Proceedings of the 2012 IEEE 12th International Conference on Bioinformatics
& Bioengineering (BIBE), Larnaca, Cyprus, 11-13 November 2012

Intensive Care Cloud

Exploiting cloud infrastructures for near real-time vital sign analysis in intensive care medicine

Dr. Nikolas Stylianides
Dept. Computer Science
University of Cyprus
Nicosia, Cyprus
nstylianides@gmail.com

Dr. Harald Gjermundrod
Dept. Computer Science
University of Nicosia
Nicosia, Cyprus
gjermundrod.h@unic.ac.cy

Dr. Marios Dikaiakos
Dept. Computer Science
University of Cyprus
Nicosia, Cyprus
mdd@cs.ucy.ac.cy

Dr. Theodoros Kyprianou
St. George's Medical School
University of Nicosia, UNIC
Nicosia, Cyprus
drtheo@cytanet.com.cy

***Abstract*— This paper introduces a novel, open architecture cloud oriented framework named Intensive Care Cloud, ICCloud. ICCloud main objective is to provide a) a common repository to store anonymized vital sign parameters retrieved from intensive care bedside medical devices and b) exploit massive computational resources to analyze vital sign parameters. ICCloud uses a simplified data model (*cloud tables*) to reduce the complexity of the deployment and the cost of usage of the infrastructure. These *cloud tables* are accessed only within the cloud infrastructure (cloud jobs) using only inbound network traffic. Pre-allocated working nodes (virtual machines used only for processing) are used to execute jobs implementing scoring functions or data analysis algorithms.**

***Index Terms*—** *Intensive Care, Vital sign analysis, Cloud medical databases, Open Repository, Near real time vital sign processing*

I. INTRODUCTION

Medical information data repositories serve a critical function in healthcare, including the areas of patient care, research and education. The quality and structure of information collected into existing data repositories varies. Within the acute environment of the intensive care unit, medical support and patient treatment could be further advanced by the development of open, structured data repositories able to process the enclosed medical records and vital sign parameters utilizing massive computational resources. Accurate and complete medical records are very important for early diagnosis [1].

Apart from the data that may be used for local assessments or evaluations within a healthcare system, such as for specific outpatient conditions or inpatient hospital events, it is also important to record patient vital signs parameters in high frequency and use them to extract information vital for early diagnosis and prognosis. Medical information must be used regionally, nationally and/or internationally for medical assessment, research, etc.

Furthermore, current medical device and medical sensors produce large amount of data using proprietary data definitions and medical information definition languages. Lack of interoperability among medical devices and/or third party hospital information systems could be addressed with open schemas and shared data repositories. More data means also increased demand in storage and computational resources. The problem becomes sound in cases where increased patient records from interregional health institutes are imported into the system as input in data analysis algorithms leading to conditions where on-premise computational resources fail to meet the requirement for near real-time data analysis.

In order to address these challenges we use Cloud infrastructures (both storage and computational), which enable to move computing and data away from desktop and portable PCs into large data centers. It also reduces complexity as far as data organization is in concern and provides a cost effective application deployment model [3].

II. MOTIVATION

The introduction of ICT in the healthcare industry has resulted in a multitude of improvements. From the real-time monitoring of patients by devices with very high fidelity to the hospital information systems. With the additional information, progress has been made in medical research in order to improve health care treatment and reduce mortality, malpractice and mistakes. One important innovation from the medical research is the various scoring systems, like APACHE II [4], SAPS II [7] and SOFA [8] that are used to determine the expected outcome for inpatients or the seriousness of their condition. By combining more accurate and higher fidelity information about the patient with the various scoring systems, the appropriate treatment for the patents can be applied quicker and hopefully with a greater successful outcome for the patient.

978-1-4673-4358-9/12/$31.00 ©2012 IEEE

It is important to note that in the intensive care setting there is, at the moment, lack of a real integrative monitoring technology that will acquire the physiological parameters from the different organ systems and analyze and present them in correlation to each other in an interactive environment that will facilitate the physician to extract the best conclusions for the patients treatment. The use of information technology aims to cover at least a part of this lack with the use of clinical information systems. However, the huge amount of the required resources associated with the purchasing, the deployment and the maintenance of the commercially available systems is currently a prohibiting factor for the wide spread in the intensive care.

