# TESTIMONY of JAMES W. STEGEMAN PRESIDENT OF COSTQUEST ASSOCIATES

## Before the SUBCOMMITTEE ON TELECOMMUNICATIONS AND TECHNOLOGY

### UNITED STATES HOUSE OF REPRESENTATIVES

September 11, 2019



#### Contents

Introduction    3      Summary of The Broadband Mapping Initiative    4      Key Findings    5      The Issue—The Current State of Broadband Coverage Data    6      Overview of the Proof of Concept    7      The Fabric—How the Broadband Serviceable Location Fabric is Created    8      Findings: What the Proof of Concept Revealed and How the Fabric Can be Used    9      Baseline Assumptions    9      Rural Location Counts    10      Rural Distance Differences    10      The Fabric in Action to Reveal Unserved Locations    11
Summary of The Broadband Mapping Initiative    4      Key Findings    5      The Issue—The Current State of Broadband Coverage Data    6      Overview of the Proof of Concept    7      The Fabric—How the Broadband Serviceable Location Fabric is Created    8      Findings: What the Proof of Concept Revealed and How the Fabric Can be Used    9      Baseline Assumptions    9      Rural Location Counts    10      Rural Distance Differences    10      The Fabric in Action to Reveal Unserved Locations    11
Key Findings.5The Issue—The Current State of Broadband Coverage Data.6Overview of the Proof of Concept7The Fabric—How the Broadband Serviceable Location Fabric is Created.8Findings: What the Proof of Concept Revealed and How the Fabric Can be Used.9Baseline Assumptions.9Rural Location Counts10Rural Distance Differences10The Fabric in Action to Reveal Unserved Locations11
The Issue—The Current State of Broadband Coverage Data.    6      Overview of the Proof of Concept.    7      The Fabric—How the Broadband Serviceable Location Fabric is Created.    8      Findings: What the Proof of Concept Revealed and How the Fabric Can be Used.    9      Baseline Assumptions.    9      Rural Location Counts    10      Rural Distance Differences    10      The Fabric in Action to Reveal Unserved Locations    11
Overview of the Proof of Concept7The Fabric—How the Broadband Serviceable Location Fabric is Created8Findings: What the Proof of Concept Revealed and How the Fabric Can be Used9Baseline Assumptions9Rural Location Counts10Rural Distance Differences10The Fabric in Action to Reveal Unserved Locations11
The Fabric—How the Broadband Serviceable Location Fabric is Created
Findings: What the Proof of Concept Revealed and How the Fabric Can be Used
Baseline Assumptions 9   Rural Location Counts 10   Rural Distance Differences 10   The Fabric in Action to Reveal Unserved Locations 11
Rural Location Counts10Rural Distance Differences10The Fabric in Action to Reveal Unserved Locations11
Rural Distance Differences
The Fabric in Action to Reveal Unserved Locations11
Lessons Learned and Areas of Improvement14
Next Steps15
Creating the National Fabric15
Using the National Fabric
Conclusion

APPENDIX A: The Broadband Mapping Initiative Report	19
APPENDIX B: Current state of Broadband Coverage	20
Congressional District Dashboard Samples:	22
APPENDIX C: Unserved Summary	37
APPENDIX D: The Reveal	41
APPENDIX E: Polygons	42



#### Table of Figures:

Figure 1: Key pilot study findings	9
Figure 2: Pre-fabric Census Block based coverage polygon	11
Figure 3: Carrier coverage polygon based on geocoded addresses	12
Figure 4: Geocode address polygons versus associated Fabric points	13
Figure 5: The Fabric reveals the extent of served and unserved below the Census Block geography	14
Figure 6: Address lookup against the Fabric and coverage polygons	17
Figure 7: Terrestrial Broadband Coverage in the U.S. based on FCC 477 December 2017 v2	21
Figure 8: Pennsylvania U.S. Congressional District 18	22
Figure 9: Ohio U.S. Congressional District 5	23
Figure 10: New Jersey U.S. Congressional District 6	23
Figure 11: Oregon U.S. Congressional District 2	24
Figure 12: Arizona U.S. Congressional District 1	24
Figure 13: California U.S. Congressional District 6	25
Figure 14: California U.S. Congressional District 9	25
Figure 15: California U.S. Congressional District 18	26
Figure 16: California U.S. Congressional District 29	26
Figure 17: Colorado U.S. Congressional District 1	27
Figure 18: Florida U.S. Congressional District 2	27
Figure 19: Iowa U.S. Congressional District 2	28
Figure 20: Illinois U.S. Congressional District 15	28
Figure 21: Illinois U.S. Congressional District 16	29
Figure 22: Indiana U.S. Congressional District 5	29
Figure 23: Louisiana U.S. Congressional District 1	30
Figure 24: Michigan U.S. Congressional District 7	30
Figure 25: Michigan U.S. Congressional District 12	31
Figure 26: Montana U.S. Congressional District 0	31
Figure 27: North Carolina U.S. Congressional District 1	32
Figure 28: New Mexico U.S. Congressional District 3	32
Figure 29: New York U.S. Congressional District 9	33
Figure 30: Ohio U.S. Congressional District 6	33
Figure 31: Oregon U.S. Congressional District 5	34
Figure 32: Texas U.S. Congressional District 17	34
Figure 33: Texas U.S. Congressional District 22	35
Figure 34: Texas U.S. Congressional District 33	35
Figure 35: Virginia U.S. Congressional District 4	36
Figure 36: Vermont U.S. Congressional District 0	36
Figure 37: MO Unserved Summary, all Census Blocks	37
Figure 38: MO Unserved Summary, non-Cable Census Blocks	38
Figure 39: VA Unserved Summary, all Census Blocks	39
Figure 40: VA Unserved Summary, non-Cable Census Blocks.	40



# The Broadband Mapping Initiative and the Broadband Serviceable Location Fabric

An assessment of current broadband coverage, future improvements and the benefits of a national Broadband Serviceable Location Fabric (BSLF) dataset

Chairman Doyle, Ranking Member Latta, and Members of the Subcommittee:

My name is James Stegeman, I am President of CostQuest Associates. Thank you for holding this hearing and inviting me to testify. It is an honor to be here to discuss the status of Broadband Mapping in the U.S.

