Smart energy management for greener supercomputing

by Mohammed el Mehdi Diouri, Olivier Glück and Laurent Lefèvre
INRIA, LIP Laboratory, University of Lyon, France

A supercomputer is a system built from a collection of computers performing tasks in parallel, in order to achieve very high performance. An exascale machine is a supercomputer capable of performing more than 10^{18} floating point operations per second (1 Eflop/s). Such extreme-scale systems are needed by 2018 in order to meet with new scientific challenges, such enabling highly sophisticated genome calculations and proposing individualized patient treatments. As they will gather hundreds of millions cores, exascale supercomputers are expected to consume enormous amounts of energy (between 25 and 100 MW). Besides being very important, their power consumption will be very irregular. Furthermore, the applications that will run on such extreme-scale systems will need services such as fault tolerance, data collective operations that are energy consuming. In order to manage the execution of extreme-scale applications on future supercomputers in a sustainable and energy-efficient way, we propose a framework called SESAMES: Smart and Energy-aware Service-oriented Application Manager at Extreme-Scale.

Since the power consumption of these large scale systems is enormous and dynamic, SESAMES establishes a permanent negotiation with the energy provider (Figure 1). Through this dialog, SESAMES provides to the energy supplier an agenda of the estimated power consumption. Besides, it gathers from the energy supplier the agenda of the energy price, the agenda of the energy sources used and the agenda of the power capping. Indeed, depending on time, the energy may be provided at a certain price, from a given energy source (coal, wind, sun, etc.) and below a certain threshold limit. On the one hand, the users of supercomputers may prefer to consume energy at times when it is the cleanest and least expensive. On the other hand, the energy providers may adapt the supply according to the demand of supercomputers. In case the supercomputer is planned to experience a reduced energy consumption, the energy supplier may produce less energy. Thus, depending on the others customers, the energy provider may disable some production of energy produced from a polluting source. On contrary, energy provider can produce activate more energy production if needed by supercomputers.

Furthermore, in order to reduce the global energy consumption, SESAMES is able to act directly on the supercomputer nodes. An energy sensor is plugged to each node and measures the current power consumption. SESAMES collects these energy logs. In order to gather the execution context, SESAMES also establishes a dialog with the users of the supercomputers. This interaction with the user occurs at the moment of reserving computing nodes and just before running applications and services (Figure 1).

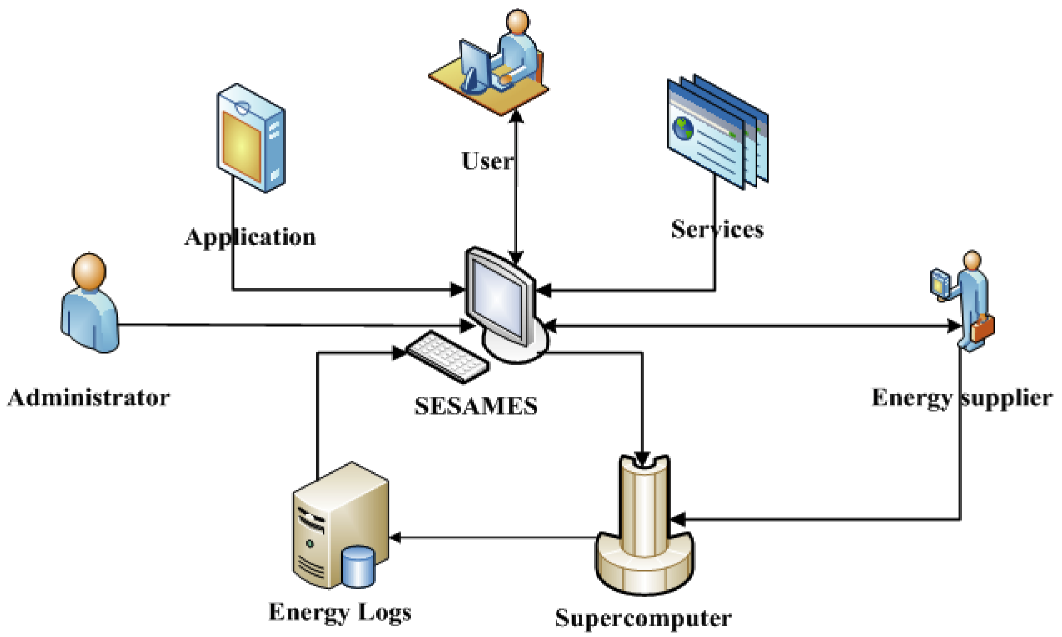

Figure 1: Global infrastructure: external interactions with SESAMES

In order to run their applications, users send to SESAMES a reservation request in order to book some of the supercomputer's nodes. At least, a reservation request consists of a number of nodes required, a reservation duration, an earliest possible start time and a latest possible start time. In order to make a reservation, SESAMES solves a multi-criteria optimization problem by taking into account several constraints. It attempts to allocate the supercomputing nodes at the moment desired by the user by consuming the least amount of energy, the cleanest energy, at the lowest financial cost and without exceeding the power capped by the energy provider. If no solution exists, SESAMES informs the user that the requested reservation is not possible. If there exists a unique solution that optimizes all the criteria, SESAMES makes the corresponding reservation and informs the user about it. Otherwise, SESAMES computes the solutions that optimize each criterion separately and asks the user to choose between the solution that minimizes the financial cost or the one that provides the cleanest energy.

Once the reservation is done, SESAMES gives the opportunity to the user to estimate and to reduce the energy consumption of the different services (like fault tolerance) that he would like to run while executing his applications. For each service, several versions are possible. The least energy consuming version depends on the applications features. Hence, the first step to consume "less" energy is to choose the least energy consuming version for each service. Thanks to the interaction with the user, SESAMES takes into account the application features and the user requirements in order to provide an energy estimation of the different versions of the services requested by the user.

To reduce the energy consumption of supercomputers, SESAMES proposes to apply some green leverages at the component level: shutting down or slowing down an idle resource component (processor, memory, disk, etc.). The shutdown approach consists in dynamically turning off unused resources and turning them back only when they are needed. The slowdown approach consists in dynamically adjusting the performance level of a resource according to the performance level the application and users really need. The green leverages proposed depend on the idle periods predicted and on the rights assigned by the supercomputer administrator to the user. The energy consumptions of these green solutions are estimated by SESAMES in order to make the user aware of the energy savings generated by the green solutions suggested.

Reference :

MDiouri, O. Gluck, and L. Lefevre. **"Towards a novel Smart and Energy-Aware Service-Oriented Manager for Extreme-Scale applications**, *First Workshop for Power Grid-Friendly Computing (PGFC'12), co-located with IEEE IGCC'12*, San Jose, USA, June 2012

Please contact:

Mohammed El Mehdi DIOURI.
LIP (UMR 5668 CNRS, ENS Lyon, INRIA, UCBL).

Tel : +33472728009
E-mail: **mehdi.diouri@ens-lyon.fr**
http://perso.ens-lyon.fr/mehdi.diouri