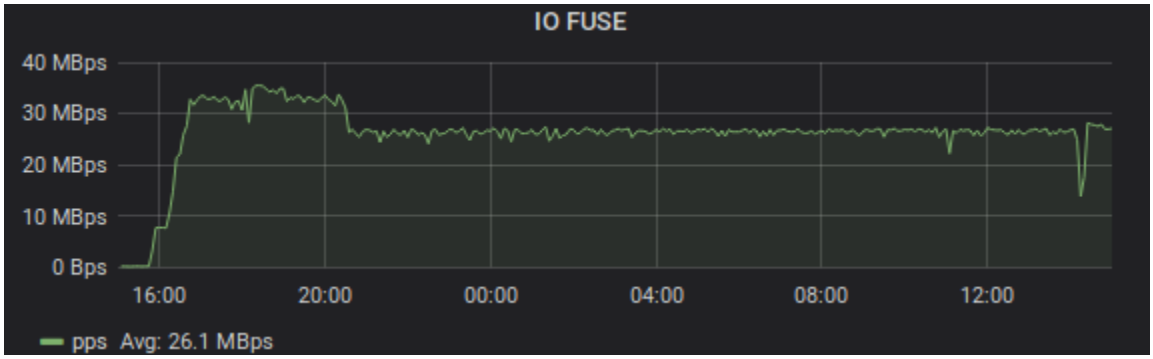
FUSEX rate test

Massimo Lamanna March 18th, 2018

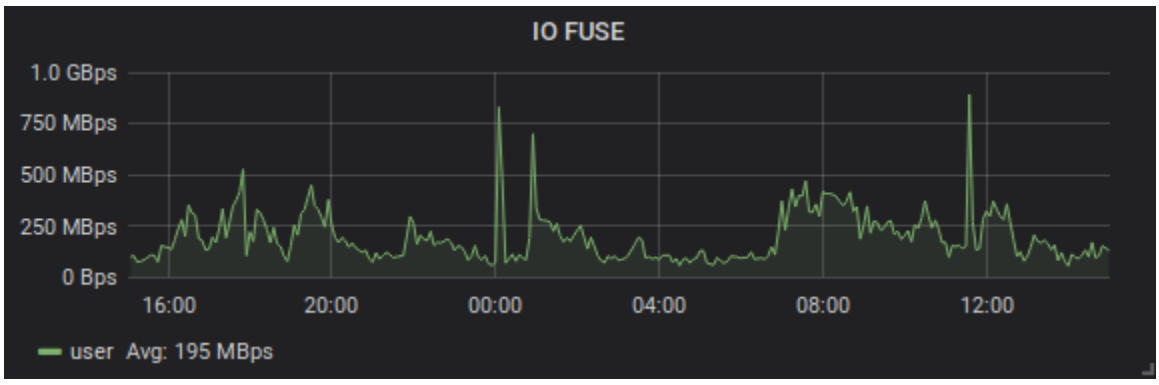
I prepared FUSEX rate test (writing) and the first one ran between 17th of March 2018 4pm and 2pm of the following day. A number of single threaded clients (5) were running from 5 machines (eospluswig).The fusex version is 4.2.17-1. EOSPPS runs version 4.2.17. This is responsible of the stable plateau from 20pm to 2pm. Some more clients were used in the initial phase (but eventually their fusex client went amok (as described in JIRA EOS-2441).

Files created during this tests are named after T1 (e.g. **/eos/pps/users/laman/run_test2/T1_eospluswig704_2694968_2_4_40_fsize:2000000_dirs:21_files:840** which is a tree of 2 levels of dirs, 4 directories in each dir. Each dir contains 40 files, 2 MB each). Files between 1MB and 20 MB have been used. Typical creation rates (per client) are around 1Hz. When measuring in rate, each client can push a few MB/s (from my small VMs).



IOFUSE (EOSPPS)

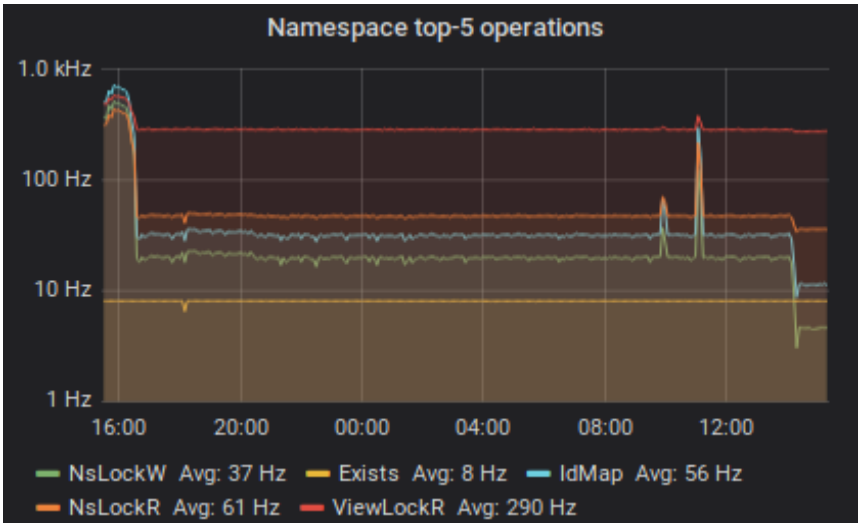
The dent around 2pm is due to a stop of T1 and a new test (T2). The only differences are that all clients try to push 50 MB and there are 4 distinct machines with 1 client each and 1 machine with 4 clients.