In my testimony today, I will provide an overview of the Broadband Mapping Initiative, an assessment of current broadband coverage and how the use of a national location-specific dataset, what I refer to as the Broadband Serviceable Location Fabric, can finally provide specificity of who has access to broadband service in America, but more importantly who does not.

#### Introduction

Let me first start with a brief introduction to CostQuest Associates.

Just over two years ago I was asked to testify in front of this body about the data needed to support the effective allocation of resources to support broadband access and use. Today I'm here to testify on the progress that has been made on this front. In its 20 years in business, CostQuest Associates (CQA) has seen, and been a small part of, tremendous improvement in data and related information that supports decision-making in telecommunications. I can tell you without equivocation that this past year has proven to be the period for which I have seen the greatest developments in understanding who has and who does not have access to adequate broadband. These developments are foundational for broadband mapping and availability reporting, which will lead to closing the digital divide.

CQA takes pride in empowering the public and private sectors with the ability to make data-driven decisions with their most critical resources, and we continue to work to improve all that we do—our models and studies— with state-of-the-art technology and processes. While CQA is known for cost expertise, geospatial design and data forms the underpinning of all our models and studies. The recent gains we have seen in geospatial data and imagery, along with machine learning and other computational logic, has led us to solutions that are materially impactful to public programs and funding under the purview of this body. I will be providing an overview of that work today.

CQA's relevant experience shows that geospatial data and models form the foundation of our most notable projects, including our continuing work with the Federal Communications Commission (FCC) and the



Universal Service Administrative Company (USAC). Below is a brief description of those projects that have relied on spatial data and modeling.

- **National Broadband Plan (FCC):** Under the direction of the FCC, CQA developed the geospatial and economic network models supporting portions of the National Broadband Plan.
- **Connect America Fund (FCC):** The Connect America Cost Model (CACM), developed by CQA under the direction of the FCC and USAC, is used to disperse over \$3 billion annually and was used to set the reserve price in last year's CAFII auction.
- New York Broadband Program: The New NY Broadband Program used CQA models to support their reverse auction for the \$0.5BIL infrastructure build-out program.
- **State Broadband Programs:** CQA has managed portions of statewide broadband mapping and planning projects for many states over the past decade.
- **City and State Fiber Builds:** CQA assists various cities and states in reviewing the business case of fiber deployment
- **Network Valuations**: CQA assists the largest ILECS, the largest Cable, and largest Wireless carriers in the valuation of their networks

In the last 20 years, the CQA team has been at the forefront of broadband mapping and costing, network modeling, economic analysis and regulatory support. We've had the privilege of working with multinational corporations, governments, trade associations and industry regulators. The support of federal and state broadband subsidy programs, the costing and mapping of over 160M locations in our cost models, and a drive to seek data-driven answers, pushed us to develop location-specific data. As such, my testimony will describe the results of our most recent work effort, the Broadband Mapping Initiative, which demonstrates the viability and benefit of creating a national **Broadband Serviceable Location Fabric** dataset.

As for my own experience, I'm a statistician by trade. As Hal Varian, chief economist at Google once said, "...the sexy job in the next 10 years will be statisticians...". As you listen to my testimony today, you'll find that it may take a bit longer, but I have hope for our day in the future.

#### Summary of The Broadband Mapping Initiative

Recognizing the need for better data and the opportunity that new data sources and technologies make possible, a coalition of leading broadband innovators launched the Broadband Mapping Initiative in April 2019. The Initiative undertook a two-state pilot effort, a proof of concept, in Missouri and Virginia to demonstrate the feasibility of identifying the precise number and location of structures that require broadband access. The resulting dataset is referred to as the Broadband Serviceable Location Fabric (Fabric or, BSLF). This Fabric of broadband serviceable structures makes it possible to precisely map and understand where broadband is available and more importantly, where it isn't. The Pilot, managed by my team at CQA, is a collaboration between USTelecom, ITTA, WISPA, AT&T, CenturyLink, Chariton Valley, Consolidated, Frontier, Riverstreet, TDS, Verizon, and Windstream.



The Pilot shows, based on the carriers participating, as many as 38%<sup>1</sup> of additional rural locations are unserved in census blocks that would have been reported as *served* in today's FCC Form 477 reporting approach. These locations are homes and businesses hidden from service providers and policymakers simply because of a lack of knowledge fueled by gaps in data—gaps that we can now fill.

Our methodology aggregates hundreds of millions of data points, applies statistical scoring, and managed crowdsourcing to pinpoint the exact locations of virtually every structure that is a candidate for broadband. Our effort seeks to contribute to an information ecosystem that is the next step in ensuring that agencies, policymakers, and providers are empowered with the data to bring the unserved online. This new information will augment our current knowledge that there are at least 21 million Americans that don't have broadband internet and help drive an efficient FCC program that has committed over \$20.4b in public funding over the next 10 years.

