QUALITY METRICS REPORT

2014

A quality metrics report of SLED 11 SP 3 & SLED 12.

TABLE OF CONTENTS

Contents

EFFECTIVE BUG RATE

Effective Bug Rate

Goal: The goal for calculating effective bug rate is to reflect on part of our bug catching quality.

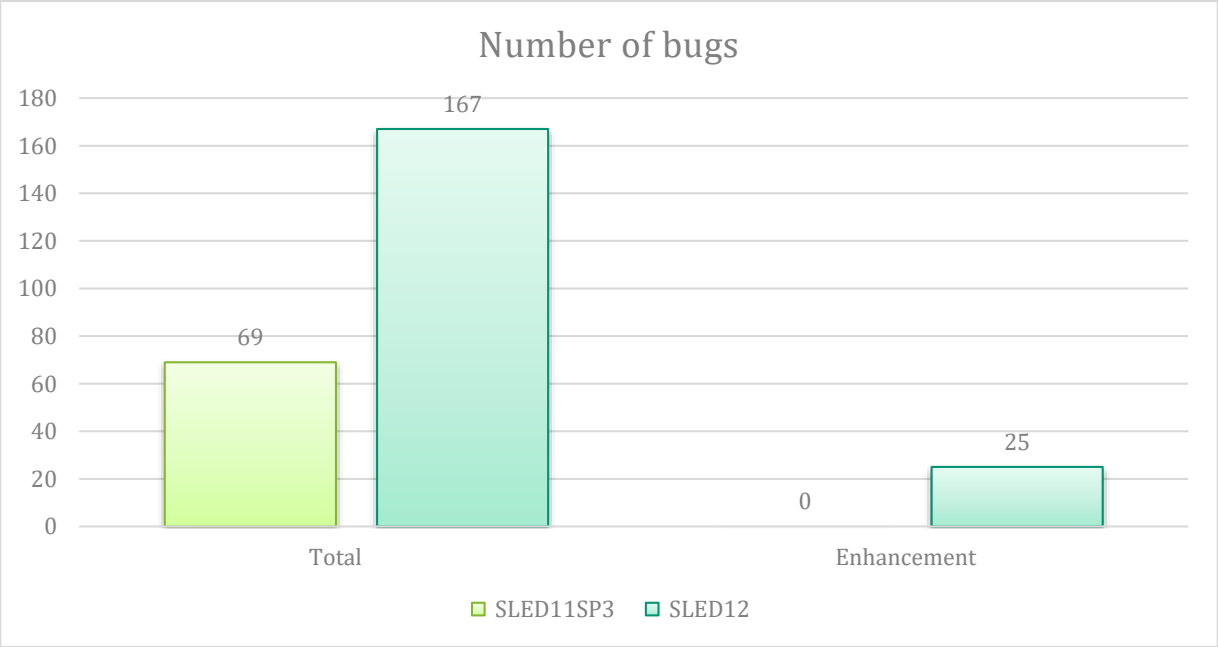
Effective bug rate is calculated by the following **formula**:

$$EBR = \frac{ALL\_BUGS - ENHANCEMENT - DUPLICATED - INVALID - WONTFIX - WORKSFORME}{ALL\_BUGS}$$

Effective bugs are defined as bugs that are not duplicated nor invalid, bugs that are to be fixed, with its solution works for most people, and are not an enhancement. All bugs are defined as bugs that are reported by the people in QA APAC I team instead of everyone as the effective bug rate is only valid when a scope of people is set.

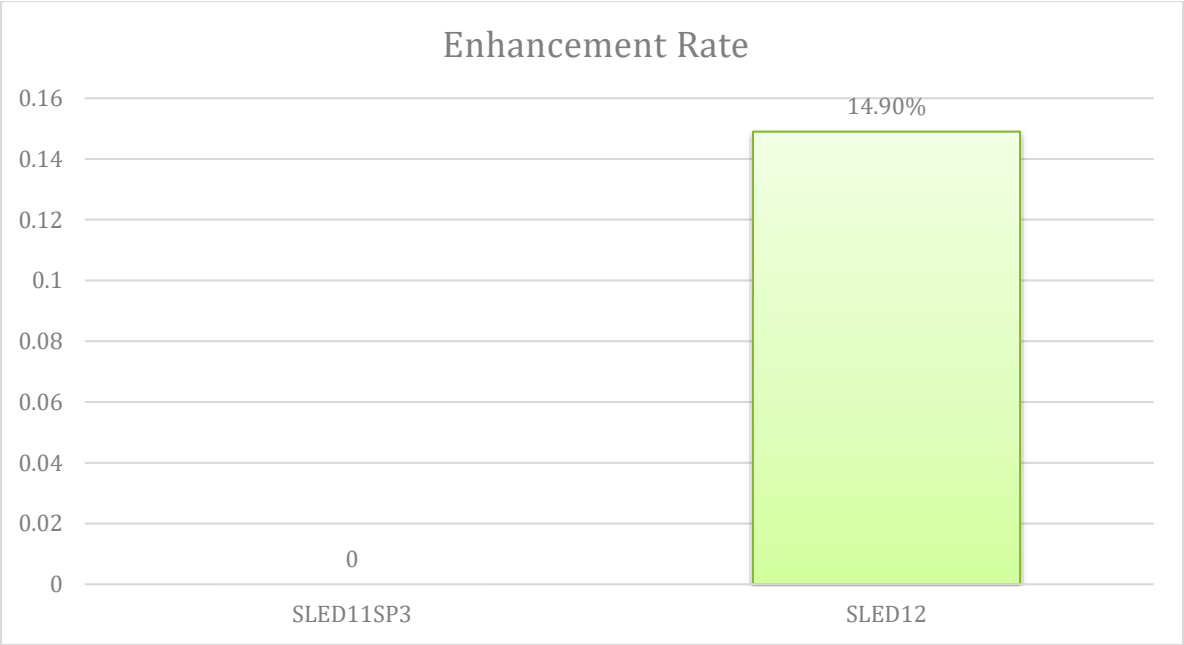
Two effective bug rates were calculated, EBR for SLED11SP3 and EBR for SLED12.

The **results** are as follows:

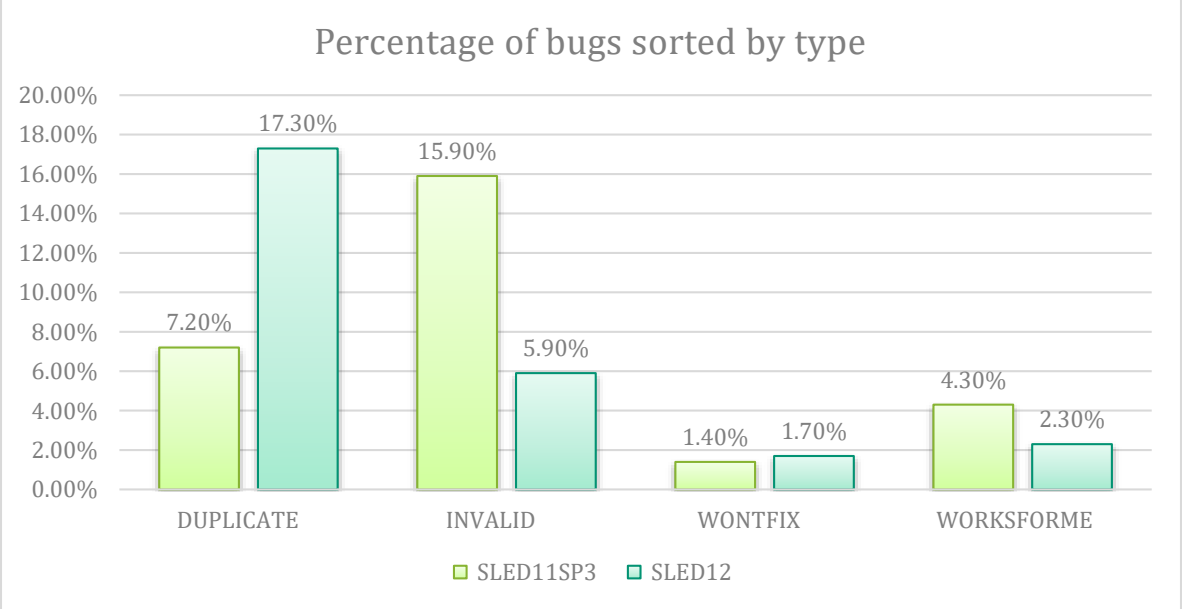


In SLED11SP3, there were 69 bugs in total, while in SLED12, there were 167 bugs altogether. The number of bugs in SLED12 was twice as many as in SLED11SP3. Among these 167 bugs, 25 of them were enhancement, making up to 14.90% of the total bugs. We believe that this is a sign of the hark work in our team.

