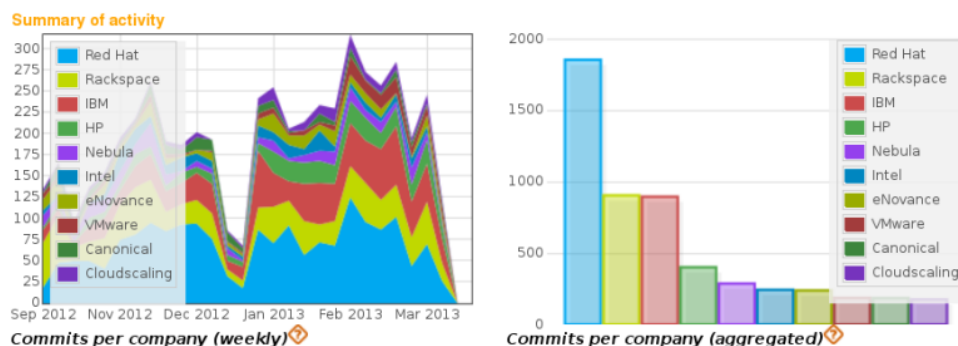


Bitergia's blog about Software Development Analytics

Open Source and Inner Source development analytics examples, tools and experiences from Bitergia staff

Companies contributing to OpenStack (Grizzly Analysis)

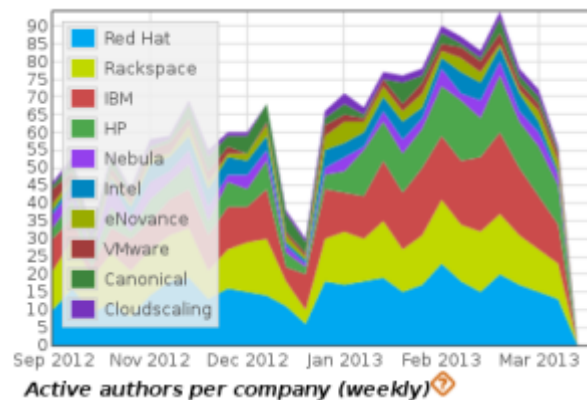
As an update of our analyses on previous releases of OpenStack (<https://blog.bitergia.com/2012/09/27/how-the-new-release-of-openstack-was-built/>), today we present our analysis of company contributions to Grizzly (http://bitergia.com/public/reports/openstack/2013_04_grizzly/), the new release of the project being published today (<https://www.openstack.org/blog/2013/04/openstack-grizzly/>). OpenStack (<http://openstack.org>) is a well known free / open source software project providing facilities for building private and public clouds. It is also a good example of a development community in which almost all participants are affiliated to some company. Therefore, it is interesting to study how companies contribute to the project by means of their employees. This is exactly what we did in our analysis.



(http://bitergia.com/public/reports/openstack/2013_04_grizzly/)
Commits by company, all OpenStack projects, Grizzly release cycle

The results are revealing, and more when you put them into context by looking at how companies were behaving in previous release cycles. The data is telling a story of how a project which was clearly dominated by a single company (RackSpace) has led to a real community with many actors. With more than 20 different companies contributing every week and a total of more than 50 companies which

contributed at some point, OpenStack is becoming one of the projects with more corporate involvement. The number of people involved is also large, and growing. In September / October, when the release cycle was starting, about 40-50 persons contributed each week. At the end of the cycle, in March, they were about 80-90.



(http://bitergia.com/public/reports/openstack/2013_04_grizzly/scm-companies.html)
Active developers by company, in OpenStack git, all projects, Grizzly release cycle

The top corporate contributors are no surprise: Red Hat, Rackspace, IBM, Nebula, HP and others are well known for their support of the project. Maybe it is interesting to see how Rackspace is slowly handing over the leadership of the project to other companies, and how quickly some of them are getting involved.

[Now, go on reading this post, or have a look at the complete analytics dashboard for Grizzly (http://bitergia.com/public/reports/openstack/2013_04_grizzly/)]

[Note (Apr 6th 00:10 CEST): Thanks to those of you who are reporting what seem to be errors or inaccuracies. We're having those reports into account and will produce a new run of the data soon. However, up to now they don't seem to change things significantly except for specific developers or maybe some company going one position up or down in some list. Please, keep those reports coming.]

The data shows how no single company is today controlling the project. The main contributor per number of commits (Red Hat) is responsible for much less than 20% of all contributions, and other metrics are providing similar results. The community is showing the usual combination of some more involved companies, and a long tail of more casual contributors. The summary charts for all companies tell the different stories of their involvement: many are growing in contributions, while some are stable and a minority are having a decreasing participation (after factoring the usual ups and downs).



