Optimizing MSS Architecture for Direct-to-Device Services

Santanu Dutta, Ph.D
ClusterSat Network Technologies, LLC
Vienna, VA 22182, USA
email: santanu@clustersat.com

Abstract—**Direct to Device (D2D) Mobile Satellite Service (MSS), whereby a smartphone connects directly to a satellite, is the flagship of New Space, attracting billions of dollars of investment. To preserve user experience, D2D imposes demanding requirements for maintaining service continuity. Industry analysts are predicting that, for D2D to become mainstream, cellular user-experience must be preserved even when it is offered from space. It is generally agreed that this means offering at least 4G LTE QoS to each user when backing up terrestrial 5G, using 5MHz of allocated channel bandwidth. This is a high bar for legacy satellite networks for two reasons: (1) typical satellite beams are much larger than terrestrial cells, and (2) legacy satellite systems use traditional beamforming for radio access. ClusterSat has shown that a better architecture for D2D would be a fractionated satellite with multiuser MIMO (MU MIMO) for radio access. This architecture synthesizes a very large antenna aperture (~400 m) using a cluster of small satellites, leading to an unprecedented, user spatial resolution of 250 m. Additionally, MU MIMO dynamically focuses radiated power on users requiring service, while protecting other users by satellite-antenna null steering. This approach maximizes capacity density rather than net capacity – an objective that is better aligned with D2D's mission. Using a numerical example, it is shown that ClusterSat's architecture, using a cluster of 1-2 m aperture satellites, offers greater throughput in cellular coverage holes than an 8 m aperture, state-of-the-art (SoA) Low Earth Orbit (LEO) satellite.**

Keywords-Direct-to-Device mobile satellite communications; fractionated satellites; multiuser MIMO from space.

I. INTRODUCTION

D2D has been a dream since the nineties. The value proposition offered by satellite to cellular was *connectivity insurance in cellular-deployed areas* (filling coverage holes) as well as *coverage of underserved areas*. Satellite hoped to gain a major economy-of-scale benefit from integration into cellular smartphones. However, despite increasing throughput, link margin and other capabilities, Mobile Satellite Operators (MSOs) failed to persuade cellular stakeholders to incorporate satellite connectivity in mass market cellphones. The reasons were (i) cell phone integration was initially attempted mostly by GEO satellites operators, wherein the satellite had a much greater link margin challenge than Low Earth Orbit (LEO) satellites; (ii) the market pull for satellite connectivity appeared lacking because of poor user experience. The main reason for the latter was the link margin gap between mobile satellite services (MSS) and cellular -- the best, extant MSS afforded a 4-10 dB link margin on a line-of-sight (LOS) link, compared to over 20 dB for cellular in a non-line-of-sight (NLOS) link. The robustness of cellular services stems partly from this abundant link margin, which compensates for various losses such as building penetration, urban clutter, multipath fading and user body loss.

In the early 2020s, with further advances in MSS and an awareness among cellular Mobile Network Operators (MNOs) of the high cost of 5G infrastructure, interest in satellite backup was rekindled under the banner of D2D. It attracted unprecedented interest/investment, not just from the MSS industry, as in the past, but also from the MNOs. In the USA today, every major MNO has a *D2D strategy* involving alliance with at least one MSO. At present, the MSO leaders appear to be AST SpaceMobile and Starlink. To increase link margin, D2D aspirants have abandoned GEOs in favor of LEOs at 400 – 700 km altitude. However, **questions still remain whether the desired user experience will be delivered by the new LEOs**.

Section 2 discusses KPIs for D2D services – specifically, why capacity-density, rather than net-capacity, might be more appropriate for D2D. Section 3 proposes a new space architecture that is better suited than classical architectures for delivering the preferred KPI. A numerical example is provided to compare the performances of the present system, ClusterSat, and the SoA LEO system, AST SpaceMobile. Section 4 describes the key features of ClusterSat. Section 5 provides some examples of PHY layer Monte Carlo simulations to demonstrate expected performance. Section 6 discusses conclusions from the presented material and outlines envisioned paths to market for the new technology.

