions, voice and video is assumed to be multiplexed with the science data.

	Mission class	# missions	Services	Data Volume	Latency
NASA Earth Observation	Class 1	1	SD, TT&C	40 Tbit/day	1 Hour
	Class 2	1	SD, TT&C	15 Tbit/day	1 hour
	Class 3	2	SD, TT&C	1-5 Tbit/day	hours
	Class 4	4	SD, TT&C	500 Gbit/day	hours
NASA Astrophysics	Class 1	1	SD, TT&C	40 Tbit/day	2 hours
	Class 2	1	SD, TT&C	2 Tbit/day	5-7 hours
NASA Heliophysics	Class 1	1	SD, TT&C	10-15 Tbit/day	1 day
NASA Human Space Flight	Class 1	1	SD, TT&C	12 Tbit/day	0 seconds
	Class 2	1	VD, TT&C	1 Tbit/day	0 seconds
NOAA	Class 1	1	Contingency	150 Mbit/day	12 hours
USGS	Class 1	1	Contingency	250 Mbit/day	12 hours

Table 4.1: Detailed user scenario

contact time and frequency of contacts for a given mission will be similar to that of today (e.g. one 5 to 15 minute contact per orbit for Earth observation spacecraft).

As a result, both the number of contacts to schedule and the total schedulable time for the network will similar to the current SN values. In contrast, the amount of data routed through the entire network will be significantly bigger, thus requiring links with increased capacity. Once these assumptions are combined with the user set from table 4.1, the resulting scenario has a total of 213 required contacts per day that account for 115 hours of total schedulable time and a daily data volume of 130Tbit per day. These represent a 10% and 273% increase with respect to the current SN operations that support, on average, 104 hours of scheduled time and 34.3Tbits per day.

4.3 Case study 1: Valuation of new technology and hosted payloads

4.3.1 Tradespace definition

Figure 4-1 presents a pictorial representation of the analyzed tradespace of network architectures. It is based on three main architectural decisions:

- Transponder selection: Four RUL transponders can be chosen with their corresponding communication payloads. For each of them, two identical copies are available so that a satellite can carry to single access transponders at the same time (except for the optical telescope). The *TDRS transponder* (TDRST) support communications at S, Ku and Ka-band with a maximum data rate of 300Mbps. The *slow RF transponder* (SRFT) provides low data rate communications at S-band (up to 6Mbps) based on the current TDRS capabilities. Similarly, the *fast RF transponder* (FRFT) only support Ku and Ka band but takes advantage of better spectrum efficiency to provide links up to 600Mbps without extra bandwidth allocations. Finally, the optical telescope (OPTel) can support up to 1Gbps.

- Antenna allocation: Once a subset of the RUL antennas has been selected, the next step is to decide how they will be deployed in space. The baseline option is to put them all together in a single spacecraft (monolithic approach), although they can also be separated into multiple smaller satellites (disaggregated approach). Based on expert input from NASA officials, a maximum of two parabolic dishes is allowed per satellites. This constraint limits the spacecraft configuration complexity.

- Contract modalities: Each of the relay satellites that has been configured through the transponder selection and allocation can be either fully procured or a flown as a hosted payload.

In figure 4-1 satellites with a blue body are owned by NASA while satellite with red body are owned by a commercial entity. Some the satellites have only one parabolic dish (or telescope) attached, while others carry two transponders at the same time. The represented options are three samples of possible architectures that can be evaluated with SNAT. The full tradespace of architectures contains 4,450 different alternative designs, a number that is reduced to 1,440 by restricting the number of deployed satellites to 9 and ensuring that none of them carries more than two antennas at the same time.

Note the large span of architectures being explored. On the one hand, TDRS-like architectures are reproduced as an instance of constellations of monolithic satellites procured and then operated by NASA. On the other hand, fully hosted payload architectures are also being considered, with the commercial satellites hosting one or two transponders. Finally, architectures that disaggregate TDRS-like multi-band payloads into a low data rate payload and a high data rate payload are also analyzed. This variety in the satellite configuration and procurement strategy becomes interesting due to non-trivial couplings between the decisions. For instance, is it better to host a low data rate payload or is it better to place a high data rate optical transponder on board a commercial satellite? Does having separate low and high data rate antennas provide greater network capacity due to greater scheduling flexibility?

Table 4.2 specifies the possible values for each SNAT architectural decision (see section 2.3). If a decision only contains one value, it then becomes a parameter that is fixed across all architectures. Note that since this case study analyzes the evolution of the SN the following assumptions are made:

- The current SN constellation pattern will be maintained, with three orbital positions in the geosynchronous belt. One of them will be placed above the Pacific Ocean, another one in the Atlantic Ocean, and the third one in the Indian Ocean.

- All satellites carry bent-pipe analog transponder than relay signals from customers directly to the ground station. No inter-satellite links are used.

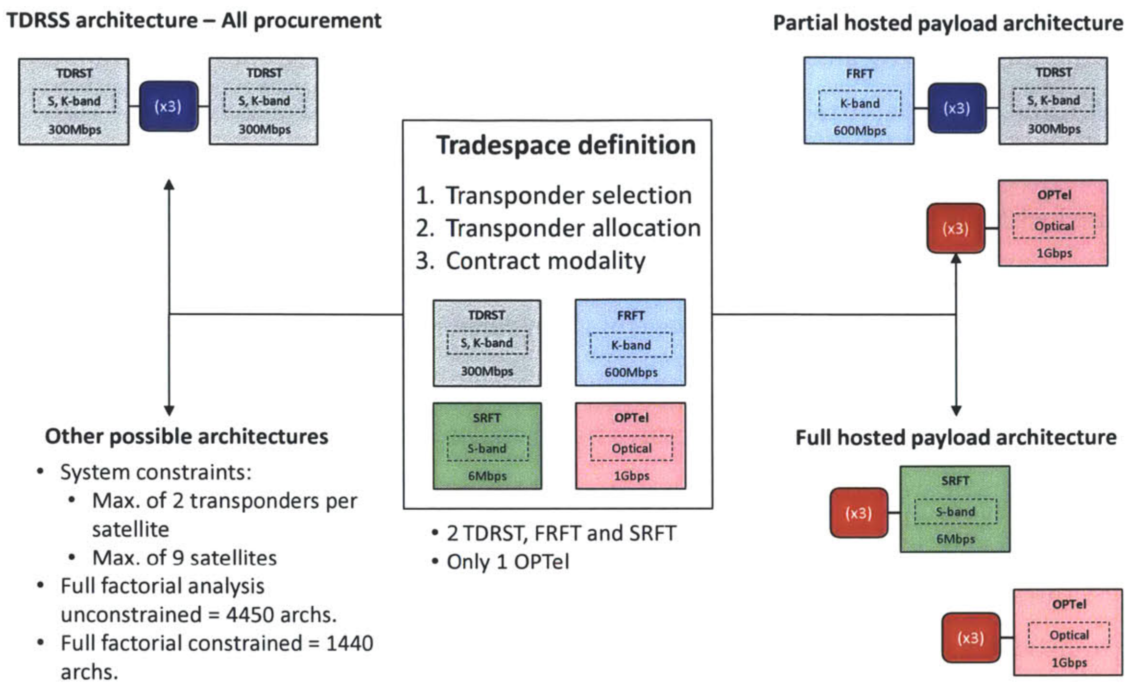


Figure 4-1: Tradespace definition

- The current 10 year SN replenishment strategy will be maintained, launching one satellite per orbital position for each TDRS generation. This means that the total number of deployed satellites in the system will always be a multiple of 3.