IOFUSE (EOSPPS)

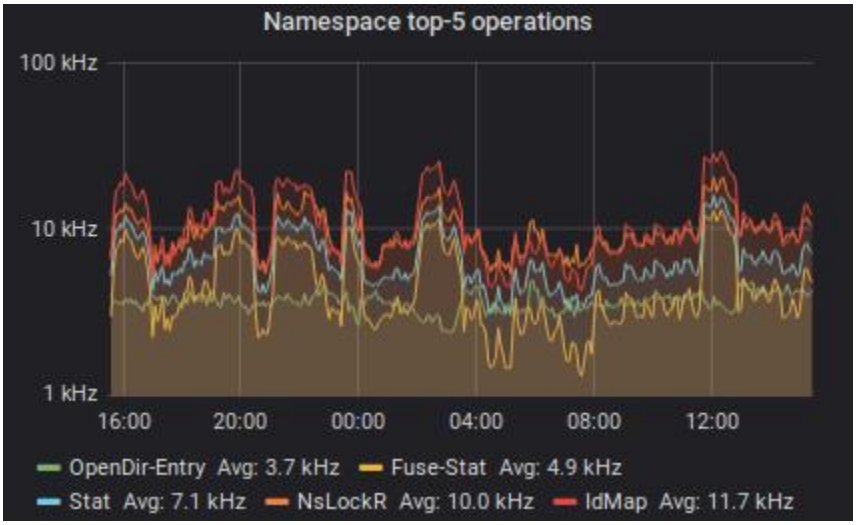
For comparison, the second plot shows the last 24h of EOSUSER (read+write activity). The monitor suggests that readers and writes activities are about 50-50 (for all activities, not FUSE only).

The namespace activity is quite different from the mix, largely dominated by ViewLockR (~290 Hz), followed by NSLockR (~60 Hz) and IdMap (56 Hz).



Namespace top5 (EOSPPS)

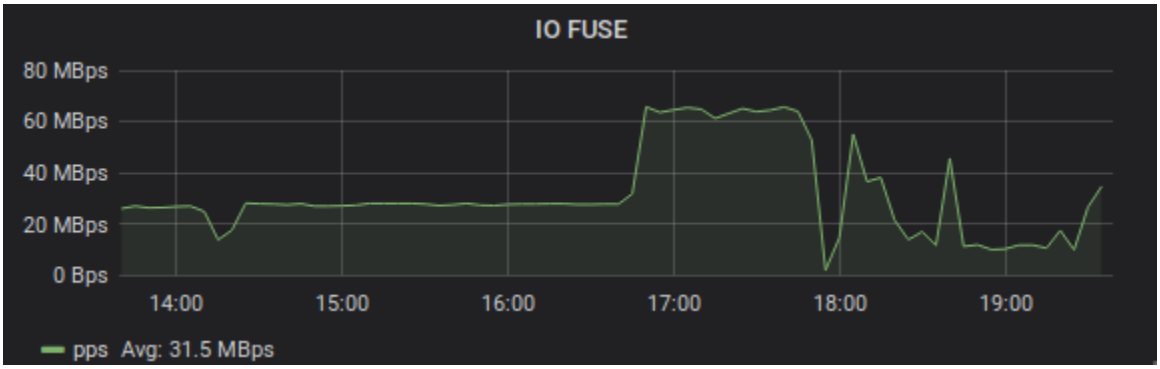
For comparison, there is no ViewLockR in the EOSUSER top-5 operations and NSLockR and IdMap are around 10 kHz (while the read+write rates are substantially higher).



Namespace top5 (EOSUSER)

The last test (T2) I have optimised the writing (cutting down the checksumming as I did in smashbox) to avoid to waste CPU in checksumming. At this point the writers are not any longer limited by the local CPU. In this second phase of T2, I run with 6 clients, 2 streams each. Each writer uses less than 10% CPU (with observed peaks of the eosxd as high as 80% CPU.

With these changes one can observe and encouraging 60 MB/s (5MB per client) although with some instabilities (to be understood. To be repeated with the newer FUSEX).



T2 test (starting at 2:30pm)

The first conclusion is that I need the mount point of EOSPPS (or EOSUAT) to meaningfully test that we can reach the same level of activity we have in EOSUSER. For some reasons my VMs cannot do much more than that (T2 test, starting at 2pm).

The second conclusion is that some of the internal operations seem to be much more "popular" on FUSEX (or the rest of the software chain) and this requires further investigation.

These tests have been performed to complement the file-creation tests which were plateauing around 1kHz (with EOSUSER sustaining ~50Hz of file creation.