pennsylvania
DEPARTMENT OF EDUCATION

May 13, 2014

Ms. Marlene H. Dortch, Secretary
Federal Communication Commission
445 12th Street, SW
Washington, DC 20554

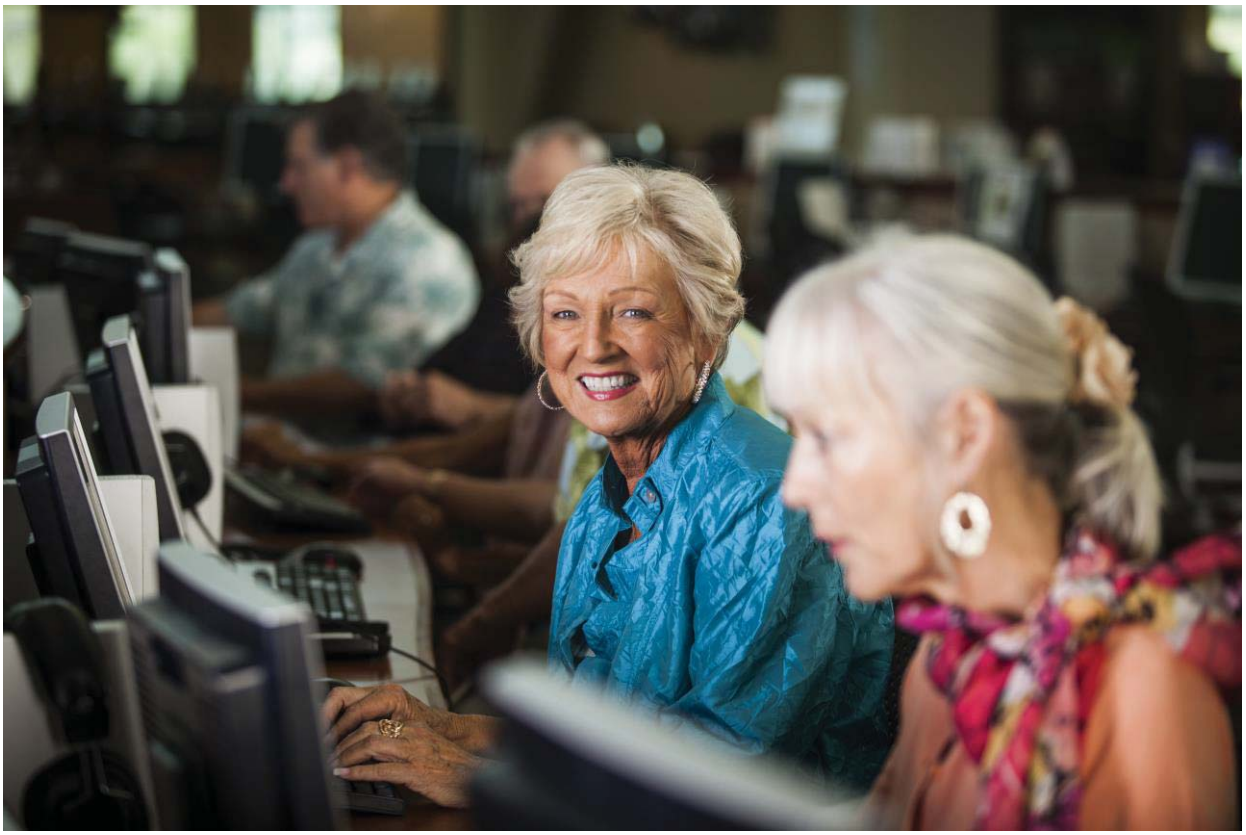
RE: Library System of Lancaster County Broadband Report

Dear Ms. Dortch,

The attached report (*Too Much Information: Bringing Clarity to Public Library Network Congestion in Lancaster*) provides data and analysis on broadband consumption and upgrades as related to the public libraries of the Library System of Lancaster County.