#### Key Findings

- The Pilot was a Success. Using innovative methods and a combination of public and commercial datasets, we have shown that it is now possible to identify and precisely locate virtually every structure in a geographic area that is a candidate for broadband service. Developing the Fabric for two states shows it is possible to do so for the entire country. The core methodology we developed works and the results of the pilot, while informative, can be improved upon with greater carrier participation.
- **Pinpointing Service Availability**. Creating the Fabric revealed that in just two states, over 445,000 homes and businesses were not reported as served by participating providers in Census Blocks that would be counted as *served* under current 477 reporting<sup>2</sup>.
- The Counts Count. We measured broadband availability by locations in a census block. The Fabric revealed that 48% of the location counts in rural census blocks are different from current estimates used by the FCC, in many cases significantly different.
- Timely and Cost Effective. A nationally developed dataset of all broadband serviceable locations consistent with the approach demonstrated in the Pilot should take no more than 5-8 months to stand up an initial national fabric for most counties and states, that could be used for testing, and 12-15 months to fully complete. The cost to do so will vary depending on the mix of open source or proprietary data sources, but a national Fabric could potentially be developed for between \$8.5-\$11 million in upfront costs and \$3-4 million in annual updates.
- Location, Location, Location. Broadband availability is about connections, but providers must know where a structure is in order to provide that link. In our pilot, the provider submitted locations for 61% of rural homes and businesses were off by over 7.6m (25 feet) and 25% are off by over 100m (328 feet), more than a football field! This distance can significantly alter our understanding of where a location is and impact the cost to deploy to an unserved location making or breaking a decision to deploy for a provider.

<sup>&</sup>lt;sup>1</sup> This estimate is at the high end of the expected count of unserved locations as all carriers in MO and VA did not participate in the study. That said, the Pilot was able to show that the one-served, all-served issue is real and demonstratable.

<sup>&</sup>lt;sup>2</sup> ibid



• **Reporting Enhanced**. Regardless of format (shapefile, propagation map, address, etc.) the quality and validity of reporting is improved using location-specific data.

We've mapped Missouri and Virginia and I'm excited to share some of our findings today. In addition to my testimony I have included, for additional detail and reference, in Appendix A the Study Report we released in August of this year. With your help, we'll expand our efforts to the entire U.S. by the end of next year. Let's continue to bridge the digital divide, together.

#### The Issue—The Current State of Broadband Coverage Data

To have a clear understanding of why the Broadband Mapping Initiative was undertaken, we need to start at our current understanding of Broadband Availability. Currently, all broadband providers must submit coverage information to the FCC through Form 477. In providing the Form 477 information, carriers provide the technology used and speeds available on a census block<sup>3</sup> basis. Under current rules, a provider can indicate a census block is covered if coverage is or can be made available in a reasonable timeframe to <u>any</u> location in the census block. In short, even if the provider only can service one location in a census block, the entire census block is reported as served—the "one served, all served"—issue you may have heard about.

It is this one served, all served approach that is the crux of the issue of who really has access to broadband service. To date, no one definitively knows how many unserved locations reside in census blocks that have been reported as served. I have only seen studies, by experts in the field, which estimate the size of the issue. One recent study<sup>4</sup> by Dr. George Ford of the Phoenix Center estimated the number of unserved households in census blocks filed in Form 477 as served at 3.45%, or in terms of additional unserved households, about 4 million nationally. While relatively small compared to the overall served count, if I look at it from the unserved perspective, we are undercounting the unserved locations in the country by almost 50%<sup>5</sup>. Another study that I am aware of estimated that the amount of unserved locations in census blocks, in more rural parts of the state in which the study was conducted, reported as served was over 30%, over 8 times higher than Dr. Ford's estimate. While these findings highlight issues with counting the unserved, they don't tell us where the unserved actually are. And from a policymaker, consumer and carrier position, understanding what locations remain unserved is crucial.

To that end, the Broadband Mapping Initiative was undertaken to start the process of understanding a more accurate count of the unserved and more importantly, the locations of the unserved. In the next section, I will start with an overview of the Broadband Mapping Initiative that was undertaken for MO and VA to demonstrate the "Proof of Concept."

Before I move on, to assist parties in understanding our current knowledge of broadband coverage, CQA has released publicly:

<sup>&</sup>lt;sup>3</sup> Census Blocks are the smallest unit of geography from the Census Bureau. There are over 11 million Census Blocks in the U.S., of which over 7 million have homes within.

<sup>&</sup>lt;sup>4</sup> http://www.phoenix-center.org/perspectives/Perspective19-03Final.pdf

<sup>&</sup>lt;sup>5</sup> Please see Appendix B which provides the latest FCC 477 data on unserved housing units.



- The State Broadband Dashboard, <u>here</u><sup>6</sup>
- The Congressional District Dashboard will be released in the near future. For current reference, sample images of the Congressional District Dashboard are provided in Appendix B.

The dashboards provided do <u>not</u> count unserved locations in Census blocks reported as served. Rather, they depict a view of the latest 477 terrestrial coverage data by density and by speed availability that only captures those Census Blocks not reported as served.

In addition to the current coverage information, the dashboard also provides an estimate of the cost to build a fiber to the location network from scratch to the homes and businesses within the specific row. These cost estimates are consistent with the Connect America Cost model values.

As you can see in the Appendix, the coverage data currently shows that 8.5 million housing units are unserved. So even Dr. Fords estimate of 4 million additional unserved households<sup>7</sup> is significant. This coverage issue is not only detrimental to the potential customer who can't get service, but also to, the carrier who is unaware of these unserved locations, and to policymakers who are working on programs to help fund the buildout to unserved areas and whose estimate of the size of the issue could be off by almost 50% or more.

