EXHIBIT 5

Multifeed Last modified Thursday, August 27, 2009 at 12:25pm by Scott Stra

The best resource for troubleshooting/fixing multifeed downtime is Fixing Multifeed.

FAQ

What is multifeed?

The new feed backend system introduced in july 2008. it is named because it takes the approach of "multigetting" your friends' feeds and then aggregating/ranking/sorting/filtering them on the fly before finally displaying your home page.

How does it work?

There are several pieces of the puzzle:

1. The aggregators. An aggregator is a server that takes requests when the home page (PHP) is loaded, connects to the leaves, fetches all your friends data, aggregates/ranks/filters the stories, and sends them back to PHP via thrift. It takes "actions" from the leaves and converts them into "stories".
2. The leaves. A leaf server stores in memory all the recent actions for some subset of the users on the site. It initially loads in the actions off log files on disk, and then receives new actions via RPC's sent by the tailers.
3. The tailers. These just tail a bunch of log files and send the actions they read in to the leaves. The log files are written to by Scribe whenever a user takes a public action on the site.
4. View state. This is a memcache tier dedicated to storing each user's most recent view of their News Feeds. It is fetched in PHP before making the request to the aggregator in order to maintain consistency of stories over time.

How do I build Multifeed?

Check out fbcode, and type "fbconfig multifeed; fbmake dbg".

How do I run my own instance of Multifeed?

(generally, you want to set up an SMC tier for your service. then you can point a dashboard to your setup if you pass the get params leaf_tiername, agg_tiername, & tailer_tiername)

Aggregator

1. clone this tier: http://www.dev.facebook.com/intern/smc/smc.php?tier=multifeed.aggregator
2. optionally modify the smc_leaf_tier property to point to your own leaf tier
3. start the aggregator server via "./_bin/multifeed/aggregator_server --smc_tiername=TIER --smc_server_port=PORT" where PORT is something like 2220 and tiername is the name of your cloned tier
4. to generate feed with your aggregator, modify lib/feed/multifeed/tiers.php to return your aggregator tier name in _multifeed_get_aggregator_production_smc_name()

Leaf

1. clone this tier: http://www.dev.facebook.com/intern/smc/smc.php?tier=multifeed.leaf_0
    1. Rename it whatever you prefer (say <user>.leaf)
    2. If you get the error: "could not get tier: <user>.leaf", then run the following 3 commands from phpsh:
        1. php> $smc_manager = new ServiceManagerHandler();
        2. php> $a=$smc_manager->getTierByName('<user>.leaf');
        3. php> $smc_manager->deleteTier($a->id);
    3. Try cloning again to <user>.leaf. It should work this time. Retry a couple times if you still see the error.
    4. If that doesn't work, email dreiss or peter.
    5. Once the tier is created, delete all registered services
        1. fbcode> php tools/smc_util/smc_util.php --cmd=rm_tasks --tier=<user>.leaf
    6. Go ahead and add a new service - set the server the same as from where you plan to start up your leaf server, and choose any port (say 2220)
    7. Add property with name as "smc_task_id", set value to 0, type INT
2. start the leaf server via: "./_bin/multifeed/leaf_server --smc_tiername=TIER --smc_server_port=PORT" where PORT is something say 2220 (using the example above) and tiername is the name of your cloned tier (<user>.leaf)
    1. If you see errors indicating that /mnt/multifeed_logs do not exist, do the following:
        1. go to one of the multifeed servers running a leaf server.
        2. vi /etc/fstab, copy the line with /mnt/multifeed_logs mount information
        3. back on your server: sudo vi /etc/fstab. paste the above copied line
        4. sudo mkdir /mnt/multifeed_logs
        5. sudo mount /mnt/multifeed_logs
    2. you might want to sudo run leaf_server if you see permission denied errors as the leaf tries to read the logs
3. (you will need to have an aggregator running and pointed at your leaf tier in order to use it)

Tailer

1. clone this tier: http://www.dev.facebook.com/intern/smc/smc.php?tier=multifeed.tailer
2. Remove all services associated with this tier as explained in the leaf cloning section above.
3. choose a machine to run tailer on (say multifeedio10.snc1)
4. add the machine as the single service in the tier on whatever port
5. edit smc_leaf_tiername in the tailer tier to pt to your leaf tier
6. fbdeploy <tier> start to get the package onto the tier and start it

7. fbdeploy tail_log should give you the progress and errors. It takes a few minutes to open all 1000 logs and start the tailing.

How do I deploy a new version of Multifeed to the tier?

We use fbpackage and fbdeploy for all our tiers. That means starting/stoping servers is just an fbdeploy call. The tiers are configured with all the needed parameters to start up correctly.

Deploying to a single machine (in the case of hardware failure), is therefore trivial. The challenge is deploying to an entire tier while still serving traffic.