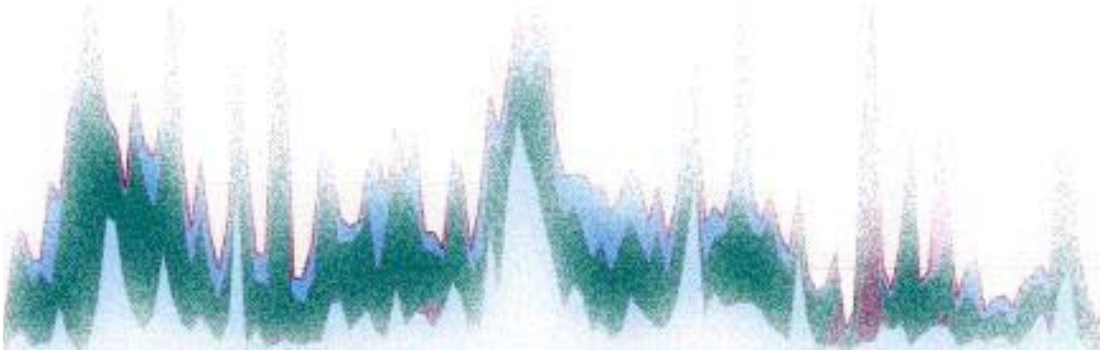
Respectfully submitted,

Stacey A. Aldrich
Deputy Secretary



TOO MUCH INFORMATION

BRINGING CLARITY TO PUBLIC LIBRARY NETWORK CONGESTION IN LANCASTER COUNTY



BY DAN COLEMAN, MANAGER OF INFORMATION SERVICES
LIBRARY SYSTEM OF LANCASTER COUNTY, LANCASTER, PA
MARCH 2014

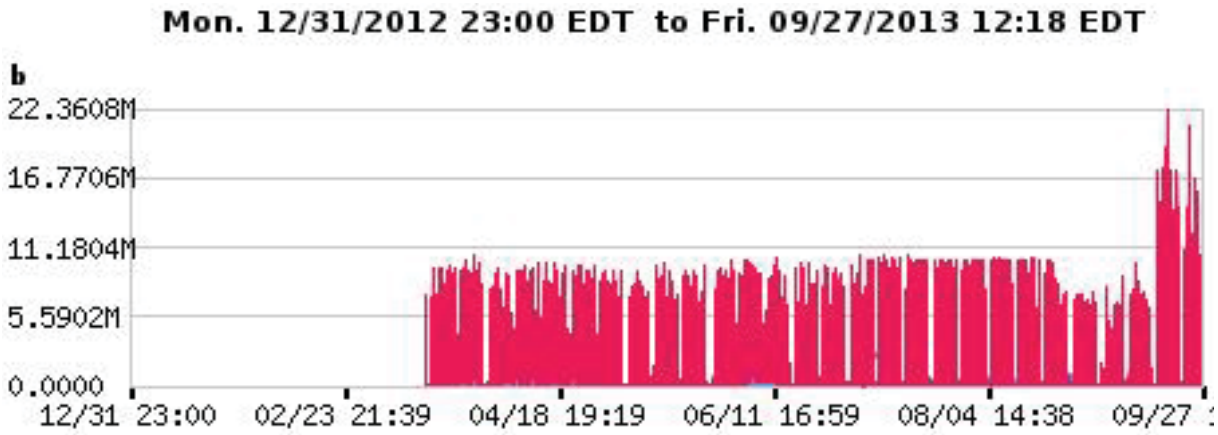
LIBRARY SYSTEM OF LANCASTER COUNTY – LANCASTER PUBLIC LIBRARY NETWORK OPTIMIZATION PROJECT

This report focuses on the network congestion challenges that faced libraries across Lancaster County in 2013. Much of the data focuses on Lancaster Public Library, because it has the highest volume of traffic and best represents a broad set of challenges that all libraries in Lancaster County dealt with in varying degrees of intensity. This report is intended for stakeholders who want to learn more about the problems and solutions that were implemented, as well as future changes to policy and network planning.

