The Most Active Commit Authors of the OpenStack Nova project

In this report, I present an overview of the commits found in the openstack/nova project in GitHub. For this report, I consider the most active commit authors as the ones who committed more frequently to the project. First, I identified the 15 most active commit authors from the last six months (Since 2021-02-01 to 2021-07-29) and how many commits have they contributed in that period. Second, I present the distribution of the number of commits by a commit author during the same time period. For the rest of this report, I refer to a commit author as an author.

Table 1 Number of commits by the three most active commit authors

Author	Number of Commits
Zuul	296
Stephen Finucane	116
Lee Yarwood	43

During the past six months, 40 authors committed nearly 590 commits in total. Among them, the three most active authors are listed in Table 1 (to see all 15 most active authors, run the python project.). Accordingly, *Zuul*, a Gerrit user, shows the highest number of commits. Furthermore, Gerrit is a code review system. Therefore, for a commit from other authors (contributors of the project), a corresponding commit by Zuul could be seen in the repository, For example, Stephen Finucane's commit as shown in Fig. 1, there is a corresponding commit from Zuul.

Unsurprisingly, the data I collected for the last six month shows that Zuul committed 50.25% of all the commits during that time period.

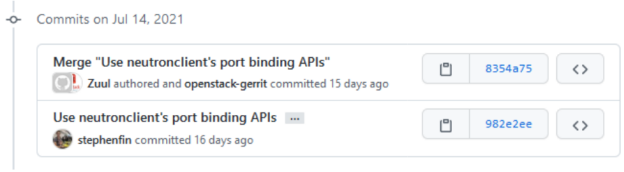

Figure 1: Stephen Finucane's (who's GitHub username stephenfin) commit "Use neutronclient's port binding APIs" and it's corresponding commit from Zuul.

Now I discuss the distribution of the number of commits per author. It is shown in Fig. 2: the x-axis represents the number of commits and the y-axis represents the number of authors. According to Fig. 2, the distribution of commits is right skewed. Fig. 3 shows a zoomed-in version of Fig. 2. According to Fig. 3, most of the authors had committed less than 10 commits.

The minimum number of commits by an author was 1, which is also the median number of commits by an author. Then, the average number of commits by an author was 14.7 and the standard deviation was 49.7. Since Zuul accounts for nearly half of the number of commits, it should be removed to get a better sense of the average and standard deviation values. Thus, removing Zuul's data from the dataset resulted in an average number of commits by an author as 7.5 and a standard deviation of 19.9.

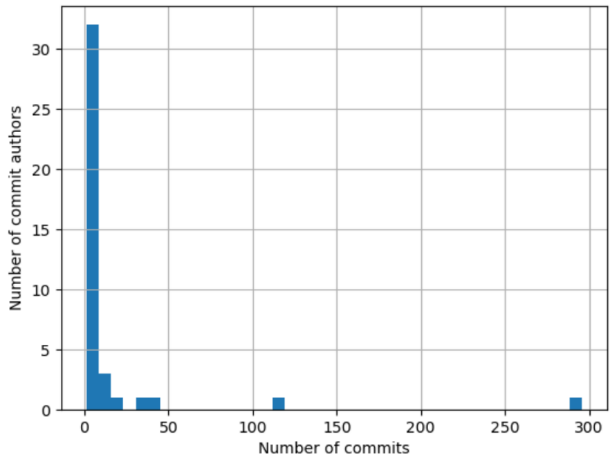

Figure 2: The distribution of number of commits since 2021-02-01 by the authors

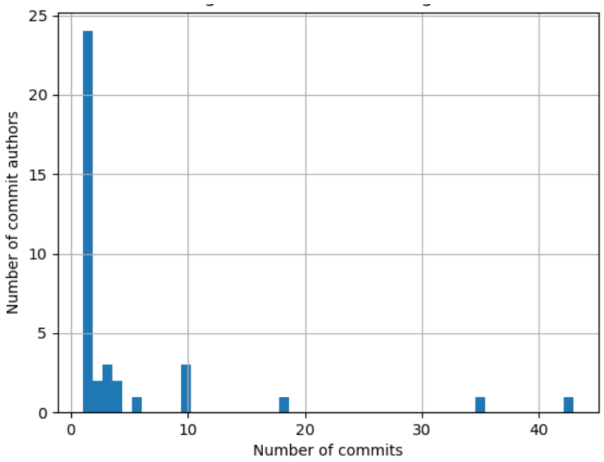

Figure 3: A zoomed in version of Fig. 2