II. KEY PERFORMANCE INDICATORS (KPIS) FOR D2D: *CAPACITY DENSITY IS PRIME*

A satellite-cellular hybrid network has two major purposes – (i) filling holes in existing cellular networks and (ii) providing cellular services in areas lacking terrestrial cellular coverage. For this solution to be acceptable, the handovers in (i) must be *truly seamless*. There appears to be consensus [6] that providing 4G LTE in dead cells would meet the seamlessness requirement from a throughput perspective. In (ii), 4G LTE capacity should be available to all users in the designated coverage area, regardless of

ISBN: 978-1-68558-391-0

location – with beamforming, this means blanketing the entire area with sufficient Power Flux Density (PFD) to provide 4G LTE *everywhere*, regardless of user locations.

Industry analysts believe that preserving user experience is essential for D2D's mass market acceptance [1]. Based on the above, we have posited the following **minimum KPI for D2D**.

> ***Each satellite-backed user must be offered at least 4G LTE QoS****, both in the holes of existing cellular networks and in areas without cellular coverage*.

This requirement positions ***capacity density* as a prime KPI, distinct from *net capacity***. In operational terms, achieving high capacity-density requires selectively deploying high capacity (which requires high PFD) in small sub-areas (250 m diameter for ClusterSat) of a satellite beam, where there are users needing satellite service. This requires narrow PFD peaks *inside* a beam at user locations, which is not possible with traditional satellite beams because they create uniform PFD within a beam's footprint.

Unfortunately, the demand-densities of most MSS applications are geographically non-uniform -- they do not match the uniform PFD inside traditional satellite beams. D2D user demands in both urban and rural areas fit the *non-uniform* description. For example, in urban areas, the dead cells form islands of poor- or no-service inside deployed cellular networks, as shown in Figure 1. In rural areas, as shown in Figure 2, the islands may be larger, some exceeding LEO beam sizes. However, even in this case, satellite users are typically distributed randomly in clusters. An example would be users visiting the national park areas in Figure 2. If such users were distributed more uniformly and with greater density, they might motivate terrestrial cellular deployment. In summary: *uniform demand density within a beam is the exception, rather than the norm*, in MSS applications.

Technically, achieving high capacity-density is challenging for a legacy satellite network for the following reasons. Cellular base stations, by virtue of being close to the ground (tens of meters) and equipped with advanced radio access technologies like Multiple-Input-Multiple-Output (MIMO), can deliver extremely high capacity-density (throughput per square kilometer) to user equipment (UEs). In contrast, traditional LEO satellites operate hundreds of kilometers from the ground and have antenna apertures which are (today) up to 8 m for AST SpaceMoble. Such satellites, using traditional beamforming, have a beam footprint of approximately 15 km and frequency reuse distance (for N=3) of over 20 km. In contrast, a macro-cellular network has a frequency reuse distance of a few kilometers (frequency reuse distance is indicative of the deployed capacity density). **The present capacity density gap between traditional MSS and present cellular networks makes it difficult for traditional MSS to deliver 4G LTE to more than a few active users**. This is illustrated with an example use case in Section C.

III. A BETTER ARCHITECTURE TO DELIVER HIGH CAPACITY-DENSITY: *FRACTIONATED SATELLITE AND MU MIMO*

One- and two-meter satellite antenna apertures are used as examples of FR-1 LEOs in 3GPP's NTN reports [2]. As 3GPP reference examples are based on industry consensus, they represent the *knee* rather than an *extremity* of the satellite cost-complexity curve. Today, D2D contenders are exceeding these apertures by orders of magnitude, as indicated above. In ClusterSat's architecture, 1-2 m is the planned aperture of individual satellites (there is a dependence on the operating frequency). As shown in Section C, nine such satellites, operating as per ClusterSat's architecture, offer greater capacity density than the current (State of the Art) SoA LEO with an 8 m antenna aperture. They are competitive even with Starlink, with an estimated 48 m aperture satellite antenna and 1.6 km diameter beam mfootprint.

