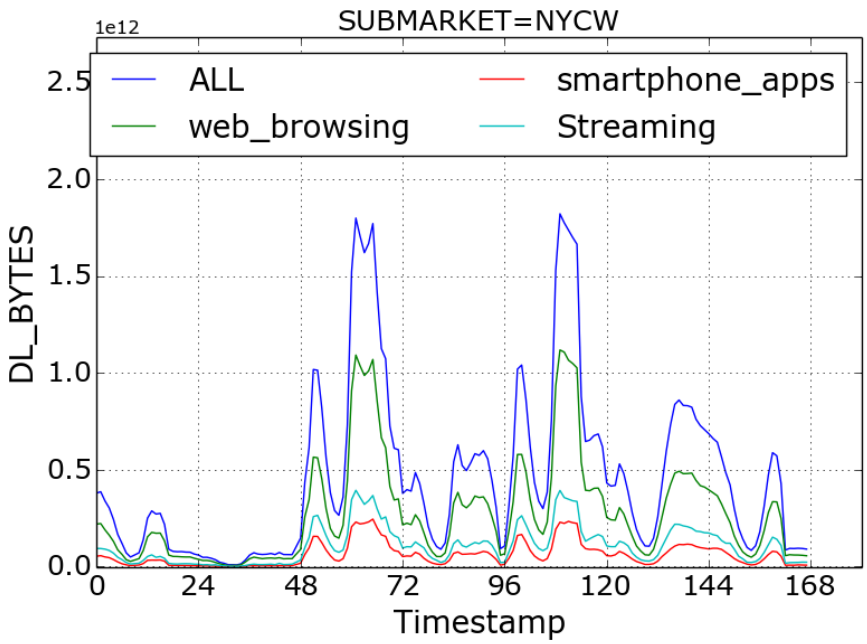
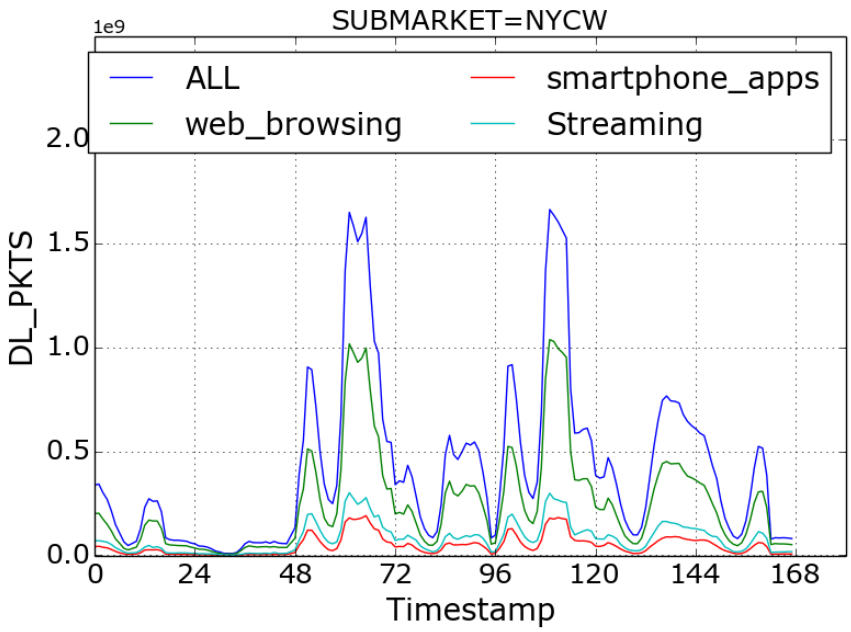
04/07 using correct schema

For below plots, we focus on a submarket: NYCNJ/NYCW

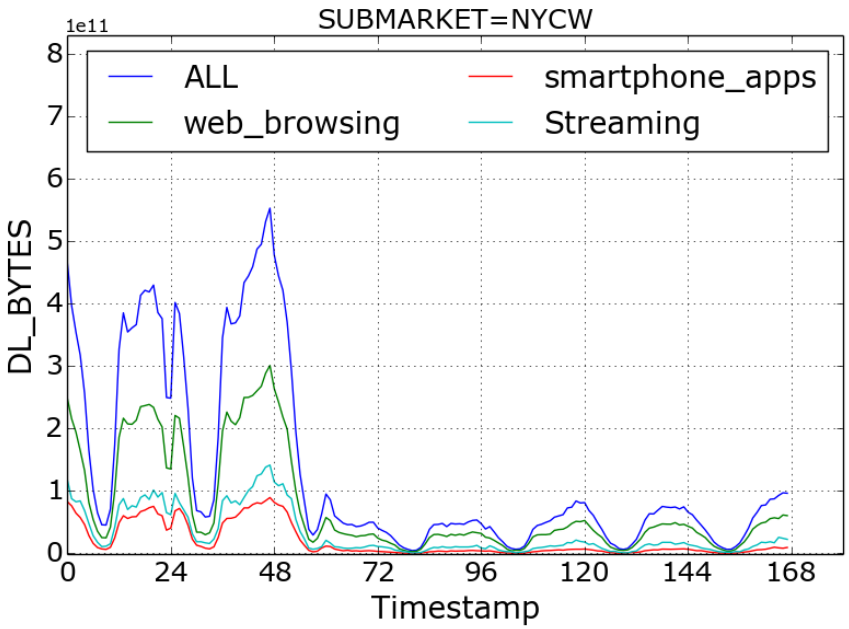
Time: 12/01/2014-12/07/2014 **Field:** DL_BYTES **Takeaway:** Although we see similar traffic volume for each day (compared to previous plots), streaming takes more portion, since right now it's Downlink and previous plots are for Uplink...	
Time: 12/01/2014-12/07/2014 **Field:** DL_PKTS **Takeaway:** DL_PKTS correlates nicely with DL_BYTES.	

Time:
11/03/2014-11/09/2014
(still Monday to Sunday)

Field:
DL_BYTES

Takeaway:
I suspect data for some days are not complete. I've ran 12/01/2014-12/07/2014, and also 11/03/2014-11/09/2014. I found that 11/03-11/09 runs much faster than 12/01-12/07, indicating much smaller data files.

For 12/01-12/07 data (above two plots), traffic peaks at 12/03, I found that 12/03 *LTE*BA* contains 51752352 rows in 24 files, however 12/02 only contains 868567 rows in 24 files (60x rows!) I checked two files (one for 12/03 and one for 12/03), the 12/03 file contains 2991055 rows while 12/02 file contains 48797 rows (61x!) but they contains similar amount of unique ENODEB ids (2929 vs. 2424). I didn't go deeper, my guess is that good days' data contains more rows for each individual ENODEB than bad days' data, although good days' data and bad days' data covers similar amount of ENODEBs.



Time:
11/03/2014-11/09/2014
(still Monday to Sunday)

Field:
DL_PKTS