EFFECTIVE BUG RATE

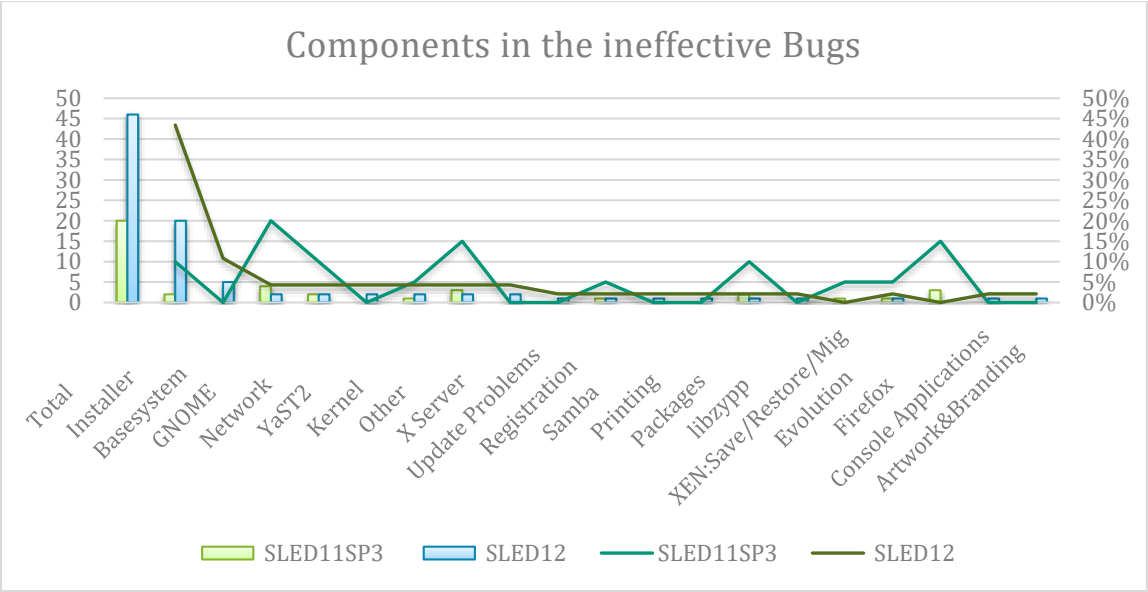


In SLED11SP3, we filed no enhancement bugs, while in SLED12, 14.9% of the bugs we filed were enhancement bugs.

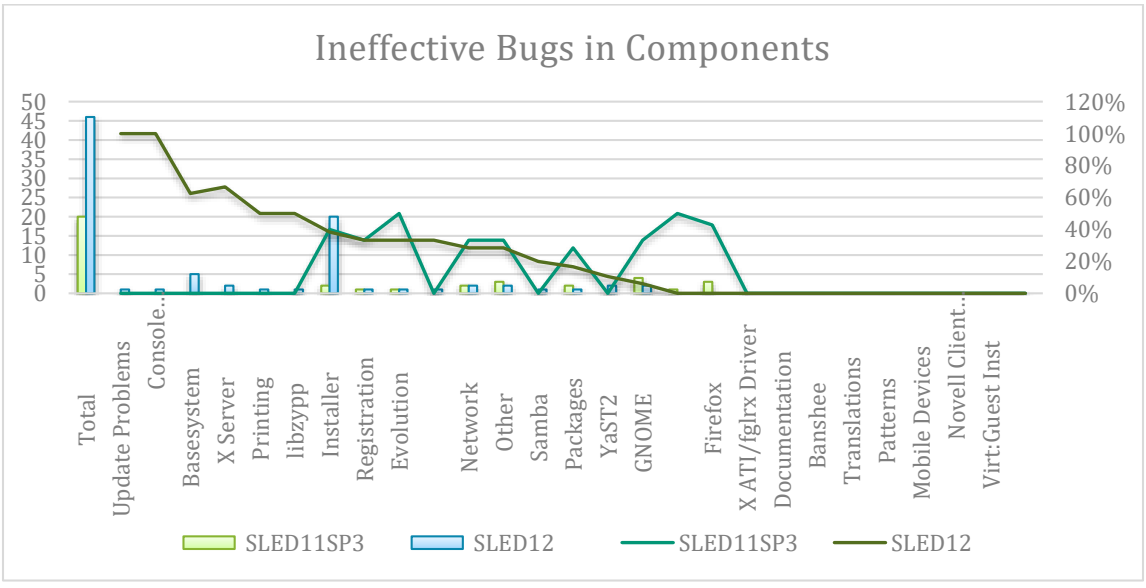


In the graph presented above, we can see that in SLED12, we filed a lot more (around 10%) duplicate bugs while in SLED11SP3, invalid bugs took up to 15.9%.

EFFECTIVE BUG RATE



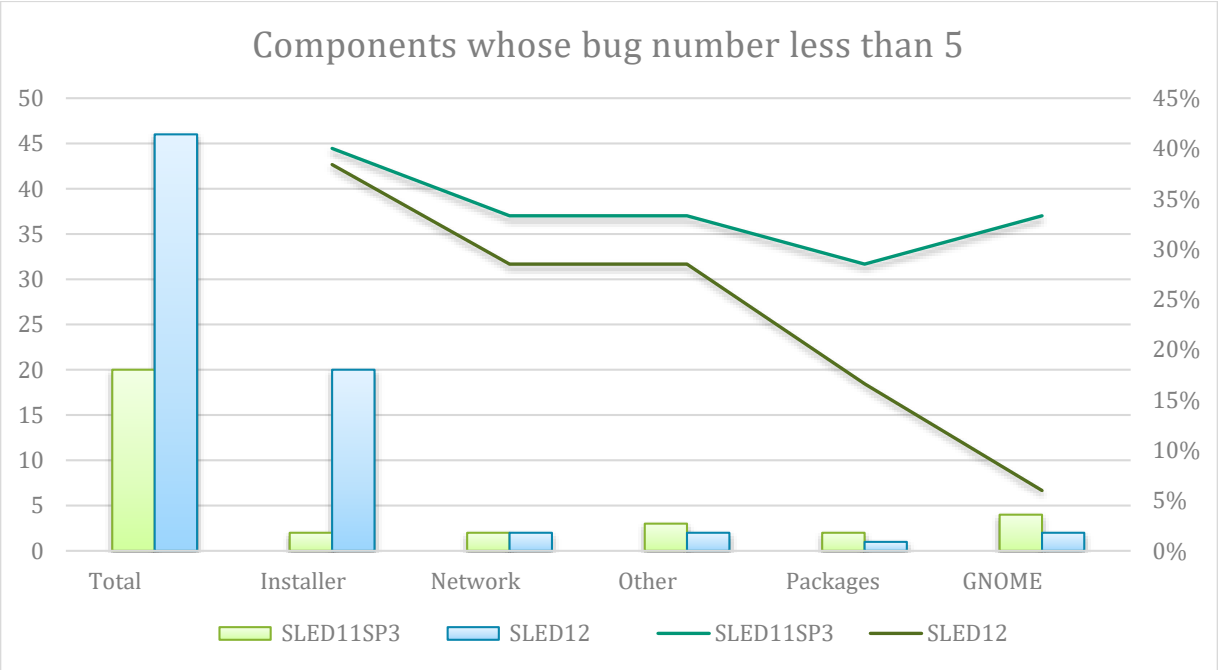
Among all the ineffective bugs, installer took the most part. In our daily testing, we do validation test, including installation test and upgrade test, approximately every two weeks. In these rounds, installer was the most tested, therefore, we believe that this is the reason for installer to be the No. 1.



Also, in the bugs filed per components, installer of SLED 12 had the most bugs. We believe this has the same reason with the graph above.

These two chart shows that installer has the most ineffective bugs. We believe this is due to the every-round installation and upgrade test.

EFFECTIVE BUG RATE



This chart indicates the components with less than 5 bugs. Because the denominator is too small, we think this is not quite useful for our analysis, so we isolated them out.

Backlog Management Index

Backlog Management Index (BMI) is a metric to manage the backlog of open, unresolved, bugs.

