| | |
|---|---|
| **Application from** | **Velliengiri, Praveen** |
| **E-mail Address** | praveenvelliengiri@gmail.com |
| **Job** | CERN openlab Summer Students 2018 (Member States and Non MS) / SUM-OPENLAB-2018 |
| **Document Type** | Application Form |
| **Application date** | 18/02/2018 20:34 |

## Personal Details

| | |
|---|---|
| **Title** | Mr. |
| **Family Name** | Velliengiri |
| **First Name(s)** | Praveen |
| **Maiden Name (if applicable)** | |
| **Gender** | Male / Homme |
| **Date of birth** | 15/05/1998 |
| **Nationality** | Indian (IN) |
| **Second Nationality (if applicable)** | |
| **Country of Birth** | INDIA |
| **Town of Birth** | Tiruppur |
| **Home Address (line 1 - max 32 chars)** | 5/223 pachakattupalayam, |
| **Home Address (line 2 - max 32 chars)** | Chengappalli |
| **City** | Tiruppur |
| **Country** | INDIA |
| **Postal Code** | 638812 |
| **Landline Phone Number (with international prefix)** | 91 04294 266773 |
| **Mobile Phone Number (with international prefix)** | +91 7538888190 |
| **What is your mother tongue?** | Other |
| **Please rate your level of English** | C1 |
| **Please rate your level of French** | I don't speak/understand French |
| **Please select any other languages you may speak** | Other |
| **Do you have a valid driver's licence?** | No-Non |

## Education

| | |
|---|---|
| **Country** | INDIA |
| **Level of Education** | OTHER - Master |
| **Title of Diploma/Qualification Note: Please give the full title in their original language (using Latin characters)** | FIVE YEAR INTEGRATED MSc DATA SCIENCE |
| **Attended From** | 06/2015 |
| **Attended To (planned end date for current studies)** | 05/2020 |
| **School/University Name** | PSG COLLEGE OF TECHNOLOGY, COIMBATORE |

| | |
|---|---|
| **Country** | INDIA |
| **Level of Education** | OTHER - Secondary education |
| **Title of Diploma/Qualification Note: Please give the full title in their original language (using Latin characters)** | SECONDARY EDUCATION SPECIALIZATION IN SCIENCE & MATHEMATICS |

| | |
|---|---|
| **Attended From** | 06/2013 |
| **Attended To (planned end date for current studies)** | 03/2015 |
| **School/University Name** | ERODE HINDU KALVI NILAYAM , ERODE |

| | |
|---|---|
| **Country** | INDIA |
| **Level of Education** | OTHER - Secondary education |
| **Title of Diploma/Qualification**<br>**Note: Please give the full title in their original language (using Latin characters)** | 10 th Grade ( Secondary School Leaving Certificate ) |
| **Attended From** | 06/2001 |
| **Attended To (planned end date for current studies)** | 03/2013 |
| **School/University Name** | KONGU MATRICULATION HIGHER SECONDARY SCHOOL, UTHUKULI |

## Specific Information (Summer Students)

| | |
|---|---|
| **Main Field of Study** | Information Technologies / Informatique |
| **Secondary Field of Study** | Mathematics / Mathématiques |
| **Tertiary Field of Study** | |
| **What is your motivation for applying for this position?** | My interest in CERN started in 2013 right after I read an article about Higgs boson discovery in "The Hindu" National Newspaper. Believe me, At first, I thought it was an atomic bomb explosion. I remembered it was a mind-blowing article with the pictures of particle collision. Even Though I didn't understand the physics behind the discovery, it was really fun to discuss the picture of collision with my friends and to paste the article into our class information board. At the time I was a 14-year-old boy with a huge interest in physics and research experiments which unravel the mysteries of our universe. But I ended with Computer science degree in university, so I didn't even know how I could serve to physic discoveries or some exploration experiments, but when came to know about Computing field at CERN. I hope this is the place I want to be in to serve the humanity. I'm very much excited to work with the truly unique organization in the world to share my knowledge and ideas in the field of Computer Science with great experts. Willing to do challenging work and to learn new things every day. I'm very proud to be a part of Organization which leads the whole humanity one step ahead to understand who we are? Where do we come from? Why things behave as it is? I want to contribute to the welfare of humanity and be helpful to solve the mysteries of the universe with the help of computer science. This internship will be a great and valuable opportunity for me to learn and enjoy the scientific community and to help the complex experiments in the world. It will be the great opportunity for me to gain international scientific experience, and to make friends all over the world.<br>Being a part of CERN, I can ably serve the humanity.<br>Thanks |