- The current SN ground stations will be maintained and upgraded to properly support the space segment. No new sites will be constructed.

- The current SN frequency and bandwidth allocations will be available in S, Ku and Ka-band.

- The default modulations for the links will be M-PSK, $M = \{2, 4, 8\}$. The current SN supports both 2-PSK (BPSK) and 4-PSK (QPSK), so it is assumed that the modems and beam formers on the ground stations will be modified to support 8-PSK. This upgrade is already being implemented under the Space Network Ground Segment Sustainment (SGSS) project [19]. Additionally, no

Decision	Range of values
Orbit selection	Geosynchronous
Constellation design	Number of planes: 1 Number of orbital positions per plane: 3
Inter-Satellite Link payload allocation	No
Transponder-to-spacecraft allocation (disaggregation)	All the possible partitions of N transponders into $1 \leq N_{sat} \leq 2$ satellites
Ground station	White Sands and Guam
Contract modality	100% procurement or hosted payloads
Transponder selection	TDRST (x2), SRFT (x2), FRFT (x2), Optical (x1)
Transponder technology selection	bent-pipe

Table 4.2: Case study architectural decisions

coding scheme is assumed for the links (coding gain equal to $0dB$).

4.3.2 Results: Infusion of new RF and optical technology

The transponder selection and allocation problem looks at what are the best frequency bands to be included in the system and how to best allocate them in relay satellites. In order to provide a non-biased comparison, this study assumes that all satellites are procured and then operated by NASA. Figure 4-2 presents the tradespace obtained with SNAT and flags the set of architectures that achieve the best performance for a given total life cycle cost (LCC). These architectures have the following characteristics:

1. Monolithic architecture with one SRFT per satellite ($\leq 6Mbps$).

2. Monolithic architecture with one TDRST ($\leq 300Mbps$).

3. Monolithic architecture with two transponders per satellite, a SRFT ($\leq 6Mbps$) and a FRFT ($\leq 600Mbps$).

4. Monolithic architecture with two transponders per satellite, a TDRST ($\leq 300Mbps$) and a FRFT ($\leq 600Mbps$).

5. Disaggregated architecture with two sets of 3 satellites. The first set carries two transponders, a SRFT and a FRFT. The second set also carries two transponders, a FRFT and an OPTel ($\leq 1Gbps$).

Based on the analysis of the overall tradespace, the following conclusions can be drawn:

- An all optical architecture achieves a very low score in benefit (0.1 approximately) but is potentially less expensive than most RF and RF/optical architectures. This is due to the fact that the mission scenario described in section 4.2 only included two missions with data volume requirements high enough to necessitate optical communications.

- All architectures that select one or two transponders to be flown rely on monolithic spacecraft (see architectures 1, 2, 3 or 4). SNAT currently does not include any penalization for spacecraft that carry more than one transponder and are therefore subject to electromagnetic interference (EMI). As a result, the model always favors architectures that minimize the number of launches by putting all the transponders together in single monolithic satellites.

- The incremental benefit of having S-band for TT&C services is approximately 40%. Similarly, having Ku and Ka-band for science data return services provides an extra 50 to 55% of the benefit score.

- The current SN architecture is able to provide approximately 84% of the total benefit, although the same level of performance can be obtained with the less costly architecture (4).

- Architectures that obtain maximum benefit support all RF bands and also provide optical communications. Nevertheless, the addition of the optical payload in the system results in disaggregation of the monolithic satellites which, in turn, dramatically increases the cost of the overall system (25% approximately).

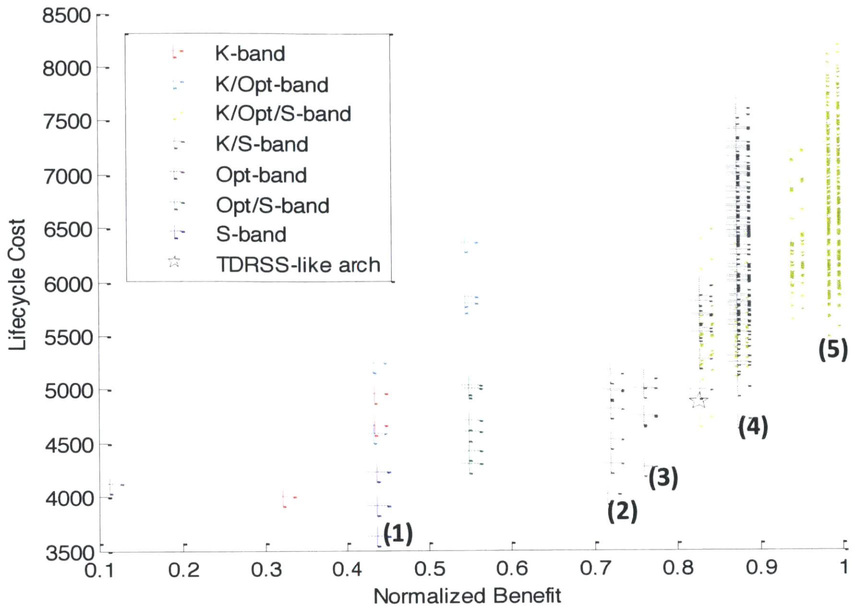


Figure 4-2: Transponder selection

Figure 4-3 shows the same tradespace but color coded according the how communication payloads are allocated into relay satellites. If only three satellites are in orbit, then the architecture is monolithic and all communication payloads are put together in a single spacecraft. The results indicate that monolithic architectures are, in general, better from a cost perspective since the cost of launching more satellites outweighs the savings of reduced mass, power and volume for the satellites. Nevertheless, higher network capacity can be achieved if the TDRST transponders (simultaneously supporting S, Ku and Ka-band) are broken into separate SRFT (S-band) and FRFT (Ku and Ka-band) transponders since S-band and K-band links can then be scheduled independently.