Another restrictive factor is the fact that the commercially available systems use "*closed*" data repositories which are not made available to the customer and in some cases may lead to further complexities and hidden costs especially when interoperability with other clinical systems is required. Furthermore, the clinical systems commercially available at the moment focus more on the management of the everyday clinical work instead on the in-depth analysis of the acquired physiological parameters. Although this is very useful for the everyday clinical practice in an ICU, it does not offer the in-depth interactive analysis environment often required by the physicians.

The aforementioned restricting functionality motivated us to develop a cloud based framework solution that will be used a) as a common data repository for ICU medical information and b) provide data analysis and information extraction software tools to medical researchers to develop more effective and accurate evaluation systems and probabilistic models.

The proposed framework has been developed under the Venus-C [2] platform, which is used as an abstraction layer to overcome cloud provider borders.

III. RELATED WORK

This section will first introduce ICU scoring systems. Then the ProSafe and ICW/ICGrid projects are discussed and compared to ICCloud.

There are several ICU scoring systems that are currently being used by ICU units worldwide. The scoring systems are designed to evaluate the severity of an inpatient's condition, or to benchmark the mortality of patients. A short introduction to the most common used systems is provided below:

- APACHE II (Acute Physiology and Chronic Health Evaluation) [4]: Provides a morbidity score of an inpatient, which is an integer value in the range of 0 to 71. Higher value indicates a more sever condition and higher risk of death. The score can be used to decide what kind of treatment should be used.
- GCS (Glasgow Coma Scale) [5]: Provides the conscious state of an inpatient and it is an integer value between 3 and 15 (in the original version the max value was 14). Lower values indicate unconsciousness while higher values indicate consciousness. This score is often used as part of other scoring systems as it cannot be used to determine the medical state of the patient.
- PIM2 (Paediatric Index of Mortality) [6]: Provides a predicated mortality value specifically designed for children. It is designed to be used as a benchmark to determine the morbidity of a patent compared to the outcome with other patients.
- SAPS II (Simplified Acute Physiology Score) [7]: Provides a score of the severity of a patient that is admitted to the ICU. The score is calculated 24 hours after admittance and it consists of an integer value between 0 and 163 and predicted mortality between 0% and 100%. It is designed to be used as a benchmark to determine the morbidity of a patent compared to the outcome with other patients. SAPS III also exists, which is a supplement to SAPS II.
- SOFA (Sequential Organ Failure Assessment) [8]: Provides a simple daily score of a patient status as it evolves over time. The score consists of six integer values between 0 and 4, with one value for each of the following systems: respiratory, cardiovascular, hepatic, coagulation, renal and neurological.

The ProSafe (PROmoting patient SAFEty and quality improvement in critical care) project [9] is a EU framework project that ran from April 2008 until February 2012. The consortium consisted of 7 partners from Cyprus, Germany, Hungary, Italy, Poland, Slovenia, and United Kingdom in addition to external collaboration with Brazil. The goal of the project was to provide ICU units with a method to evaluate their weaknesses and strengths by comparing to other ICU units and therein identify opportunities and threads for improving critical care. To accomplish the goal it was identified that the ICU must continuously, easily, rigorously, and confidently measure its own performance against others. The project therefore defined what information needed to be collected from the ICU and what the format of this information should be specified in. Case Report Forms (CRF) were defined and collected from the participating ICUs. Two categories of forms were developed. The first was a set of forms that are used to describe the services, facilities, medical staff, and capabilities of the ICU unit and the hospital that ICU unit resides within. The second category comprises forms used to report on the patients as they are admitted and treated in a specific ICU unit. A multi language web application was developed to collect these forms electronically. The data is used for comparison of statistics and performance of the participating ICU units. The following ICU scoring systems were used to compare/benchmark the ICUs performance: GCS, SAPSII, APACHEII, SOFA, and PIM2. By providing these statistics to the participating ICUs good practice can be shared and exchanged between the ICU sites, as well as to other ICU, in order to improve critical care.

The ICCloud can complement the SafePro project in many ways. The main purpose of SafePro is to provide an environment for ICU to benchmark their overall performance in order to improve. This is in contrast to the ICCloud project, which provides the ability for the physician to search for similar incidents in near real time as she is treating a patient in the ICU ward.

