Geophysical Research Abstracts
Vol. 13, EGU2011-3548-1, 2011
EGU General Assembly 2011

Modelling Climate Change Effects on Species Distribution using a Cloud Computing Infrastructure

Valerio Angelini (1), Paolo Mazzetti (1), and Michele Pierri (2)
(1) CNR - National Research Council of Italy, (2) University of Florence at Prato - Piazza Ciardi 25, Prato - Italy

Climate change threatens to commit 15-37% of species to extinction by 2050 accelerating a mass extinction precipitated by widespread land use changes. The need to assess these impacts and recommend solutions to policy-makers is correspondingly acute and has been highlighted by the Fourth Assessment Report of the Intergovernmental Panel for Climate Change (IPCC, 2007) [1]. Such analyses require robust infrastructures capable of integrating huge volumes of data from biodiversity archives, satellite remote sensing, and climate change data. The main requirements for such infrastructures are: a) the support of multidisciplinary interoperability in order to harmonize services interfaces and data models; and b) high scalability to enable running of complex models with different climate change and species distribution scenarios. The first issue is currently addressed by several initiatives aiming to provide advanced infrastructures for geospatial and Earth Science resource sharing, such as projects funded under the 7th Framework Programme, standardization activities from international bodies like ISO and OGC, European and global initiatives like INSPIRE, GMES, GEOSS. The high scalability could be provided, in principle, by the Distributed Computing Infrastructures (DCIs) designed and developed in the last decade. The integration of geospatial infrastructures and DCIs has been investigated by several initiatives. For example, recently, the implementation of standard OGC services on top of different Grid middlewares has been analyzed in the context of several initiatives such as FP6/FP7 Projects (e.g. CYCLOPS, DEGREE, ENVIROGRIDS, etc.), standardization bodies activities (e.g. OGF-OGC MoU, OWS Phase 6) and international working groups (e.g. G-OWS).

We present an experimentation done integrating the architectural solution proposed by the Climate Change and Biodiversity Working Groups of the GEOSS Architecture Implementation Pilot Phase 2 and Phase 3 (AIP-2, AIP-3) [2], with a Cloud Computing Infrastructure. In this experimentation the system is based on open standards hiding all the complexity of the distributed computation to the user. The Ecological Niche Model (ENM) processing implemented using OpenModeller [3], an open-source niche modelling project, is exposed through an OGC WPS 1.0 standard interface. The process is able to retrieve the requested input datasets from external sources publishing standard interfaces, such as OGC WCS 1.1 for climate change environmental layers and GBIF access service for species occurrences. The WPS server we implemented makes use of a batch system to launch and control the execution of the required jobs on a remote pool of computing nodes. The computing nodes run on a mainstream Infrastructure-as-a-Service (IaaS) Cloud Computing platform (Amazon EC2/S3). An additional software component is able to re-size the working nodes pool by dynamically instantiating or destroying the dedicated computing instances on the Cloud infrastructure. The desired size of the pool is constantly balanced basing on the length of the batch system job queue. To be able to run multiple instances of the ENM process we developed an additional WPS that, accepting ranges of different parameters launches multiple instances of the model by computing all the required combinations of the parameter values. Finally, we developed a light AJAX GUI to launch and follow the execution workflow and then visualise the results.

Tests performed on the developed prototype demonstrated the possibility to run ENMs calculations and projections maintaining the execution time almost independent of spatial resolution, spatial coverage, number of scenarios, etc. exploiting the capabilities of on-demand allocation of computing power and storage space.

In conclusion, the solution proposed in our experimentation presents various advantages:

- Interoperability, through the adoption of open standards for resource sharing.
- Scalability, through the adoption of IaaS solutions, enabling the typical scientific use cases with possible peaks of intense computing resources utilisation.
- Cost-effectiveness, since for the entire prototype development and the experimentation we paid approximately 50€ for about 700 CPU hours.

References

[1] S. Nativi, P.Mazzetti, H. Saarenmaa, J. Kerr, É. Ó Tuama, "Biodiversity and climate change use scenarios framework for the GEOSS interoperability pilot process", Ecological Informatics 4 (2009) 23-33.

[2] GEOSS Architecture Implementation Pilot, available at http://www.ogcnetwork.net/AIpilot

[3] OpenModeller Home Page, available at http://openmodeller.sourceforge.net/