Finally, the evolution of the SN as a NASA owned system can be derived from combining figures 4-2 and 4-3. Results indicate that the first step to improve the network requires replacing one of the old TDRS transponders for a new FRFT transponder that can provide communications up to 600Mbps. This increases the overall benefit of the system by 5% while slightly reducing its cost (-2%). The cost reduction is achieved thanks to the spacecraft mass reduction when fewer S-band transponder

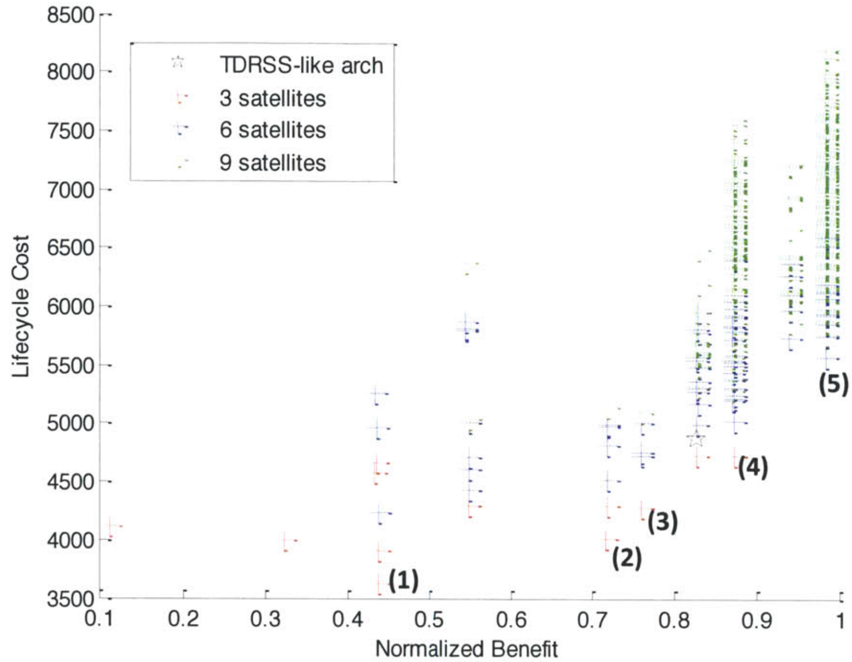


Figure 4-3: Transponder allocation

electronics are put into orbit. Note that this evolutionary step of the SN architecture has the advantage of leveraging mature TDRS technology to maintain a reliable and well proven S-band service.

Nevertheless, a network that is able to satisfy the requirements of the most demanding users requires optical communications. In that case, six satellites per generation have to be launched, three of them carrying an optical telescope (see figure 4-4). This alternative results in an 18% increase in the overall architecture score, but also incurs in an additional 25% cost.

4.3.3 Results: Procurement vs. hosted payloads

This analysis focuses on the impact of including hosted payloads as an alternative to fly part of the network asset. Figure 4-5 presents an extended representation of the previously presented tradespace where hosted payload architectures are included. The first immediate conclusion is that, based on the cost model from [13], hosted

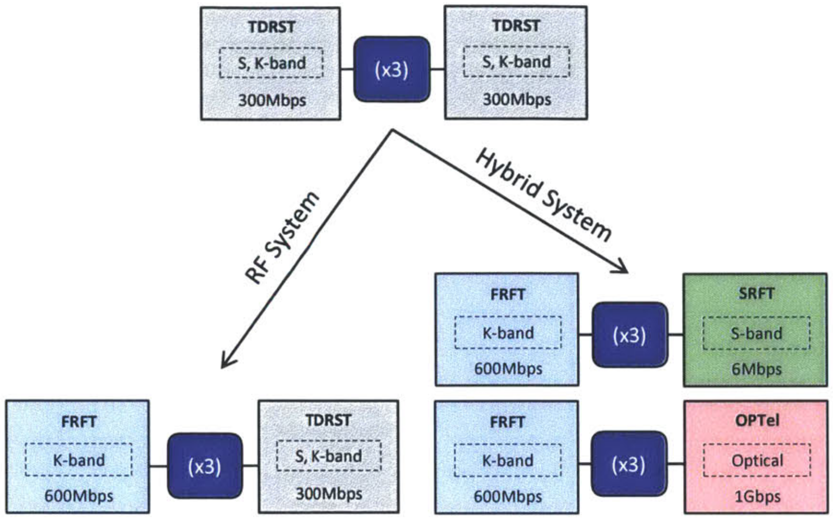


Figure 4-4: Evolution of a procured Space Network

payloads offer the possibility of significant cost reductions (between 15% and 30%). As a result, all architectures on the Pareto front rely solely on hosted payloads.

On the other hand, concerns about using hosted payloads for sensitive applications such as astronaut communications or contingency operations have been flagged as a potential problem. Being that the case, an optimal solution would entail using a portfolio of hosted payloads and privately owned satellites, where sensitive contacts would always be scheduled through the latter ones. Thus, building this portfolio can be done by analyzing what communication payloads result in higher life cycle cost savings when placed in a commercial satellite.

Figure 4-6 presents the tradespace of architectures that mix hosted and procured transponders. Results indicate that including hosted transponders in the system does not have a major impact on its performance but requires disaggregating the monolithic satellites identified as optimal in the previous case study. That being the case, the next questions arise: Which payload obtains better savings when being hosted? How many payloads do you want to fly in a hosted payload approach?

The answer to the first question can be obtained by analyzing the non-dominated architectures from figure 4-6 and estimating the cost savings of hosting different types of transponders. Results indicate that hosting an optical payload can potentially