NETWORK CONGESTION – 10 MBPS TO 20 MBPS

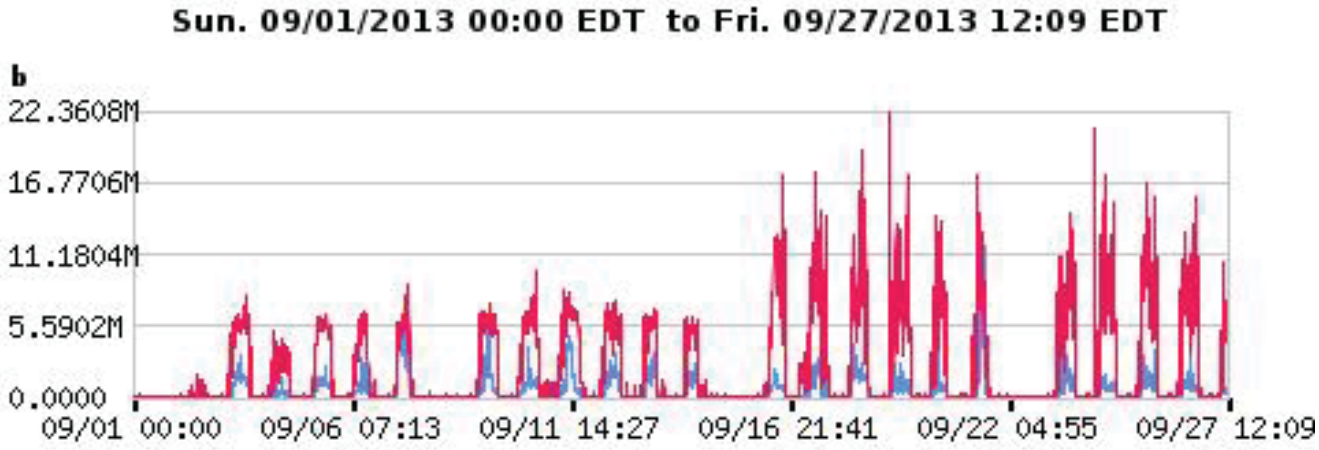
In early 2013, Lancaster Public Library (LPL) had public and staff PCs in use that were 5-11 years old. These PCs had physical hardware limitations which prevented users from processing high definition graphics and limited the ability to process multiple data consuming tasks simultaneously. In February 2013 about 25% of the staff PCs were upgraded. Then on July 4, 2013, all the public PCs were replaced, making a total of 43 PCs which were upgraded in one building. This dramatically increased network congestion.

When the year began the fiber optic link to LPL delivered 10 Mbps UP/DOWN. This link was taxed before the upgrade and significant network congestion followed the hardware upgrade. The following graph shows that between mid-March and July 4, there was heavy network use with some fluctuation. After the July 4th upgrade that data flow plateaued at 10 Mbps, or 100% of capacity. This severe congestion caused significant performance impacts to the Integrated Library System (ILS), the core application that essentially functions as a POS for all transactions. The Library System of Lancaster County (LSLC) Information Technology Services (ITS) Department did not have the tools to diagnose or analyze the data properly.



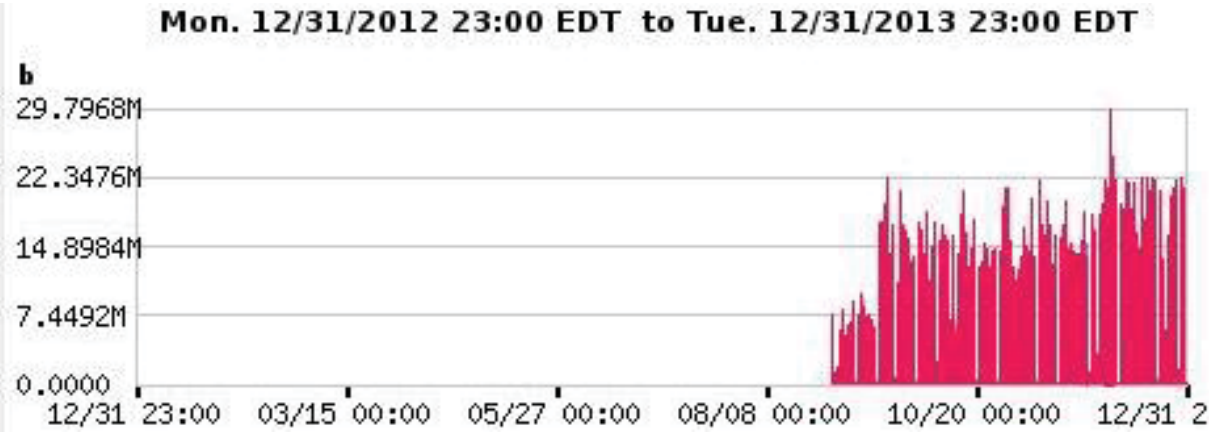
After much troubleshooting by LSLC ITS, it became clear in August that the source of the ILS performance problems was network congestion. On August 21st, LSLC requested an immediate/emergency upgrade in network capacity at LPL and other two other sites. After the Summer Reading Program Ended in August, congestion lessened, but the request for an increase in capacity continued and the ISP finally turned up in mid-September.

The following graph shows September 2013 traffic before and after the upgrade to 20 Mbps UP/DOWN. Almost immediately after the increase the additional capacity was occasionally peaking at 100%.

