**Description**

It would be useful to have several additions to the crunch job logging system:

The ability to set at least one more set of rate limits (i.e. another byte limit / interval specification) to rate limit on larger time intervals. In our system, we are using the current limits to avoid overwhelming crunch-dispatch (job) with vast amounts of instantaneous log output, so we have the limit set to 64kB over 1s.

However, other parts of the system are more effected by longer term limits. For example, at our current short term limit of 64kB/s our jobs could still log over 230GB of log per hour. In order to avoid filling the postgres log table too quickly, we might set a much lower hourly limit (perhaps 100MB over 1h).

Currently, our only longer term option is to set the absolute job log limit, which means once that limit is exceeded we never hear from the job again. It is not a good user experience to be wondering if a job is still running ok after 5 days because it stopped logging after it reached the limit on day 2.

It could also be useful to have an excluded time at job startup during which one or more of the interval rate limits don't apply.

Finally, it would be good if there was an option such that crunch-dispatch (jobs) own log files were not effected by the rate limit (or limits), as if we have external means to rotate those logs it should be possible for a sysadmin to check on job status by examining those log files on disk.

**History**

#1 - 08/29/2017 01:45 PM - Tom Morris

- Target version set to Arvados Future Sprints