Networking in Heaven as on Earth

Tobias Klenze  
ETH Zurich  
tklenze@inf.ethz.ch

Giacomo Giuliani  
ETH Zurich  
giacomog@inf.ethz.ch

Christos Pappas  
ETH Zurich  
pappasch@inf.ethz.ch

Adrian Perrig  
ETH Zurich  
aperrig@inf.ethz.ch

David Basin  
ETH Zurich  
basin@inf.ethz.ch

ABSTRACT
The Internet will undergo a major transformation as satellite-based Internet service providers start to disrupt the market. Constellations of hundreds to thousands of satellites promise to offer low-latency Internet to even the most remote areas. We anticipate exciting business and research opportunities.

Motivated by the potential of the new satellite networks, we describe business and interconnection models for space-operating ISPs and how they could be integrated into the backbone of today’s Internet. At the same time, in view of the high risk of these ventures, we study constellations under partial deployment. We take the SpaceX constellation as an example and show that even at 10% deployment it offers a high level of connectivity to most areas. However, this connectivity is intermittent, which raises challenges for integrating satellite networks into the Internet backbone.

1 INTRODUCTION
We are witnessing a new space race—this time not among countries or ideologies, but between private corporations. The advent of cheap re-usable rockets, coupled with the miniaturization and mass production of space hardware has been dubbed “NewSpace” [1]. Recently launched CubeSats constellations have already disrupted the Earth imaging industry [2], and “NewSpace” companies have now set their eyes on disrupting the Internet market. These companies plan to launch mega-constellations of hundreds to thousands of communication satellites into low Earth orbit (LEO) in the coming years. Their proposals are already past the first stage of regulatory approval: SpaceX, OneWeb, Space Norway, and Telesat have secured RF spectrum from the FCC for their constellations [14, 16–18].

Satellite networks (SNs) will bring low-latency broadband connectivity even to the most remote areas. This is possible due to appealing properties of the proposed constellations.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

HotNets-XVII, November 15–16, 2018, Redmond, WA, USA
© 2018 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-6120-0/18/11.
https://doi.org/10.1145/3286062.3286066

First, placing communication satellites in LEO (335-1,400 km altitude), compared to the more common geostationary orbit (GEO; 35,786 km altitude), reduces the propagation delay immensely, achieving a round-trip-time between Earth and satellites in the sub-10ms range, in stark contrast to the 240ms offered by today’s GEO satellites. Second, most of the proposed SNs have high-capacity intersatellite links, creating a network in space; existing satellites only act as repeaters between two ground stations (a bent-pipe architecture). Since light travels faster in vacuum than in fiber, satellite latency will beat fiber for longer distances between communicating end points [22]. While the bandwidth of each satellite is at most just a few Gbps (for now), the aggregate capacity of a SN is already expected to reach multiple Tbps.

Universal low-latency broadband access will significantly disrupt the Internet market. Yet, the effect of SNs will not be restricted to providing connectivity to remote users. An intriguing idea is to integrate SNs into the Internet backbone to establish an interconnection market that extends and operates in space similar to its terrestrial counterpart. We devote Section 2 to describing a possible space-enabled interconnection market, where SNs sell connectivity to terrestrial ISPs, or act as a globally reachable Internet exchange point.

While integrating SNs into the Internet backbone offers exciting opportunities, it also leads to risks and new challenges. To offer a near-global footprint (area of coverage), an entire satellite constellation is needed, since the footprint of a single satellite in LEO is limited and covers a ground station only for a few minutes. However, requiring a fully-deployed constellation before offering any service significantly raises the bar for success [32] as previous business failures in the market have shown, e.g., Globalstar and Teledesic. As a response, the aerospace community has looked into reducing the risk by optimizing deployment strategies and incorporating multiple deployment stages [25].