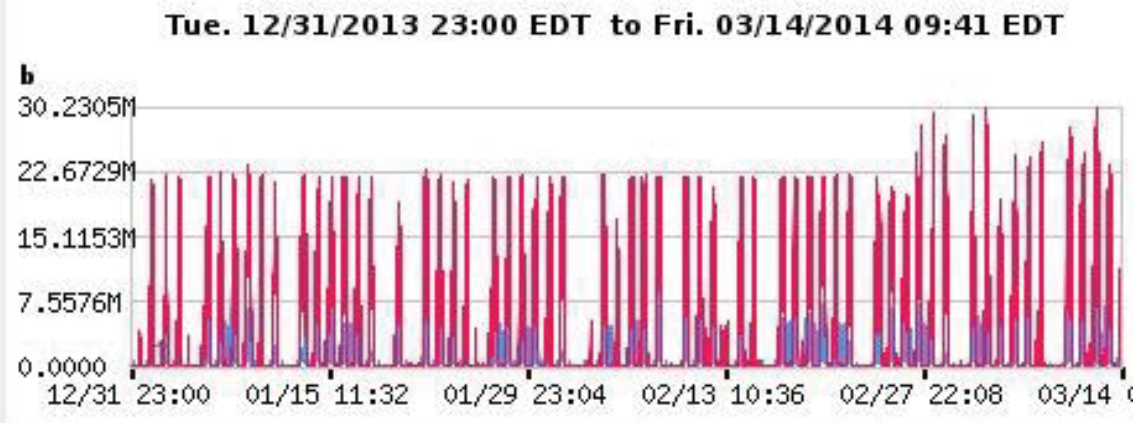


CONTINUED CONGESTION – 20 MBPS TO 30 MBPS

In September and October, LSLC ITS put the hardware upgrade project on hold and focused all resources into finding a solution to manage traffic congestion so that the ILS traffic would not be effected and to put better tools in place to analyze and understand the problems across the network. Several methods of analysis were performed by IT staff and contracted vendors. It showed that some sites were experiencing network congestion which was impacting the ILS data in a way that cause performance problems at most libraries across the county. This was impacting all sites differently, which contributed to the difficulty assessing the problem. In November LSLC IT purchased and installed a network optimization solution (http://www.exinda.com/products/wan-optimization-suite/) that allowed extremely detailed analysis of network traffic and congestion, as well as extremely effective prioritization of ILS traffic. That tool went online December 3, 2013 and analysis began immediately. LSLC ITS changed prioritization to ensure compliance with computer use policies which the previous technology was not able to do. This tool analyzed data at the application layer. This allowed ITS to identify traffic which was previously undetectable, and violated computer use policies, to now be inspected and countermeasures put in place. However, it became immediately clear that the additional capacity of 20 Mbps at LPL was not going to be adequate in the near future. The following graph shows the data use increase from mid-September through the end of 2013. There was some variability early on, but by December traffic was again plateauing. On December 4, 2013, LSLC IT requested an increase from 20 Mbps to 30 Mbps.



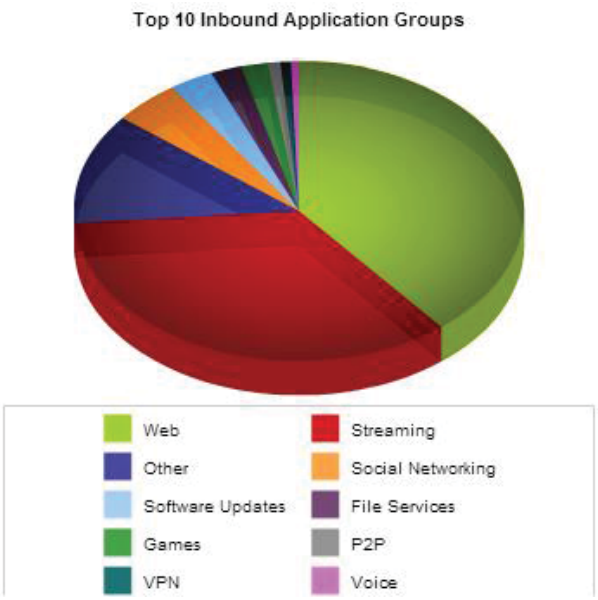
By December, a month of generally lower data use, the traffic was consistently at 100% capacity, again. Because of delays from the ISP, the upgrade from 20 Mbps to 30 Mbps was turned up on Feburary 27, 2014.



DATA USE ANALYSIS

Total Traffic:

Here's a summary of all the traffic that moved from member libraries across the Internet between December 2013, and March 2014.



- Web Traffic: 40%
- Streaming: 35%
- iTunes: 8%.
- Cloud File Storage: 7%
- Social Networking: 3%
- Software Updates: 3%
- Bittorrent: 1%
- Gaming: 2%
- Other: 1%

(All data is compiled from all libraries and represents only inbound traffic.)