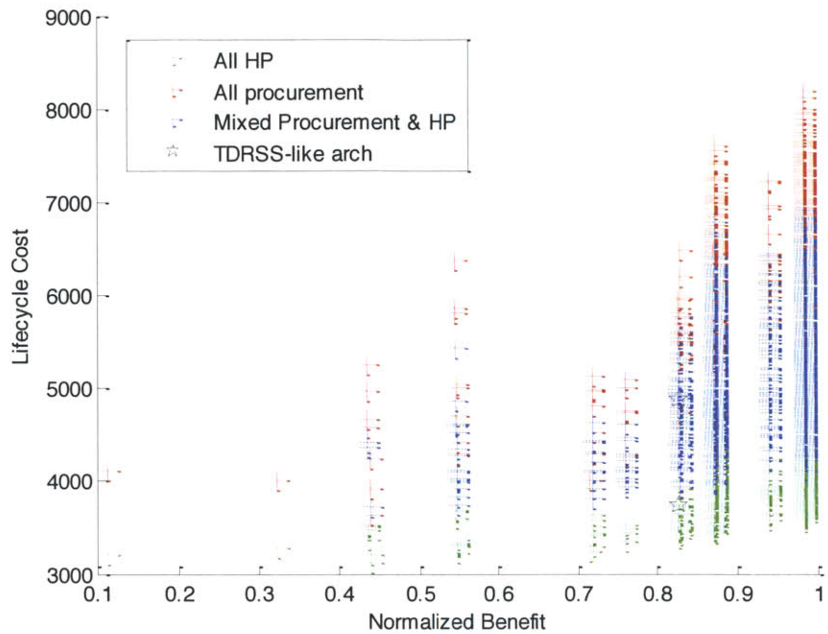


Figure 4-5: Procurement vs. Hosted payloads

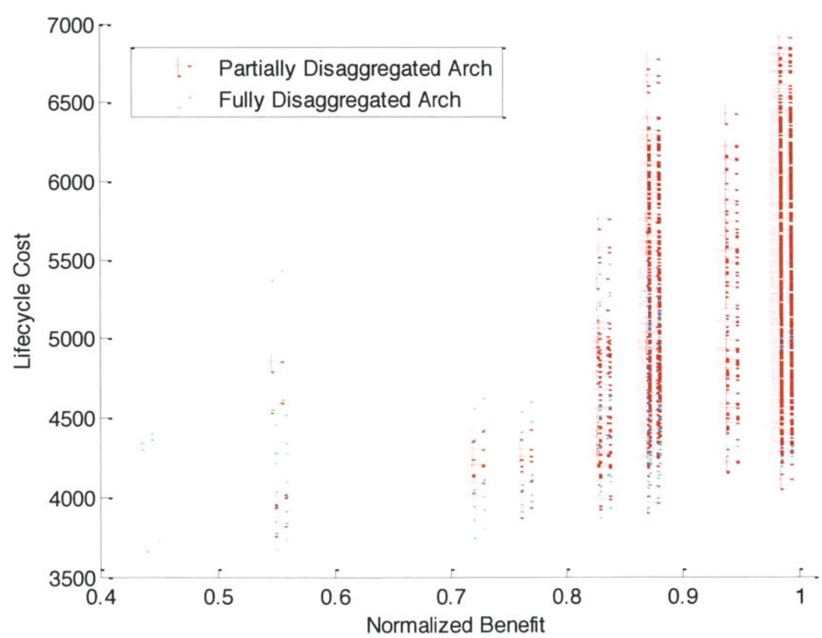


Figure 4-6: Combinations of procured and hosted payloads

save up to 28% of the cost, while a low data rate payload (S-band) only obtains a 16% cost reduction. Therefore, it seems to be more advisable to put high data rate payloads like optical terminals in commercial satellites and retain control of the S-band communications. This conclusion is particularly interesting because it partially addresses the open problem of handling sensitive information in a network with hosted transponders. In particular, since the hosted terminals should only be high data rate payloads, the network scheduling system ought to prioritize critical contacts through the owned asset even if they use lower data rates, and then use the hosted transponders as an augmentation system to relay other types of data.

Finally, figure 4-7 presents the tradespace of hosted payload architectures color coded by the number of antennas that are being hosted. The Pareto front indicates that almost all optical architectures host one single transponder as opposed to two in order to minimize the burden to the host platform. This burden is also contingent on the mass and power of the hosted communication payload. This is the main reason why optical terminals become so attractive for hosted payload architectures as opposed to low frequency systems that require big parabolic antennas and massive high power amplifiers.

4.4 Case study 2: Valuation of inter-satellite links

4.4.1 Tradespace definition

Table 4.3 lists the different SNAT architectural decisions for the the valuation of inter-satellite links. The basic assumptions for this case study are as follows:

- All satellites will be placed in geosynchronous orbit.
- Satellites can be deployed independently from one another, i.e. no replenishment cycle strategy is assumed.
- All satellites are procured, launched and operated by NASA.

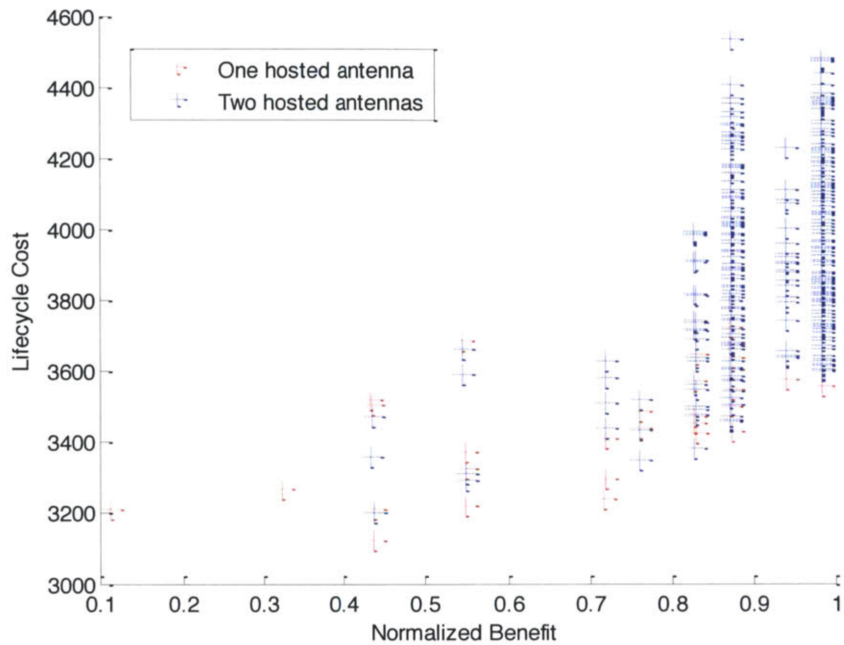


Figure 4-7: Number of hosted payloads per satellite

- All satellites use bent-pipe transponder technology. If a satellites is not in view of a ground station, then the bent-pipe transponder communicates to and from another relay satellite that can downlink the information to the ground.

- The current SN frequency and bandwidth allocations will be available in S, Ku and Ka-band.

- The default modulations for the links will be M-PSK, $M = \{2, 4, 8\}$. The current SN supports both 2-PSK (BPSK) and 4-PSK (QPSK), so it is assumed that the modems and beam formers on the ground stations will be modified to support 8-PSK. This upgrade is already being implemented under the Space Network Ground Segment Sustainment (SGSS) project [19]. Additionally, no coding scheme is assumed for the links (coding gain equal to $0dB$).

Additionally, three extra constraints were added in order to reduce the size of the tradespace and, at the same time, make sure to only evaluate sensible architectures. First, architectures with only one ground station must use White Sands, while architectures with two ground stations can only select White Sands and Guam. This constraint was added since the SN is already operating these sites and therefore it would not make sense to close an existing ground station to build another one in Madrid. On the other hand, Madrid was chosen as a possible site for a SN ground station since it provides excellent coverage of Europe and the east region of the Atlantic Ocean, and NASA already has a Deep Space Network complex that could host the SN antennas.

Second, architectures with satellites placed in orbital positions that neither have support of a ground station, nor have visibility with any other satellite are eliminated. As an example, consider an architecture with only White Sands and a *GEO-1-2* constellation where the first satellite is placed above New Mexico and the other at the Indian Ocean. The spacecraft above the Indian Ocean is never in view of White Sands, nor does it have a viable link to the other spacecraft in the network. Therefore, with bent-pipe transponder technology data sent to that relay can never be downloaded to the ground.