Deploying the aggregators

If multifeed is currently running, the sitevar MULTIFEED_AGGREGATOR_PRODUCTION_TIERS show which tiers are getting the load. Assume it's "multifeed.aggregator_1". This means that there are 2 tiers right now that are getting requests: multifeed.aggregator_1_west_v1, and multifeed.aggregator_1_east_v1. The region and version are auto-appended by the frontend. What we need is to deploy the new aggregators on another tier, and slowly move the load over. Then shut down the current one.

1. build the correct package (from multifeed release branch run fbpackage build multifeed.aggregator"
2. deploy the package to tiers that are not being used (multifeed.aggregator_2_[east,west]_v1). Just run "fbdeploy multifeed.aggregator_2_west_v1 start; fbdeploy multifeed.aggregator_2_east_v1 start.
3. Check that the dashboard shows the tiers as up
4. edit sitevar MULTIFEED_AGGREGATOR_LOADTEST_CONFIG and send some load to the new tiers
5. monitor scribe for multieed fatals. you may see a few in the very beginning, but after a few minutes, there should be none.
6. slowly raise the load to 100% on the new tier. continue to monitor scribe logs to make sure there are no fatals
7. change the loadtest load back down to 0 for the new tiers
8. shift 1% of production load to the new tier (edit MULTIFEED_AGGREGATOR_PRODUCTION_TIERS)
9. if everything looks fine, shift the rest of the load to the new tier
10. after a few minutes, it should be safe to turn off the old aggregator tiers by doing "fbdeploy stop" on them. make sure they are not getting any requests by looking at the multifeed_samples log in scribe.
11. Good job. Get yourself a cookie.

- Note: Check the the launch tool at http://www.intern.facebook.com/intern/launch_tool.php to see what percentage load is at. MULTIFEED_AGGREGATOR_LOADTEST_CONFIG will only divide load as a fraction of what is set in launch tool.

Deploying the leaves

Leaves should be redeployed 1 rack at a time. So use fbdeploy with the --tasks argument to only stop/start a rack of leaves at a time. Once the leaves start up, use the dashboard to monitor their loading progress. When the whole rack is back up and fully loaded "green", you can do the next rack.

Deploying the tailers

To update the package run: 'fbpackage build multifeed.tailer' from fbcode directory

To start them on the west coast, run: 'fbdeploy multifeed.tailer start' from fbcode directory.

This will start two tailers with the latest package.

If you are restarting, you can use --tasks=1 or --tasks=2 to restart a single tailer. Waiting a few minutes in between assures no loss of data. Both tailers are configured to run on all buckets, so only one needs to be running at a time, and the second is for redundancy.

For the east run: fbdeploy multifeed.tailer_east_0 start

What to do when a leaf goes down?

Just restart the leaf. It will take a little while for the leaf to fully load in all the data, but it should eventually start working again.

What to do when an aggregator goes down?

Restart it. Aggregators are stateless and so can be restarted at will (as long as there are enough running to handle the current load).

What to do when a tailer goes down

Restart the tailer. The bad news is that the leaves will be missing anything that happened while the tailer was down, so if the tailer was down for any substantial amount of time, you've got to cycle the leaves so that they can load in all the data. Do not take down all the leaves at once or there will be trouble!

How to check if a specific leaf or aggregator is working?

What machines are running multifeed?

- see dashboard for up-to-date machine-specific info: http://www.dev.facebook.com/intern/feedtools/multifeed_dashboard.php
- west coast
  - node group: multifeed.snc1 - west coast aggregators/leaves
  - node group multifeedi.snc1 - west coast tailers + scribe node
    - see dashboard for info on which are using tailers
  - multifeedmc.snc1 - view state machines (000-010)
  - mfscribe001.snc1 - scribe node
- east coast
  - node group multifeed.ash1
    - multifeed{001..105}.ash1 - east coast aggregators/leaves
    - tailers running on multifeed118.ash1/multifeed119.ash1
  - node group multifeedmc.ash1 - view state machines (000-010)
  - mfscribe001.ash1 - scribe node

- cdb: multifeed.settings

Multifeed Host counts in 'aggregator and leaf' tiers

	sys. multifeed	sys. multifeed.ash1	sys. multifeed.snc1	sys. multifeed.ash1	sys. multifeed.snc1
SMC TOTALS	410	439	400	439	400
multifeed.aggregator_2_east_v2	120	319		319	
multifeed.aggregator_2_west_v2	280		120		120
multifeed.leaf_east_3	120	319		319	
multifeed.leaf_3	240		120		120

What is the configuration of those machines?

