Develop a cloud computing adoption advisory tool

Products and Services Committee
Zhipeng Gui, Jizhe Xia, Nanyin Zhou, Chaowei Yang
Center of Intelligent Spatial Computing for Water/Energy Science, George Mason University

Tags/Keywords: advisory tool, cloud computing, evaluation, recommendation, testbed
Stakeholders: ESIP Federation Testbed
Themes: ESIP testbed cloud computing adoption advisory tool
Technologies: Web-based advisory system

Overview

The rapid advancements of cloud computing technologies witness the creation of many cloud vendors and services with their own unique strengths and shortages. Similar to many other applications, geospatial applications can greatly benefit from the virtually unlimited resources of cloud computing. Many scientist and geospatial application providers are considering transforming their current computing infrastructure into clouds (IaaS and PaaS). However, it is a big challenge to select the most suitable cloud platforms and configuration solutions for the cloud novices and even for experienced cloud users due to the following reasons:

- Different capacity and advantages of each cloud product;
- Various requirements and constraints for computing resources of the application; and
- Trade-off between cost and performance.

These factors make cloud solution selection oriented complex and difficult task. Therefore, it is highly desirable to design and develop a web-based advising tool to assist domain experts from earth science to compare and select the most suitable cloud solution. Such a tool should integrate computing experiences and knowledge and is capable of recommend cloud solutions to achieve cost-efficiency and high performance. The ESIP P&S committee announced the project to develop a cloud adoption advisory tool. The objectives of the project are to
1) Help cloud novices understand the basic concepts and potential applications of cloud computing providers, services and technologies;
2) Assist cloud computing early adopters easily and effectively select the best solutions based on their unique application requirements; and
3) Periodically collect/update the mainstream cloud platforms' information and build an expert system and database.

We proposed and were awarded to develop such a tool includes
1) A database/warehouse that stores pricing strategies and configuration schemes of mainstream Infrastructure as a Service (IaaS) and Platform as a Service (PaaS) level cloud platforms;
2) A general cost assessment model and a cloud function measurement criteria model for different cloud platforms and application types; and
3) A web-based Graphical User Interface (GUI) to be integrated with ESIP Testbed portal for help users select cloud solutions by evaluating their application requirements through interactive and visual manners.

Problem Statement

The diversity of cloud products provides many choices for cloud adopters. At the same time, it also introduces the complexities and challenges in cloud product selections for potential adopters. The complexities can be grossly summarized as the following three perspectives.

1) From the application's perspective, different application types (e.g., simple web /computing applications, concurrent intensive applications, communication intensive borderless applications, and computing/data intensive e-science or enterprise level applications) have different requirements on computing, storage and network resources (e.g., data volume size/data transfer velocity, communication/access frequency, computing intension).
2) From the cloud provider's perspective, different IT technologies (e.g., virtualization, storage and network) as well as different tenancy and pricing models (e.g., on demand/reserved/biding mode) are adopted in a mix fashion. Meanwhile, different commercial and open source cloud platforms (e.g., Amazon EC2, Microsoft Azure, Google App, GoGrid, Rackspace, CloudStack) have shown their own unique strengths and in computational capacities.
3) From the cloud adopter (application providers)'s perspective, the business expectations and cost constraints are various. The trade-off between the diverse computing provisions and the application requirements, combines the terminologies and criteria related to both the businesses and technologies together, and complicates the decision-making for a most suitable cloud solution.

Therefore, it is not an easy job to make a wise choice for the novices and even for experienced cloud users. Furthermore, the cloud providers usually post their own configuration schemes, tenancy and pricing models, and provide simple fee calculator on their own websites. The cloud adopters need go through these websites individually to gather information. It is not convenient to make intuitive comparisons and selections. Currently, there is no such a mature advisory tool available online to help users in earth science domain to select suitable IaaS and

PaaS level cloud platforms and concrete configuration solutions for their applications based on their requirements.

Solutions for Addressing the Problem

A well-designed and user-friendly advisory tool can help cloud adopters significantly reduce time and effort spent on information collection, and make wise selections through human-computer interaction. Many web-based systems have shown their power on multi-sources information and function mashup (integration) and presentation. Especially, with the rise of the visual analytics, it provides a new way to conduct information/data visualization and human-computer interaction to solve problems, by enabling a synergetic work of humans and computers. To fully assist decision-making of cloud adopters, the following functions are developed:

1. Cloud platform information collection
2. Fee cost estimation and capability evaluation models
3. Application requirement wizard
4. Solution visualization and comparison

Cloud platform information is collected by an information collector and stored in a database/warehouse. The information collector periodically crawls and parses the configuration schemes and the capability declarations of cloud resources (e.g., Virtual Machine, storage, and network) and relevant pricing rules of mainstream commercial IaaS and PaaS level cloud providers from their own website. It then stores or updates these information into a database. The out-of-date information will be moved into a warehouse as historical information for further analysis in the future. The information serve as the data foundation for conducting solution evaluation, recommendation and visualization.

Cost estimation model provides a method to estimate the potential cost of a cloud resource tenancy. Since uncertainty exists in the usage of applications (e.g., the user access number will result in different data transfer volume and storage size), the model provides a reasonable estimation and outputs a potential cost range to describe cost uncertainty.

Cloud capability evaluation model creates possible cloud solutions based on users' application requirement and cost constraints, then sorts and selects the best solutions by leveraging the computational capability and the potential cost. The outputs of evaluation model are the recommended solutions and the sorted all feasible solutions.

An application requirement wizard enables cloud adopters to describe and specify their application requirements and constraints on the web pages.

Visualization and comparison functions provide tables and charts based multiple visualization methods to show and compare cloud configuration solutions through user interactive ways.

Remaining Challenges

There are two challenges remain:
1) The currency of the cloud platform information in the database. Since the cost information is frequently updated upon the periodical usage and cloud providers' business strategies. Without the updated information, the recommendation solutions and their configurations are unreliable. Therefore, how to keep the data up to date in the database is a big issue. We will try to frequently update this information and automate this process as much as possible. We also utilizes the subscription and notification functions provided by the cloud providers to obtain the information update events and then trigger the crawling and updating process.
2) The accuracy of both fee cost model and cloud capability evaluation model is another issue. Without correct and high quality evaluation and calculation models, the results of the advisory tool are not trustable. We will continually adjust and refine the models upon the further investigation progresses and user feedbacks as well. Meanwhile, we will build multiple models to fit different application types. The combination of these strategies may expect to produce an acceptable result.

Resources

- ESIP Federation Member Organizations (with contact information)
- Third party cloud monitors, analytic and evaluation reports and websites
- Websites (swp.gmu.edu/esiptestbed) / Documents / publications
- Commercial and open source cloud platforms and their websites
- Development Technologies (Drupal, Dojo, Google Charts, JavaScript, Java, Apache Tomcat, MySQL)

References

Martens, Benedikt; Teuteberg, Frank; and Gräuler, Matthias, 2011. Design and implementation of a community platform for the evaluation and selection of cloud computing services: a market analysis. *ECIS 2011 Proceedings.* Paper 215. http://aisel.aisnet.org/ecis2011/215
Andrzej Goscinski, Michael Brock, 2010. Toward dynamic and attribute based publication, discovery and selection for cloud computing. Future Generation Computer Systems, 26(2010) 947-970.

Jonas Repschläger, Stefan Wind, Rüdiger Zarnekow, Klaus Turowski, 2011. Developing a Cloud Provider Selection Model. In Proceedings of the 4th International Workshop on Enterprise Modelling and Information Systems Architectures, EMISA 2011, Hamburg, Germany, September 22-23, 2011.