The ICCloud project is a continuation of the ICW (Intensive Care Window)[10] and ICGrid (Intensive Care Grid) [11] projects. The ICW is an extensible component-based medical devices extraction software. It is used to retrieve the raw data that are produced/collected by medical devices, which exist in an ICU unit. In order to support a new medical device, a small plug-in for connecting and converting the data must be provided. The ICGrid project provided for the functionality for a physician to annotate data that were collected from the ICW (1st generation) and then upload this information to the EGEE Grid Infrastructure. The main reason to move from the Grid to the Cloud is that the Cloud provides for a much better predictability of resource availability.

IV. SYSTEM OVERVIEW

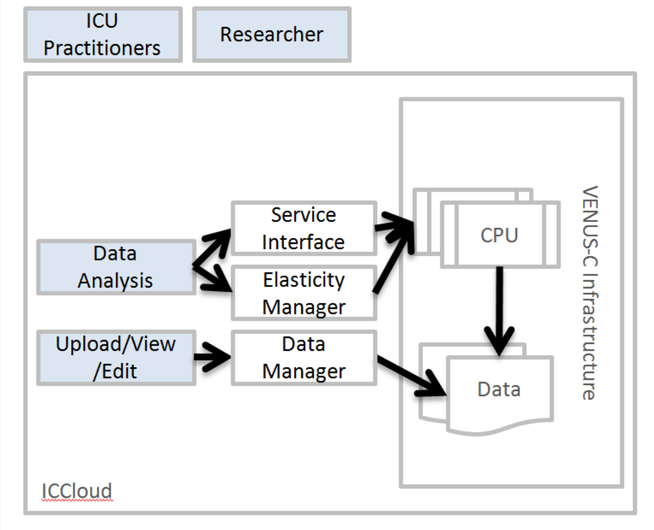

Figure 1: ICCloud major functions based on User Roles

A. System Description

ICCloud has been designed and implemented based on the architecture and provided services of the Venus-C platform. The Venus-C platform provides a level of abstraction for interdisciplinary cloud infrastructures providing seamless access to computational and storage resources. Venus-C infrastructure mainly uses resources from Microsoft Azure [12] and Barcelona Supercomputing Center [13]. The current version of ICCloud has been developed as a desktop application using .net languages VB.net and C#. The next generation ICCloud will be deployed as Software-as-a-Service solution being pure cloud solution to reduce maintenance (code and runtime) overhead.

Figure 1 depicts the architecture of ICCloud based on the major functionality and user roles. ICCloud major functionalities include:

Upload/View/Edit Functionality enables a) *real time patient monitor*, which is a solution build upon former proved and tested software services such as the Intensive Care Window and could be installed in ICU bedside medical device networks to acquire patients vital sign parameters in real time (vital signs are depicted both in tabular and graphical format) and b) *real time patient vital sign repository synchronization*, a service used to upload patients' vital sign parameters to cloud repositories in real time. Additional data such as evaluation scores and/or probabilistic model results are manually imported into the system.

Data Analysis functionality. Common repositories serve the need to create large datasets that can enrich information quality. Vital sign parameters are stored in tables in the form of time series. Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. Time series analysis is a CPU demanding procedure. In order to reach the near real time analysis requirement the system takes advantage of cloud computational capabilities to deploy CPU demanding programs within the cloud infrastructure.

Major user roles are: a) Medical researchers that process vital signs to extract interesting properties that would lead in novel or improve existing scoring algorithms and/or probabilistic and b) Intensive Care Unit practitioners that could compare and analyze patient's data versus others' patients' episodes stored in common repositories in near real time.