In this paper, we look into partial deployment of constellations and its effect on connectivity for ground stations; by partial deployment, we mean any state of deployment prior to reaching the final fully deployed constellation. Our results are based on the SpaceX constellation at 10% deployment of its initial plan of 1,600 satellites. Then, based on our connectivity analysis, we look at the potential effects on the Border Gateway Protocol (BGP).
Overall, our results indicate that even for early deployment stages, highly-populated areas will enjoy a high level of connectivity on average. On the flip side, connectivity is intermittent with short—yet frequent—disconnection bursts that would become a problem if exposed to the interdomain environment.

With this paper, our goal is to point the community to potential research opportunities that arise from the latest advancements in the aerospace industry. To this end, we consider making SNs an integral part of the Internet backbone infrastructure and conduct a small feasibility study under the constraints of partially deployed satellite constellations. We conclude by discussing open problems.

2 INTERCONNECTION MODELS

We envision the Internet market to extend into space and operate under similar principles as it does on Earth. With this paper, we do not suggest a single modus operandi for satellite-based Internet. Rather we believe that there are exciting—foreseeable and unforeseeable—business opportunities for Internet stakeholders. In this section, we explore how existing Internet businesses could use space and how SNs can play multiple roles in the interconnection market in an effort to attract traffic and customers.

Figure 1 illustrates an Internet ecosystem with integrated SNs, with the focus on the control plane. We refer to sessions between border routers of different autonomous systems (ASes) as exterior gateway protocol (EGP) sessions and between border routers of the same AS as interior gateway protocol (IGP) sessions. The SN consists of a ground segment (gray shaded circles) and the space segment, i.e., the constellation. The routing logic (reachability information and policies) of the SN is implemented in a logically centralized manner on the ground; the SN only forwards traffic between ground stations (GSTs). We will use this figure to explain the following interconnection models.

**Last-mile providers.** The most common business model for satellite-based Internet is to act as a last-mile provider (not shown in Figure 1). Connecting customers at locations that are remote or do not have sufficient network infrastructure is already the business of satellite operators. Existing systems rely on satellites in GEO or medium Earth orbit, but proposed LEO systems can also act as last-mile providers. Internet service is also offered to end points in mobile environments, such as airplanes and ships when communication over terrestrial networks is not possible.

From the interdomain’s routing perspective, there are no major hurdles to implementing this business model. The SN acts as a typical eyeball ISP today, i.e., it owns IP prefixes and an AS number. Since it must connect its customers to the rest of the Internet, it must peer with other ISPs and buy transit from ISPs that provide global reachability; then the SN advertises its IP prefixes to the other networks. The routing between GSTs and between satellites is managed through custom intradomain routing protocols (see Section 5).

**Transit providers.** A more sophisticated interconnection model is when the SN acts as a transit provider that sells connectivity to other ASes, instead of end customers. This provides opportunities to regional networks or enterprises that have no option but to buy service from a single ISP in their area. Since the SN acts as a transit provider, it must itself connect with ISPs that have global reachability (Tier-1).

This approach has the potential to disrupt today’s Internet market. Today, lacking alternatives, single-homed networks do not participate in interdomain routing. However, given the nearly global coverage of SNs, single-homed networks in remote areas could become multi-homed. The benefits are manifold: the increased path diversity and redundancy translates to higher availability and performance, and the network operators gain more traffic engineering capabilities since such networks can now participate in interdomain routing and influence how to reach and be reached by other networks.

We illustrate the above with an example based on Figure 1. AS C is multihomed and has two providers: AS E which is an SN in LEO and AS A that is a provider with a satellite in GEO. Whenever the low-latency path via AS E is lost, AS C could still be reachable over AS A. Consider, for instance, a large cruise ship that could operate its own AS and offer better performance and higher reliability to its passengers. Cruise ships today operate large networks and host an enormous IT infrastructure (up to a few datacenters) [35], yet all satellite traffic is tunnelled through a single (or very few) terrestrial base stations across the globe, contributing to very high end-to-end latencies.
Internet exchange points. An exciting business case that we envision is that of a globally reachable Internet exchange point (IXP). The SN can become a globally reachable rendezvous point where parties can interconnect and exchange traffic.