PEAK USE

The total bandwidth use only tells part of the story. Peak use is the largest cause of network congestion. For most sites this is between 10 am – 11 am, and again from 3 pm – 5 pm. When usage spikes there are generally many users consuming on small or moderate amounts of data. This is then combined with only a few users, generally 1-3 devices, who are responsible for causing congestion due to large file transfers of software updates, gaming downloads, or Bittorrent downloads. The Bittorrent traffic was reduced after applying policies to align the data use with computer use policy limitations. Also, there exceptions to this rule when occasional traffic spikes due to a

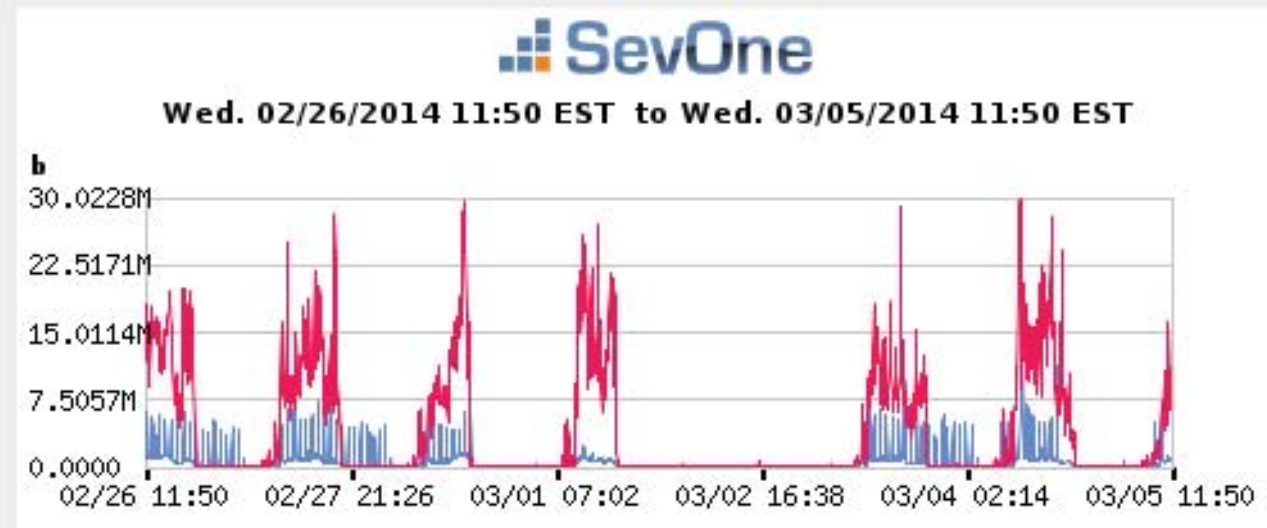
large volume of streaming content. One example is the Monday after the super bowl when LPL traffic was spiking for long periods due to at least 15 simultaneous HD streaming feeds, however this is general the exception and not the rule.

KEY CAUSES OF PEAK BANDWIDTH USE:

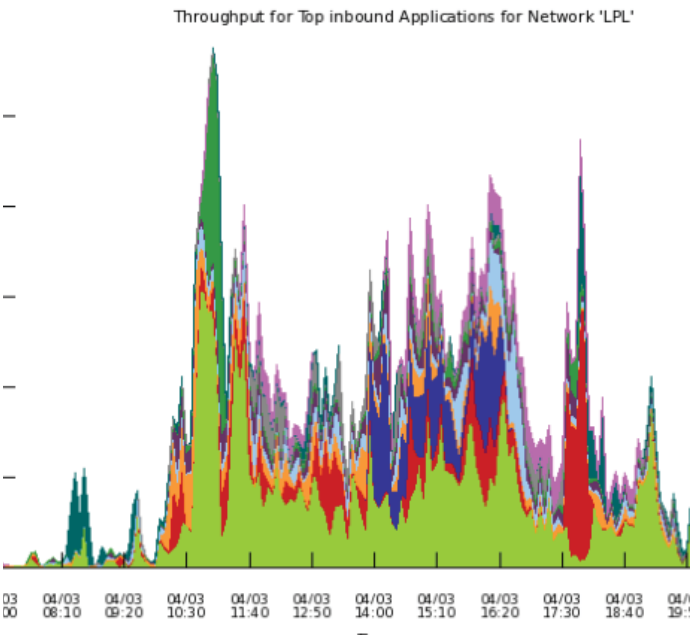
- SOFTWARE UPDATES: Patron owned devices over the wireless, usually iOS devices with large software updates for phones or tablets.
- GAMING: Patron owned devices over the wireless, generally laptop downloads of 3GB-6GB of data at one sitting.
- BITTORRENT TRAFFIC: Patron owned devices over the wireless, generally laptop downloads of 1GB-6GB of data at one sitting.
- CLOUD BASED FILE STORAGE: Patron owned devices over the wireless, generally laptop downloads of .5GB-3GB of data at one sitting.