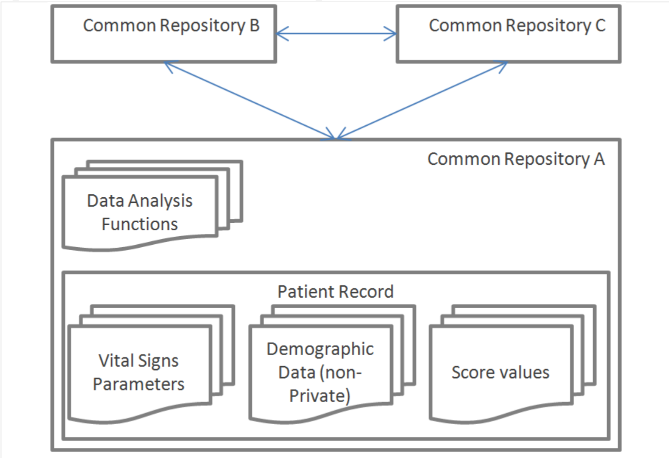

Figure 2: ICCloud Data Model

B. System Resources

As aforementioned, the system has been developed also to exploit Venus-C processing capabilities and elasticity feature in order to provide end-users (Doctors and Nurses) with computer-aided diagnosis capabilities. The data analysis techniques and time series processing though was kept outside the scope of this work since data analysis for the purpose of the evaluation was applied on a small set of data that was adapted to the needs of the test. In a full feature system the user will be able to import on demand and based on preconfigured method data analysis algorithms into the system.

In Venus-C platform as in all cloud infrastructures we distinguish two major resources, data and computation.

ICCloud has been designed to fit into these models seamlessly under Venus-C infrastructure.

Common Repositories. ICCloud data are organized in common shared repositories (see *Figure 2*). These repositories uphold data from an organization such as a hospital or research organization. These repositories follow the data model and basic rules to export tables and relationships in a formal manner. Authentication and therefore access in these repositories is achieved using public key certificates architecture. The infrastructure utilizes RESTful services to access core elements such as storage entities (Binary Large Objects, BLOB and Tables). The basic storage entity that the ICCloud solution utilizes in these repositories is cloud table. A cloud table stores data in entities (similar to table rows except that each entity can define its own fields), can uphold million of entities and span in several storage elements. Each ICCloud repository composes of a) administration tables, b) patient records and c) data analysis function references (the actual function is stored in BLOBs folders). Patient vital signs entities are stored in a single table for each patient. Each table entity is composed of a) vital sign identifier, b) timestamp, c) vital sign value, d) vital sign value type (real number or textual) and e) value sign state (normal | alarm). It is important to point out that the vital sign identifier is used as a partition key in order to keep identical vital sign types in the same storage partitions and limit down interregional data transfers among different resource zones which introduce processing delays and increase usage cost. Scores table entities consist of a) score identifier, b) score parameter identifier, c) value (textual | numeric) and d) timestamp.

Computation. Each computation, either using one computational node or many, has the characteristics of a simple experiment. Each experiment can be a set of one or more experiments which in turn produce there results to complete the experiment result. This type of configuration works well with the map-reduce programming model. In its simple form, the one used for scenario testing and evaluation, an experiment consists of a) a set of time series (vital signs) and other medical reports to analyze, b) a large pool of data sets to compare and c) the processing functions to use in the experiment. Each experiment is executed within a cloud "Job". Venus-C architecture uses job queues to maintain job submission requests from users. Each Venus-C computational node (virtual machine instance) runs a monitor service named Generic Worker (see *Figure 3*), which extracts jobs from the queue and executes them. These instances group together to compile computational services forming one-to-many relations. Experiments run within services using the hosted VM instances.

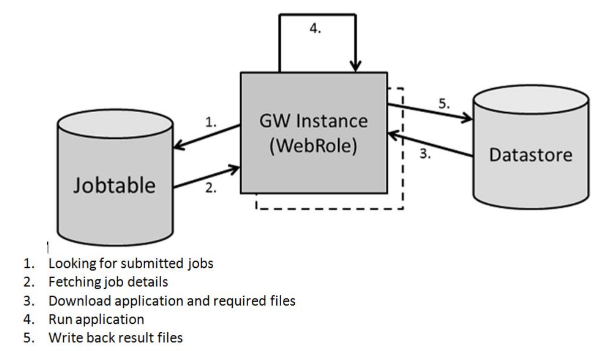

Figure 3: Generic Worker walkthrough

The researcher is able to submit a batch of M jobs where M is greater than the number of VM instances in each service. The jobs are served sequentially in a first come first served order by the Generic Worker instance. The report of each job is stored in text files under a uniquely identified cloud directory. The complete set of these files form the result of the experiment. The result could be downloaded for further process or may be used online.

