A Cloud-aware distributed parallel compiler

CloudBook’s main entity is its advanced cloud compiler, that eases the deployment of resource hungry applications to run on large scale distributed deployments as for example a circle of cloud resources which communicates each other. CloudBook solution split original programs into a set of code pieces (deployable units) freeing the programmers from thinking of distributable designs and provides an easy access to cloud resources with a reduced required knowledge. This on one side speeds up the usage of cloud resources whilst keeping the development time for resource hungry applications low.

CloudBook benefits are especially useful to match the requirements of the industry 4.0 where smart applications must face the processing of big data and complex IA algorithms which requires the use of distributed and scalable applications. CloudBook makes easier the creation of these distributed designs, provides reliable parallel-distributed software that manage the available resources dynamically and require a reduced knowledge. Therefore, CloudBook speedup the design time, the development time and avoids the common re-design of software to make it more scalable.

Main focus

The main focus of the CloudBook project is to democratize supercomputing. CloudBook unifies the cloud and the compiler in one piece, freeing programmers from thinking of distributable designs and taking benefit from non-reliable hardware and parallel computation mechanisms.

Approach

1. The state-of-the-art in cloud computing, grid computing, remote monitoring, HPC (High Performance Computing), HTC (High Throughput computing), security, scripting, etc. has been studied in order to find innovative solutions that could help or be the base for this project. The CloudBook project fills the gap in the existing solutions, providing a full solution for a distributed parallel compiler, splitting the programs and distributing the pieces

2. Design, definition, and implementation of a graph analyser component that will analyse the input program (dependencies between functions) in order to assist the splitting process.

3. Taking the output of the graph analyser as input, a cloud-aware splitter has been created. The splitter splits the input program into different pieces of code (deployable units) to be executed by agents that will be installed in distributed machines.

4. A deployer has been created with algorithms which decide on the fly when, where, and how to distribute the deployable units in order to achieve an optimum execution performance. A monitoring tool of the execution of the deployable units allows to dramatically increase the performance based on real statistics of invocations among deployable units.

5. Different use cases have been developed to run on top of the CloudBook platform (serverless online video game, stochastic processes, videowall, parallel video-processor, intrusion detection system, high-
The main results of the project are as small companies company covering the Celtic governmental EUREKA network. Celtic applications & services focusing on a telecommunications, new media, future Internet, and projects in the area of telecommunications. Even companies outside the EUREKA countries may get some possibilities to join a Celtic-Plus project under certain conditions.

**Achieved results**

The main results of the project are open source developments and public documents:

- **CloudBook platform**: composed of graph analyzer, splitter, deployer, stats monitor, agents, and security component. This platform allows programmers to split their original programs into distributable pieces of code to take advantage of parallel execution. Additionally, it is the first tool that solves the two problems of parallel computing. This was not attended because the different agents communicate to each other without a "orchestrator". In other distributed solutions (such as BOINC) agents communicates with a central server and cannot communicate directly between each other.

- **CloudBook programming manual**: 25 pages with all the knowledge that any programmer needs to develop and take maximum benefit from CloudBook capabilities. One of the main advantages of CloudBook is the reduced required knowledge to design parallel programs

- **A set of different Proof of Concepts (PoCs) and examples:**
  - PoCs: N-Body, Hanoi, barber, classes programming: well-known classical problems that challenge any parallel code to measure computation vs usage of resources.
  - Video wall: screen composition using distributed synchronized software
  - Video processor: video processing in parallel
  - IDS: intrusion detection system
  - Serverless online video game: online games without server, server is running among players.
  - Stochastic processes: an example of parallel computation
  - High performance computing
  - Parallel data computing

**Impact**

CloudBook results have been materialized in a "cloud compiler", available from the Internet and testable by any scientific, industrial, social, or educational actor. It is a cloud-aware, open-source platform that brings the following advantages to developers:

- Allows executing programs in a distributable way without the need for a distributed programming design or high required knowledge. This execution is neither coordinated by any central node nor follows the master-slave paradigm, which is an unprecedented innovation.

- Accesses a higher pool of computing resources that could be required by programmers’ needs (supercomputers, commercial clouds, individual user resources) attached to CloudBook and forming a “circle” of cloud resources.

- Splits the code into deployable units taking into account the available computation resources and the programmer requirements in terms of performance.

- Dynamically manages the available resources, allowing the non-stop execution of a program while new computers are added to the circle and others are removed. Additionally, analyzes in real time the statistics of communication among machines to provide improved splitting and to dramatically increase the performance (more than 10x).

- Regarding use cases, the most relevant impacts are new possibilities for video games user experiences such as video-wall visualization, and mobile phones resource sharing. In data processing and artificial Intelligence, CloudBook allows speeding up programs. In HPC projects it allows combining high-performance computers with low-end machines, improving rational use of HPC resources.

The designs and developments carried out as part of CloudBook are based on open de facto standards.

The potential of this tool for creating distributable software without the need for experienced software designers capable to design distributable software, and the possibility to automatically re-make the software to scale and distribute from few to thousands machines, makes CloudBook a valuable tool for increasing throughput, not only for CloudBook partners but also for the whole software industry, introducing the new concept “code as a service”.

---

**About Celtic-Plus**

Celtic-Plus is an industry-driven European research initiative to define, perform and finance through public and private funding common research projects in the area of telecommunications, new media, future Internet, and applications & services focusing on a new “Smart Connected World” paradigm. Celtic-Plus is a EUREKA ICT cluster and belongs to the intergovernmental EUREKA network. Celtic-Plus is open to any type of company covering the Celtic-Plus research areas, large industry as well as small companies or universities and research organizations. Even companies outside the EUREKA countries may get some possibilities to join a Celtic-Plus project under certain conditions.

**Celtic Office**

c/o Eurescom, Wieblinger Weg 19/4
69123 Heidelberg, Germany
Phone: +49 6221 989 381
E-mail: office@celticnext.eu
www.celticnext.eu