Backlog Management Index is calculated by the following **formula**:

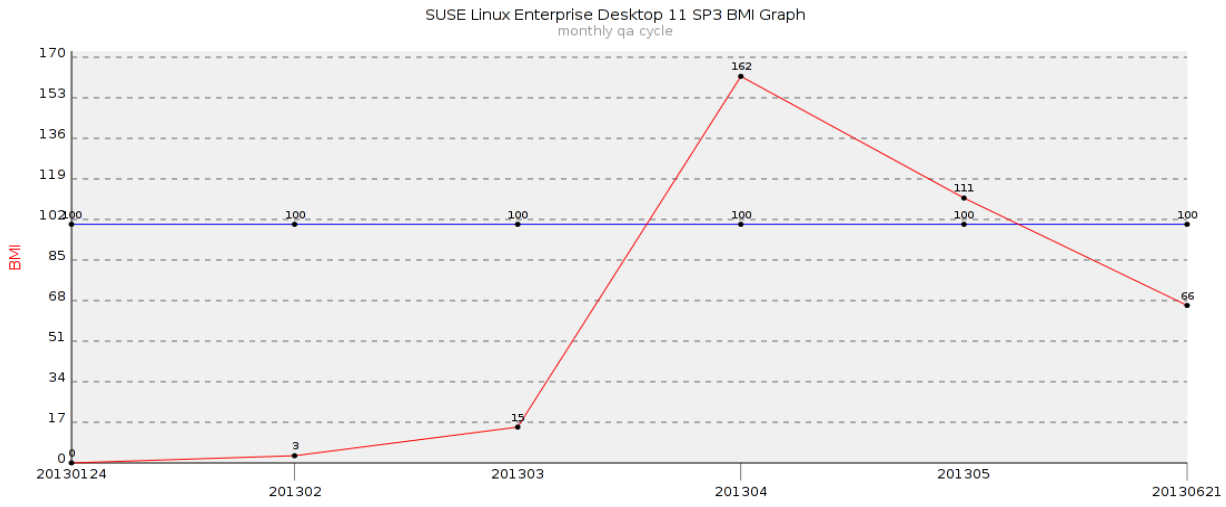
$$BMI = \frac{Number\ of\ bugs\ closed\ during\ the\ month}{Number\ of\ bugs\ opened\ during\ the\ month} * 100\%$$

As a ratio of number of closed, or solved, bugs to number of bugs arrivals during the month. If BMI is larger than 100, it means the backlog is reduced. If BMI is less than 100, then the backlog increased.

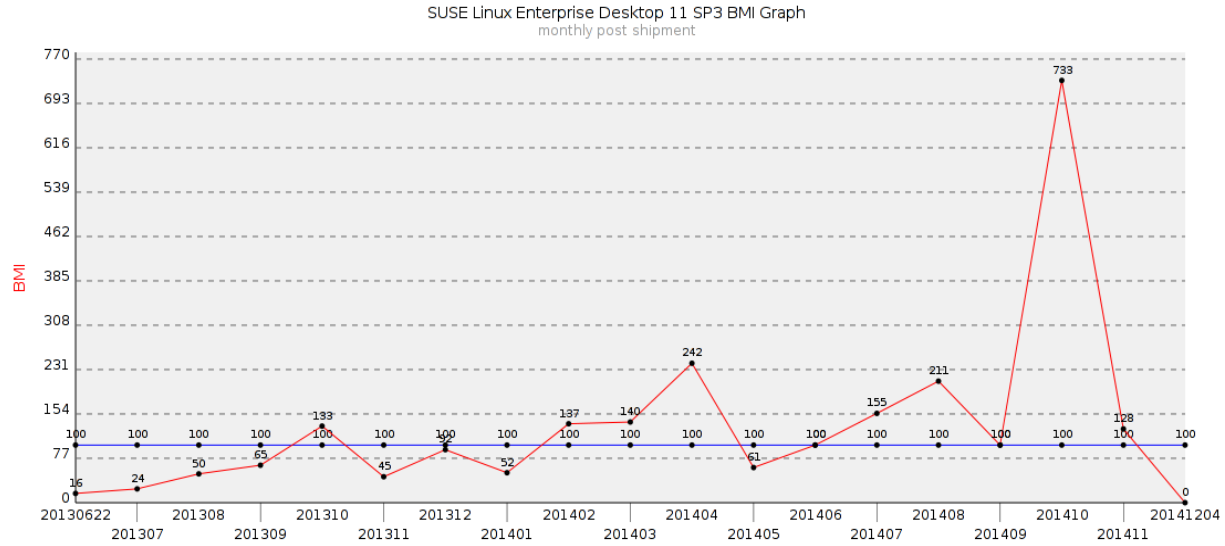
The **goal** is always to strive for a BMI larger than 100. A BMI trend chart should be examined together with trend charts of bugs arrivals, bugs fixed, and the number of bugs in the backlog.

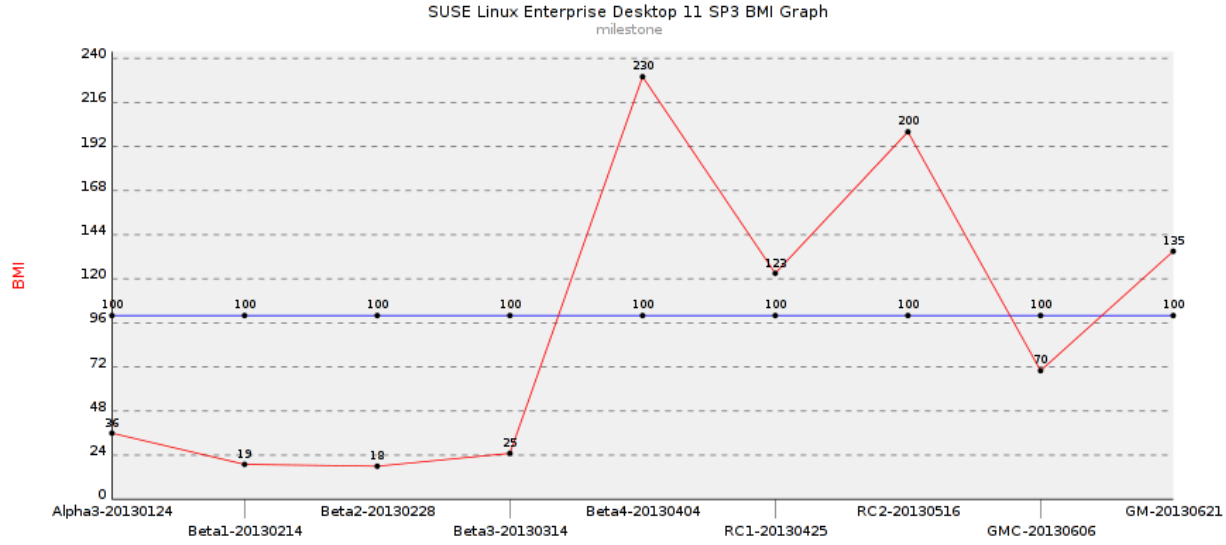
We have implemented two sets of BMI calculating tools using Perl and Python. Their source code can be found at:

https://github.com/mitiao/bmi
https://github.com/zxdvd/scripts

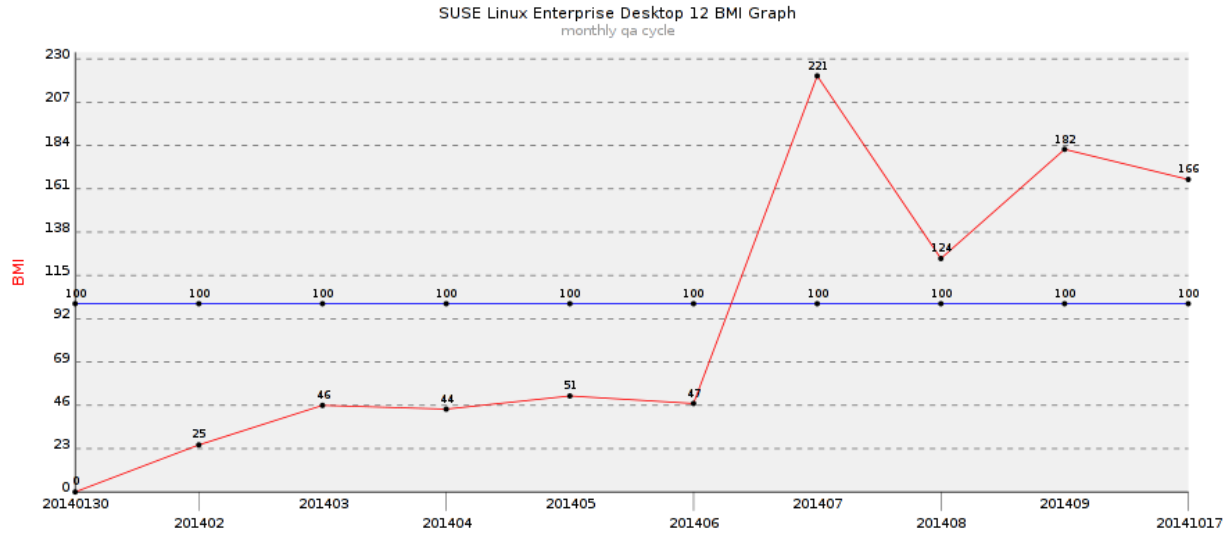


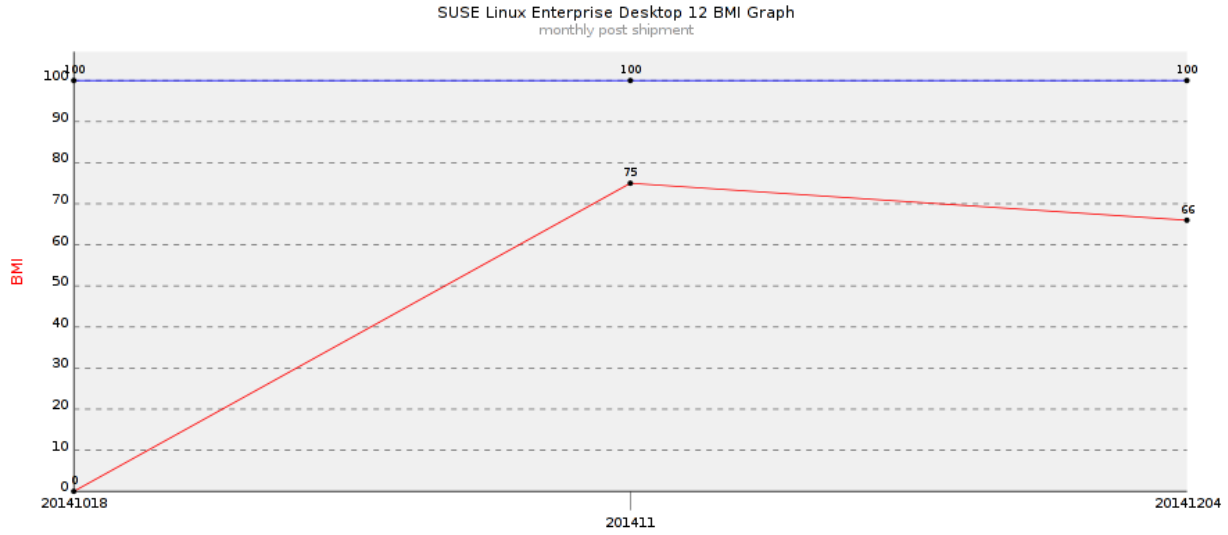
BACKLOG MANAGEMENT INDEX



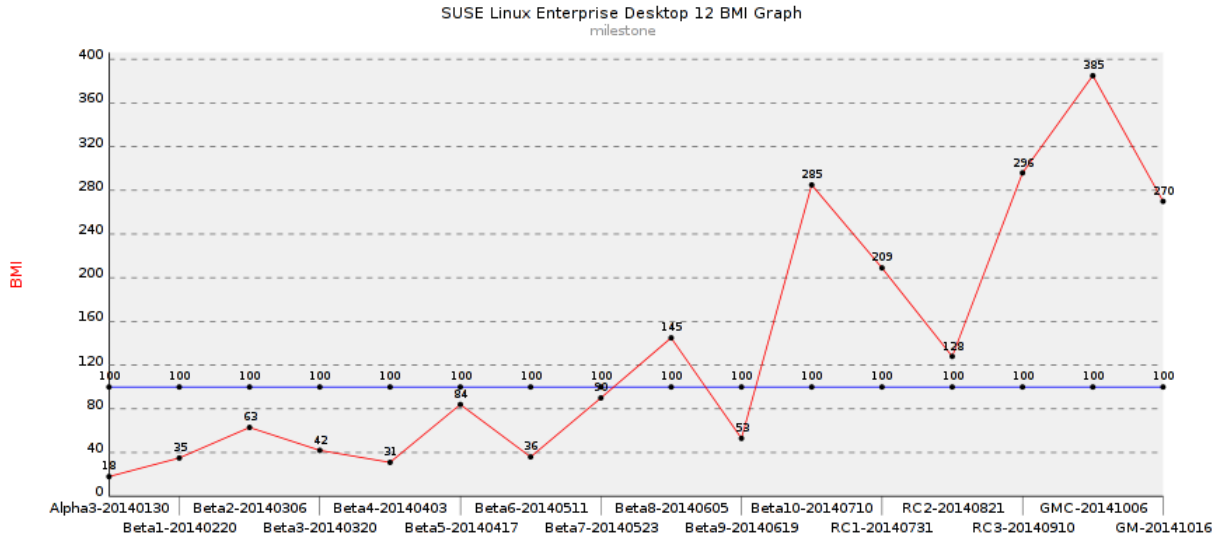


BACKLOG MANAGEMENT INDEX





BACKLOG MANAGEMENT INDEX



Above are the charts indicating the BMI of different stages of SLED11SP3 and SLED12.

All BMI charts shows the BMIs are going to be larger than 100 during the periods, indicating that the capability of the fix process was functioning normally. The bug charts explains the rise and fall of the bug arrivals and closures. Arrived bugs are more than closed bugs in early stage, and closed bugs are more than arrived bugs in later stage, which results the trend of BMI are going to be larger. For SLED12 BMI & Bugs Charts (Post-Ship Monthly), the BMI doesn't meet the goal (under 100) because the given periods was not enough.

PRODUCT LEVEL BUGS DATA COMPARISON STATISTICS

Product level bugs data comparison statistics

The product level comparison (SLED 12 VS SLED 11), could have 2 perspectives (formula as above):

1. The statistics of each components' valid bug report rate

2. The statistics of each components' component bug report rate

First we get the 2 group data for bugs reported by all colleagues.

Then we change the data range, which means we just pick bugs that reported by our teammates, and calculate the 2 group data use the same method as above.

The purpose of us making our team's bug reporting status statistics is that, through the statistic we could see our test points and where we differ with other QA colleagues and we also could see our workload and efficiency.

Notes and formula:

Valid bugs: bugs (exclude whose status is *ENHANCEMENT-DPLICATED-INVALID-WONTFIX-WORKSFORME*)

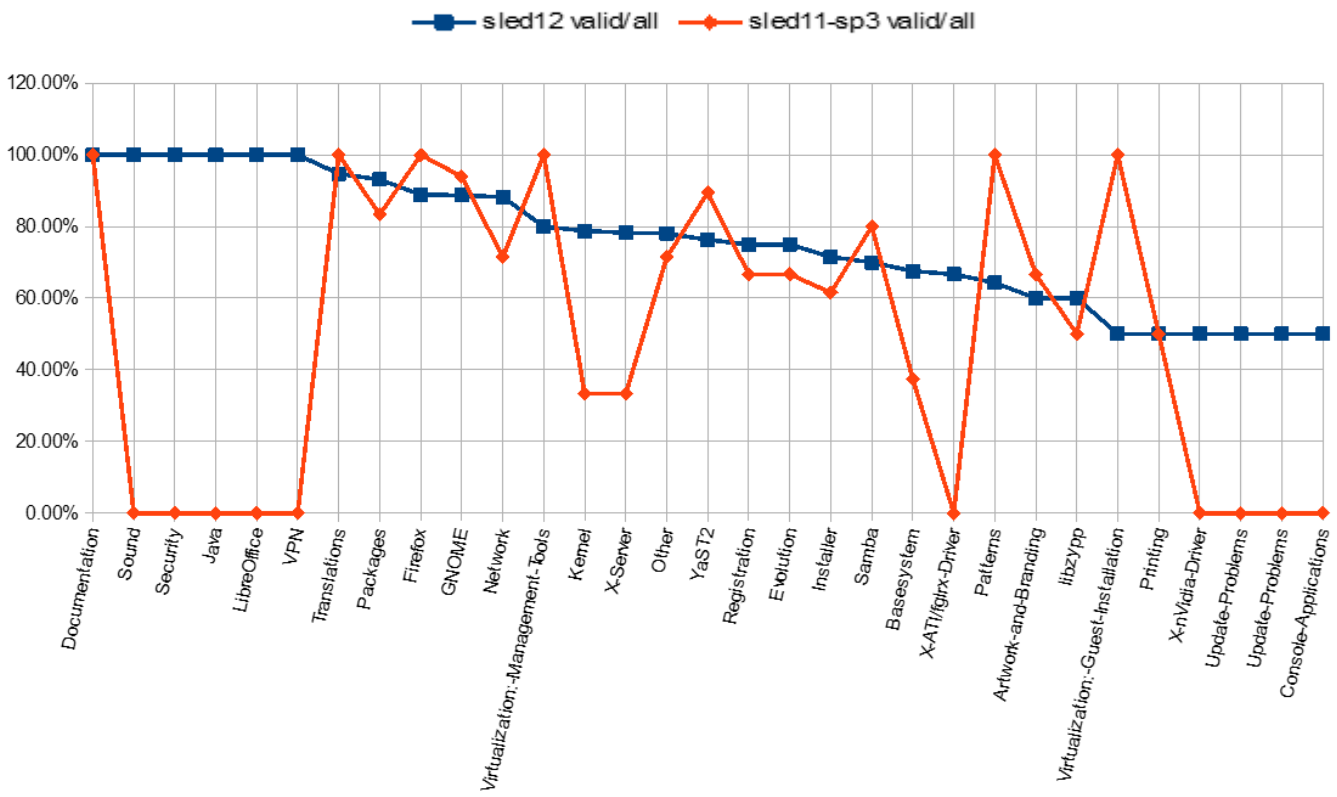
Valid bug report rate: component valid bugs number /component all bugs number

