

Received 11 February 2015; revised 6 August 2015; accepted 21 September 2015. Date of publication 01 October 2015; date of current version 13 October 2015.

Digital Object Identifier 10.1109/JTEHM.2015.2485268

Health Informatics for Neonatal Intensive Care Units: An Analytical Modeling Perspective

HAMZEH KHAZAEI¹, (Member, IEEE), NADJA MENCH-BRESSAN^{2,3}, (Member, IEEE),
CAROLYN MCGREGOR², (Senior Member, IEEE), AND JAMES EDWARD PUGH^{2,3,4}

¹Canada Research and Development Center, IBM, Toronto, ON L3R 9Z7, Canada

²Faculty of Business and Information Technology, University of Ontario Institute of Technology, Oshawa, ON L1H 7K4, Canada

³Department of Neonatology at Hospital for Sick Children, Toronto, ON M5G 1X8, Canada

⁴Division of Neonatology at McMaster Childrens Hospital, Department of Paediatrics, McMaster University, Toronto, ON L8S 4L8, Canada

CORRESPONDING AUTHOR: H. KHAZAEI (hamzeh.k.h@ieee.org)

This work was supported in part by the Canadian Foundation for Innovation under Grant 203427, in part by the Canada Research Chairs Program under Grant 950-225945, and in part by Southern Ontario Smart Computing Innovation Platform Consortium.

ABSTRACT The effective use of data within intensive care units (ICUs) has great potential to create new cloud-based health analytics solutions for disease prevention or earlier condition onset detection. The Artemis project aims to achieve the above goals in the area of neonatal ICUs (NICU). In this paper, we proposed an analytical model for the Artemis cloud project which will be deployed at McMaster Children's Hospital in Hamilton. We collect not only physiological data but also the infusion pumps data that are attached to NICU beds. Using the proposed analytical model, we predict the amount of storage, memory, and computation power required for the system. Capacity planning and tradeoff analysis would be more accurate and systematic by applying the proposed analytical model in this paper. Numerical results are obtained using real inputs acquired from McMaster Children's Hospital and a pilot deployment of the system at The Hospital for Sick Children (SickKids) in Toronto.

INDEX TERMS Health informatics, data management, real-time analytics, analytical modeling, capacity planning, cloud computing.

I. INTRODUCTION

High speed physiological data together with high speed data from other medical support devices such as ventilators, infusion pumps and in the case of neonatal intensive care, incubators, is a largely untouched resource in healthcare [1]. While monitoring these forms of data has origins in the support of critical care medicine only, the growth of personalized medical sensing devices available in increasing numbers to a diverse range of consumers is quickly changing this paradigm. Opportunities abound for significant medical discovery and quality improvement in healthcare that lead to improvements in productivity together with reduced morbidity and disability rates through the establishment of systemic approaches for the acquisition, transmission, processing, analytics and storage of these forms of Big Data. In addition, the maturing of cloud based platforms naturally lends itself as the technology of choice to provision Big Data platforms for use in healthcare. While Big Data in healthcare exists beyond critical care, critical care still presents one of the most complex settings where multiple high speed streams of data are generated per patient and need to be brought

together in various ways concurrently for new approaches to bedside care. To enable the use of Big Data platforms at bedside, we must understand the nature of their usage requirements and create a model for how to determine the potential load requirements for their provision in a cloud computing setting. One of the most complex settings to build such a model is that of neonatal intensive care.

Premature birth defined as birth before 37 weeks gestational age, has been identified as one of the most important health problems resulting in a high chance of long term morbidity impacting not only the child and caregivers but wider society in industrialized nations. Neonatal Intensive Care Units (NICUs) provide critical care for premature and ill infants. Premature infants in NICUs can be as young as 23 weeks gestation [2].