#### Overview of the Proof of Concept

In order to address the current state of broadband coverage data and the "one served, all served" issue, CQA was asked to develop a Proof of Concept to first illustrate the ability to accurately identify the location of homes and businesses that are receiving or could require broadband. Once illustrated, the next step was to showcase how the location data could be used with carrier information to help identify served locations, but more importantly to identify the locations that are unserved.

In effect, we were asked to show the ability to create, use, and demonstrate the necessity of more granular data. With more granular data we can, with more confidence, locate those areas that remain unserved.

The Proof of Concept consists of two interconnected parts: The Broadband Serviceable Location Fabric and the data submitted by the participating carriers. Let me briefly cover both parts.

The Fabric is a robust dataset that uniquely links parcels data, tax attribute data, building footprints, roads, and address data. With this linkage, CQA had the information to begin the process of identifying serviceable locations in Missouri and Virginia. I will discuss in more detail in the next section how this data was linked and why it is only now possible.

The carriers' data played an important role to validate the importance of granular data and the need for the National Fabric. Each carrier submitted either shapefiles or address data of locations they have served, are currently serving, or are capable of serving. For those carriers that provided addresses, they also indicated for each address: the source of the address, latitude and longitude, and whether the address would have been filed in the 477. Once the address data was received, CQA validated their data and began

<sup>&</sup>lt;sup>6</sup> See <u>https://www.costquest.com/state-broadband-dashboard</u>

<sup>&</sup>lt;sup>7</sup> A housing unit is a place where a person can or could live. A household is an occupied housing unit.



a textual matching process to match carrier address data to the Fabric. The carrier received the Fabric Key and the latitude and longitude of the matched Fabric Location. With the results of the carrier address matching process, CQA was able to identify potentially unserved locations that would have been considered served based on the current Census Block based 477 filings. The carriers were informed of the unmatched Fabric points in the Census Blocks they reported data in. We also were able, using the carrier provided geocoded latitudes and longitudes, to see the inaccuracies of carrier provided geocoded addresses.

#### The Fabric—How the Broadband Serviceable Location Fabric is Created

As a first step, we reviewed nationally available datasets for parcels, building footprints, tax assessor data, addresses, and roads to determine if there was a primary source to identify where all the broadband serviceable locations are. However, in reviewing these datasets, we could not find a single dataset that would provide the basis of the national Fabric. Each had shortcomings and few focused on the unique characteristics of the rural parts of the two states.

In reviewing these datasets, we were, however, able to determine that working with the combination of these datasets would provide the information required to build a complete Fabric dataset that yields a final product that is more accurate than any of the source datasets are on their own. As such, my team undertook the effort to develop the Broadband Serviceable Location Fabric (BSLF).

To develop the BSLF, it was necessary, after the required datasets were identified, to first develop a robust linkage between tax assessor data and parcel data. With this linkage in place, the tax assessor information provides an understanding of what the parcel is used for while the geometry of the parcel provides the geographic constraint to process the entire collection of data considered. Once a good link between assessor data and parcels has been established, other datasets including building footprints, addresses, and roads are brought in through a variety of geospatial processes. The resulting combination forms the basis of a statistical model that is calibrated and validated by human reviewers. Using the result of the statistical model, building footprint locations on each parcel of land are then able to be identified as likely or unlikely to be broadband serviceable.

In the case that the model presents an inconclusive result for a parcel, a crowdsourced visual verification approach was employed working with our partner firm CrowdReason, who is a recognized leader in effective use of on-demand labor solutions. In this effort, a trained crowd workforce visually reviewed over 140,000 records across MO and VA. These visually reviewed records filled in the gaps where confidence of knowledge of a location was low.

As a final step in identifying serviceable locations, the carrier provided address data was used to help improve the proper identification of locations. That is, if a carrier has provided service to an address, clearly the location is broadband serviceable.



Before leaving this section, it is important to note that this effort to create the Fabric could not have been accomplished a few years ago. The missing key was the building footprint<sup>8</sup> data. This data only became available in the last few years.

# Findings: What the Proof of Concept Revealed and How the Fabric Can be Used

The Proof of Concept was a success. First and foremost, we were able to show that National Fabric can be constructed and that with carrier provided data we can actually identify unserved homes and businesses. The results of the Fabric Pilot also highlighted some key areas where location-specific data can provide value to broadband programs at both the federal and state level. We've summarized the key findings in Figure 1 below and will walk through them in this section.

#### Key Pilot Findings - Rural Missouri & Virginia



Figure 1: Key pilot study findings

#### **Baseline Assumptions**

- These findings focus on the rural census blocks of Missouri and Virginia, but our overall analysis extends to all areas of both states
- 14 carriers submitted data (addresses, latitude/longitude coordinates, etc.) to our Pilot program and indicated whether they would file a location in their Form 477 filing (i.e. indicate that service is available here or could be made available without an extraordinary effort)
- Our estimations of unserved can serve as the top end starting point and will be improved as carrier data is added to further indicate where coverage exists



• When I say *location,* we're referring to a structure (home or business) that is a candidate for needing broadband service

#### **Rural Location Counts**

<u>"38% of Total Rural Locations in Census Blocks Reported to be Served are UNSERVED - by the carriers in the</u> <u>Pilot study"</u> – This means that in the rural census blocks of Missouri and Virginia, where carriers that participated in our Pilot would have indicated they provided service, we found that 38% of those locations were not reported by the carriers in the study, which amounts to 445,000 homes and businesses. While this is an upper bound of the ground truth of the unserved locations within currently reported served Census Blocks, as we didn't have all carrier data in our analysis, it still highlights that there are locations in SERVED census blocks that were previously unseen and thus potentially unserved. If we were to exclude from this analysis those Census Blocks served by cable carriers who were the primary carriers who did not participate in the study, we still find over 200,000 unserved locations.