An assumption of ClusterSat's architecture is that users needing satellite service will (i) have much lower geographical density than cellular users (averaged over a cellular market area), and (ii) be distributed randomly in the said area. Below, we examine the validity and implications of these assumptions using empirical coverage data from FCC databases.

A. Filling coverage holes in deployed cellular markets

A real-world example is shown in Figure 1, corresponding to San Francisco and Washington D.C.

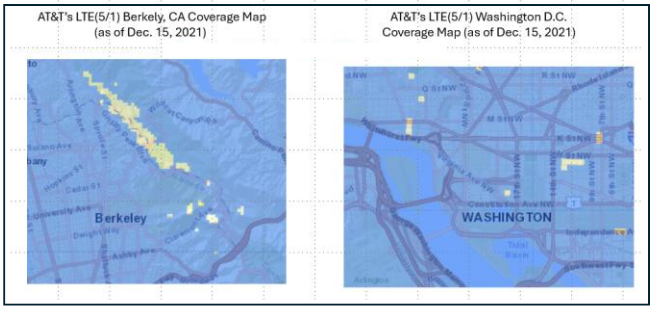


Figure 1. Typical "holes" in urban/suburban cellular coverage

If a regional satellite beam had a diameter of 15 km (like the SoA LEO), much of the downlinked PFD would be wasted because the beam is generally bigger than the coverage hole. In contrast, ClusterSat can focus PFD (and thereby capacity density) on the relatively few UEs in the coverage holes. Note that ClusterSat does not need to cover the entire footprint of regional beam footprint as most of that footprint is covered by the terrestrial cellular network.

Copyright (c) IARIA, 2026. ISBN: 978-1-68558-391-0

B. Providing cellular services in large unserved/underserved areas

A real-world example is shown in Figure 2, corresponding to the Western U.S.A.

In remote areas such as national parks, the beam size *may* be comparable to contiguous areas needing satellite coverage. As these are remote areas, the user density is likely to be low, and the users widely dispersed. As in the case of urban coverage holes discussed above, ClusterSat would have the advantage of dynamically focusing its downlink power on the few users who need service; as the number of users are likely to be small, the number of satellites required would not become huge (the number of satellites is lower bounded by the number of independent hotspots). A traditionally-beamformed single satellite would waste power by radiating it over areas not needing service.

Note: **a traditional satellite can create high capacity-density only if it *also* provides high net-capacity**. This is because its **capacity-density is constant** and its net-capacity is given by net-capacity = (capacity density).(regional beam area). Increasing capacity-density increases net-capacity proportionally. In contrast, **ClusterSat can provide variable capacity-density independently of the net-capacity of its regional beam,** wherein the capacity-density is shaped to match the user density inside the beam**.** To explain further: in the case of ClusterSat, net-capacity = Σ(satellite capacity deployed in each cell inside the beam, or cell-capacity). For ClusterSat, the cell-capacity is non-zero for dead cells and small for live cells, whereas for a traditional satellite, it is constant and high for all cells. The presence of a few dead cells in a beam does not raise the net-capacity to the same level as serving all cells in the beam. **ClusterSat's adaptive capacity-density better optimizes the network relative to its mission, which is to serve dead cells.**

C. Performance comparison: SoA LEO vs. ClusterSat

Figure 3 shows link budget analysis results for a specific use case where the performance of a state-of-the-art (SoA) traditional LEO system (approximately resembling AST SpaceMobile) is compared with ClusterSat for the same usage scenario.

The RF characteristics of AST SpaceMobile were obtained from [7]. The throughputs of both systems were calculated from 5MHz-bandwidth LTE link budgets.