Component bug report rate: component valid bugs number /sum of all component valid bugs number

PRODUCT LEVEL BUGS DATA COMPARISON STATISTICS

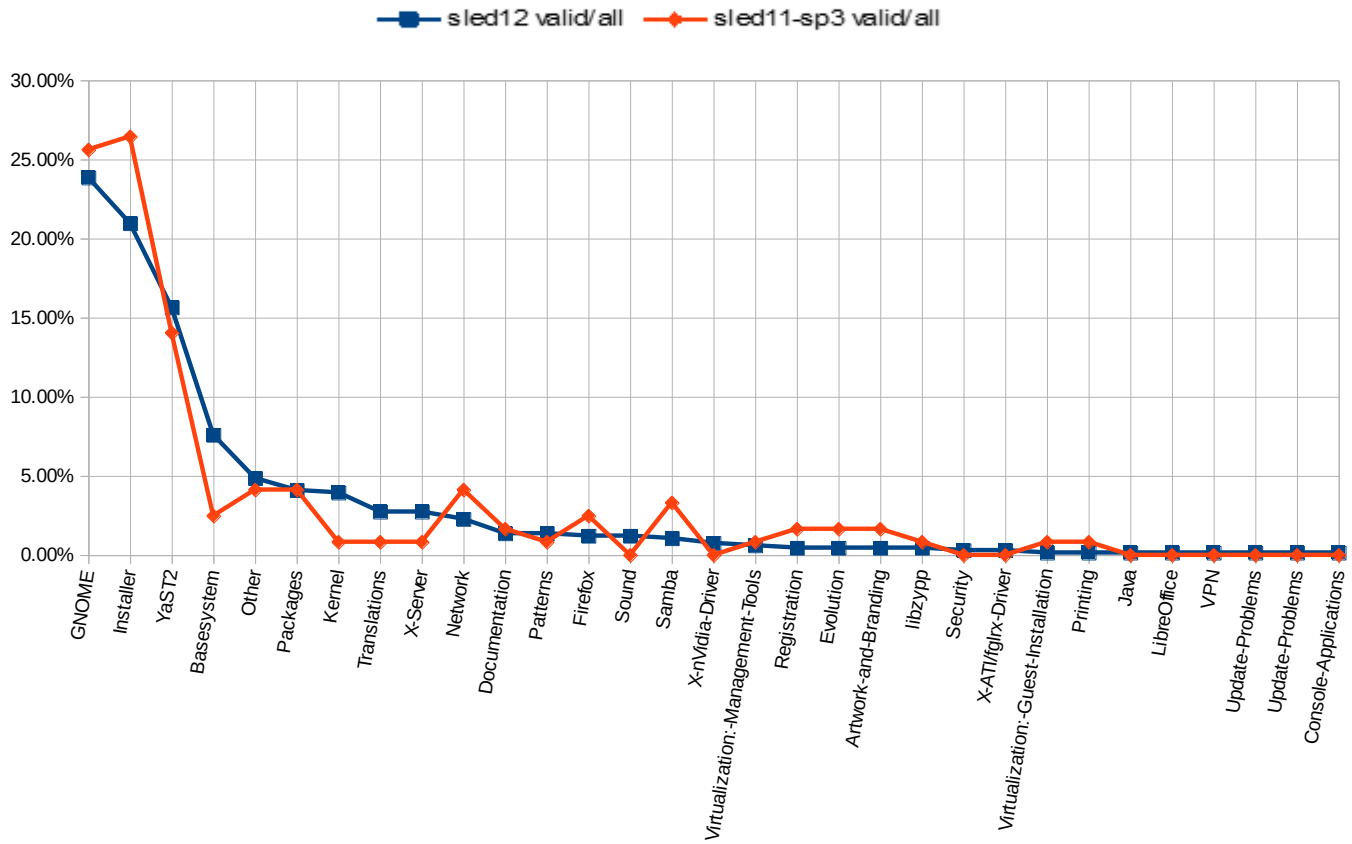
PRODUCT LEVEL BUGS DATA COMPARISON STATISTICS FOR P1 BUGS

1. SLED 12 VS SLED 11 SP3 valid bug report rate:



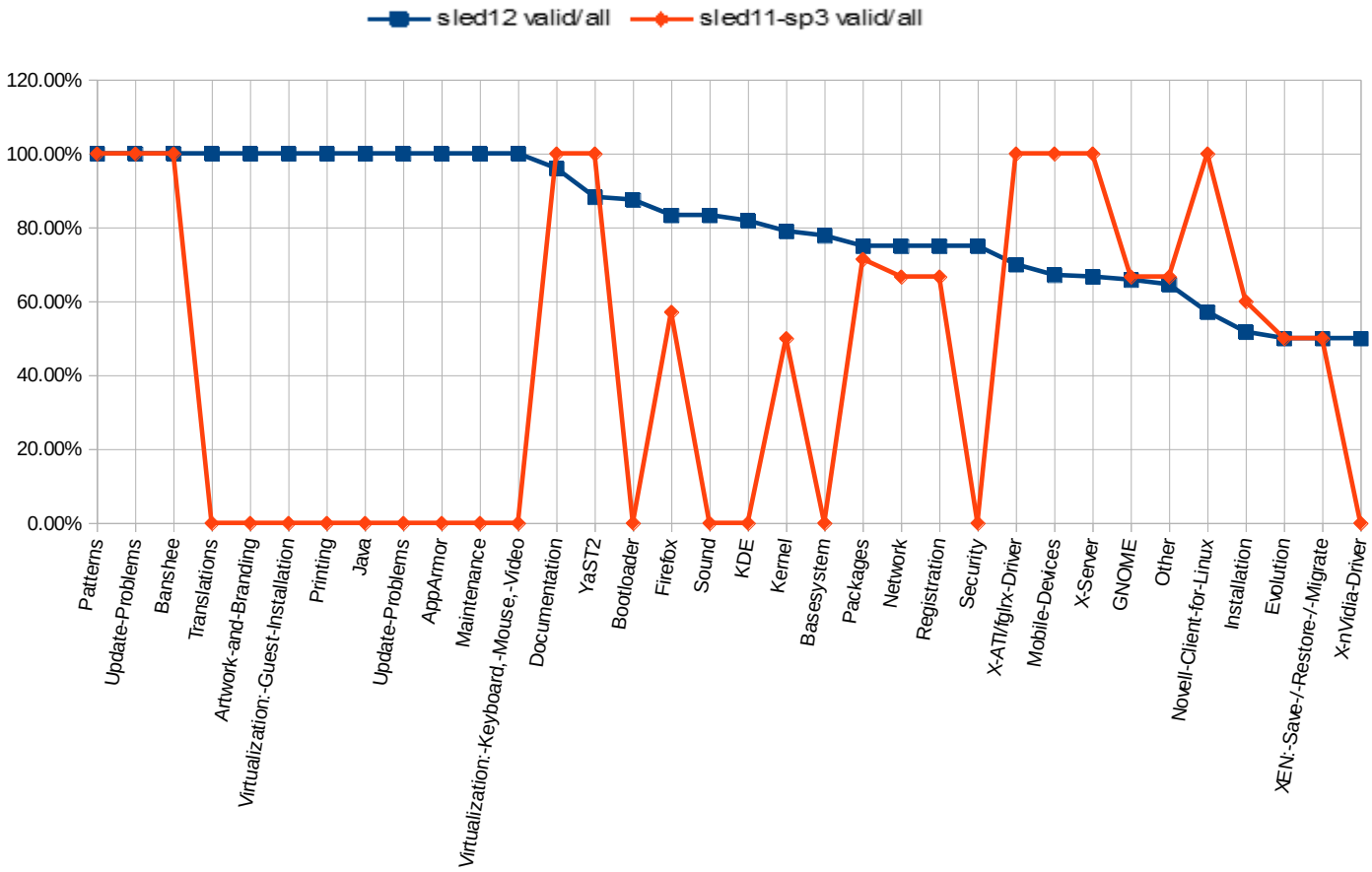
PRODUCT LEVEL BUGS DATA COMPARISON STATISTICS

2. SLED 12 VS SLED11 SP3 component bug report rate:



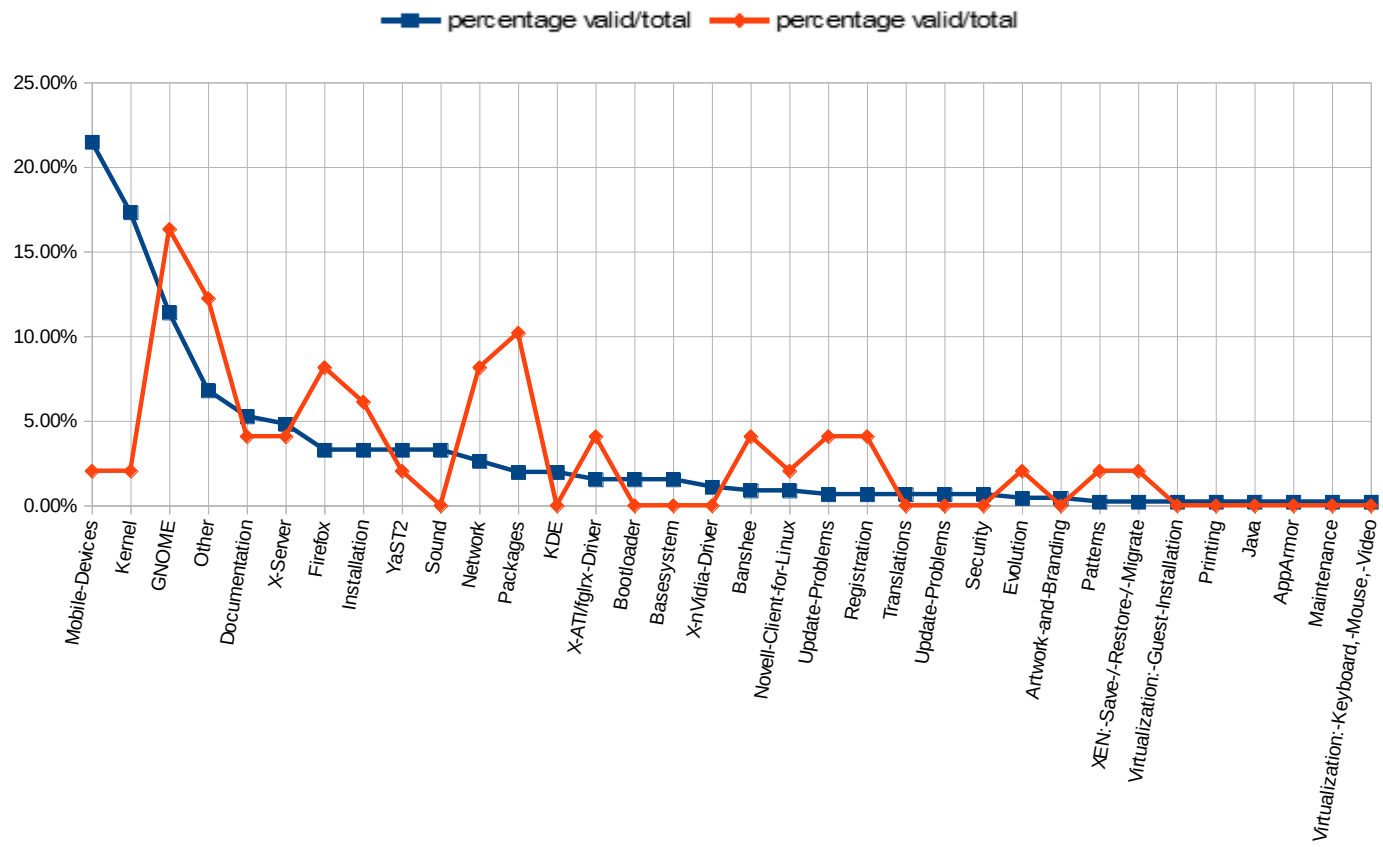
PRODUCT LEVEL BUGS DATA COMPARISON STATISTICS

3. SLED12 VS SLED11 SP3 valid bug report rate (APAC team I)



PRODUCT LEVEL BUGS DATA COMPARISON STATISTICS

4. SLED12 VS SLED11 SP3 component bug report rate (APAC team I)

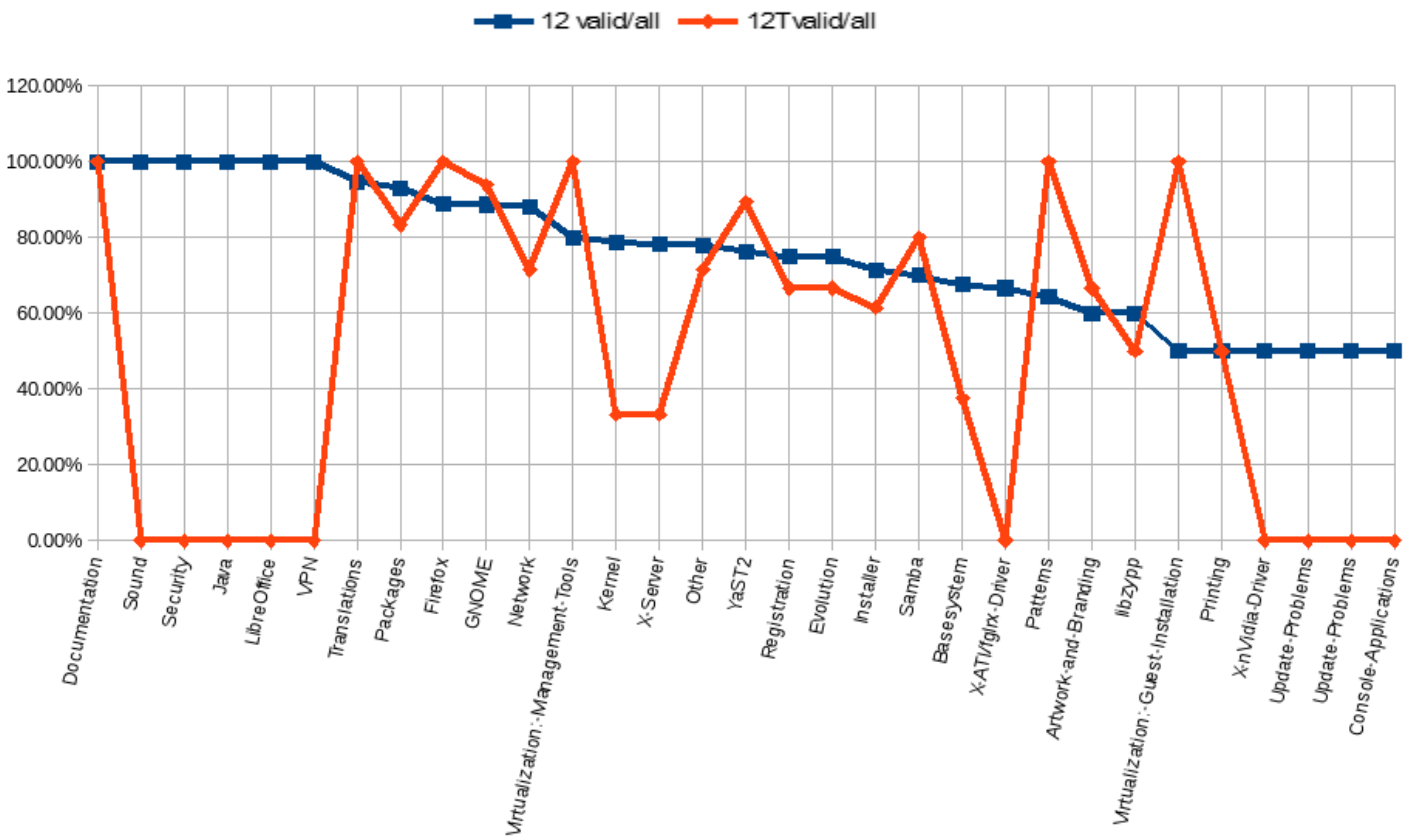


The above charts show the comparison of the valid / invalid bugs filed by our team / bugs filed respectively.