Each job or group of jobs are identified based on a uniquely identifier and are associated with the researcher account. Since the Venus-C infrastructure does not support in depth description of each job or group of jobs the ICCloud maintains a record under the researcher account. This record provides information regarding the experiment parameters (data input, analysis function, scores to include in the analysis procedure, common repositories used, etc) and the overall result. These results could be fine grained with extra analysis at the researcher side.

V. SCENARIO TESTING & EVALUATION

As aforementioned the ICCloud can address the computational requirements of CPU demanding functions and experiments faced both in everyday ICU practice and ICU research. During the evaluation and testing phase though scenarios must be kept simple enough and complicated scenarios must be tested at later time in more matured implementation phases. In order to evaluate both ICCloud solution and Venus-C infrastructure a scenario that tests and evaluates both data and computation is selected.

Scenario Description. Our goal is to identify clinically interesting episodes in real time from within the ICU. To achieve that we need to monitor patients' vital signs in real time and analyze these parameters base on historic data. In order to achieve that we need a) to select a set of vital signs b) select a time window in which an exception is occurred and c) using a comparison function compare the data against a large pool of data sets. What we seek is similarities (patients with similar characteristics) within these large historic data sets that could be used in the diagnosis procedure.

Dataset. The scenario needs real vital sign parameters, which were captured from ICU bedside medical devices. These vital signs are retrieved and stored as parameters. The capture frequency is one snap-shot of the patient state every 15sec. A group of patients with a variety of admission reason have been selected. The admission reason set consists of the following: cardiogenic shock, cardiac arrest, coronary artery bypass graft, acute abdomen, respiratory failure, brain injury, multi-injury and pelvic abscess have been selected. The vital sign parameters of each patient were uploaded to the storage element in cloud tables (one table per patient). The common repository used for this scenario was also amended with patient data that were created (copies) using these patients vital sign parameters. Doing so, we were able to simulate the behavior of the system with small to medium size repositories (less than 1000 records). Using the graphical view of ICCloud the researcher can load in a single view one or more vital sign parameter for analysis. Vital signs are grouped and vertical aligned and can be processed concurrently with a single user action. Using this tool the researcher is able to browse the vital signs (time-series), move forward and backward in time and search for clinically interesting episodes within these time-series from a single view.

Computation. Once a clinically interesting episode is identified the area of interest is selected also (a group of vital signs and the time window). The researcher selects the comparison function, the vital signs of the patient to analyze and compare against the pool of data sets and fires up the experiment. To prove the feasibility of the scenario we have selected a CPU demanding comparison function such as the cross-correlation. We use the cross-correlation function because of its characteristics in time series analysis. In time series analysis, as applied in statistics, the cross correlation between two time series describes the normalized cross covariance function and is used to estimate the degree to which the time series are correlated. The computation takes place within a cloud service. The ICCloud will submit M jobs, where M is equivalent to the number of patients in the common repository. Therefore each job will process the area of interest against the data set of one patient within the common repository. The result of each process is stored in the experiment record and it will be used along with other data (score tables, admission reason, etc) to form the final experiment result (a list of similar episodes). The outcome of this experiment using the specific data set is of low importance. This is because the data set is just a copy and reproduction of real time series, therefore the result will be a perfect match which is not normal for a real scenario. The evaluation of the scenario focuses on two parameters: a) the experiment processing time of the experiment (how long would it take to process the request?) and b) the actual cost of each experiment.

Processing time. This is an important factor for each experiment especially for this scenario where the experiment results should be completed within a few hours. Each patient stays in ICU on average for 72 hours. Using small VMs (single core PC) the process of two time series cross correlation completes in less than 10 seconds (this includes submission, copy data to local storage from cloud table, processing and report the result). Based on the size of the common repository we use for this scenario (500 patients) we would need approximately 30 hours to process the data on a single core PC using a single thread application. Using ICCloud and a cloud service with 15 single core VMs (small service) the experiments complete in less than 2 hours which is comparable to other laboratory tests.

Cost of usage. This factor is also very important to take in concern especially in times where hospital budgets for IT resources and personnel are limited. The selected scenario budget includes storage and computation cost. *Storage cost* is divided in storage volume and transactions. The cost to upload and maintain for one year a common repository such as the one selected for the evaluation scenario (store 500 patients' medical records using cloud tables and cloud blobs only) is less than 250 USD. The c*omputation cost* for the evaluation scenario is less than 4 USD and is calculated as the product of the cost of 15 small VM (0,115 USD / hour) 2 hour computation.