In Appendix C, I provide the detail, by state, behind the Unserved counts. As part of this detail, I have included a summary of non-cable served Census Blocks in both states.

<u>"48% of Rural Census Block Fabric Location Counts Don't Match Currently Used Estimates of Location</u> <u>Counts"</u> – We compared location counts in rural census blocks, between our Fabric location counts to values consistent with those currently used in FCC CAF efforts. Almost half of these rural census blocks had different counts. This is meaningful when assessing the scope of the unserved problem, determining build-out requirements, and, ultimately, identifying how much budget is needed to remedy.

#### **Rural Distance Differences**

"61% of Rural Pilot provided geocoded Locations NOT at the correct structure location...25% of Rural Pilot Locations are off by over 100m" – When examining the supplied latitude and longitude coordinates provided by the participating Pilot carriers as compared to our Fabric location for the address, we found that most submitted coordinates missed where the Fabric structure was by at least 7.6<sup>9</sup> meters (25 feet), with 25% off by more than a 100 meters, or more than a football field. If the majority of the coordinates provided to our Pilot carriers were off, that means that locations being built to could be assigned to the incorrect census block and estimated build-out costs could vary substantially from what will actually be incurred.

**"23% of Rural Pilot Locations NOT geocoded to Correct Census Blocks"** – This is a deeper dive of distance differences that builds on the point above. First, if carriers are using these geocoded locations as the basis of current Census Block based FCC 477 filings, there could be Census Blocks that are mis-identified as served. Second, Carriers that receive subsidy funds have build-out requirements to provide service availability to a certain number of locations within a given geographic area. In cases where a structure has an incorrect latitude, longitude or Census Block assignment, there exists potential for built out locations to be counted towards the wrong totals or not counted at all.

<sup>&</sup>lt;sup>9</sup> 7.6 meters is the margin of error accepted by the HUBB portal (where carriers must submit locations they've made service available to) to verify that the location is in a census block eligible for funding.



#### The Fabric in Action to Reveal Unserved Locations

The Figures below, which are also in Appendix D, demonstrate, using approximately 10 populated sample Census Blocks in Missouri, how the Fabric is able to identify the exact coordinates of served locations AND unserved locations. As these Figures will demonstrate, we can collectively use the Fabric location-specific data as the foundation upon which we can layer coverage data, funding areas of different programs, and additional data captured in the future. As data is added, we can continue to refine our understanding of where, exactly, unserved Americans exist.

In Figure 2, I highlight what our current FCC 477 based understanding of broadband coverage would like for these Census Blocks (outlined in dark blue). Using the pilot providers' data, these Census Blocks would have been reported as served (shaded in light blue).



Figure 2: Pre-fabric Census Block based coverage polygon

In Figure 3, I demonstrate what polygons might look like under the FCC's new Digital Opportunity Data Collection ("DODC") broadband mapping program that will replace the current FCC 477 effort. In the new DODC, carriers will file polygons that capture where they provide service. These polygons are intended to capture coverage below the Census Block level. However, the specific guidelines on how to form the



polygons is out for public comment. As such, my team implemented a potential approach that created the polygons (light blue bounded areas) based on the carriers' provided geocoded latitude and longitude. Using these referenced points, we identified associated road segments and then created a 150-foot buffer about the identified road segments. This is one approach to polygon creation, but there are others. I provide in Appendix D a discussion and samples of how the polygons can be formed to help one understand the concept of polygon reporting.



Figure 3: Carrier coverage polygon based on geocoded addresses

In Figure 4 I show how the geocoded address polygons shown in Figure 3 compare to the Fabric latitude and longitudes for the addresses used to create the polygons formed from the carrier provided geocodes. As you can see, many of the addresses from the Fabric fall outside the service polygons. The key takeaway is that while polygon coverage reporting does provide insight into coverage below the Census Block, the quality of the coverage polygons is highly dependent on the underlying data used to create the polygon. In this Figure, it is clear that our understanding of coverage based on the polygons created from the carrier geocoded addresses is distorted as compared to the fabric locations associated with these same addresses.





Figure 4: Geocode address polygons versus associated Fabric points

In Figure 5, using the Fabric, I am able to show the extent of served (green dots) and more importantly the unserved locations (red dots) in these Census Blocks. Of any finding or demonstration of the Fabric, this clearly demonstrates why the Fabric is needed. Specifically, polygon reporting in the new DODC efforts will only provide knowledge of areas claimed to be served<sup>10</sup>. The Fabric is needed to then provide knowledge of the unserved locations.

<sup>&</sup>lt;sup>10</sup> Depending on the final rules of DODC polygon creation, 'served polygons' may contain a mix of served and unserved locations.





Figure 5: The Fabric reveals the extent of served and unserved below the Census Block geography.