C.1 Scenario assumptions

- User distribution
    - 10 users, non-uniformly distributed in a 15 km diameter circle
- SoA LEO
    - Frequency 2000 MHz
    - Allocated spectrum 5 MHz
    - Tx power in allocated spectrum 11 dBW
    - Altitude 720 km
    - Satellite antenna gain: 41 dBi (corresponds to 8 m aperture)
    - Radio access: Beamforming with 8 m antenna
- ClusterSat
    - Frequency 800 MHz (scalable to any frequency in 800 – 3000 MHz band)
    - Allocated spectrum 5 MHz
    - Tx power in allocated spectrum 4 dBW
    - Altitude 600 km
    - Gain of each satellite antenna: 22 dBi (corresponds to 2 m aperture at 800 MHz)
    - Radio access: MU MIMO with 9 satellites

C.2 Discussion of results

In the scenario shown in Figure 3 both SoA LEO and ClusterSat have identical throughputs (13/3 Mbps, DL/UL) for a single active user. **With 10 active users in the beam, the per-user throughput dilutes to (1.3/0.3 Mbps) for SoA LEO** as, unlike ClusterSat, it lacks the benefit of SDM**. For ClusterSat, thanks to MIMO/SDM, the throughputs can be peaked to their maximum values (13/3 Mbps) individually for each UE** (provided the UEs are more than 250 m apart; otherwise, the throughput is shared by UEs inside a 250 m diameter hotspot).

When compared on the basis of aggregate throughput in a regional beam, SoA LEO shows a net beam-throughput of (13/3 Mbps), while ClusterSat, with 9 satellites, can support up to 7 hotspots, resulting in an aggregate beam-throughput of (91/21 Mbps) -- a **7x net throughput advantage**.

IV. CLUSTERSAT'S ARCHITECTURE

ClusterSat creates small hotspots (250 m diameter) on UEs scheduled to be serviced during a RAN Scheduler scheduled time-frequency block while protecting other, own-network UEs using the same time-frequency block by means of antenna null steering.

The network architecture is shown in Figure 4. The system comprises a primary satellite surrounded by a cluster of smaller ancillary satellites. The number of reuses of an allocated channel scales with the number of satellites. Simulations show that, with 9 satellites, at least 7 users can reuse an allocated channel at a spatial separation greater than 250 m with greater than 10 dB output signal-to-noise-and-interference-ratio (SNIR).

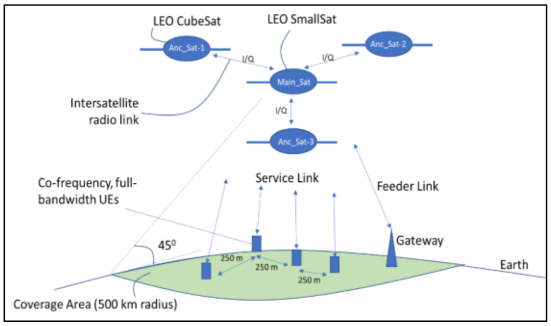


Figure 4. ClusterSat System Architecture

ISBN: 978-1-68558-391-0

The cluster of satellites flying in formation creates a large-baseline, adaptive, phased array antenna. Each satellite acts as an element of the array, transmitting and receiving signals that are tightly synchronized and processed coherently to synthesize a highly directional, high-capacity communications channel.

The main satellite (Main_Sat) may be a small LEO satellite at 600 km altitude, surrounded by a cluster of ancillary smaller satellites. Size and cost reduction of the Ancillary satellites is possible by allocating most of the digital processing to the Main satellite. The digital processing includes digital-hybrid-beamforming (a combination of MIMO, and traditional, regional beamforming), as well as the typical processing performed by a 3GPP base station onboard a regenerative satellite. Centralizing the bulk of digital processing in the Main satellite has the potential to reduce the role of the Ancillary satellites to remote antennas/transceivers, hence having lower size/weight/power (SWAP). Some similarity with the Remote Radio Head (RRH) functions in Open Radio Access Network (ORAN) architectures may be noted but it should also be noted that all satellites need to be synchronized in their transmit and receive functions, unlike in ORAN. The satellite spacing is approximately 100 m (scaling with frequency). The number of satellites depends on deployment scale (a few to hundreds). The operating frequency will typically be below 3 GHz.

Figure 5 shows an example-configuration of 9 ClusterSat satellites. The cluster comprises a three-dimensional, flexible polyhedron in space. The larger the volume of the polyhedron, the greater the UE spatial resolution is, because a larger polyhedron creates a larger array aperture. In addition to a large array aperture in the x-y plane, the polyhedron also has a reasonable component in the z-direction, radial to the Earth, providing greater angular resolution at lower elevation angles. The satellite cluster will operate in a selected Low Earth Orbit (LEO), but the specific orbital parameters are not critical to the underlying ClusterSat architecture.