VI. CONCLUSIONS

The results from the evaluation scenario depict the usability and cost effectiveness of the proposed solution. Common repositories, as mentioned in previous sections, have been created from several research organizations and have been used mainly as medical data sharepoints. Data are kept in databases or in binary files (available through URLs) and are available to researchers for processing. ICCloud aims to complement this solution with formal structured common repositories, interregional share points (for data and processing functions) and the processing elements (computational resources utilizing cloud virtual machines and application services and data analysis functions to process data). The solution is cost effective and eliminates the need for host IT infrastructures and IT personnel. Furthermore to the scenario used for evaluation a medium size ICU (20-30 patients) would need 100 GByte to store one year's data (vital sign parameters only). If we project the data needs to a 10 year scenario the cost to maintain a common repository is less than 10K USD.

ICCloud is an early effort to bridge the gap among research organizations in the field medical research. Future work includes the following: 1) develop a cloud version of ICCloud user interface and deliver the application as S-a-a-S, b) further develop the data share model and access rules, c) invite more ICUs to share data and knowledge and d) develop (or reuse) and share a complete set of mathematical functions and processing tools.

ACKNOWLEDGMENTS

We would like to express our gratitude to all the partners of the Venus-C project [2] and especially the resource providers. ICCloud has been one of the sixteen pilot applications selected to participate in the evaluation and testing phase of the project. Venus-C provides its storage and computation resources for our experiments.

REFERENCES

[1] B. Misset, D. Nakache, A. Vesin, M. Darmon, M. Garrouste-Orgeas, B. Mourvillier et. al., "The Outcomerea Database Investigators Reliability of diagnostic coding in intensive care patients", Critical Care, vol.12, no. 4, 2008.

[2] "VENUS-C", http://www.venus-c.eu

[3] M. D. Dikaiakos, G. Pallis, D. Katsaros, P. Mehra, A. Vakali. "Cloud Computing: Distributed Internet Computing for IT and Scientific Research", Guest Editorial in IEEE Internet Computing, vol. 13, no. 5, pp. 10-13, Sept.-Oct. 2009.

[4] W. Knaus, E. Draper, D. Wagner, J. Zimmerman, “APACHE II: a severity of disease classification system”, Critical Care Medicine, vol.10, pp. 818–829, Oct. 1985.

[5] G. Teasdale, B. Jennett, “Assessment of coma and impaired consciousness: A practical scale”, vol. 304, no. 7872, pp. 81-84, July 1974.

[6] A. Slater, F. Shann, G. Pearson, Paediatric Index of Mortality (PIM) Study Group, “PIM2: a revised version of the Paediatric Index of Mortality”, Intensive Care Med, vol. 29, no. 2, pp. 278–285, February 2003.

[7] J. Le Gall, S. Lemeshow, F. Saulnier, “A New Simplified Acute Physiology Score (SAPS II) Based on a European/North American Multicenter Study”, Journal of American Medical Association (JAMA), vol. 270, no. 17, pp. 2957-2963, December 1993.

[8] JL. Vincent, R. Moreno, J. Takala, S. Willatts, A. De Mendonça, H. Bruining, et. al., “The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure.” On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. Intensive Care Med, vol. 22, no. 7, pp. 707-710, July 1996.

[9] “ProSafe Project”, http:// http://prosafe.marionegri.it

[10] N. Stylianides, M. D. Dikaiakos, H. Gjermundrød, G. Panayi, T. Kyprianou, "Intensive Care Window: Real-Time Monitoring and Analysis in the Intensive Care Environment," Information Technology in Biomedicine, IEEE Transactions on, vol.15, no.1, pp. 26-32, January 2011

[11] H. Gjermundrod, M. D. Dikaiakos, D. Zeinalipour-Yazti, G. Panayi, and Th. Kyprianou (2007) “ICGrid: Enabling intensive care medical research on the egee grid”, HealthGrid 2007, pp. 248-257, Geneva, Switzerland.

[12] Microsoft Windows Azure: http://www.windowsazure.com

[13] Barcelona Supercomputing Center: http:// www.bsc.es