#### Lessons Learned and Areas of Improvement

In building the foundational layers of the Fabric it was clear that weaving together the multiple sources of data makes *all* the data better. It is also clear that some threads of this Fabric can be made stronger. A large part of our work was related to the deep investigation of these various threads of data at a very granular level. In that process we found the following:

- 1. Land Use Data is incomplete and lacking consistent standards. County assessors typically record the land use of a property, such as "Residential Single Family", "Business", "Industrial", and "Vacant Land". For our purposes, land use identification helps us to identify and determine *Serviceable Locations* (e.g., residential). Some counties provide good, detailed information. However, some jurisdictions do not have, or make available, land use data or simply use non-informative descriptions. Others do not collect or track such data in a consistent manner. A national effort to produce guidelines and encourage use by assessor's use would lead to an improved fabric effort.
- 2. <u>Parcel Boundaries</u> in some jurisdictions are missing or lacking consistent standards. Some areas of the country lack public parcel information. These parcel boundaries constrain processing of all the



various layers of data. A national effort to create a complete national parcel layer would lead to an improved fabric product.

- 3. <u>Address Data</u> is consistently inconsistent. Various sources of addresses were used as we built the Fabric. This address level data is key to linking addresses from carriers to the Fabric locations. What we found is that regardless of the source of addresses, there were inconsistencies in format and in the counts of dwelling units. An effort to improve the National Address Database could improve the quality of address data.
- 4. <u>What is a Broadband Serviceable Location?</u> Having worked in the industry for over 30 years, I have a sense, but there is no current, clear definition of what constitutes a Broadband Serviceable Location. Is it a barn, the farmhouse, both/neither? To avoid contention, the FCC needs to clearly define the term.
- 5. <u>Visual Verification</u> is important for to correcting data issues. Where and when data inconsistencies present themselves in the process of creating the Fabric, *Managed Visual Review* can help correct these issues. The process of carefully guided human review of quality imagery leads to a much higher quality Fabric.

While the development of the Pilot showed that the Broadband Serviceable Location Fabric is a monumental improvement over the disparate individual data sources in terms of determining and locating served and unserved structures, continual maintenance and improvement of the Fabric should be considered a core tenant of such an approach. Collaboration across data owners, public or private, to normalize and improve source data, will make the Fabric stronger.

#### Next Steps

The Broadband Mapping Initiative and the resulting Pilot for the **Broadband Serviceable Location Fabric** have given us all a large step toward a vastly improved foundation for service availability identification and closing the digital divide. However, if the Fabric is to provide the critical foundation for these efforts, it will need to be made available across the U.S. and used as part of the new Broadband Mapping effort conceived in the Digital Opportunity Data Collection Order ("DODC"). The order has specifically called out the potential need for a National Fabric, and programs like the forthcoming Rural Digital Opportunities Fund (RDOF) will clearly benefit from the improved targeting of funds as a result of its use.

#### Creating the National Fabric

Building the National Fabric is a clear possibility that is on the immediate horizon. The Pilot has provided the blueprint to getting it built. Completing the Fabric will require the following:

- 1. **Data Collection**. Data needs to be collected, for all sources contributing to the Fabric, for each State and territory. Contractual agreements need to be in place and derivative use needs to be defined within those agreements.
- 2. Data Assessment and Normalization. Collected data needs to be assessed for completeness, normalized, and linked together to form the Fabric. This process is variable by state and jurisdiction.



- 3. Visual Verification. Managed visual review leads to a higher-quality Fabric. The visual verification is crowdsourced and carefully managed. This process requires gathering the labor, defining the crowd's tasks, managing the review process, and incorporating the results into the Fabric.
- 4. **Producing the Fabric**. The data needs to be combined, processed, and tested to create the initial National Fabric.

Creating the National Fabric, if starting from where the Pilot left off, should take no more than **5-8** months to stand up an initial national fabric for most counties and states, that could be used for testing, and 12-15 months to fully complete and would cost between \$8.5-\$11M. This assumes that the data from which the Fabric is built remains proprietary, recognizing that the third-party data is restricted in use. While proprietary, that does not mean the Fabric is not viewable by the public or useable by Federal Agencies or carriers. Rather, proprietary would restrict who has access to the underlying data and how that data can be used. Building a fully *open dataset source* Fabric would take longer and rely heavily on Visual Verification which could more than double the budget.

#### Using the National Fabric

The Digital Opportunity Data Collection Order (DODC) envisions a process whereby broadband providers submit polygons—shapes of coverage—showing current service availability. Even though this process both simplifies and improves provider filings, it requires the solid foundation of accurate locations in order to be truly useful.

While technical standards related to submitting polygons and other data showing service availability have not yet been established, there is general agreement that the Broadband Service Location Fabric could provide the backbone for these filings through precise location identification. Examples of uses of the Fabric, within the framework of the DODC, includes, but is not limited to the following:

- The Fabric as the foundation for served and unserved locations. Layering in coverage polygons over the top of the Fabric gives us a precise view of what is served and unserved by carriers.
- The Fabric as the basis for polygons. If carriers file address locations of served (or serviceable) demand, the Fabric locations can be used as the basis to make those filings, thus creating polygons of service that are founded on precise locations.

I invite you to review the examples of polygon-creation using the Fabric in Appendix E.

Regardless of the approach for incorporation of the Fabric, using precise location-based filings gives us a clear view of the world of the unserved, allowing us to move to bring residences and business online.

I will leave this section with one additional benefit that the fabric combined with new DODC efforts will provide. That is, consumers could be given a clear view of the reported coverage to their home or business and will have solid information from which to challenge the reporting, if required. In Figure 6, I illustrate an address lookup, with the Fabric in the background (red dots = unserved, green dots = served) along with coverage polygons based on the Fabric and carrier data (yellow bounded areas using 500-foot road buffer).