This approach enables ASes that want to exchange traffic, but have geographically dispersed points-of-presence (PoPs), to peer directly. The interconnecting ASes would have to pay the SN for offering its infrastructure, similar to today’s business model. From a technical perspective, the SN is not acting as a separate AS, but rather as a switch (or multiple switches) in space. For example, in Figure 1, AS C can peer with AS D and directly exchange traffic instead of using their upstream ISPs.

3 FEASIBILITY STUDY

In this section, we assess the feasibility of incorporating SNs into the Internet backbone infrastructure. We start by presenting challenges (Section 3.1) and then we quantify the effect of partial deployment (Section 3.2).

3.1 Challenges

Incorporating SNs into the core of today’s networking infrastructure comes with challenges.

At the intradomain level, the constantly moving constellation results in a dynamic network environment. The relative movement between satellites in different orbits requires customized intersatellite routing protocols for optimal path discovery between source and destination satellites. Similarly, the link-layer handover between GSTs and satellites is frequent, since GSTs are within the footprint of a certain satellite only for a few minutes.

At the interdomain level, this dynamic environment will hit BGP where it hurts most: its (in)stability and long convergence time in the face of unstable paths [23, 24]. A single event such as a GST-satellite link failure can trigger a cascade of BGP updates globally as BGP speakers are exploring alternate paths. In addition, certain design choices and errors in BGP implementations have often led to an alarming rate of unnecessary BGP updates. As a response, route-flap dampening has been introduced to limit the bandwidth and CPU overhead of update floods [33].

Exposing SNs to the Internet backbone could further aggravate the already fragile situation. Satellite networks would introduce new sources of path instability due to the volatile nature of satellite connectivity. A source of path instability that we study here is that of partially deployed satellite constellations. Assuming that a satellite constellation is fully deployed, it would offer full-time connectivity to the promised area of coverage.¹ However, a SN could start offering Internet service much earlier. For example, the SpaceX constellation has a first deployment stage with 1,600 satellites in orbit [18], out of the 4,425 planned for the final constellation. The timeline for the first deployment stage is six years, but limited connectivity could be offered with far fewer satellites [6, 31].

The partial constellation deployment translates to a reduced footprint resulting in intermittent connectivity. A GST may be connected most of the time, but with disconnection bursts lasting from seconds to minutes. If this link intermittency is exposed to the interdomain ecosystem, paths through the SN must be withdrawn and then reannounced for every reconnection leading to path fluctuations. In the rest of the paper, we quantify link intermittency due to partial deployment and study the effect that it has on BGP announcements.

3.2 Evaluation

We evaluate the interplay between partially deployed SNs and BGP through simulation. Our first step is to analyze the connectivity between GSTs and the constellation (Section 3.2.1). Then, based on our connectivity analysis, we examine the effect of frequent connect-disconnect events on BGP announcements (Section 3.2.2).

For our evaluation, we have built a simulation framework for SNs. It can be used to simulate large constellations, their footprint on Earth, GST-satellite and intersatellite links, as well as the connectivity and latency of satellite paths. Our framework is generic and can simulate any constellation of satellites with circular orbits; here, our results are based on the SpaceX constellation. We have chosen SpaceX since their goals are not restricted to last-mile connectivity, but include routing the “majority of long distance Internet traffic” [8], for which interdomain routing challenges become interesting.

3.2.1 Ground-to-Satellite Connectivity. In a partially deployed constellation, depending on the location, a GST may experience full-time, partial, or no connectivity at all. By connectivity, we mean that a GST at a certain location is in the footprint of at least one satellite, and the connectivity profile for a given location is its connectivity over time. Since differences in longitude only introduce a constant time-shift in a location’s connectivity profile, we ignore it and reduce location to latitude.