The data generated by medical devices in neonatal intensive care is a big data problem. Vital sign monitoring together with ventilation support and nutrition or drug titration through smart infusion pumps all generate large volumes of data at high frequency. A two channel electrocardiogram can generate 1000 readings a second. Heart rate,

respiration rate and blood oxygen are displayed each second resulting in 86,400 readings each per day. A premature newborn infants heart beats more than 7000 times an hour which is approximately 170,000 times a day. A newborn infants neurological function could also be monitored resulting in multiple waveforms each generating tens of millions of data points per patient per day. Drug and nutrition infusion data from smart infusion pumps can yield more than 60 different fields provided every 10 seconds. Given that these infants can have more than 10 infusions concurrently, infusion can generate more than 1GB of drug infusion data from a single patient per day [3].

Infusion drug data has been used for real-time analyses for the past 10 years most commonly in the anesthesia field with the integration of physiological data and drug dose to automatically control anesthesia for short term surgical windows of time. However, the use of infusion pumps, is not just for short term operations but can continue for months or even years in neonatal, paediatric and adult intensive care units as well as in paediatric hospitals where they are used for all infusions. This real-time data in the clinical environment represents a relevant aspect if properly translated into information to advise clinicians and health practitioners during day-to-day care [4], [5].

Recent research is building a strong case for the benefits of real-time data analysis, with clinical events such as late onset neonatal sepsis (LONS) [6] exhibiting early warning signs in physiological data before clinical signs become apparent. However, that research takes a condition specific, patient specific or physiological data stream type specific approach [7], [8].

Through our research we have proposed the Artemis platform that provides data acquisition and storage of physiological data and clinical information for the purpose of real-time analytics, retrospective analysis and visualization. Artemis is not an acronym; it is named after the Greek Goddess associated with protecting child-bearing women and young children. Artemis enables concurrent multi-patient, multi-stream and multi-diagnosis through temporal analysis to support real-time clinical decision support and clinical research [7]. We have designed an expanded Artemis Cloud platform to service multiple healthcare facilities. This has many benefits especially given the new computing tools such as stream computing that are used to analyze the data in real-time and the lack of skills at each hospital to support such platforms. However, to be able to correctly size implementations on a per hospital basis, based on the number of beds and patient characteristics, the creation of an analytical model to enable capacity planning for the usage of such a platform is required. Analytical models within the context of cloud based big data solutions is currently an under researched area, especially within the context of its use in healthcare [9].

In this paper we present a method to design and evaluate an analytical model to enable more accurate estimation of storage, memory and computation power for the real-time and retrospective analytics components of our Artemis Cloud.

The model utilizes realistic patient population distribution based on gestation age characteristics and condition onset probabilities within those contexts. Both of these variables dictate the predicted length of stay for that infant. In this work we model the Artemis Cloud deployment at McMaster Childrens Hospital taking into account all exogenous parameters including all types of patients, infusion pumps, online/offline analytics and other requirements specific to this environment. The results of this work will be an important input for the final deployment of the platform at McMaster Childrens Hospital NICU and an important aspect for translational engineering research for the deployment of Big Data solutions in healthcare.

II. RELATED WORK

Current cutting edge health informatics research projects aim to discover new condition onset behaviors that are evident in physiological data streams earlier than traditional detection of conditions in critical care data [7]. To do this, some hospitals may participate in pilot programs that aim to collect real-time patient data from network enabled monitoring devices. This collected data is then analyzed to extract relevant temporal behaviors and usually stored for future data mining and analysis operations.

Historically physiological stream monitoring of ICU patients has been provided by regulatory body approved medical devices located at the patients bedside. While there has been a growing body of biomedical engineering and clinical research over the past 20-30 years proposing newer approaches for advanced physiological waveform monitoring they still predominantly have either a physiological stream, clinical condition or patient centric approach [7]. Zhang et al. [10] have discussed the implementation of a Health Data Stream Analytics System called the Anaesthetics Data Analyzer (ADA). The ADA has been developed to provide anaesthetists with the ability to monitor and query trends in physiological signals data, a kind of stream data from the health care domain.