| | |
|---|---|
| **Within your studies, which are your preferred topics for gaining work experience?** | I'm interested in solving problems which greatly requires strong computing infrastructure and programming abilities. I'm passionate about High-performance Computing and Compiler optimization, these areas help to increase the efficiency and speed of any modern complex scientific software. My favorite language is Modern C++. I also familiar with Java, C, PL/SQL. I'm also excited to work on Heterogeneous Computing Platforms to optimize code for different processors. I think contributing to ROOT or Runtime Systems would be a good match for me to improve programming constructs to use GPGPU resources effectively. I think adding executors to C++ parallel algorithms in ROOT would be a great advantage to use GPGPU, CPU's, FPGA's effectively. I hope executors concept may be extended to CERN cloud platforms like OpenStack to schedule and run tasks remotely.<br>I'm started to use Root recently and I really like its cling interpreter. I'm super excited to work with clang based-cling interpreter to improve its AST and Optimization.<br>I'm always interested in learning new technologies and languages because they help me to push my limits and to explore the vastness of Computer Science. I have gone through the "FUTURE ICT CHALLENGES IN SCIENTIFIC RESEARCH " white paper by CERN Openlab. I hope "Computing Performance and Software R&D and Database Technologies" topics will be the suitable place for me to contribute.<br>I have done a couple of toy projects using kubernetes. Exploring Cloud Infrastructures and Data-Centric Architecture will be helpful for my studies. I'm open and willing to learn new technologies as the part of this internship. Excited to collaborate with mentors to find new and effective ways to solve complex problems. I will learn from others and my mistakes, to improve myself. |
| **Describe any relevant work or social experience obtained during your studies, training periods or visits abroad** | I'm a Google Summer Of Code 2017 Student with Systems Technology, Emergent Parallelism, and Algorithm Research Group,(STE\|\|AR Group) Louisiana State University. STE\|\|AR Group actively develops High-Performance Parallelx (HPX) Runtime Systems and many libraries to tackle the future need of "Exascale Performance". I worked on a unique project "Distributed Component Placement" with my mentors Patricia Grubel and parsa amini. During GSoC 2017 I have learned various new exciting technologies like High-Performance Computing, Generic Programming Models. I have also learned how to contribute to the opensource and how to learn from feedback. I have successfully completed my first two evaluations but I can't able to complete my third evaluation because my university schedule is too bad. I can't able to find time for my GSoC Project. However, I learned from my mistakes. Now I'm good at finding time to work on my Pet Project along my university schedule. During the GSoC 2017, I came to know about many parallel and distributive computing libraries such STAPL and DASH. GSoC provides me a great opportunity to learn from exciting open source projects and people. I learned how to collaborate with people to improve myself and my skills.<br>Junior Research Student - As a Research Student in my College I learned about Operating Systems and Complex Memory Models. I also got familiarity with OpenCL. I have contributed to open source projects like chapel programming language and DASH to improve their documentation.<br>Participated in Shaastra 2017  Finalist - Jarvis Machine Learning Contest and attended hands-on workshops in computer vision, game development. I didn't have abroad internship experience, but I'm willing to have one to improve my skills and cultural knowledge.<br>Thanks |
| **Have you ever worked at CERN before?** | No |
| **If yes, for how long (in months)?** | |
| **How many years of full time study at university level will you have completed by the summer of your stay at CERN?** | 3 |
| **Applied physics** | |

| | |
|---|---|
| **Describe the projects where you used the selected applied physics topics and/or any others that are not listed** | |
| **Architecture** | |
| **Describe the projects where you used the selected architecture topics and/or any others that are not listed** | |
| **Surveying** | |
| **Describe the projects where you used the selected surveying topics and/or any others that are not listed** | |
| **Chemistry** | |
| **Describe the projects where you used the selected chemistry topics and/or any others that are not listed** | |
| **Civil engineering** | |
| **Describe the projects where you used the selected civil engineering topics and/or any others that are not listed** | |
| **Programming Languages** | C<br>C++<br>Java<br>SQL, PL/SQL or similar |
| **Describe the projects where you used the selected programming languages and/or any others that are not listed** | Tensor is a project for representing multi-dimensional array as STL Container in C++. It is different from C style arrays " Type ( )( )... ". Tensor mainly includes three classes<br>Region class - define the layout (row or column major) and size of the data structure.<br>Memory handles template class - It allocates a linear chunk of memory using an allocator and helps to index them as a multi-dimensional storage.<br>Tensor template class - It acts as a bridge between Region and Memory handle class, this class is implemented using PIMPL idiom. It includes two C++ smart pointers, one point to Region object other points to Memory handle object. Strong typing is used for elegant naming.<br>Computation in tensors can be offloaded to GPGPU's. It is stored as OpenCL Image Objects. Computation offloading to GPGPU's is based on the cost function "Which tells whether the offloading gives significant performance improvements" then data transfer and operations are enqueued in device command queue takes place.<br>Work Flow is a Java GUI application. Functionality is to ease the development steps of Java programmers. The application creates Dependency Graphs for the classes under certain conditions - (one class depends on another) when Inheritance or composition is used. It helps to understand the class hierarchy easily. It provides some meaningful warning messages when some Standard rules(capitalization of class names, breaking the large function into multiple smaller ones) are not followed in the program.<br>CURE is a Hospital and Patient service system, which manages patients medical records and treatment history. I have used Oracle database and wrote PL/SQL procedures for functionalities like register, update, and deletion of patient records. |
| **Databases** | Oracle<br>NoSQL |