Finally, the third constraint is related to the spacecraft configuration complexity. Similar to the previous case study, a maximum of two RUL transponders are allowed per satellite. With these constraints in place, the resulting tradespace contains 2808 candidate architectures.

4.4.2 Results

Figure 4-8 presents the resulting tradespace of evaluating the 2808 architectures defined in 4.4.1. Architecture (1) achieves the maximum benefit (94%) through a combination of two constellations of three satellites. The first only one carries a TDRS SA antenna while the other one combines it with an optical telescope. The space segment is supported from the ground through a single ground station, White Sands, that provides service to the entire fleet of satellites through inter-satellite links. This last feature is shared by all the architectures in the Pareto front (see table 4.4).

On the other hand, the vertical clusters in the tradespace are mainly driven by the selection of payloads and number of deployed satellites. Architectures that are

Decision	Range of values
Orbit selection	GEO
Constellation design	Number of planes: 1 Number of sats per plane: 2, 3 or 4
Inter-Satellite Link payload allocation	Yes or no
Transponder-to-spacecraft allocation (disaggregation)	All the possible partitions of N transponders into $1 \leq N_{sat} \leq 3$ satellites
Ground station	Subset of White Sands, Guam or Madrid
Contract modality	100% procurement
Transponder selection	SA (x2), Optical (x1)
Transponder technology selection	bent-pipe

Table 4.3: Case study architectural decisions

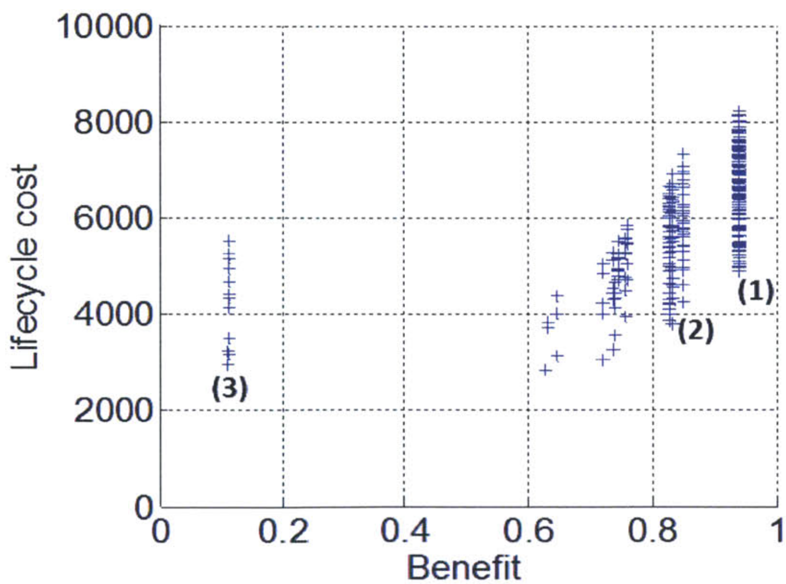


Figure 4-8: Benefit-cost tradespace

	Space segment	Ground segment	ISL
(1)	[SA + Optical] (x3) [SA] (x3)	White sands	All satellites
(2)	[SA + Optical] (x4)	White sands	All satellites
(3)	[Optical] (x3)	White sands	All satellites

Table 4.4: Pareto front architectures

missing any of the frequency bands, or that have very few satellites, are incapable of satisfying all the customer requirements. In particular, all-optical architectures do not provide service to the vast majority of users that demand S-band, Ku-band, and Ka-band contacts. As a result, this architecture only obtains a normalized benefit score of 11%.

Figure 4-9 presents the same tradespace color-coded according to the number of satellites. Additionally, the marker type is used to indicate the ground stations that are supporting the constellation of relay spacecraft. All the architectures in the resulting Pareto front have between 3 and 6 satellites, thus indicating that launching more relays only decreases the utilization of the communication transponders without improving the service provided to the customers. This is consistent with the findings from section 4.3.2 and figure 4-4, where results proved that at most six satellites should be launched in order to satisfy the expected demand for the 2020-2030 time frame.

Note also that there is a consistent stratification between the architectures that use one, two and three ground stations. In particular, it can be seen that the cost of adding the new ground station in Madrid renders all the architecture that contain this site completely dominated. This fact indicates that the extra cost of building a new ground station vastly exceeds the development cost and extra on-orbit mass of a system that uses inter-satellite links. This is further reinforced by figure 4-10, where high performing architectures mostly rely on White Sands as the only operating ground station and then use inter-satellite to provide downlink capabilities for all satellites in the network.