EXAMPLE 1: LANCASTER PUBLIC LIBRARY (CURRENT CAPACITY 30 MBPS UP/DOWN)

- This site has a large number of users with a variety of data use patterns.
- Note the spike early in the day on the large graph of 3/4/14. This shows peak use spike when a single user was downloading Google Encrypted traffic. This was caused by one user downloading large encrypted files, which consumed approx. 10 Mbps for about 20 minutes. This pushed the total traffic spike to the 30 Mbps limit.



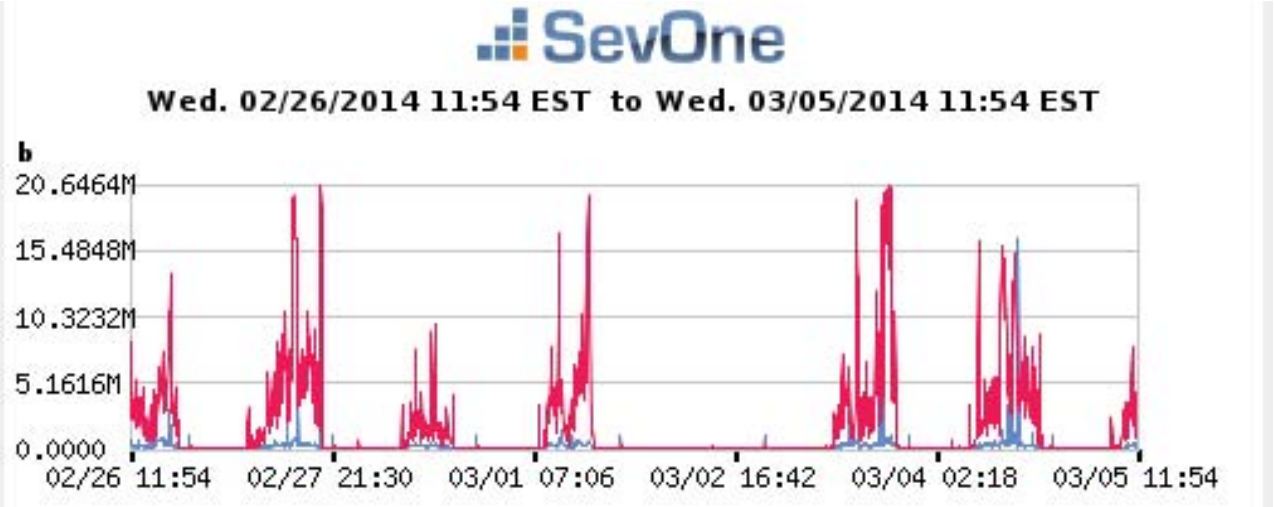
on Network 'LPL' | 07:00AM 04 Mar 2014 - 08:00PM 04 Mar 2014



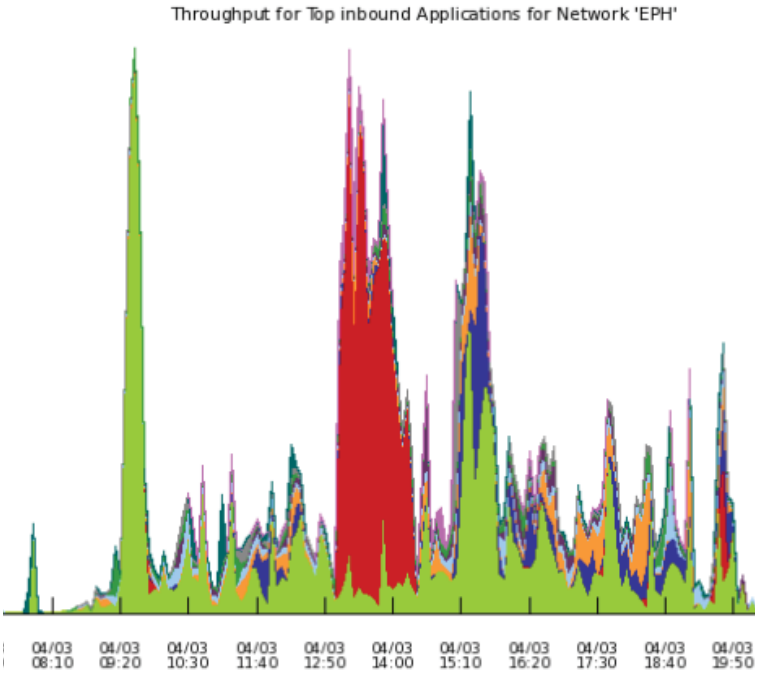
Top 100 Inbound Applications on Network 'LPL'			
	Name	Packets	Data (MB)
	HTTP	17015912	19783.745
	Quicktime	4209008	6047.840
	SPDY	3338672	4309.538
	HTTPS	4413338	4077.915
	Flash	2538563	3567.990
	Facebook	2985694	3202.911
	Google Encrypted	2333655	2935.473
	MPEG	1644772	2366.380
	Windows Updates	1297762	1864.036
	Google Docs	1129067	1564.121

