

Google Summer Of Code 2019 Proposal

For

Project:

Analyze and improve JSON performance in
Syslog-ng



Parth Wazurkar

GSoC 2019 Proposal: Analyze and improve JSON performance (parthw1)

Abstract

Syslog-ng is a famous log management software with content-based filtering, rich filtering capabilities, flexible configuration options and many other features. Syslog-ng provides many possibilities to parse the messages, or parts of the messages. One of these is JSON parsing/formatting. JSON is widely used amongst syslog-ng users. Huge amount of data is processed with syslog-ng using JSON everyday, so increasing the performance of JSON parsing in syslog-ng will be important and beneficial.

Currently syslog-ng uses an external library `json-c` for JSON parsing. This project aims to improve overall performance of JSON parsing in syslog-ng by analysing the currently used library `json-c` and compare it with other available json libraries. The library providing the best performance will be chosen and integrated into syslog-ng.

Project Goals

- To begin with, proper parameters will be defined for analysing the current JSON performance in syslog-ng.
- Current implementation will be properly analyzed based on these parameters, and possible improvements and shortcomings will be found out.
- Comparing other available json libraries based on the parameters and finding the best library out of them so that the shortcomings are addressed.
- Integrating the new library with syslog-ng, thereby improving the performance of JSON parsing in syslog-ng.

Deliverables

- A detailed report of analysis of current JSON implementation in syslog-ng, which will make it crystal clear about choosing a new library.
- A detailed analysis report of comparison of candidate json libraries, making it clear to choose the best one.
- Implementation and Integration of new library into syslog-ng.
- A new and improved JSON parser for Syslog-ng (Final work product).

Benefits to the community:

Syslog-ng is a powerful tool, that gives solution for wide range of logging problems, but technology is evolving faster and faster, and new problems comes alive every day. syslog-ng has to keep pace with these claims, and the improvement in the performance of the JSON parsing which is used by a lot of syslog-ng users around the world, will be highly benefited with this performance improvement.

Knowledge Area

- Syslog-ng codebase and how syslog-ng uses JSON - **Beginner (Improving)**
 - I have been tweaking with the codebase for quite some time now, and have also contributed by solving some issues. I have a basic understanding of the working of the codebase and will improve with time.
- C Programming - **Proficient**
 - I have been doing programming in C for many years now and my GSoC 2018 project with FreeType was also completely based in C, so I am comfortable with C.
- Git and Github - **Proficient**
- Profiling and how to use performance analysis tools - **Beginner**
 - I am currently learning about performance analysis tools and will improve with time.
- How to manage a library replacement in an existing project - **Proficient**
 - I have previously worked on developing and integrating new modules into an existing project during my GSoC 2018 project. So I am confident about replacing library into an existing project

Why the project is interesting to me?

I have always been passionate about optimizing and improving performance of my code and projects. Also, when I came across syslog-ng and read more about it, it was really fascinating to know how logging is essential in software development. My previous GSoC project in 2018 with FreeType was to create new modules and integrate them into FreeType, so doing a analysis of modules and integrating them into the codebase has been very interesting to me. And this was the reason of choosing this project.

Project Timeline

Project Division: I have divided the project into 3 phases and 2 milestones.

- **Phase 1:** Analysis phase.
- **Phase 2:** Comparison phase.

Milestone 1: Best JSON library is chosen.

- **Phase 3:** Implementation and Integration phase.

Milestone 2: Chosen library is integrated into syslog-ng.

Project Plan: This is the project plan, in which I have tried to schedule work as per the timeline of GSoC.