Our analysis is based on the SpaceX constellation in its first stage of deployment. At the end of the first stage, the constellation will be composed of 1,600 satellites in circular orbit on 32 planes, at 1,150 km altitude, and with 53° inclination [30]. Since 1,600 satellites cannot be launched overnight, we simulate different partial deployment scenarios. For our evaluation, we present results that are obtained with the SpaceX constellation at 10% of its stage-one deployment, i.e., 160 satellites. We choose this setting since it is a realistic goal that can be achieved in the foreseeable future. Even this initial partial constellation has almost twice as many satellites as any existing constellation [4] and already provides connectivity for most areas and, in particular, those that are highly populated. Moreover, at 10% the constellation will already cover, on average, 82% of its final footprint. We distribute the 160 satellites into 8 equi-spaced orbits of 20 satellites each. We experimented with different settings for a 10% deployment and found that

¹Note that even a fully deployed constellation does not necessarily cover the Earth at all latitudes.
one to optimize coverage. Other parameters to the simulations are taken from SpaceX’s FCC application [30]. We abstract from the low-level details of connection initialization and handover protocols between satellites, and simply assume that GSTs within the footprint of a satellite are connected.

Results. Our results for the 10% partial deployment of the SpaceX constellation are shown in Figures 2–4. Figure 2 shows a snapshot of the global footprint of the constellation, given by the (partially overlapping) footprints of each satellite.

Figure 3 draws the connectivity profile of the earth’s surface by latitude. The left y-axis shows the portion of time during the day that regions (identified by latitude) are in the footprint of at least one satellite. The right y-axis shows the number of connect-disconnect events that a single GST would observe in a day (by latitude). Note that this includes only events in which a GST’s connectivity to the SN as a whole changes—the much more frequent handovers between satellites are not considered here. We highlight two observations: First, some of the most highly populated latitudes have connectivity most of the time (> 90%), while the other areas have an up-time of at least 50% (excluding latitudes entirely beyond reach of the SN). Second, as expected, there are no connect-disconnect events for fully covered areas. However, there can be hundreds of such events for areas that are partially covered, even where there is more than 90% connectivity. Figure 4 shows the connectivity profile for a latitude of 40° (solid line). While the uptime is on average more than 90%, there are many connect-disconnect events.

3.2.2 Effect on BGP announcements. We next study the effect of the connectivity profile on the number of announcements that a BGP speaker must make. We assume that a terrestrial AS has only one GST and therefore interdomain connectivity with the SN depends only on the up-time of one satellite link. Whenever the connectivity between the GST and the SN is lost, there is a cascade of BGP messages: the SN withdraws all prefixes of the terrestrial AS customer; the terrestrial AS withdraws all prefixes that were advertised through the SN and then announces new paths to its own customers. The number of BGP messages depends on the location of the failed link in the AS graph and on specific implementation details of the BGP speakers (e.g., batching withdrawal messages for multiple prefixes).

We present results for a GST that is positioned at 40° latitude. We chose this setting, as the 40th northern parallel traverses some of the most populated areas in the world (see Figure 2). As Figure 3 shows, such a GST has more than 150 connect-disconnect events per day. Even if each event triggered only a single BGP message, the fact that they occur in bursts (cf. Figure 4) will trigger route-flap dampening, preventing the propagation of BGP announcements.

To reduce the number of BGP messages, we apply a simple filtering mechanism. The connectivity profile of a GST is predictable based on the satellite orbits and can therefore be computed in advance. Connect events that are shortly followed (e.g., within minutes) by disconnect events are thus ignored, and the satellite path, while available, is not announced.
(a) The number of connect-disconnect events at 40° latitude and the percentage of connectivity waste as a function of the filtering threshold.

(b) The number of connect-disconnect events and the percentage of connectivity waste by latitude. The filtering threshold is set to 6 minutes.