Bressan et al. demonstrated the use of drug dose data when synchronized with physiological data streams in a contextualized system had a strong correlation with heart rate variability, weigh and maturation in the premature infants population. The design and deployment of a Pharmacokinetic/Pharmacodynamics simulator for the Artemis platform combining it with physiological data will allow for the development of an advanced decision support tool to aid clinicians in developing personalized drug dosing for infants in the neonatal intensive care unit [11].

The multi-patient, multi-stream and multi-diagnosis structure of Artemis Cloud enable the inclusion of infusion drug data into medical algorithms. In a recent study [12], the late onset Sepsis algorithm by [13] was improved with the addition of morphine concentration. The infusion drug data contextualized the physiological data resulting in 100% of accuracy to identify false positives in the late-onset of neonatal sepsis (LONS) Risk Score. The analytical modeling

described in this paper provides a better understanding of the computational requirements to include infusion devices into the streaming process of Artemis Cloud and how efficiently deploy infusion drug data into medical algorithms.

Cloud computing has attracted considerable research attention, but only a small portion of the work completed so far has addressed performance issues, and a rigorous analytical approach has been adopted by only a handful among these [14]–[17]. Many research works have carried out a measurement-based performance evaluation of the Amazon EC2 [14], [18], IBM Blue Cloud or other cloud providers in the context of scientific computing [14], [16], [19]. Khazaei et al. [20] proposed various analytical models for performance evaluation of cloud computing centers; however, the authors investigated performance metrics related to generic cloud centers as opposed to a cloud-based solution for a specific domain.

Hayes et al. [21] proposed an analytical model for a proposed infrastructure which is supporting an in-house deployment of Artemis. Khazaei et al. [9] modeled a version of Artemis project deployed at SickKids Hospital in Toronto; however that modeling is specific to SickKids’s NICU which has different types of patients, monitors, and analytics

compared to McMaster Children’s Hospital. In addition, that work did not incorporate load testing that included smart infusion pump data.

Artemis is a unique system that permits multi device analysis and interpretation of the data at the rate it is generated. Previous approaches have been either data stream centric such as only processing electrocardiogram, condition centric, such as focusing on a condition such as late onset neonatal sepsis or patient centric to enable monitoring of the onset of patient instability. Artemis is a Big Data platform that is multi-dimensional catering for all these requirements simultaneously. In addition, it is unique as nearly all other systems use down sampled data and can not combine data from multiple sources. e.g. monitor, pumps and ventilator to produce new information that has not previously been available.

III. METHOD

Artemis is an approach for online health analytics of high speed physiological data bootstrapped with historical analytics that had been performed on various patients’ medical data. Table 1 describes the acronyms and terms that we used in Figs. 1, 2 and in the text. We also describe the studies