Days	Work to be done
May 6	Students Projects Announced
May 6 - May 27	Community Bonding Period
May 6 - May 19	<ul style="list-style-type: none">• Get acquainted with the workflow of Syslog-ng• Set-up the required environment for the project• Discussion on weekly status update method.<ul style="list-style-type: none">◦ Blog/Wiki?◦ Email?• Setup of wiki/blog/mailing list as per above.• Establish clear timings for meetings, code reviews and discussions.• Increase familiarity with mentor and community.• Fill in the gaps in the understanding of Syslog-ng codebase.• Go through the concepts and improve knowledge of the current JSON implementation and get doubts cleared from mentor.
May 20 - May 27	<ul style="list-style-type: none">• Create extensible list of candidate JSON libraries and ruling out those which have been previously reported buggy, or have no users, checking for potential reviews, other.• Finally after this, coming out with properly examined candidate libraries for analysis.• Here, the currently used library 'json-c' will also be a candidate library.
May 27 - June 28	First Coding Period

	Phase 1: Analysis Phase.
May 28 - June 14	<ul style="list-style-type: none">Defining parameters and test cases to analyze the current implementation of JSON parser.Analyzing the current implementation based on these defined parameters, test cases.Creating a detailed report.
June 15 - June 22	<ul style="list-style-type: none">Checking if the shortcomings can be improved by some way in the current implementation itself.Based on the outcome of the analysis define parameters for comparative analysis of candidate libraries.
June 22 - June 26	Buffer Period
June 26 - June 28	<ul style="list-style-type: none">Midterm evaluation preparations.
June 24 - June 28	Midterm Evaluations
June 28 - July 22	Second Coding Period
	Phase 2: Comparison Phase.
June 28 - July 12	<ul style="list-style-type: none">Based on the parameters, a comparative analysis is done.A comparative Analysis report is created and the best library is chosen.
<i>Milestone 1 Achieved!! (Best library is chosen.)</i>	
	Phase 3: Implementation and Integration Phase.
July 13 - July 24	<ul style="list-style-type: none">Starting integration of the chosen library into the syslog-ng codebase.
July 24 - July 26	<ul style="list-style-type: none">Second midterm evaluation preparation
July 22 - July 26	Second Midterm Evaluations
July 26 - August 12	<ul style="list-style-type: none">Integration into codebase continues.Setting up test cases and documentation.Changes according to the review comments.

	<ul style="list-style-type: none"> Analyzing performance. Fine tuning the performance. Immediate bug fixes (if any).
July 12 - Aug 19	<ul style="list-style-type: none"> Buffer Period
Milestone 2 Achieved!! <i>(Best performing library is integrated into Syslog-ng)</i>	
Aug 19 - Aug 26	Students Submit Code and Final Evaluations
Aug 26 - Sept 2	Mentors Submit Final Evaluations
Sept 3	Results Announced
Onwards	Keep contributing to Syslog-ng and will keep resolving the issues faced by the users while using any of the above features.

Note:

- Open to change the timeline according to mentor.
- GSoC timeline events have been marked in grey, and merged, for better representation. Start of phases are marked as light grey and achievement of milestones has been marked green.

Personal Information

Name	Parth Wazurkar
Country	India
Email	parthwazurkar@gmail.com
Contact Details	+919552283766/+918668509601
Time Zone	Nagpur, India UTC+5:30
Website	https://parthw1.github.io
Github	https://github.com/parthw1

About Me

I am Parth Wazurkar, pre-final year student from Indian Institute of Information Technology, Nagpur (M.H) India, pursuing computer science and engineering. I had participated in GSoC 2018 with The FreeType organisation and worked on a project “Integrating VFLib’s TeX format drivers into FreeType”. I love to participate in open source software development, I also regularly do competitive programming. I have been working with C, Linux Bash Shell, and Unix build tools from quite a long time.

I have 24/7 internet access and do not have any work commitments during the summer. I can easily work 30 hours per week or even more according to the requirements.

Contributions to Syslog-ng

I am an active participant in gitter conversations. I am also contributing to the syslog-ng's issues. Currently I have submitted one PR and wish to submit more in the near future.

Pull Requests:

[Syslog-ng: 2661](#)