PRODUCT LEVEL BUGS DATA COMPARISON STATISTICS

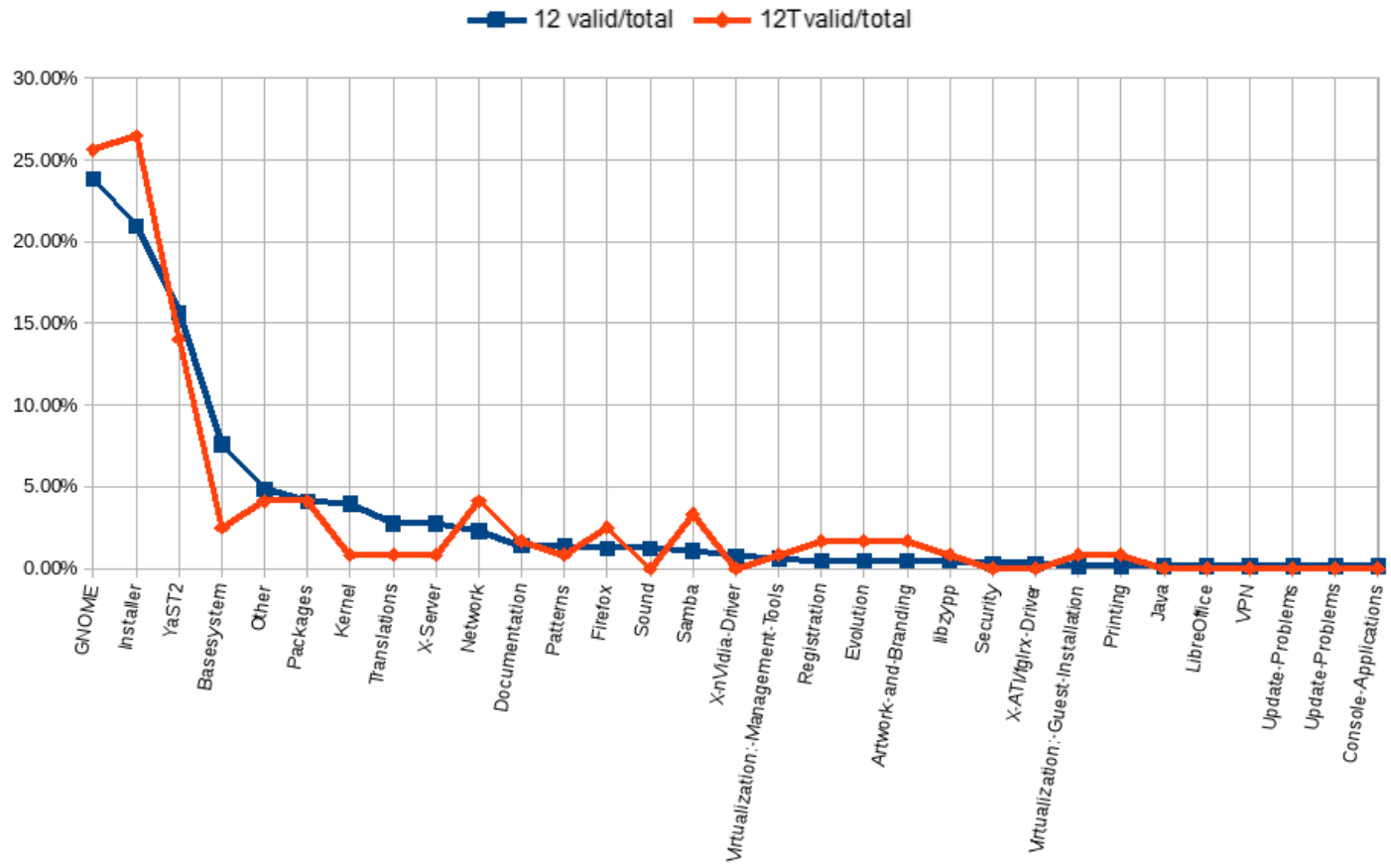
PRODUCT LEVEL BUGS DATA COMPARISON STATISTICS FOR P2 BUGS

1. SLED 12 valid bug report rate comparison



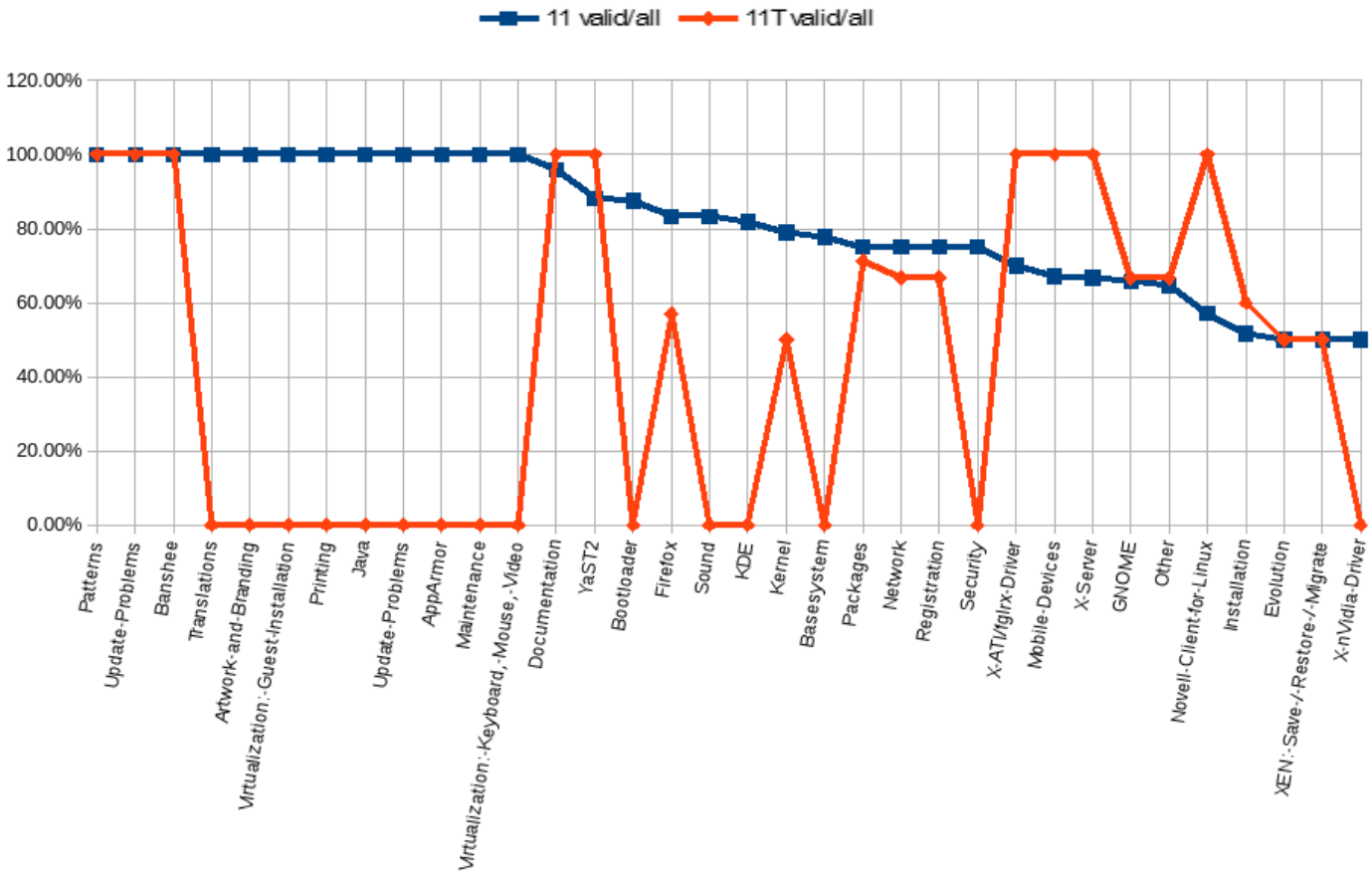
PRODUCT LEVEL BUGS DATA COMPARISON STATISTICS