Figure 6: Address lookup against the Fabric and coverage polygons.

#### Conclusion

I have worked in the telecommunication industry for over 30 years. My firm has been at the forefront on advancing industry knowledge for the last 20 years. In that time, there have been drastic changes in both technology and regulation. This is one of the most dynamic industries in the world and it has been an exciting industry to be involved in.

My testimony today focused on broadband mapping data and the Broadband Mapping Initiative that validates the viability of building a nationwide Broadband Serviceable Location Fabric that, tied to the FCC's update of the broadband coverage data, will allow all to finally understand the unserved issue in the country.

While efforts to create the Fabric and to understand Broadband coverage at the location rather than at a census block may be tedious, cumbersome, time, and data intensive, we must make every effort to collect accurate locational and coverage information and analyze that information in a way that contributes to a



wise and efficient allocation of resources. In short, as with all my client work, we need strive to collect the best information to help make informed decisions.

To continue moving forward, I urge members of the Subcommittee to consider the following:

- 1. We need a national Fabric dataset. The proof of concept has shown that the national Fabric can be constructed, helps reveal the unserved home and businesses in the country, and can be accomplished in a reasonable timeframe at a modest budget.
- 2. We need to link the fabric to the upcoming DODC efforts. Without the Fabric, the DODC polygons will only depict images of what is served. There is no reporting of the unserved.
- 3. We need to maintain the Fabric. The Fabric needs to be a living dataset that improves over time and recognizes the changes in locations for homes and businesses.

The national goal to expand coverage to all citizens will be challenging. As a first step, knowing who needs broadband coverage and leveraging the FCC's new DODC efforts requires a national Fabric dataset. In short, a national Fabric dataset is paramount to achieving the national goal.

Thank you for your time today.



#### APPENDIX A: The Broadband Mapping Initiative Report

James W. Stegeman – Appendix A to Testimony to House Committee Energy and Commerce, Subcommittee Communications and Technology Sept. 11, 2019 – See separate document with unique page numbering



#### APPENDIX B: Current state of Broadband Coverage

In August, the FCC adopted<sup>11</sup> a Report and Order and Second Notice of Proposed Rulemaking to improve the accuracy of broadband coverage data. The Report and Notice lays out a reasonable set of questions and areas for comment. CQA is supportive of the FCC rule making effort and look forward to seeing the improved data.

Even with the ongoing FCC rule making, it remains important to use Form 477 data to understand, as best as possible, the current state of Broadband coverage.

This appendix will first provide an overview of national broadband coverage and then provide detail for selected Congressional Districts.

As part of the Connect America Fund orders, the FCC currently defines benchmark, fixed location (nonmobile), broadband service as the ability to obtain service that provides a downstream bandwidth of 25Mbps and upstream bandwidth of 3 Mbps<sup>12</sup>. The state of New York, in their recent broadband auction defined Unserved as access to service less than 25Mbps. This aligns with the FCC. However, New York added in an Underserved category for areas that have access to service speed between 25Mbps and 100Mbps. *Served* is defined as having access to speeds equal to or above 100Mbps. I believe this New York distinction is informative to understanding broadband coverage nationwide and is consistent with measuring success under the FCC's National Broadband Goal No. 1<sup>13</sup>.

Goal No. 1: At least 100 million U.S. homes should have affordable access to actual download speeds of at least 100 megabits per second and actual upload speeds of at least 50 megabits per second.

Using this New York classification, my team reviewed the status of the terrestrial based, non-mobile provider<sup>14</sup> broadband deployment in the U.S. using the FCC's latest 477 data<sup>15</sup>.

<sup>&</sup>lt;sup>11</sup> <u>https://docs.fcc.gov/public/attachments/FCC-19-79A1.pdf</u>

<sup>&</sup>lt;sup>12</sup> Please see <u>https://www.fcc.gov/reports-research/reports/broadband-progress-reports/2015-broadband-progress-report</u>

<sup>&</sup>lt;sup>13</sup> Please see page XIV in the FCC's National Broadband Plan *available at* <u>https://transition.fcc.gov/national-broadband-plan.pdf</u>

<sup>&</sup>lt;sup>14</sup> For purposes of terrestrial service, I include copper, fiber, fixed wireless and coax technologies identified for consumers.

<sup>&</sup>lt;sup>15</sup> For the analysis, I used December 2017, v2 data, which is the latest available and posted at

<sup>&</sup>lt;u>https://www.fcc.gov/general/broadband-deployment-data-fcc-form-477</u>. As presented, these data do not reflect affordability measures.



	<u>Served</u>	<b>UNDERserved</b>	<u>Unserved</u>
AK	70.0%	6.7%	23.3%
AL	82.0%	5.3%	12.7%
AR	59.3%	19.2%	21.5%
AS	0.0%	0.0%	100.0%
AZ	85.5%	1.8%	12.8%
CA	94.5%	0.9%	4.7%
СО	87.5%	3.5%	8.9%
СТ	99.4%	0.0%	0.6%
DC	99.1%	0.0%	0.9%
DE	97.8%	0.0%	2.2%
FL	95.4%	1.7%	2.9%
GA	87.9%	5.2%	6.8%
GU	0.9%	99.1%	0.0%
HI	96.2%	0.0%	3.8%
IA	84.4%	5.0%	10.6%
ID	79.6%	1.9%	18.5%
IL	92.3%	1.1%	6.6%
IN	87.1%	2.3%	10.6%
KS	80.8%	6.9%	12.3%
KY	85.7%	6.7%	7.6%
LA	83.8%	5.4%	10.8%
MA	97.9%	0.5%	1.7%
MD	96.7%	0.0%	3.3%
ME	84.6%	10.4%	5.0%
MI	87.8%	1.5%	10.8%
MN	87.1%	4.1%	8.7%
MO	80.9%	4.6%	14.5%
MP	0.5%	0.0%	99.5%