| | |
|---|---|
| **Describe the projects where you used the selected databases and/or any others that are not listed** | Stock Market Analysis :<br>The functionality of this application is to allow the user to buy and sell the stocks from and to the Stock Exchange. I have used Document-Oriented Database (MongoDB) for this application and web frontend was done in JavaScript and HTML. The application has four main collections namely customer, portfolio, stocks, transaction.<br>Customer - It maintains the user personal information, bank details, and their portfolio. The relationship between customer and personal information, bank details are embedded and the relationship between customer and portfolio is referenced.<br>Portfolio - contains the information stocks, bonds, and cash equivalents.<br>Documents in customer collection contain references to Portfolio collection using MongoDB DBRefs. The relationship between customer and portfolio is one-to-one.<br>Stocks - It maintains a list of stocks in the market and their attributes like price, owners of the stocks, availability. To avoid scanning every stock in DB. I have created an index on the stock_id field.<br>Transaction - This collection stores the transactions occurred (buying/selling of stocks) between the customer and owners of the stock like a number of stocks sold, price per stock. Documents in portfolio collection contain references to collection transaction. A relationship between portfolio and transaction is one-to-many.<br>We(team project) have also done some Stock Market Data Analysis, to discover when and what shares one can buy to gain more profit.<br>CURE :<br>The functionality of this application is to maintain patient and their medical records in the hospital. I have used Oracle database for this application and wrote PL/SQL procedures and triggers for insertion, update, deletion queries. |
| **Information Technologies** | Building web applications (e.g. with jQuery, HTML5)<br>Developing distributed computing systems (e.g. clusters, batch systems)<br>Using software development tools (e.g. Git, Jira, Trac) |
| **Describe the projects where you used the selected information technologies and/or any others that are not listed** | I Designed a feature in HPX for distributing and allocating data objects over the cluster of compute nodes. I have used Chapel programming language domain maps - an abstract concept for distributing the Multi-dimensional index-based data structures over a cluster of computers. Domain maps represent the layout of the data structure like row-major, column-major, compressed row format when they are used in a single compute node. Domain maps represent the data distribution policies like block, block-cyclic, cyclic when they are created within the cluster of compute nodes. This duality of domain maps makes them a good abstraction for distributing and laying out the data structure in HPC frameworks. These are modeled as HPX_COMPONENT's - an abstraction for creating C++ class/struct remotely by specifying the compute node id(hpx::locality_id). Each component object has a unique gid(global identifier). Functions on the remote objects are carried over HPX_ACTION's which serialize the function send it over the network and invokes on correct object. Domain Maps receive distribution policy as a parameter and partition the domain using the policy. Domain class represent the size and shape of the data structure and locations(nodes) where each chunk should be allocated are retrieved from domain maps class. The domain class is the factory class which creates the multi-dimensional data structure (Arrays) in the associated compute nodes. I also designed an Embedded Domain Specific Language using "C++ Expression Templates" for specifying the distribution policies.<br>I'm using Git, Gitkraken to contribute to opensource projects. I have done some web applications using HTML, CSS, JavaScript for my course in college. |
| **Theory of electrical engineering** | |
| **Describe the projects where you used the selected theory of electrical engineering topics and/or any others that are not listed** | |
| **Networks and systems** | |
| **Describe the projects where you used the selected networks and systems and/or any others that are not listed** | |

| | |
|---|---|
| **Low and high frequency engineering** | |
| **Describe the projects where you used the selected low and high frequency engineering topics and/or any others that are not listed** | |
| **Experimental Physics** | |
| **Describe the projects where you used the selected experimental physics topics and/or any others that are not listed** | |
| **Materials and experimental techniques** | |
| **Describe the projects where you used the selected materials and experimental techniques and/or any others that are not listed** | |
| **Mathematics** | Optimization |
| **Describe the projects where you used the selected mathematics knowledge and/or any others that are not listed** | I have done a simple telescope scheduling application. It is NP-Hard problem. The main functionality of the application is to effectively schedule the telescope time slots for various space observations. Each observation has a priority value based on the scientific importance and specific time interval for the observation of distant space object. If events do not overlap then the telescope is scheduled freely. If events overlap then telescope time is allocated for the observation which has more scientific importance and a good probability of collecting information about the object. Sometimes space is cloudy this affects the scheduling of telescope for observation which has more importance but less probability of being successful observation. I have used the ideas from the Research Paper "Optimization of telescope scheduling, Algorithmic research and scientific policy" by A.I Gomez de castro and J. Yanez. I have used C++11 and it's numerics library (for probability distribution)  to solve this problem. |
| **Mechanical engineering** | |
| **Describe the projects where you used the selected mechanical engineering topics and/or any others that are not listed** | |
| **Safety** | |
| **Describe the projects where you used the selected safety topics and/or any others that are not listed** | |
| **Choose a date** | from 18 June to 17 August 2018* |
| Y | |