Cloud Computing Considerations for Biomedical Applications

Richard Rauscher
Department of Computer Science & Engineering and College of Medicine
Pennsylvania State University
University Park and Hershey, Pennsylvania, USA
rauscher@psu.edu

***Abstract*—This poster considers the practical barriers to public cloud use for biomedical applications and the advantages of private cloud use for such applications. In addition, it discusses operating environment statistics that are relevant to correctly allocating resources in a private cloud.**

Keywords-platform virtualization; biomedical computing

I. INTRODUCTION

Cloud computing has received significant attention in the popular press, in academic publications and in the market. Most of the attention has been focused on so-called "public clouds" but a popular alternative are private clouds. There are unique challenges to the use of public cloud computing in biomedicine, particularly with regards to the transmission and storage of protected health information (PHI). These challenges are non-trivial and increase the attractiveness of private cloud computing solutions. This poster examines important operating environment considerations when sizing a private cloud using the REDCap application database server as an example.

II. PUBLIC AND PRIVATE CLOUDS

There are public and private clouds. Public clouds are analogous to a power company or a telephone company. A central organization operates the public cloud computing infrastructure and charges the users for access. Presumably, the efficiencies realized by the central company are such that it is able to reduce its costs to the point where it can offer the services at a reasonable price while maintaining a positive margin. Private clouds often use similar or precisely the same technology but are created and controlled by (typically) a single entity. They are analogous to private telephone networks and on-site generated power.

III. HEALTHCARE AND CLOUD COMPUTING

Schweitzer discussed the regulatory concerns and contractual scaffolding required to manage HIPAA-classified PHI using public cloud computing at length [1]. However, there are pragmatic issues, particularly associated with "infrastructure as a service" models that were not discussed. These are described below.

A. Commodity Cloud Computing

In the traditional out-sourcing model (which was it was referenced in [1]), establishing the contractual relationship between the external provider ("business associate" in HIPAA parlance) can be non-trivial or simply impossible. The Penn State Milton S. Hershey Medical Center and College of Medicine (PSMSHMC and CoM) include shared liability/risk as the cornerstone of its business associate agreement (BAA). Negotiating the language in the BAA has introduced significant legal costs and contractual delays. Although our experiences are relatively small using the most recent version of our BAA, we estimate that the shared liability/risk clauses have increased contract negotiations by forty-five to sixty days. This is due, in part, to the unknown risk being assumed by the business associate. The risk of mismanagement of PHI could result in liabilities that are several fold the benefit to the business associate. This assumes that the business associate is willing to enter into a BAA. Many are not willing to consider entering into a HIPAA-related BAA due to the legal responsibilities that such an agreement would impart.

B. Private Cloud Usage

For organizations that wish to reap the benefits of cloud computing, given the list of issues with the proceeding approaches, the cost-optimal solution may be the implementation of a private cloud. In this scenario, the covered entity builds its own set of servers to provide the cloud service using a virtualization system such as OpenStack or VMWare. If the covered-entity is to create its own private cloud, it's important that it be able to make intelligent decisions about which systems it can virtualize.

IV. PERFORMANCE CHARACTERIZATION

This poster discusses resource requirements and performance characteristics that dictate a properly-sized private cloud for biomedical applications. As an example, we walk through determining the cloud resources required by a "REDCap" biomedical database server [2].

ACKNOWLEDGMENT

The project described was supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant UL1 TR000127. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

REFERENCES

[1] Schweitzer, E., "Reconciliation of the cloud computing model with US federal electronic health record regulations," J Am Med Inform Assoc. 2012 Mar-Apr;19(2):161-5.

[2] Harris, P. et al, "Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support", J Biomed Inform. 2009 Apr;42(2):377-81