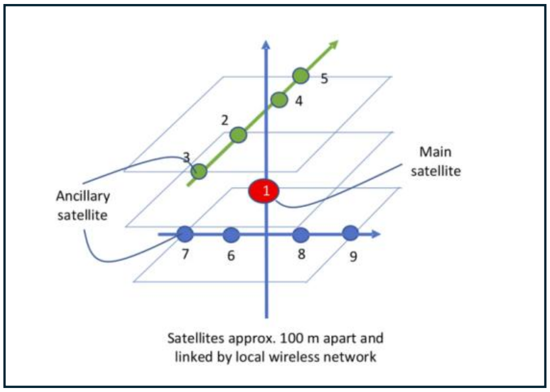


Figure 5. An example of ClusterSat's cluster configuration

The system's core concepts are orbit-agnostic and broadly applicable across a wide range of LEO constellation configurations. As the satellite cluster moves over the face of the Earth, UEs are handed over between different clusters using techniques standardized in 3GPP NTN. It is not necessary that the satellite separations within the cluster be identical or remain strictly constant. In fact, a non-uniform distribution makes the array more robust against grating lobes, which are also avoided by the motion of the cluster in its orbit at 7.5 km/s. The minimum supportable separation between UEs depends on the aperture of the satellite cluster and the system design.

The air interface simulations were performed using a proprietary, OFDM waveform. The latter was similar, but not identical to, 3GPP's MIMO waveform. A simpler (more expedient) waveform than 3GPP was used as the objective of the simulations was to demonstrate the art of the possible, not to emulate the performance with a specific air interface. It is recognized that, for reasons of scale, it may be desirable to leverage 3GPP's MIMO waveforms, although the present, proprietary waveform has performance advantages over 3GPP as it is optimized for a near-LOS satellite channel whereas 3GPP's waveform is designed for a non-LOS cellular channel.

The weight adaptation, based on LMSE optimization, is trained by known pilot signals in both the downlink (DL) and uplink (UL). For the service DL, the complex signals to be transmitted by all satellites, main and ancillary, are created by Main_Sat. The complex baseband (I/Q) signal for each ancillary satellite is communicated to that satellite by the main satellite, together with a GPS time stamp (GPST), over an intersatellite radio link. The GPST indicates the scheduled time of DL transmission, selected by a RAN Scheduler in the main satellite. The DL signals are transmitted by all satellites at the scheduled GPST. **This process ensures that signals transmitted from all satellite antennas are GPS time-synchronized and therefore coherent**. For the service UL, the signal flow is reversed relative to DL. UL signals received by the individual satellites (Ancillary and Main) are down-converted to complex baseband and sent to the Main satellite for processing over the ISL (inter-satellite link). Blocks of signal packet data are accompanied by GPST data (a 100 ms signal block is bound to synchronously sampled GPS UTC data), indicating the time of arrival of the signal data block at the receiving satellite antenna. At the Main satellite, the signals received from the ancillary satellites and the Main satellite are time-aligned and subjected to synthetic beamforming.

V. SIMULATION RESULTS

A large volume of PHY layer Monte Carlo simulations has been performed using representative received SNR. Some sample results are presented in Figure 6. The baseline scenario involves a case of close UE spacings, i.e., 9 UEs were clustered within a 1 km diameter circle, with inter-UE spacings (within the cluster) between 250 m and 1000 m. The simulations involved full-orbit runs, which start from the overhead case of 0^0 nadir angle and end at the nadir angle of 60^0, corresponding to an elevation angle of 18.6^0. Each run involved 135 s of real time operation. The key performance indicators (KPIs) considered were (SNIR) and signal-to-interference-ratio (SIR). SNIR is the prime KPI of the present system as it determines the BER (bit error rate) of digital modulations. The simulated system operated with an input SNR at each satellite of approximately 14 dB. The presented SNIRs show the net result of ClusterSat's MIMO processing on the end-end communications performance. In contrast, the SIR shows how well null steering towards interference sources (a key technology of MIMO) performed. SIR may not always correlate directly with the SNIR, e.g. when the input noise is dominant. For simulations, which involve exclusively the pilot signals, the signal bandwidth is the 15 kHz (bandwidth of one OFDM subcarrier). The results are shown in Figure 6 and the observations are summarized below:

- **SIR reaches high values in all cases** (more than 20 dB when the number of UEs is 4 or more). Even in the theoretically maximum case of 9 UEs, the SIR is above 15 dB for elevation angles above 45^o and above 10 dB for elevation angles above 30^o. When fewer UEs (e.g. 4) are involved, the SIR is above 20 dB. The high values of SIR indicate that SDM is working – cochannel, simultaneous signals from multiple UEs are experiencing mutual isolations more than 20 dB. The SIR is useful in investigating causes for poor SNIR.
- SNIR shows **no loss over the input SNR of 14 dB in all cases of interest** (elevation angles of 45^o or greater). At a 30^o elevation angle, insertion loss of from the MIMO process is limited to 3- 4 dB. It should be noted that the "insertion loss" is not really an *operational loss* as it is enabling full-bandwidth channel sharing among all UEs.
- The **SNIR is relatively constant in the region of interest** – elevation angles from 90^o to 45^o. Most of the SNIR drop occurs between 45^o and 30^o. This is corelated with the SIR, which also begins to drop in this range; this is expected because the angular resolution required to discriminate between two given UEs increases with decreasing elevation angles.

A. Observations re: simulation results

ClusterSat's system can enable at **least 7-fold frequency reuse** (7 co-frequency, simultaneously transmitting UEs) with **separations on the ground between 250 m.** The degradation beyond 7 UEs is graceful, even while preserving full bandwidth channel sharing. The above spatial frequency reuse is unprecedented for MSS. Altogether, simulations strongly support the core innovations underpinning ClusterSat's concept of operations, comprising:

- Robust MU-MIMO performance using a large-aperture, phased array antenna synthesized from a fractionated satellite cluster, demonstrating no adverse effects from grating lobes, despite longstanding concerns associated with wide spatial baselines.
- Accurate forward-prediction of channel coefficients over time – a patented ClusterSat innovation enabling significant gains in spectral efficiency and reducing reliance on real-time channel estimations.

VI. CONCLUSIONS AND FUTURE WORK

All legacy MSS architectures involve monolithic satellites and traditional beamforming. ClusterSat has shown via simulations/analyses that, except Starlink, present monolithic satellite architectures may be suboptimal in providing D2D services.

D2D requires at least 4G LTE Quality of Service (QoS) for backup service, typically represented by (5/1 Mbps Dl/UL) in 5 MHz of channel bandwidth. Link budget analyses show that, with 8 m satellite antenna aperture, it may be challenging to offer the above QoS to more than a few active users. ClusterSat can meet this challenge for 10 active users with 9 satellites.

Starlink, by pushing the satellite antenna aperture to extreme limits (~48m) *may* be able to deliver the objective QoS as a byproduct of providing Primary, MNO-like services, where net capacity is the primary objective. However, the required capitalization is also extreme. If the primary objective is Supplementary Coverage from Space (SCS), as per the FCC mandate [5], ClusterSat's architecture may be more cost effective and power- and spectrum-efficient.

Currently, ClusterSat does not intend to become a Mobile Satellite Operator (MSO). It plans to license its patented technology, which could comprise the lower-PHY layer of an otherwise standardized protocol like 3GPP NTN NR. Potential licensees could be MSOs, terrestrial cellular operators, or base station platform vendors. The licensing could involve a combination of joint development and supplying ASICs encapsulating the company's beamforming IP. The ASIC would interface with a 3GPP-compliant satellite base station via ORAN interfaces.

REFERENCES

[1] Analyses Mason Direct-to-Device Webinar June 24, 2025.

