CT Broadband Mapping

Data Processing Report

Supplement

Submission 6

October 1, 2012

CONNECTICUT PROGRAM OVERVIEW

In response to the Notice of Funds Availability published in the Federal Register on July 8, 2009 (NOFA), the State of Connecticut Department of Public Utility Control (CT DPUC) submitted a grant application for consideration under the National Telecommunications and Information Administration's (NTIA) State Broadband Initiative Grant Program (SBI), for broadband mapping. The CT DPUC, pursuant to Executive Order 32-A, was designated as the single Connecticut state entity eligible to apply for funds under this program.

In July of 2011, the CT DPUC was merged with the CT Department of Environmental Protection to form a new agency called the Department of Energy and Environmental Protection (CT DEEP). CT DEEP will now be the lead agency coordinating with NTIA on this program.

The State has long been committed to broadband delivery and enhanced use as a fundamental goal. The State has developed a planning strategy to marshal the State's resources and stakeholders and establish Connecticut as a leader in broadband usage, in addition to being a leader in "e-Government" and other broadband-dependent endeavors.

The State entered its SBI initiative not possessing any data related to broadband service, availability, or infrastructure that could readily support the requirements of the Broadband Data and Development grant program. Due to technical considerations, DEEP has partnered with Applied Geographics Inc., and subcontractor Sanborn, to support the data collection and mapping efforts.

So far CT has been very successful in acquiring the requested information from the broadband service providers, and is utilizing this information on our own http://CT.gov/Broadband website as well as providing the needed information up to NTIA to support the national map.

FALL 2012 SUBMISSION OVERVIEW

According to both our research and lists provided to use by NTIA, there was the potential for CT to have up to 141 broadband providers:

We contacted every provider on this master list.

51 Companies stated they do not provide any type of broadband service in CT. Many of these are either national carriers without a CT presence, or they file 477 reports because they provide VOIP or Video Teleconference services (but not broadband).

360 Networks
8x8, Inc.
Accessline Communications Corporation
Acecape Innovative Networks
AlphaStar
American Fiber Network, Inc.

American Fiber Systems, Inc.

Apptix, Inc

Aptela, Inc

Bellsouth Long Distance, Inc.

Broadcore, Inc.

CIMCO Communications, Inc.

Custom Network Solutions

Cybershore

Echostar

Global Crossing North America, Inc.

GlobalPhone Corp.

GreatCall, Inc

Hickory Tech Corporation

i2 Telecom International, Inc
IDT Corporation
inContact, Inc

InPhonex.com, LLC

Intra Global Communications Inc.

IP Communications, LLC

ITC^DELTACOM Communications

Kosmaz Technologies LLC

M5 Networks, Inc

Matrix Telecom, Inc

New Global Telecom, Inc

Ooma, Inc.

Phone.com, LLC

Qwest Interprise America, Inc.

RCN Corporation

RingCentral, Inc.

Sage Telecom, Inc

SBC Long Distance, LLC
SkyTerra LP
SkyWay, USA

Software Cellular Network Ltd.

Stella Communications

Tata Communications (America) Inc.

Telefonica Data Corp SA

Telefonica USA, Inc.

University Corporation For Advanced Internet Devel

VoiceINC.COM Corporation

VoIPnet Technologies

VoIPStreet, Inc.

Vonage Holdings Corp

Zayo Enterprise Networks, LLC

23 Company names turned out to be a DBA or legal holding names for another firm that is listed in another category. So these duplicates were dropped from our list.

A-R Cable Investments, Inc.
AT&T Corp.
AT&T Services, Inc.
Broadwing Communications, LLC
Cablevision Lightpath CT
Cablevision Systems Corporation
Cellco Partnership
COMCAST CABLE COMMUNICATIONS, INC
Connecticut DataNet, LLC. dba Lightower Fiber Netw
DataNet Communications Group, Inc.
Deutsche Telekom AG
DSLnet Communications, LLC
DSLnet Communications, LLC (Megapath)
Enventis Telecom Inc.
Harron Communications LP
Hudson Valley DataNet, LLC.
Hughes Communications, Inc.
New Cingular Wireless Services, Inc.
Saturn Telecommunications Services, Inc
Verizon Business Global LLC dba Verizon Business
Verizon Communications Inc.
WilTel Communications Group, LLC
Yipes Holdings, Inc