2. SLED 12 component bug report rate comparison



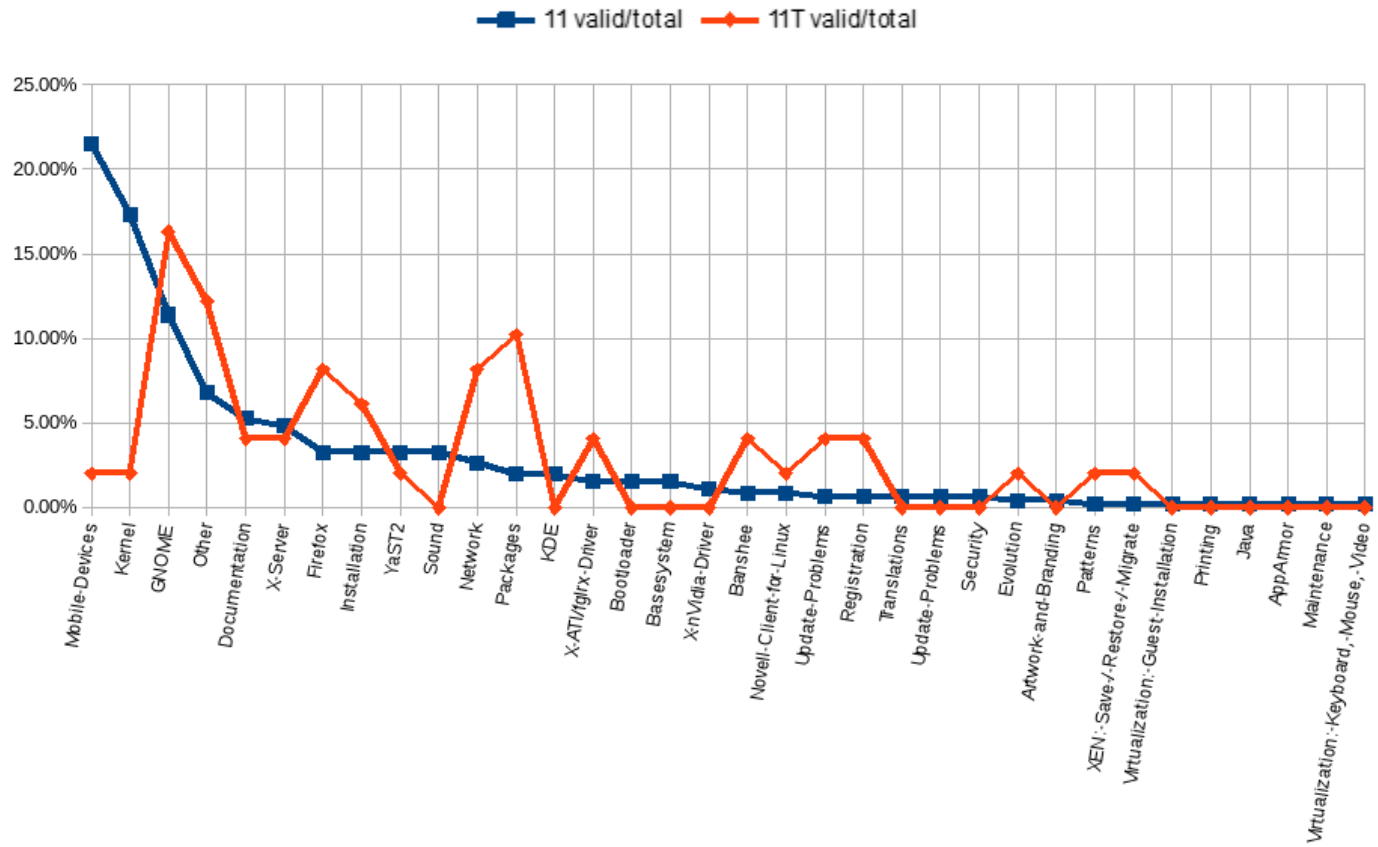
PRODUCT LEVEL BUGS DATA COMPARISON STATISTICS

3. SLED11 SP3 valid bug report rate comparison



PRODUCT LEVEL BUGS DATA COMPARISON STATISTICS

4. SLED11 SP3 component bug report rate comparison



Quality metrics questionnaire

We have also developed a questionnaire to subjectively assess the quality of our products.

The questionnaire is based on the research of Stephen H. Kan and Gary Wilson Jr. .

We have divided the factors that will affect the quality of our products into 9 groups and assigned each group with a weight.

The questionnaire is grouped by scenarios of various kinds. Please be noticed that the questionnaire is not intended to be completed by one person. Everybody should pick the most familiar parts to answer.

INSTRUCTIONS OF USING THE QUESTIONNAIRE

- Quality Factor: Separate/Classify each questions into one of 9 Factors.
    - Correctness x 45
    - Reliability x 30
    - Usability x 22
    - Maintainability x 19
    - Testability x 15
    - Efficiency x 14
    - Flexibility x 14
    - Reusability x 12
- Score system: First answer will get the highest score.
    - Totally Agree x 5
    - Partially Agree x 4
    - Medium x 3
    - Partially Disagree x 2
    - Totally Disagree x 1

Questions are categorized into quality factors which are referred to assign the weigh to the questions.

The answers scoring distributed evenly among the questions. In any of the questions, the first answer is

scored as 5 and the last one is scored as 1. With n questions altogether, the final result of the questionnaire will be:

$$\sum_{1}^{n} Weigh \times Score$$

REFERENCES

1. Stephen H. Kan, *"Metrics and Models in Software Quality Engineering (2nd Edition)"*, 2002.
2. Gary Wilson Jr. *"Software Quality Survey Results"*, http://thegarywilson.com/blog/ 2013
3. *Software Quality from Wikipedia* , http://en.wikipedia.org/wiki/Software_quality

QUALITY METRICS QUESTIONNAIRE

QUALITY SLE DESKTOP QUESTIONNAIRE

Please give your general opinion through the usage of system.

1. Compare with different "Operation System"(including Linux, Windows, and Mac OS). Do you agree it's a high quality "Operation System"? [Usability]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

2. Compare with different "Linux (or Unix-like) Desktop Distribution"(including SLED (or openSUSE), Fedora, Ubuntu, Debian, and Mint). Do you agree it's a high quality "Desktop Distribution"? [Usability]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

3. Compare with different Operating System regarding to particular hardware environment. Booting Time consumption is satisfying for your general/regular usage. Do you agree? [Efficiency]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

4. Do you agree "Coverage of Function/Feature" is sufficient for your general/regular usage? [Usability]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

5. Do you agree "User Experience (UX)" is intuitive for your general/regular usage? [Usability]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

6. How many crashes/bugs have you encountered during the whole period of using the system? [Reliability]

[ ] 0 [ ] 1 [ ] 2-3 [ ] 4-5 [ ] more than 5

Scenario A: SLED 12 was released, I guess you cannot wait to have a try with it already. Let's install it and actually feel of what are the new features and functions there.

1. Do you agree the installation progress is smooth enough? [Reliability]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

2. In an average hardware setup, the installer timing consumption is acceptable. Do you agree? [Efficiency]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

3. It's very easy to achieve your partitioning/file system plan by using the layout. Do you agree? [Usability]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

Scenario B: Once installation is done, please login to Gnome Desktop environment, overview the whole layout and try with some special features. Such as, moving mouse to top left corner or selecting "applications"--> "Activities Overview", overall opening applications are shown on current desktop, and launch the applications on one or multiple desktops from "Favorite bar" or "Quick search bar" or "Applications" menu.

1. Do you agree there is no difficulty of finding application/system configure? [Usability]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

2. Do you agree all applications can be launched correctly without any problems? [Correctness]

[ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

Scenario C: Surfing Time! Launch Firefox to look at what are the headlines today, and launch Rhythmbox/Video to play some audio/video files separately. Meanwhile, checking the reliability and stability of wired network and wireless network with all kinds of encryption methods as possible you can. (e.g. none/wep/wpa/wpa2, hidden/unhidden, etc.).

1. Do you agree all web pages are rendered correctly? [Correctness]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

2. Do you agree all supported media files are played well and the quality is acceptable? [Usability]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

3. Connection stability with both wireless and wired network in office environment is pretty good. Do you agree? [Reliability]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

4. Do you agree all needed wireless encryptions work normally? [Usability]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

Scenario D: Please say "hello" to your colleagues, friends, or customers via Empathy or Pidgin (need install it firstly) by using your existed accounts. (e.g. AIM/MSN/GoogleTalk/GroupWise/IRC...etc.). And check/receive/send mails via Evolution.

1. Do you agree the login for every IM protocols you used are working smoothly? [Usability]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

2. Do the functions of Empathy/Pidgin/Evolution fit all your working needs and with high productivity? [Usability]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

Scenario E: Do some paper work by using "LibreOffice", open/create some files (e.g. "LibreOffice", "Base", "Calc", "Draw", "Impress", "Writer", "Document Reviewer"), edit and save them. Such as create a slides with 10 pages, using "Calc do the math statistics, edit and print some "Writer" pages, etc.

1. Is LibreOffice sufficiently compatible with opening/saving all kinds of your daily documents? [Usability]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

2. Do you agree using Libreoffice is reasonably productive to daily work? [Usability]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

Scenario F: What should I do when I need to install some specific software or package? That's very easy, the built-in package management tools called "Yast" and "Zypper". Such as, you can search your target software/package in "Yast" or by using "zypper", like "zypper se xxx", and install it "zypper in xxx". Of course, system/software/package upgrade can be easily done as well.

1. How easy is it to conduct an upgrade of system or application? [Usability]

   [ ] Very easy [ ] Easy [ ] Medium [ ] Hard [ ] Very Hard

2. How often do you meet an unexpected error (e.g. package conflict/dependency) during installation and upgrading? [Reliability]

[ ] Never [ ] Occasionally [ ] Sometimes [ ] Often [ ] Always

3. Do you agree SCC online registration are easily and quickly done?

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

Scenario G: Finally, do some system level checking, such as, checking booting messages, hardware detecting, driver/module loading, and enable power saving model (Dim screen/Automatic suspend).

1. Do you agree all drivers/modules are loaded correctly? [Correctness]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

2. Do you agree the behaviors of common external devices are supported? [Correctness]

   [ ] Totally Agree [ ] Partially Agree [ ] Medium [ ] Partially Disagree [ ] Totally Disagree

3. Do you meet any problems when power saving is triggered in accordance with the specific time interval settings? [Correctness]

   [ ] Never [ ] Occasionally [ ] Sometimes [ ] Often [ ] Always