[2] 3GPP TR 38.821 V16.2.0 (2023-03), Solutions for NR to support non-terrestrial networks (NTN), Release 16.

[3] FCC Report and order and further notice of proposed rulemaking, ADOPTED: MARCH 14, 2024, FCC 24-28, Subject matter: Single Network Future: Supplemental Coverage from Space (SCS).

[4] D. Zlatev, "Starlink set to launch 5G carrier service with drastic satellite constellation expansion", Notebookcheck, September 20, 2025, https://share.google/cea6PvqrIe6ZFSpBT.

[5] J. Brookin, "FCC APPROVES Starlink plan for cellular phone service, with some limits" arsTECHNICA, November 11, 2024,

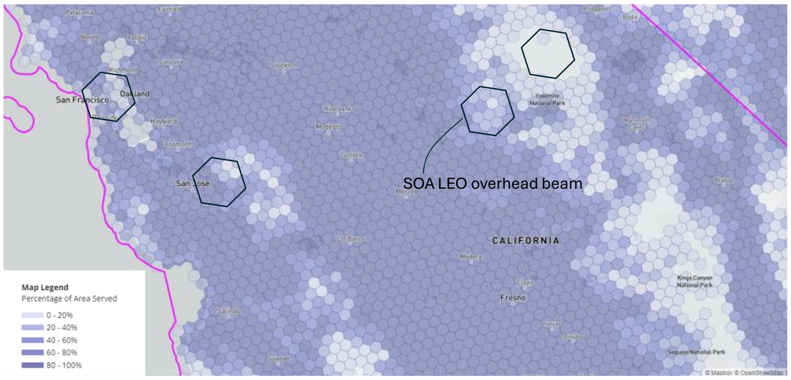
ISBN: 978-1-68558-391-0

https://arstechnica.com/tech-policy/2024/11/fcc-approves-starlink-plan-for-cellular-phone-service-with-some-limits/.

[6] The Future of Starlink Direct to Cell, Starlink Update, September 8, 2025.

[7] AST&Science LLC Petition for Declaratory Ruling Granting Access to the U.S. Market for a Non-U.S. Market for a Non-U.S.-Licensed Non-Geostationary Orbit Satellite Constellation, Call Sign S3065, IBFS File Nos. SAT-PDR-20200413-00034, SAT-APL 20200727-00088, & File No. SAT-APL-20201028-00126, April 13, 2020.

AT&T's LTE(5/1) N. California Coverage Map (as of Dec. 31, 2022) [FCC database]



Note: Each cell (small hexagon) has a diameter of approx. 4 km. The large hexagon has a diameter of approx. 15 km and represents a SOA LEO beam with overhead satellite.

Figure 2. Cellular coverage in remote areas of Western U.S.A.