31 Companies reported that they are strictly resellers (which we are not including in our submission).

ACN Communication Services, Inc
Airespring, Inc.
Bandwidth.com, Inc
BCN Telecom, Inc.
BullsEye Telecom, Inc.
Caused Based Commerce Incorporated
Cypress Communications, LLC
Direct TV
Dish Network
Earthlink
Ernest Communications, Inc.
Fionda VOIP, LLC
Granite Telecommunications LLC
Lightyear Network Solutions LLC
Metropolitan Telecommunications Holding Company
New Edge Holding Company
One Communications Corporation

PAETEC Communications, Inc.
Prescient Worldwide
Proximiti Communications
Reallinx, Inc
Smart Choice Communications, LLC
Stage 2 Networks, LLC
Telesphere Networks Ltd
Trans National Communications International
Transbeam Inc.
TW Telecom Data Services
VCOM Solutions, Inc
Wholesale Carrier Services
Wholesale Carrier Services, Inc
Windstream

8 Companies may be broadband providers, but either they indicated they are not willing to provide data, or were completely unresponsive to multiple attempts of contacting them.

Advanced Corporate Networking, Inc.
DSCI Communications, Inc.
Great Auk Wireless (GAW Communication)
Interglobe Communications, Inc.
Meriplex Communications, Ltd.
One Communications Corporation
SkyWay USA

27 Broadband providers submitted new or updated data:

AT&T Inc.
Broadview Networks, Inc.
Cablevision
Charter Communications
Clearwire
Cogent Communications, Inc.
Comcast
Connecticut Educational Network /CEN
Cox Communications
CSC Holdings, Inc
Fibertech Networks, LLC
Groton Utilities
HNS License Sub, LLC
Level 3 Communications, LLC
Light Tower Fiber Long Island, LLC
Megapath/Covad Communications Grou,p Inc.
METROCAST COMMUNICATIONS OF CT
Reliance Globalcom Services, Inc.
Sidera Networks
Skycasters
Sprint Nextel Corporation
StarBand Communications, Inc.
T-Mobile USA, Inc.
Verizon New York Inc.

Verizon Wireless
Wave2Wave Communications Inc.
ViaSat/Wild Blue Communications, Inc.
XO Holdings, Inc.

For the Fall 2012 submission (S6), roughly 60% of the state providers submitted either entirely new or significantly revised data sets. This is slightly down from the last submission where approximately 65% of the providers submitted either entirely new of significantly revised data sets.

In general, the submission 6 processes followed the same basic approach that was used in earlier submissions. This document summarizes the following:

- Submission 6 Processing Assumptions
- Reference Data Creation
- Processing of new provider data
- Additional automated quality control checks
- Improved validation techniques
- NTIA quality control scripts
- NTIA Submission Data Model Schema Changes

An additional quality control step and additional validation were also performed on this submission to assure the quality of the data submitted.

SUBMISSION 6 PROCESSING ASSUMPTIONS

Based on NTIA feedback and information provided in NTIA webinar sessions, the submission 6 data processing workflow is based on the following assumptions to meet NTIA submission requirements.

1. All census blocks and road segments are mapped based on 2010 census data set. Any data submitted in 2000 or 2009 format was converted to 2010 for submission.

2. For this submission we again requested actual speed data from the providers in addition to max advertised and typical speeds. Approximately 80% of the providers provided this data to us. This data was then populated into an internal data model, was used to support validation efforts, and will be used to enhance the functionality of the state broadband web site.

3. Due to our NDA restrictions, last mile points are still not being submitted to NTIA.

4. Due to NDA restrictions and our inability to accurately flag service by "category of end user", address points were not submitted to NTIA for any commercial provider.

5. Some providers did not submit middle mile elevation. Wherever possible, we went back to providers to obtain their middle mile elevation information, but it is not available for every record. Due to changes in the NTIA check script, when a provider provided us with and elevation that was negative (below grade level), this value was changed to zero so the check script would not report a failure even though we feel this is inaccurate.

6. Terrestrial Mobile Wireless and Terrestrial Fixed Wireless (licensed and unlicensed) were again treated as wireless coverage and were delivered as a shape. In cases where a provider served

the same technology and spectrum with different speeds, overlapping areas were removed and the higher speed was assigned.