Figure 5: The effect of the filtering threshold on the number of connect-disconnect events and the resulting connectivity waste.

This way, the number of effective events and thus BGP announcements is reduced. In contrast, disconnect events that are shortly followed by connect events cannot be ignored, since they impede communication. Exceptions to this include predictable outages in the order of milliseconds, which can occur during satellite handovers. For such durations, traffic could be buffered with a small impact on latency and latency variation. This filtering approach can drastically reduce the number of BGP announcements, but at the same time introduces connectivity waste: time periods that the SN could be used, but remains unused due to filtering.

Results. Figure 4 shows filtering being applied (dashed line) to the initial connectivity profile (solid line). Intervals shorter than 6 minutes are suppressed. In this example, more than 10 events of short connectivity bursts are filtered out in a time period of less than 2 hours. The filtering threshold introduces a tradeoff between BGP convergence and connectivity waste (similar to the half-time timer in route-flap dampening): a higher threshold implies fewer effective events, fewer BGP messages, and thus better convergence, but at the cost of a higher connectivity waste; and vice versa.

We quantify this tradeoff in Figure 5. Figure 5a shows that the number of announcements decreases as the filtering threshold increases, whereas connectivity waste increases. The figure also indicates that for the 40° latitude, most connectivity bursts have an interval of less than 5 minutes. However, setting the threshold to this value leads to a 20% loss of usable connectivity.

The situation is similar at other latitudes, as shown in Figure 5b. We fix the filtering threshold to 6 minutes (above the critical value of 5 minutes in the above example). It is clear that filtering greatly reduces the number of effective events, but with the connectivity waste ranging from 15% to 45%.

Overall, our results demonstrate that while the connectivity percentage is high on average, it is also intermittent and therefore it is challenging to incorporate SNs as an integral part into the backbone infrastructure. Even if one sacrifices 15-45% of uptime, close to 20 connect-disconnect events could occur per day for a single GST. This rate of connectivity interruptions would introduce a significant overhead for BGP speakers and negatively impact BGP convergence times. An important next step in our research agenda is to explore recent trends in networking (e.g., path-aware networking) and alternate interdomain networking architectures that improve on today’s architecture [29].

4 OPEN PROBLEMS

Many problems remain open or require better approaches to seamlessly integrate SNs into the Internet. We discuss a few interesting areas.

Dynamic Topologies. SNs have highly dynamic topologies, but we explored them only from the viewpoint of partial deployment and its effect on connectivity. Another aspect that contributes to the dynamic environment are failures, expected and unexpected. Satellites in LEO have a significantly shorter expected lifetime than their GEO siblings (5 to 7 years [15] as opposed to 15 years or more). Similarly, failures can also occur unexpectedly and can range from a failed networking component to physical satellite damage. Satellites in use today have been built with redundant systems to provide high reliability, but at a significant cost. The mass production of satellites can reduce the cost, but also start new trends such as the commoditization of space-based hardware. While reducing cost, the frequency of unexpected failures will increase. Given that satellite replacement is a non-trivial task that may take time due to legal, financial, or technical difficulties, the
networking community can contribute by designing resilient protocols that adapt to dynamic topologies.

**Fluctuating Bandwidth.** Another interesting topic is bandwidth fluctuation, which is attributed to two factors: the bursty nature of Internet traffic and the dependency of satellite link capacity on natural phenomena. Terrestrial network operators deal only with the former factor, but as satellite network operators compete over bandwidth spectrum, they are driven to bands beyond 10 GHz. At these frequencies, rain fade becomes a serious problem [28] and can significantly reduce bandwidth of GST-satellite links. Rain fade can be partially mitigated by increasing the power or gain of antennas, but this greatly increases the cost of the satellites or GSTs. An interesting alternative for ISPs would be to place GSTs in diverse locations and incorporate reliable weather predictions into their routing policies. A network architecture that is able to mitigate the effects of (predictable) bandwidth fluctuations caused by phenomena such as rain fade is by itself an interesting problem for future research.