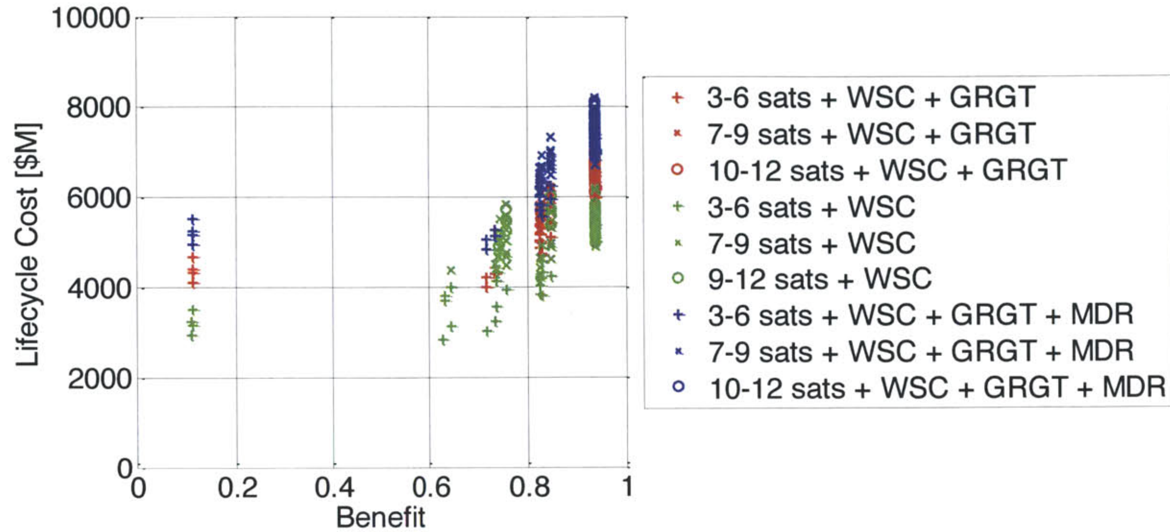


Figure 4-9: Detail of the number of satellites and ground stations

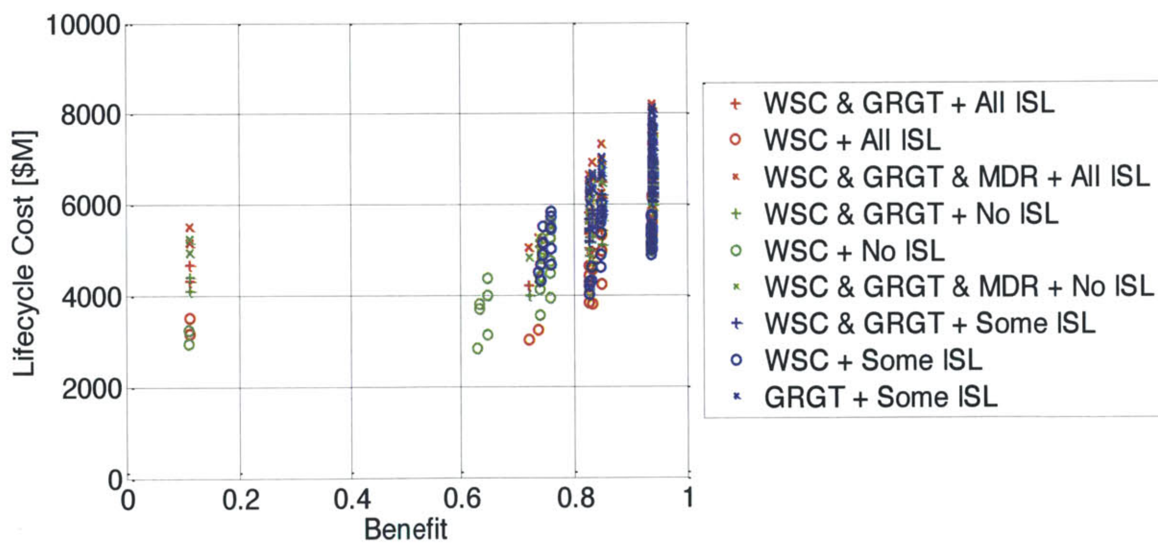


Figure 4-10: Detail of inter-satellite links and ground station selection

However, it is important to note that most architectures with that configuration are heavily constrained by the RF space-to-ground bandwidth allocation. Since satellites with an operational SGL connection have to downlink both their information and data coming from other relay satellites (through the ISLs), they have to support an extremely capable SGL. In most cases, this forces the link to take advantage of optical technology, therefore rendering it highly sensitive to atmospheric attenuation and cloud coverage. Previous studies have proven that in order to guarantee a 99.9% SGL availability 10 small ground stations have to be scattered to achieve spatial diversity [20]. This fact is currently not modeled in SNAT and could potentially increase the cost of the ground segment by several orders of magnitude, thus changing the findings of this case study.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 5

Conclusions

5.1 Thesis summary

This thesis introduced a new tool to architect space communication networks. The first chapter started by motivating the need for this kind of tools in order to assist in the evolution of current space networks. It argued that architecting a network can be viewed, at the highest level, as a decision making process where both technical and non-technical factors have to be taken into account in order to capture the complexity of the problem.

With this framework in mind, a literature review of past approaches to architect and synthesize space network architectures was conducted. It highlighted the trade-off between modeling breadth and simulation depth, and identified four main approaches to the design of space networks: First, network simulators, where the main emphasis is to model the elements of the network as accurately as possible in order to approximate the exact operations and performance of the system. Second, point designs, where requirements for the network are quantified and benchmarked against a handful of alternative architectures. Third, architecture studies, where the primary goal is to state the different feasible alternatives to architect the system without necessarily providing quantitative evaluation of their performance or cost. And fourth, tradespace exploration, where the space of feasible network architectures is enumerated and evaluated against a set of predefined metrics in order to identify optimal designs.

Chapter two provided a summary of the Space Network Architecting Tool (SNAT), a computational tool that is tailored to architecting future space communication networks within the near Earth domain. It first discussed the set of decisions that are available to the network architect from a generic standpoint and then formalized them into eight architecting decisions to be used in the tool. Then, it provided a detailed description of the different modules that are used to size the communication payloads and antennas that compose the space and ground segment. Next, it introduced a rule-based scheduler that was specifically developed to rapidly approximate the operations of the network and, in doing so, assess its overall capacity. And finally, it provided an overview of the cost estimation module and the search engines that are respectively used to provide estimates of the system life cycle cost and explore the space of alternative network architectures.

The third chapter of this thesis was devoted to the validation of SNAT. SN operational information was first gathered and analyzed in order to understand the current performance of the system. Based on this analysis, a typical day of operations and a high load day of operations were identified and accurately modeled as inputs to the tool. The results of using SNAT to model the SN were then benchmarked against the actual performance of NASA's network. They demonstrated that the tool can successfully approximate the capacity and life cycle cost of the system with a 10% and 15% error respectively. Thus, it was argued that the tool was ready to explore the space of alternative network designs and provide recommendations from its results.

Finally, the last chapter was devoted to two case studies related to the evolution of the SN. First, SNAT was used to evaluate the importance of including new RF and optical technology as a means to increase the network capacity and accommodate new customer communication requirements. This study was coupled with the possibility of flying network assets as hosted payloads in order to reduce the cost of the system. On the other hand, the second case study focused on the valuation of inter-satellite links. The current SN is based on a bent-pipe approach where all relay satellites act as mirrors that directly downlink the information to a supporting ground station. This forces NASA to maintain and regularly upgrade ground stations outside the

Continental US (CONUS), therefore drastically increasing the cost of the ground segment. This second case study analyzed the possibility of using inter-satellite links between relay spacecraft as an alternative to alleviate this problem and reduce the number of supporting ground stations.

5.2 Main contributions

The main objective of this thesis was to identify space network architectures that better address the needs of future near Earth space missions by (1) characterizing their communication needs, (2) identifying the set of decisions to architect a space communication network, and (3) exploring the space alternative of network architectures. The accomplishment of theses goals has led to both methodological and case-specific contributions.

5.2.1 Methodological and modeling contributions

The following contributions can be identified:

- **Tradespace definition**: The space of alternative architectures for space communication networks has been characterized as a three axis decision making process. The first axis defines the network topology depending on the degree of connectivity between the different relay assets. The second axis is related to the contractual agreement between the service provider and the network customer. Finally, the third axis characterizes the communications architecture, from the physical to the network and transport layers.

- **Problem formulation**: The space network architecting problem has been formulated as a multi-objective optimization problem coupled with decision making process. Two approaches have been used to solve the problem: A full factorial search of the design space, and a heuristic search based on a genetic algorithm. In both cases, the decision making process has been automated by codifying the decisions into numerical variables and letting the optimization algorithm assign values to each decision in order to generate new valid network architectures.