<u>State</u>	<u>Served</u>	<b>UNDERserved</b>	<u>Unserved</u>
MS	74.6%	6.2%	19.2%
MT	67.4%	9.9%	22.7%
NC	92.3%	2.5%	5.2%
ND	92.9%	2.2%	4.9%
NE	77.6%	9.5%	12.9%
NH	93.2%	4.1%	2.6%
NJ	99.3%	0.0%	0.7%
NM	79.1%	3.7%	17.2%
NV	90.9%	1.5%	7.6%
NY	97.4%	0.3%	2.2%
ОН	93.3%	1.2%	5.5%
ОК	72.0%	6.4%	21.6%
OR	89.9%	1.3%	8.9%
PA	93.7%	1.4%	4.9%
PR	87.0%	3.4%	9.6%
RI	98.4%	0.0%	1.6%
SC	86.6%	4.9%	8.5%
SD	82.9%	6.2%	10.9%
TN	88.4%	2.6%	9.0%
ТХ	85.8%	3.6%	10.5%
UT	90.2%	4.3%	5.5%
VA	88.9%	2.1%	8.9%
VI	100.0%	0.0%	0.0%
VT	80.7%	14.9%	4.4%
WA	92.6%	2.4%	5.0%
WI	81.6%	5.3%	13.1%
WV	76.6%	0.5%	22.9%
WY	66.8%	9.9%	23.3%

Source: December 2017 v2 Produced by: CQA

Figure 7: Terrestrial Broadband Coverage in the U.S. based on FCC 477 December 2017 v2



#### Congressional District Dashboard Samples:

For members of this Subcommittee, the figures below provide snapshots from our upcoming Congressional District Dashboard that will be released publicly. These images provide broadband speed coverage by Congressional District. In addition, my team has provided the Total Investment column which provides an estimate of the cost to build a fiber to the location network from scratch to the homes and businesses within the specific row.



Figure 8: Pennsylvania U.S. Congressional District 18





Figure 9: Ohio U.S. Congressional District 5



Figure 10: New Jersey U.S. Congressional District 6





Figure 11: Oregon U.S. Congressional District 2



Figure 12: Arizona U.S. Congressional District 1















Figure 15: California U.S. Congressional District 18



Figure 16: California U.S. Congressional District 29









Figure 18: Florida U.S. Congressional District 2





Figure 19: Iowa U.S. Congressional District 2



Figure 20: Illinois U.S. Congressional District 15





Figure 21: Illinois U.S. Congressional District 16



Figure 22: Indiana U.S. Congressional District 5





Figure 23: Louisiana U.S. Congressional District 1



Figure 24: Michigan U.S. Congressional District 7





Figure 25: Michigan U.S. Congressional District 12



Figure 26: Montana U.S. Congressional District 0





Figure 27: North Carolina U.S. Congressional District 1





32









Figure 30: Ohio U.S. Congressional District 6





Figure 31: Oregon U.S. Congressional District 5



Figure 32: Texas U.S. Congressional District 17

34





Figure 33: Texas U.S. Congressional District 22



Figure 34: Texas U.S. Congressional District 33





Figure 35: Virginia U.S. Congressional District 4



Figure 36: Vermont U.S. Congressional District 0



#### APPENDIX C: Unserved Summary

In this Appendix I provide detail information on our unserved findings.

In Figure 37, I provide the detailed summary of our Unserved summary for Missouri. In this analysis we matched the Pilot Carrier address data that they indicated would be filed in the FCC 477 as served against the Fabric locations for these addresses. In those Census Blocks that contained carrier matched Fabric points, we counted the number of Fabric points not matched. This count is shown in the Figure as the Unserved count. Again, as noted in the body of my testimony, these unserved counts represent an upper bound as all carriers did not participate in the study.



#### Location Fabric Data and Carrier Pilot 477 Data: MO Fabric

Figure 37: MO Unserved Summary, all Census Blocks



1–25 26–50 50–75 Percent of Unserved Fabric Locations within a Census Block

Filed 62.63%

47 R

75-<100

In Figure 38, I provide the similar information for Missouri as in Figure 37. However, in this Figure, I exclude all Census Blocks shown as served in the latest FCC 447 by cable technologies. I exclude these Census Blocks to estimate the impact if Cable providers had participated in the Pilot and, by chance, had reported all the addresses in their 477 served Census Blocks as served.

45K

40k

3.0k

!0k Jns( 15K

104

0k

All 477 Pilot Locations vs Unserved Fabric Locations

Unserved 37,37%

Fabric



#### Location Fabric Data and Carrier Pilot 477 Data: MO Fabric

Figure 38: MO Unserved Summary, non-Cable Census Blocks



In Figure 39, I provide the Virginia summary in the same fashion as provided in Figure 37 for Missouri.

#### Location Fabric Data and Carrier Pilot 477 Data: VA Fabric



Figure 39: VA Unserved Summary, all Census Blocks



In Figure 40, I provide the Virginia summary in the same fashion as provided in Figure 38 for Missouri.

#### Location Fabric Data and Carrier Pilot 477 Data: VA Fabric



Figure 40: VA Unserved Summary, non-Cable Census Blocks.



#### APPENDIX D: The Reveal



#### APPENDIX E: Polygons