7. If a cable based wireline provider can provide both DOCSIS 2.0 and DOCSIS 3.0 service to the same area, the block or road was listed only once with a technology code of 40.

8. Providers were only willing to indicate on a general level if they severed business, residential or both, so we did not get any providers that broke down the type of service by block. Only if the provider stated they only serve business to business customers did we fill in the "category of end user" with a code of 2, otherwise this field was left blank.

9. The submission 6 Provider data model is currently based on the NTIA June 2012 data package.

SUBMISSION 6: REFERENCE DATA

This section describes the reference data used in submission 6.

BLOCK REFERENCE SETUP

For Submission 6, Census 2010 data was utilized. The data was set up as follows:

- Block size (AREA) is calculated combining the 2010 land area (ALAND) and water area (AWATER)
- AREA is converted from square meters to square miles to calculate square mileage (SMI).
- If the SMI of a block is less than or equal to 2, then the less than or equal to 2 square mile indicator (LE2SMI) is set to true.

ROAD REFERENCE SETUP

2010 Tiger Line IDs (TLID) were used for data processing in S6.   The data was set up as follows:

- The GT2SMI (Greater Than 2 Square Mile) indicator is set to True when:
    - The 2010 road segment is completely within a block that is NOT less than 2 square miles
- Only minimum and maximum address ranges and a single zip code for each road segment is maintained.

SUBMISSION 6: PROCESSING OF NEW DATA

For submission 6, AppGeo started data collection on July 5, 2012 by sending out data update requests and technical data specifications to all providers. This incorporated all the NTIA changes released as of August, 2012. These were sent to a large list of companies which were compiled from past collection efforts, and the revised FCC 477 list.   All new data was requested using Census 2010 geography whenever possible.

We then actively followed up with the providers. As we had discovered in the past, many of the providers listed on the FCC 477 list are either resellers, or not involved in the actual delivery of broadband. (Many are VOIP or teleconference service providers that utilize existing broadband connections.)

In our solicitation for data updates, we told known past providers that if we didn't hear from them by a certain date, we would default to using their data from Submission 5. We contacted them after the due date a few times but for six providers, we eventually had to just reuse Submission 5 data.

All data received went through the following processing steps:

1. **Triage**: All new data was quickly reviewed to understand what was received, and in what format. We also made sure we had all the required components for NTIA's data model, such as their FRN and advertised speed information. We also screened for any known issues that we might have seen before (such as Excel 2003 spreadsheets that cut off at 32k rows.)
2. **Ingest**: At this time the data is actually brought into our systems. Each provider is set up with a unique file geodatabase to store their information. Record counts of what was received is logged so that we can validate we did not drop anything in processing.
3. **Data Processing**: This is where the data goes through a number of ETL routines to convert the raw proprietary information into a format similar to the NTIA format. The exact routine utilized depends on how the data is received:

a. When a wireline provider submits a service boundary, we select all the blocks and roads inside that shape.
b. If a wireline provider submits a customer address list, the points are geocoded, and then the appropriate block or road segment is selected.
c. If a wireline provider submits block and road information using Census data, we just make sure everything is formatted to the appropriate specifications
d. If the wireline provider submits any type of road or line data that does not direct correlate to the TIGER data set, we convert the lines to TIGER by selecting the road centroid and spatially selecting the closed segment in our data set. If the road is in a block less than 2sqmi, than the block is selected. Some manual cleanup is also applied to make sure we do not accidentally drop any road segments that should have been processed.
e. Wireless provider data is formatted to ensure that there are no any overlapping polygons with the technology type. In addition the data is cropped to the state boundary.
f. After each round of processing, we make sure that we only keep unique records. A unique record is defined as having a one of a kind combination of FRN, Block/Road ID, and technology type. If there are multiple records with different speeds, but all else is equal, than we select the maximum of the advertised speeds.

4. **QC Review**: All data is then sent to a different analyst to perform a through quality control review on the processed data set. Record counts are compared to what was submitted. The QC staff also make sure the ETL scripts and routines populated all of the right fields.

5. **QA Review:** Data is then sent to another team for Quality Assurance Review. In this step the data is not only double checked against what was originally submitted, but it also brought up inside standardized MXD templates that allow us to make sure our results make sense. This often involves comparing the new data set with prior submissions, as well as looking for any possible technology or speed anomalies. At this stage we also start in on our validation process. This includes looking at the provider data in comparison to things such as speed test results, franchise boundaries, siting information, and feedback from the planning surveys.