- snc1 machines: /mnt/multifeed_logs => nfsutil302.snc1:/vol/multifeed01
- ash1 machines: /mnt/multifeed_logs => nfsutil202.ash1:/vol/multifeed01
- currently running kernel 2.6.21

How to check on the health of multifeed as a whole?

Check out the dashboard - http://www.dev.facebook.com/intern/feedtools/multifeed_dashboard.php and Ganglia - http://ganglia001.sf2p.facebook.com/ganglia-multifeed/

Check the fatals log - tail -f /mnt/scribe/multifeed_fatal/multifeed_fatal_current

Check the ods latency graph - http://www.fburl.com/?key=999116 If latency is not below 20-30ms, or is ramping up, there is likely a problem. Recently restarting the leaves has taken care of some of these issues.

Check the ods request rate - http://www.fburl.com/?key=999118

Error logs

- Most important is fatals: all fatals (=> retried the max number of times and given up/fallen back to old feed) are logged in /mnt/scribe/multifeed_fatal/multifeed_fatal_current. Tail this log to see how many users are failing requests. format described here: https://trac.facebook.com/trac/browser/tfb/trunk/www/lib/feed/multifeed/multifeed.php#L294
- 1 out of 100 requests get logged to /mnt/scribe/multifeed_sample/. format is as described here: https://trac.facebook.com/trac/browser/tfb/trunk/www/lib/feed/multifeed/multifeed.php#L309
- all errors are recorded in /mnt/scribe/multifeed_error/ in format described here: https://trac.facebook.com/trac/browser/tfb/trunk/www/lib/feed/multifeed/multifeed.php#L209

There is a script running on devrs004.sctm that monitors the fatals log and sends alerts when they get too many errors. The script is in fbcode/multifeed/scripts/logmon.py and is itself monitored by chronicle.

What do I need to know about machine failures/upgrades?

It's important to not restart too many machines at a time, if it's at all avoidable.

For the multifeedi machines (two of which are running the tailers), please only restart 5 at a time and *make sure* the 5 do not include both tailers. Ie. one tailer should always be left running. You can check status of the tailers on either the multifeed.tailer smc tier or the dashboard to see where and whether they're running. (TBD: Make tailers run on start-up). If tailers don't run on start-up, we need to be careful to restart them after each reboot.

What tools are available for working with multifeed?

- http://www.dev.facebook.com/intern/feedtools/multifeed_log.php - this can show you what an employee's feed looked like for any load of their feed, and some meta-data about the request that was made to get that feed. one thing to note about multifeed_log: it "caches" your days worth of feed in a separate file (ie it does "grep uid allemployeesforthedayfile > yourcachefortheday") so if you load it and then reload your feed you have to click the "force reload" checkbox to have it regenerate your cache for the day. this is pretty necessary since the log file is like 1-2 GB/day and grepping through it every page load is slow (as you can see by how slow it is the first time you load the page)...
- http://www.dev.facebook.com/intern/feedtools/multifeed_console.php - this lets you do queries against the multifeed backend with custom parameters or for custom lists of users.
- http://www.dev.facebook.com/intern/feedtools/multifeed_dashboard.php - this shows the state of the multifeed tier (which machines are running, how caught up the leaves are, etc.)
- http://www.intern.facebook.com/intern/ods/dashboard.php?dashboard_id=149 - monitor latency
- http://www.dev.facebook.com/intern/smc/smc.php?tier=feed - smc page for feed, allows you to create your own multifeed tier or modify parameters for the live multifeed tier.
- http://www.dev.facebook.com/intern/feedtools/stats.php - the usual feed stats page; you can filter it to see different stats about multifeed vs old feed.
- http://www.dev.facebook.com/intern/launch_tool.php - we're using this to control which users we have launched to.
- http://ganglia001.sf2p.facebook.com/ganglia-multifeed/ - standard ganglia monitoring
- http://nms006.sf2p.facebook.com/megarrd/cgi-bin2/viewMetrics.cgi?namespace=/storage/netapp/nfsutil302.snc1.facebook.com/volumes/vol/multifeed01 - megarrd monitoring for filer
- http://syslog01b.tfbnw.net/cgi-bin/dash.cgi?d=15 - show the port channels on the core switches that the top of rack switches uplink to. It also shows the gigabit ports that make up the port channel.

Who do I contact if I need help with multifeed?

- multifeed@lists.facebook.com is the best point of contact
- multifeed-errors@lists.facebook.com is for automated emails about stuff like machines going down
- if necessary, you can contact david braginsky, serkan piantino, lior abraham, ari steinberg, or any of these people: http://lists.facebook.com/mailman/roster/multifeed

How does the algorithm work?

not really answering this question, but various settings for the algorithm (such as app weights, story type weights, etc) are set in the cdb multifeed.settings.

This article has been marked as Evergreen. The Reaper will duly disregard this article as it makes the rounds.