EXAMPLE 2: EPHRATA PUBLIC LIBRARY (20 MBPS MAXIMUM SPEED)

- This site has more wireless users than wired users, but the total number of users is moderate and has a variety of data use patterns.
- Note the spike the afternoon on the large graph of 3/4/14. This shows peak use spike when a single user was downloading iTunes data. This was caused by one user consuming 4Gbps-12Gbps for about 3 hours.



Network 'EPH' | 07:00AM 04 Mar 2014 - 08:00PM 04 Mar 2014

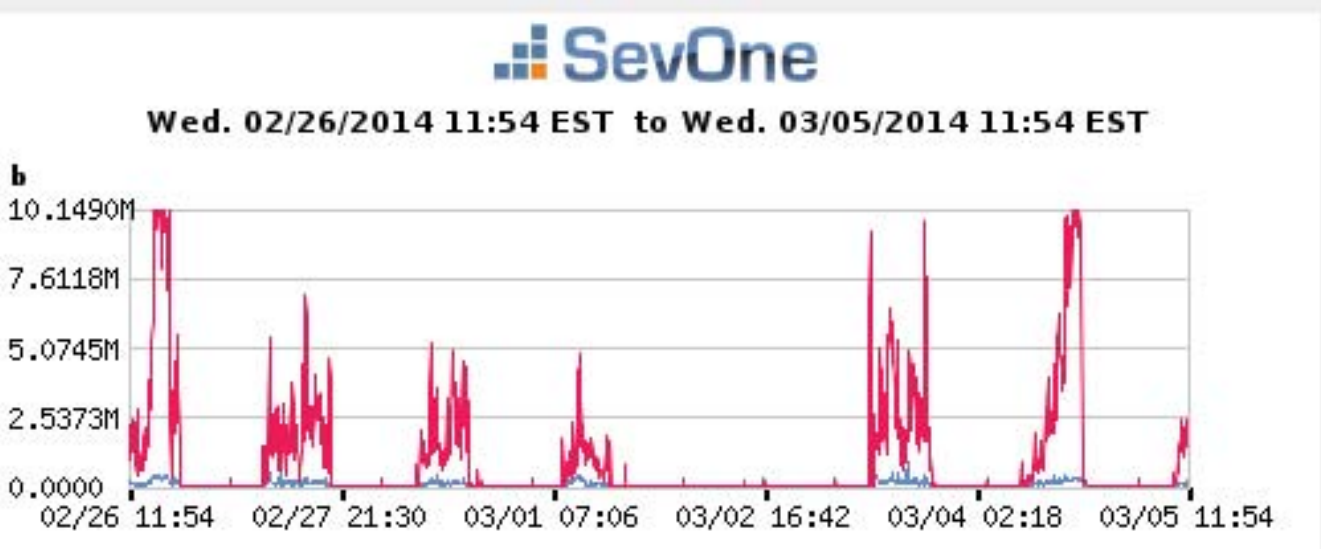


Top 100 Inbound Applications on Network 'EPH'