(http://bitergia.com/public/reports/openstack/2013_04_grizzly/company.html?company=Red%20Hat)

Activity of Rackspace during the Grizzly release cycle

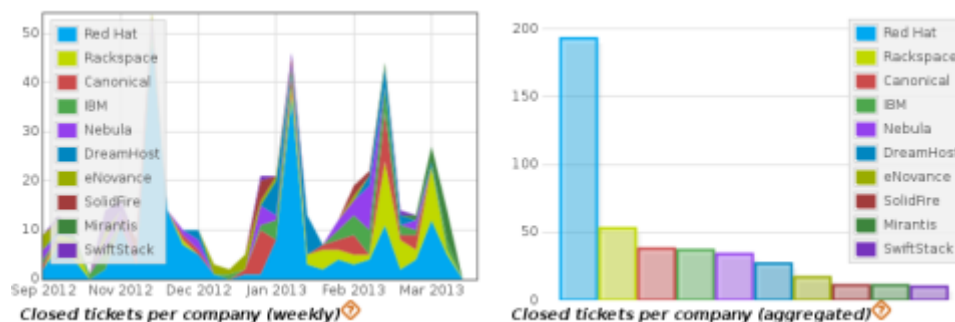
When looking at the larger contributors by commits, those stories are apparent. Red Hat (http://bitergia.com/public/reports/openstack/2013_04_grizzly/company.html?company=Red%20Hat) has become the first contributor, with about 15-20 persons involved in development. IBM (http://bitergia.com/public/reports/openstack/2013_04_grizzly/company.html?company=IBM) is the rising third (or second, depending on the metric), with a clear increase during the period, up to maintaining a team of 15-20 during the last weeks. Rackspace (http://bitergia.com/public/reports/openstack/2013_04_grizzly/company.html?company=Rackspace) is the other main contributor, stable during the period with a team of about 10-15 active developers.



(http://bitergia.com/public/reports/openstack/2013_04_grizzly/company.html?company=Rackspace)

Commit activity of Rackspace during the Grizzly release cycle

The different charts show how companies are not only involved in shaping the source code, but are also contributing by working to close bugs, and are participating in the mailing lists. All in all, the history of these six months of OpenStack development is one of growth, increased involvement by most of the companies participating in it, and increasing participation by number of companies.



(http://bitergia.com/public/reports/openstack/2013_04_grizzly/its-companies.html)

Tickets closed by company, all OpenStack projects, Grizzly release cycle

The raw data about several metrics of activity in the git repository for all identified companies can be downloaded as an spreadsheet [ODF

(http://bitergia.com/public/reports/openstack/2013_04_grizzly/data/spreadsheets/OpenStack-Grizzly-activity.ods), Excel

(http://bitergia.com/public/reports/openstack/2013_04_grizzly/data/spreadsheets/OpenStack-Grizzly-activity.xls) or rendered as PDF

(http://bitergia.com/public/reports/openstack/2013_04_grizzly/data/spreadsheets/OpenStack-Grizzly-activity.pdf) formats]

Methodology notes:

The analysis has been performed by retrieving data from the development repositories of the project: git for changes to the source code (official repositories (<https://github.com/openstack/>) are hosted in GitHub (<http://github.com>)), Launchpad for tickets (<https://launchpad.net/openstack>), including bug reports, and Mailman for mailing lists (<http://lists.openstack.org/cgi-bin/mailman/listinfo/>). For the retrieval, MetricsGrimoire (<http://metricsgrimoire.github.com/>) was used, and vizGrimoire (<http://vizgrimoire.github.com>) is the basis of the analysis and visualization. We have identified affiliations for developers based on data provided by the OpenStack Foundation and on our own research. For source code changes (git), we have considered authors of commits (as opposed to committers of commits, which for OpenStack are in many cases bots such as Gerrit). For tickets, we have considered those closed by authors of commits (using their identifiers in git and Launchpad to do the matching) and for email messages we have considered messages sent by authors of commits as well (in this case matched using their address).

No single metric can show all aspects of the participation of a company in the project. In particular, number of commits or number of tickets closed are just proxies for activity and contributions, and cannot be considered as exact measures of any of them: commits and tickets may be very different from each other, and have a very different value for the project. However, the analytics used in the report is usually considered as a good approximation to the overall activity.

April 4, 2013 by Jesus M. Gonzalez-Barahona Categories: bitergia, community, opensource, reports | Tags: openstack | 30 Comments

30 thoughts on “Companies contributing to OpenStack (Grizzly Analysis)”

Joseph Heck (@heckj) says: April 4, 2013 at 11:53 pm
Soren Hansen works for Cisco these days, not Nebula

Reply

Joseph Heck (@heckj) says: April 4, 2013 at 11:55 pm
likewise Vish Ishaya is at Nebula, not Rackspace. I suspect you should do another deeper review of the organization listings...

Reply

Daniel Izquierdo Cortázar says: April 5, 2013 at 7:41 pm
Hello Joseph,

First, thanks a lot for your comments!. Indeed data was correct, at least for Vish Ishaya and Soren, but the issue was found in the queries when talking to the database and only in those parsing information from the mailing lists.