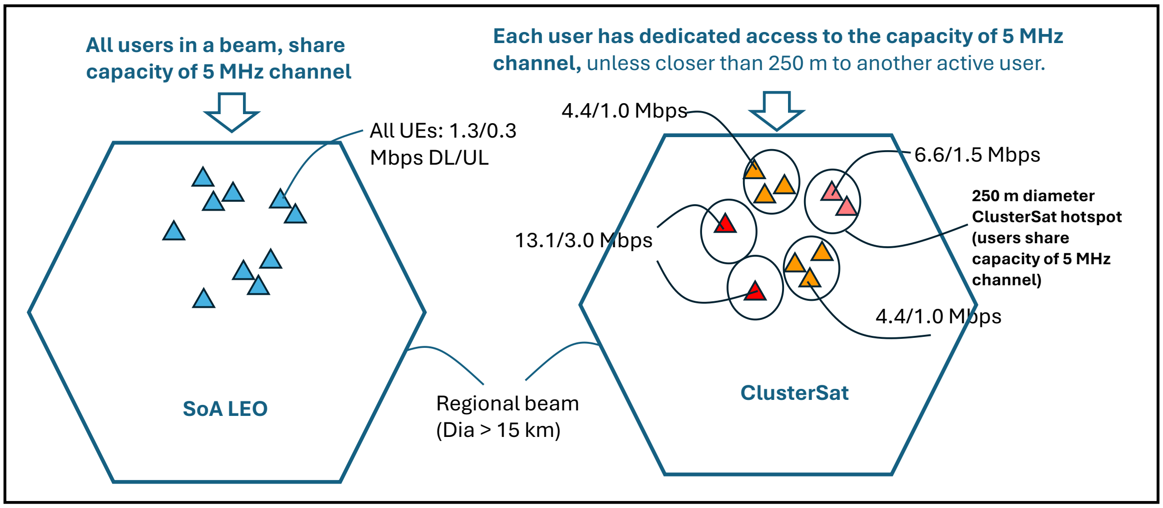


Figure 3. ClusterSat versus legacy SoA LEO

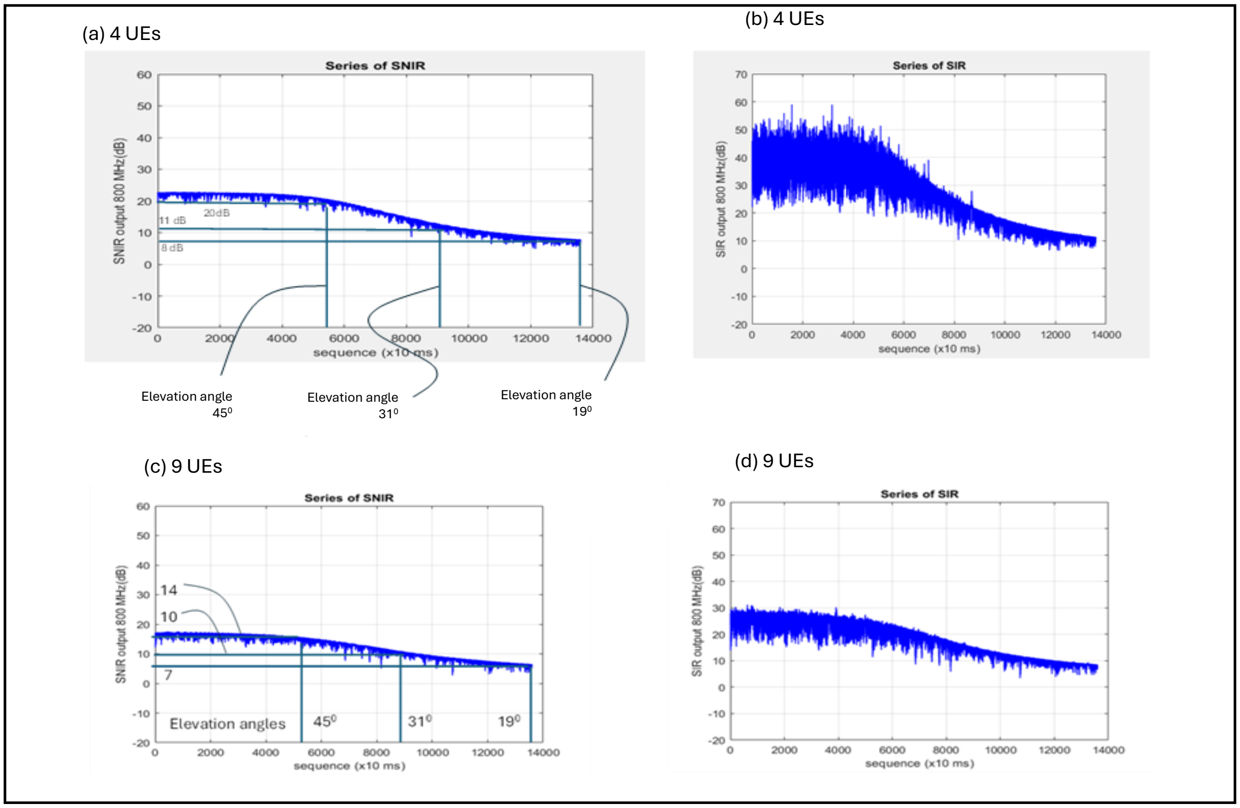


Figure 6. Simulation results

ISBN: 978-1-68558-391-0