Another source of bandwidth fluctuations is the use of narrow spot beams for the GST-satellite link, as proposed for some constellations [26]. Each spot beam provides connectivity for only a small area out of the satellite’s footprint. Depending on bandwidth saturation of each spot beam, a satellite may internally handover a GST to another spot beam. Similarly, in case of redundant satellite coverage, intersatellite handover is possible based on the available capacity at each satellite. The problem of optimal handover and of optimal alignment of spot beams to GSTs that maximizes the aggregate throughput is an interesting direction for future research.

**Effects on the Interconnection Market.** Integrating SNs into the Internet backbone raises research problems, as we have discussed so far. Moreover, additional interesting questions arise about the effects of such an integration. From an economic perspective, can global satellite coverage solve market failures that stem from the lack of competition between ISPs in certain areas? Increased competition will not only improve performance and reduce prices, but raise hopes for solving the regulatory mess around net neutrality [3]. At a technical level, what effect can satellite-based transit providers or satellite-based IXPs induce on the AS graph? Can they flatten the AS graph and shorten the average AS path length even more?\(^2\) Short AS paths can have interesting effects, such as reducing the effect of BGP prefix hijacks.

### 5 RELATED WORK

The idea of integrating a satellite system into the Internet backbone dates back to the Atlantic SATNET, interconnecting European research networks from 1979 to 1985 [21]. However, the interdomain ecosystem and the satellite landscape have changed drastically since then.

To integrate satellite constellations into the interdomain system, past work has suggested BGP modifications. Some of these proposals make strong assumptions such as no regular fluctuations in connectivity or latency between GSTs [12, 13]. Other proposals similarly assume a static connectivity profile, but for a geostationary setting [10, 11]. While geostationary satellites do not suffer from the connectivity fluctuations studied here, the high latency is an inhibiting factor for many use cases.

Previous satellite-routing proposals have focused mostly on intradomain challenges. Routing within a SN becomes challenging due to the constant topology change attributed to the relative movement between satellites and between satellites and the ground. Algorithms for mobile ad-hoc networks (MANETs) lend themselves to satellite routing since they are designed for rapidly changing topologies [27, 34]. Moreover, another line of work leverages the predictability of satellite orbits to propose special protocols for satellite routing [19, 20]. Satellites with intermittent connectivity have also been discussed in the context of delay-tolerant networking [9], and specifically for store-and-forward services.

In independent work that was developed in parallel, Bhattacherjee et al. focus on the potential of LEO constellations to reduce latency, especially for long-distance traffic [7]. Their results demonstrate that for long distances a SN reduces latency. In addition, it improves over the optimal fiber latency, i.e., the latency achieved if we were to lay a fiber cable along the geodesic between the distant locations. Furthermore, complementing our connectivity results, they show that increasing the density of the constellation reduces latency and latency variation, which is expected to be high in the early stages of deployment.

### 6 CONCLUSION

Satellite mega-constellations are in the making and their effect on the Internet market has yet to play out. While satellite networks provide exciting opportunities, they are also risky investments and therefore their viability largely depends on early revenue.

In this work, motivated by the potential of satellite networks, we describe ways that space-operating ISPs can become part of, and potentially disrupt, the interconnection market. At the same time, we evaluate the effect of leveraging satellite networks at very early stages of deployment. Our evaluation suggests that benefits are possible even for the early deployment stages as highly populated areas enjoy a high level of connectivity. However, we also show that connectivity is volatile, raising challenges for the seamless integration of satellite networks into today’s Internet.

For future work, our plan is to investigate architectural solutions that address the connectivity volatility and will enable to take full advantage of SNs. Recent trends in networking, such as path-aware networking is a promising direction to improve on the weaknesses of today’s interdomain architecture.
REFERENCES