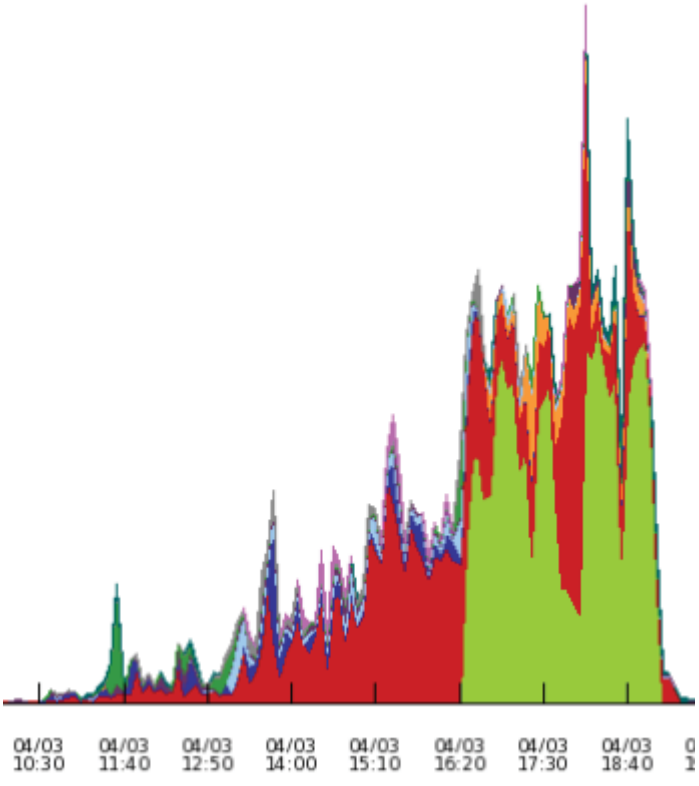
	Name	Packets	Data (MB)
	HTTP	8110922	9544.886
	iTunes	3520800	5091.423
	Quicktime	1364471	1968.296
	Flash	1268755	1731.221
	HTTPS	1327496	1131.380
	MPEG	535245	772.915
	Facebook	667125	721.176
	Google Encrypted	594981	658.289
	Windows Updates	280504	400.744
	GMail	215300	245.345

COLUMBIA PUBLIC LIBRARY (10 MBPS MAXIMUM SPEED)

- This site has a small number of users with less variety in data use.
- Note the spike late in the day on the large graph of 3/4/14. This shows peak use spike when a single user was downloading Steam Game content over the wireless network. This was caused by one user downloading large Game files, which consumed 2 Mbps-8 Mbps for about 2.5 hours. This pushed the total traffic spike to the 10 Mbps limit for over 2 hours.



Throughput for Top inbound Applications for Network 'COL'



Top 100 Inbound Applications on Network 'COL'			
	Name	Packets	Data (MB)
	Steam	4957024	7108.650
	HTTP	4723222	6405.302
	Facebook	588642	690.430
	Google Docs	351165	505.136
	MPEG	325125	467.750
	HTTPS	401592	358.795
	Google Encrypted	269975	320.208
	Flash	165510	232.570
	Windows Updates	85628	121.056
	Quicktime	67085	96.849

FUTURE PLANS AND CAPACITY BUILDING

The solution to network congestion should be analyzed for long-term and short-term success. In the short term, 2014 will require a combination of policy changes to limit patron consumption during peak times, and increases to bandwidth capacity, especially for libraries with a high volume of users that represents a large variety of data sources. In the long term, additional capacity will continue to be an ongoing issue. Identifying partners like IU13, the county, or other organizations could help reduce costs when the contract is up for renegotiation in 2017.

The following capacity increases are planned for June 1, 2014:

- Ephrata Public Library: 20 Mbps → 30 Mbps
- Lancaster Public Library: 30 Mbps → 40 Mbps
- Lititz Public Library: 10 Mbps → 20 Mbps
- Manheim Township Public Library: 10 Mbps → 20 Mbps
- Elanco Public Library: 10 Mbps → 20 Mbps
- Quarryville Public Library: 10 Mbps → 20 Mbps
- Network Operations Center: 100 Mbps → 200 Mbps

Possible network use policy changes that should be considered include:

- Limit the impact of a single user who is downloading large files from impacting the other users at a site. This can be accomplished through policy limits applied on the Exinda. It should include a maximum limit per user, which can be tied to peak times, of no more than 2-5 Mbps. This is based on recommended data transfer rate for YouTube of 2.5 Mbps to deliver 720p at 30 fps using h.264 protocol. Netflix requires far less transfer rates, generally 1.5Mbps. The Edge Initiative currently recommends between .5 Mbps and 1 Mbps.
- Eliminate Peer-to-Peer traffic from the network. This is a security risk and often moves significant amounts of data. It is also nearly always in violation of most computer use policies.
- Put lower thresholds of access for software updates to limit their impact on other users.
- Identify additional priority applications and make them a higher priority. For example, Internet telephony, Staff access to remote access tools like LogMeIn or GoToMyPC. This could help prevent future issues