- **Model implementation**: SNAT has been proposed as a systematic tool to tackle the problem of architecting a space communication network. The tool first decomposes the problem into eight architectural decisions that, when coupled, can be used to formally enumerate the space of network designs. Then, it describes the network as a set of satellite constellations, each one flying cooperatively under different configurations and with different communication technologies. Their capabilities dictate the level of service that can be provided to the network customers and, in turn, allow direct assessment of their satisfaction. On the other hand, constellations can be procured and launched privately, they can be flown as hosted payloads on-board commercial spacecraft, or they can be a commercial system from which services are leased. This determines the cost of the building and operating each constellation and, ultimately, the entire network.

5.2.2 Findings from the case studies

The conclusions drawn from applying SNAT to architecting the future NASA Space Network are summarized below:

- **Customer characterization**: The communication requirements for missions in the 2020-2030 time frame have been characterized according to the type of service required and mission class. Two types of mission are expected to be drivers for the network capabilities: Exploration missions that demand up to 40Tbit/day to achieve their scientific goals with very low or moderate latency. Human spaceflight missions that require continuous contact throughout the day at low or moderate data rates to obtain voice and video services.

  Additionally, the network will also have to support a variety of medium sized missions that, despite their limited data volume requirements, currently account for a large fraction of the SN transponder utilization. Combining them with the high demanding missions yields a customer forecast of 15 to 20 missions, with an expected total data volume of 130Tbits/day (a 270% increase from the current load). Continuous coverage of low Earth orbit will continue to be critical since both human space flight missions and exploration missions will require low latency reliable communications.

- **Technology infusion and hosted payloads**: Improved Ka-band transponders and optical technology will be required in order to return the expected amount of scientific data produced by the next decade robotic missions. In particular, the inclusion of new RF technology to double the current SN capabilities will increase the network performance by a 5%. However, the most demanding missions will require links at more than 1Gbps, a challenging data rate given the current Ka-band bandwidth allocations.

  The proposed solution to service these high demanding missions is to incorporate one optical telescope in the system. This would be coupled with a new RF transponder to obtain three high data rate relay satellite to be placed at geosynchronous orbit. They would be accompanied by three smaller satellites supporting both S-band and Ku, Ka-band communications. Nevertheless, the main problem of this network configuration is that it requires to regularly procure and launch six satellites. This increases the cost of the network by a 25%, but fully satisfies all potential customers.

  Mitigating this problem can be achieved by including hosted payloads in the system. In particular, hosting the optical terminal can reduce the cost of this asset by as much as 28%. This option has been found to be preferable to hosting legacy S-band transponders both from an economical (16% savings) and operational perspective (maintaining ownership of S-band transponders ensures reliable contingency communications to spacecraft in distress).

- **Inter-satellite link valuation**: Inter-satellite links are identified as the preferred alternative for providing support to a constellation of relay satellites. Results demonstrate that the added cost of building, operating and upgrading a ground station far exceeds the additional costs of flying a transponder that provides inter-satellite communications.

5.3 Future work

Several areas of improvement are possible in order to extend SNAT and refine its outputs. In particular, the following items have been identified as important additions to the tool:

- **Modeling deep space customer**: The motivation of this thesis introduced the concept of a unified network that could provide seamless communication and navigation services to both near Earth and deep space missions. However, SNAT's decisions and modeling capabilities were specifically tailored to architecting near Earth space communication networks. Therefore, a potential extension to the tool would be incorporate the requirements from deep space missions and use the tool to identify commonality strategies that reduce the economical burden of operating two separate near Earth and deep space networks.

- **Network financial analysis**: The SN currently operates under a ten year replenishment cycle where funds for procuring new satellites and launch vehicles are uncertain despite being thoroughly planned by the program management. Including these cycles in SNAT and quantifying the yearly expenditure for the network would allow trading architectures where budgetary spikes are required versus ones where it is possible to support the network with approximately flat expenditure profiles.

- **Network reliability analysis**: The goal of the reliability analysis within SNAT is to quantify the effects of losing network assets with respect to its quality of service. This reliability analysis could be implemented using a Monte Carlo simulation that, for a given network architecture, assigns probabilistic failure profiles to different network assets and randomly generates degraded scenarios. Each of these degraded scenarios would have part of the network inoperable.

- **Optical ground stations**: Optical links to ground stations are heavily constrained by atmospheric attenuation and clouds. The best strategy to combat their effects is to use spatial diversity, i.e. place several ground stations at multiple geographical locations where weather patterns are uncorrelated. The current implementation of SNAT does not take into account this fact, thus potentially underestimating the complexity and cost of a system where optical links to the ground are required.

THIS PAGE INTENTIONALLY LEFT BLANK

Bibliography

[1] Marco Adinolfi and Amedeo Cesta. Contributed Paper Heuristic Scheduling of the DRS Communication System. 8, 1995.

[2] National Aeronautics and Space Administration. Tdrs-k media kit. January 2013.

[3] Bjørn Roger Andersen, Odd Gangaas, and Jens Andenæs. A dvb/inmarsat hybrid architecture for asymmetrical broadband mobile satellite services. *International Journal of Satellite Communications and Networking*, 24(2):119–136, 2006.

[4] Henry Apgar. Cost Estimating. In *Space Mission Engineering: The new SMAD*, chapter 11. Microcosm, Hawthorne, CA, 2011.

[5] Larry Baranyai, Enrique G Cuevas, Scott Davidow, Christopher Demaree, and Paul DiCaprio. End-to-end network modeling and simulation of integrated terrestrial, airborne and space environments. In *Aerospace Conference, 2005 IEEE*, pages 1348–1353. IEEE, 2005.

[6] Brian Barritt, Kul Bhasin, Wesley Eddy, and Seth Matthews. Unified approach to modeling & simulation of space communication networks and systems. In *Systems Conference, 2010 4th Annual IEEE*, pages 133–136. IEEE, 2010.

[7] Kul Bhasin, Anthony W Hackenberg, Richard Slywczak, Prasanta Bose, Marcos Bergamo, and Jeffrey Hayden. Lunar relay satellite network for space exploration: architecture, technologies and challenges. In *Proceedings 24nd AIAA International Communications Satellite Systems Conference*, pages 12–14, 2006.

[8] Kul Bhasin and Jeffrey L Hayden. Space internet architectures and technologies for nasa enterprises. In *Aerospace Conference, 2001, IEEE Proceedings.*, volume 2, pages 2–931. IEEE, 2001.

[9] Kul B Bhasin, Jeffrey L Hayden, Thomas Sartwell, Ronald A Miller, and John J Hudiburg. Integrated network architecture for nasa's orion missions. In *SpaceOps 2008 Conference*, 2008.

[10] Kul B Bhasin, Chuck Putt, Jeffrey Hayden, Shirley Tseng, Abi Biswas, Brian Kennedy, Esther Jennings, Ron Miller, John Hudiburg, Dave Miller, et al. Architecting the communication and navigation networks for nasa's space exploration

systems. In *System of Systems Engineering, 2007. SoSE'07. IEEE International Conference on*, pages 1–6. IEEE, 2007.