6. **Provider Review:** Processed data is all posted to a customized web application we refer to as our Provider Portal. All providers were notified once their data was available in the site, and they were always given at least ten business days to review the data and respond. In this site, providers can log on and visually see their processed data in a map format. It also allows them to overlay their raw data to help them validate that we did indeed process things correctly. The provider portal also has a suite of markup tools that will allow the providers to edit their data, including adding or removing service areas, and making changes to the data attributes.

7. **Comment Processing:** All comments and feedback received from the provider portal, is then reviewed and applied to the processed data set. This updated data set goes back through our QA and QC processes, and if time allows, back out to the Provider Portal, for the provider to review and sign off on.

8. **Data Append:** After all of the individual data sets are processed and approved, we run an append process which merges all of the individual provider data sets into one geodatabase. This is also the point where our team will do any final transformations to get our working data model into the latest NTIA publishing format.

9. **Submission Comparison Check:** Starting with this submission an additional check was added to our quality review process. An application was written that compares the individual provider's unique data that is stored in their unique file to that which is stored in our final appended file and the NTIA submission data. Any variation in each of these data files in thoroughly investigated and resolved. This was done to assure no data loss or data transformation issues. We also compare the submission 5 dataset to the submission 6 dataset, review any variations and assure that the changes found can be documented as being requested by the provider.

10. **Final QA/QC:** A series of quality checks are run on the final appended data sets to ensure it is ready for submission to NTIA. We also run the latest version of the NTIA receipt tool at this time. If any issues are flagged as failing they are reviewed and corrected. All warnings are also reviewed and either corrected or documented in the attached document which explains that we have validated this data and it should be accepted. Any last issues are corrected, and the data is sent to the state for their review.
11. **Submission to NTIA**

As with the fifth data submission, we followed the following protocols:

1. We did not collect data from resellers
2. We collected data from satellite providers, only if they were able to provide to us all of the required information we need to pass onto NTIA: including spectrum, FRN, and advertised speeds.

COMMUNITY ANCHOR INSTITUTIONS DATA

The community anchor institutions data was primarily populated through State resources, in particular the CEN database which services many schools, colleges, and libraries. The CEN database was significantly improved for this submission by working closely with the state's BTOP team.

We also were able to get a connection survey results for all the libraries through the state library association. Location information for all other CAI points, notably, police, fire, and town halls, were obtained through the Department of Public Safety. All of this information was then populated into an online data gathering and validation web based application. Each town was contacted and asked to update their respective site information. While the web based responses have not been as high as we would like, we do feel that we are fortunate to have a good base set of data from the state.

CONNECTICUT SPECIFIC INFORMATION

Due to Connecticut's geography and population, 99.75% of the census blocks in the state are less than two square miles. The need for us to break apart coverage based on blocks versus roads leads to a lot of unnecessary confusion as well as creates some distorted pictures when you try to visualize this information on a map. For this reason, all of the maps available on the CT.gov/broadband website are published after we convert all of the data to just use blocks.

In the documentation form NTIA there has been a lot of discussion about making sure that a provider uses the same DBA and FRN consistently across all feature classes. We mentioned this to the providers, but there was some push back. Most providers complied with this request, but a few providers pointed out that while they may share a common name, they actually operate as separate organizations. Also, due to regulatory implications of the different FRN's a few providers did insist that their records not be combined.

The State of Connecticut built and maintains the Connecticut Educational Network, which is used to provide one high speed network connection to each town in the state (typically fiber, but some outliers are still on DSL.) CEN network will typically install one fiber uplink in each town, and then it is the town's responsibility to provide connection between facilities. So for example CEN may supply the board of education's office with a 10mb connection, but then the board of education will run lines to each of the schools in the district. Because of this, many towns are reluctant to report speed information as there may technically be 10mb available to the school, but reporting that speed at each school would grossly overestimate how much connectivity they have in total, when in fact there may be 15 schools sharing that same uplink. In addition, CEN's primary mandate is to provide site to site connectivity between towns, and so

they do not feel they meet the true definition of an internet provider, and as such, do not have a FRN. CEN is also limited by regulations to only support educational facilities, so they requested that their data only be shown as address points, as they cannot provide service to anyone else in that census block.