So, this is why you saw both developers with repeated values in all of the companies that they were part of at some point (as probably happening with other developers in the case of mailing lists analysis).

Now, checking at the charts and tables, this is fixed.

http://bitergia.com/public/reports/openstack/2013_04_grizzly/company.html?company=Nebula
http://bitergia.com/public/reports/openstack/2013_04_grizzly/company.html?company=Rackspace
http://bitergia.com/public/reports/openstack/2013_04_grizzly/company.html?company=Cisco%20Systems

Just to let you know a bit more about the methodology, we already made a couple of analysis on OpenStack, so this is a mix of data from the OpenStack Gitdm project in GitHub plus our own dataset plus polished data from the community and feedback in the blog.

Thanks for the comments!. Please, let us know any other concerns you may have, it's great to have feedback.

Regards,
Daniel.

Reply

Daniel Izquierdo Cortázar says: April 5, 2013 at 7:50 pm
BTW, in case you're interested in more in depth analysis of the matching between developers and affiliations, I recommend you to have a look at the databases that can be found at:

http://bitergia.com/public/reports/openstack/2013_04_grizzly/data/db/

Our purpose is to be as transparent and neutral as possible. This is one of the reasons to release all of our software (retrieval, filtering and visualization) as free software.

So any doubt about the methodology, incoherent results or features you would like to have, let us know!

Regards

Reply

Eric Windisch says: April 5, 2013 at 3:34 pm

The number of authors per company seems inconsistent. I've seen variances between what is in your spreadsheet and what is in the JSON files.

Also, the commits-per-author statistic is interesting, but it should not be a mean, it should be a median.

Reply

Daniel Izquierdo Cortázar says: April 5, 2013 at 8:20 pm

Hi Eric,

We're working on it.

The issue is related to the fact that we were not counting unique identities in a correct way in a couple of queries.

Give us some minutes and this should be fixed.

Thanks for your comment, and let us know any other incoherent data you may find!

Regards

Reply

Daniel Izquierdo Cortázar says: April 7, 2013 at 4:06 pm

This finally took longer than expected, so just to keep current version, we decided to wait till Monday to update results and to be sure that the new dataset is coherent.

Thanks again Eric!

Daniel Izquierdo Cortázar says: April 8, 2013 at 5:02 pm

Finally charts and tables are updated. Issues came from queries in some of the cases where we were counting identities and not real people (in some cases you may find more than one identity for a given developer).

So, this was happening in the case of the static info per company in each companies's page, but not in the tables in the main page.

Thanks for your patience!

Regards

Reply

Tweeks (@Tweeks_tx) says: April 6, 2013 at 4:56 am

Hehe.. great article guys.. love to see the momentum, but I'm a little surprised that in every one of your graphs, the graphing library you're using does not remove the last (zero) data point. According to your metrics, as of March, there are zero commits, zero developers, and all of your spikes and

surges are all trending to zero.

Great article, but the zero trends are a bit distracting and cheat your trending.

Just sayin..

Tweeks

Reply

Daniel Izquierdo Cortázar says: April 7, 2013 at 4:04 pm

Hi Tweeks,

Thanks a lot for the comment, as you said this is an issue. In any case, please notice that after the removal of such zeros, there will be again a decreasing trend. This is due to the fact that the last week is not complete. The methodology takes sets of seven days, and the last set contains less than seven days.

We're working now on a more general approach where the whole history of OpenStack could be visualized together with specific releases, so please keep in touch!

Regards

Reply

Pingback: OpenStack: Grizzly release is here!! | homepage of atul jha == koolhead17

Pingback: Cloud-Software OpenStack 2013.1 "Grizzly" freigegeben | virtualfiles.net

98pm says: April 9, 2013 at 4:09 pm

Reblogged this on 98pm and commented:

Informative analysis on OpenStack contributors

Reply

Pingback: Will Red Hat's OpenStack Contributions Turn To Gold? | MattsLens

Pingback: Will Red Hat's OpenStack Contributions Turn To Gold? | Ngọc Đông Hải Ngoại

Pingback: Will Red Hat's OpenStack Contributions Turn To Gold? | TechDiem.com

Pingback: OpenStack and Network Virtualization | The Console Blog - VMware Blogs

Pingback: Will Red Hat's OpenStack Contributions Turn To Gold? | BEMBA.COM

Pingback: Cloud Interoperability and The Battle For The Open Cloud | Technology

Pingback: Cloud Interoperability and The Battle For The Open Cloud | NewsCred Plugin For WPMU

Pingback: RDO: una comunità per l'uso di OpenStack con Red Hat

Pingback: Ericsson, Red Hat, and SAP Ventures Back Mirantis with Investment : VMblog.com - Virtualization News and Information for Everyone

Pingback: It's true, Rackspace and Red Hat dominate OpenStack contributions | How to Password