[11] Robert S. Bokulic, Christopher C. DeBoy, Scott W. Enger, Jeremiah P. Schneider, and Joseph K. McDermott. Spacecraft Subsystems IV Communications and Power. In *Space Mission Engineering: The new SMAD*, chapter 21. Microcosm, Hawthorne, CA, 2011.

[12] Edward F. Crawley. System architecture. http://ocw.mit.edu/courses/engineering-systems-division/esd-34-system-architecture-january-iap-2007/.

[13] Kwon D. W. & Shannon P. Davidson, A. Pricing a hosted payload. In *Aerospace Conference, 2012 IEEE*, 2012.

[14] Olivier de Weck, Richard de Neufville, and Mathieu Chaize. Enhancing the economics of communications satellites via orbital reconfigurations and staged deployment. *American Institute of Aeronautics and Astronautics*, 2003.

[15] Olivier L De Weck, Richard De Neufville, and Mathieu Chaize. Staged deployment of communications satellite constellations in low earth orbit. *Journal of Aerospace Computing, Information, and Communication*, 1(3):119–136, 2004.

[16] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and TAMT Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *Evolutionary Computation, IEEE Transactions on*, 6(2):182–197, 2002.

[17] Exploration and Space Communications Division. Space network user's guide (snug). Technical report, NASA, 2007.

[18] Committee for a Decadal Survey of Astronomy and Astrophysics. New worlds, new horizons in astronomy and astrophysics. Technical report, National Research Council of the National Academies, 2010.

[19] Thomas Gitlin and Keith Walyus. Nasas space network ground segment sustainment project preparing for the future.

[20] Optical Link Study Group. Optical link study group final report. Technical report, NASA, 2012.

[21] Space Communication Architecture Working Group. Nasa space communication and navigation architecture recommendations for 2005-2030. Technical report, NASA, 2006.

[22] Systems Engineering Handbook. *The Art of Systems Architecting.* Third edition, 2006.

[23] Esther Jennings and David Heckman. Performance characterization of space communications and navigation (scan) network by simulation. In *Aerospace Conference, 2008 IEEE*, pages 1–9. IEEE, 2008.

[24] Cyrus D. Jilla. *A Multiobjective, Multidisciplinary Design Optimization Methodology for the Conceptual Design of Distributed Satellite Systems.* PhD dissertation, Massachusetts Institute of Technology, Department of Aeronautics and Astronautics, May 2002. This is a full PHDTHESIS entry.

[25] Benjamin H. Y. Koo, Willard L. Simmons, and Edward F. Crawley. Algebra of Systems: A Metalanguage for Model Synthesis and Evaluation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 39(3):501–513, May 2009.

[26] Wiley J Larson, James Richard Wertz, et al. *Space mission analysis and design.* Microcosm, 1992.

[27] Mark W. Maier and Eberhardt Rechtin. *The Art of Systems Architecting.* CRC Press - Taylor and Francis Group LLC, Boca Raton, Florida, third edition, 2009.

[28] Kevin McCarthy, Frank Stocklin, Barry Geldzahler, Daniel Friedman, and Peter Celeste. Nasas evolution to ka-band space communications for near-earth spacecraft. In *Proceedings of the SpaceOps 2010 Conference*, 2010.

[29] NASA. Space communications and navigation (scan) network architecture definition document (add). Technical report, NASA, 2011.

[30] http://classwww.gsfc.nasa.gov/gsams/greet_spectrumAllocationSummary.htm.

[31] http://nsnam.isi.edu/nsnam.

[32] http://web.scalable-networks.com/content/qualnet.

[33] Madwaraj Rao, Sreeram Ramakrishnan, and Cihan Dagli. Modeling and Simulation of Net Centric System of Systems Using Systems Modeling Language and Colored Petri-nets : A Demonstration Using the Global Earth Observation System of Systems. *Systems Engineering*, 11(3):203–220, 2008.

[34] Marc Sanchez, Daniel Selva, Bruce G. Cameron, Edward F. Crawley, Bernard Seery, and Antonios Seas. Results of the mit space communication and navigation architecture study. In *2014 IEEE Aerospace Conference*, Big Sky, Montana, March 1-8 2014. Institute of Electrical and Electronics Engineers.

[35] Marc Sanchez, Daniel Selva, Antonios Seas, Bernard Seery, Bruce G. Cameron, and Edward F. Crawley. Exploring the architectural trade space of nasas space communication and navigation program. In *2013 IEEE Aerospace Conference*, Big Sky, Montana, March 3-10 2013. Institute of Electrical and Electronics Engineers.

[36] James S Schier, John J Rush, W Dan Williams, and Pete Vrotsos. Space communication architecture supporting exploration and science: Plans and studies for 2010-2030. In *1st Space Exploration Conference: Continuing the Voyage of Discovery*, volume 30, 2005.

[37] Graeme B. Shaw. The generalized information network analysis methodology for distributed satellite systems. Master's project, Massachusetts Institute of Technology, Department of Aeronautics and Astronautics, February 1999. This is a full MASTERSTHESIS entry.

[38] Afreen Siddiqi, Jason Mellein, and Olivier L De Weck. Optimal reconfigurations for increasing capacity of communication satellite constellations. In *46 th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, pages 1–13, 2005.

[39] Willard Lennox Simmons. *A Framework for Decision Support in Systems Architecting.* PhD dissertation, Massachusetts Institute of Technology, Department of Aeronautics and Astronautics, February 2008. This is a full PHDTHESIS entry.

[40] Philip Springmann and Olivier de Weck. Parametric scaling model for non-geosynchronous communications satellites. *Journal of spacecraft and rockets*, 41(3):472–477, 2004.

[41] http://www.agi.com/products/stk/modules/default.aspx/id/stk-free.

[42] et al. Tran, J. J. Evaluating cloud computing in the nasa desdyni ground data system. In *Proceedings of the 2nd International Workshop on Software Engineering for Cloud Computing*, 2011.

[43] Sunny Tsiao. *Read you loud and clear! The story of NASA's spaceflight tracking and data network.* Library of Congress, Washington DC, 2007.

[44] Daniel Selva Valero. *Rule-Based System Architecting of Earth Observation Satellite Systems.* PhD dissertation, Massachusetts Institute of Technology, Department of Aeronautics and Astronautics, June 2012. This is a full PHDTHESIS entry.

[45] Tim Weilkiens. *Systems engineering with SysML/UML: modeling, analysis, design.* The Morgan Kaufmann/OMG Press, Heidelberg, Germany, 2006.

[46] Haitao Yang, Xinmin Wang, and Hongli Zhao. The modeling, simulation and effectiveness evaluation for communication networks of multi-layer satellites constellation. In *Computational Intelligence and Security, 2006 International Conference on*, volume 2, pages 1055–1060. IEEE, 2006.

[47] Hubert Zimmermann. Osi reference model–the iso model of architecture for open systems interconnection. *Communications, IEEE Transactions on*, 28(4):425–432, 1980.