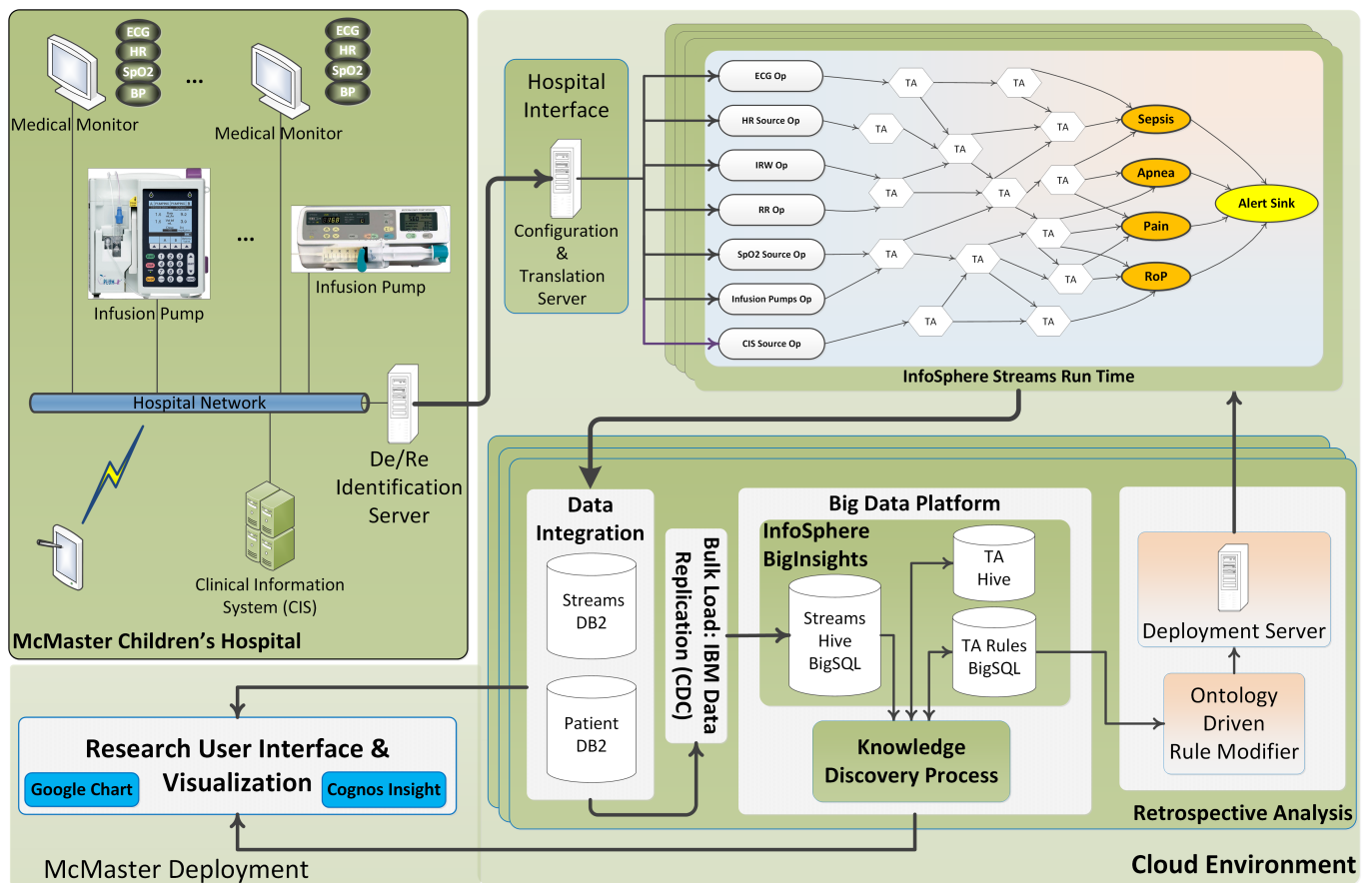


FIGURE 1. The high level architecture of Artemis Cloud. The hospital environment on the top left; realtime processing platform on top right; Data archival, data mining and knowledge discovery on bottom right; and visualization of live and historic data on the bottom left.

TABLE 1. Descriptions of acronyms and algorithms.

Acronyms	Description
TA	Temporal Abstraction
OP	Operator
CDC	IBM InfoSphere Change Data Capture
Algorithms	Description
Sepsis	Sepsis is a potentially life-threatening complication of an infection [22].
Spell	Neonatal spells are cardiorespiratory events that occur in newborn infants with variable combinations of cessation of breathing, decrease in blood oxygen saturation and decrease in heart rate [3], [23].
Apnea	Lack of breathing activity for at least 15-20 seconds or less if decrease in saturation or bradycardia are present [3], [23].
Pain	Detecting pain in neonates [24].
RoP	Retinopathy of prematurity, which can result in permanent blindness [13].

(i.e., algorithms) that have been developed by leveraging Artemis platform.

Now we describe the architecture of Artemis with respect to Fig. 1. Artemis Cloud is capable of gathering physiological and infusion pump data from a vast variety of medical devices and monitors (top left box) in a secure way. Anonymization and potential transformation are performed on the data before transmitting from the hospitals. Artemis Cloud also has an interface for communicating with a hospital clinical information management system in order to obtain complementary information about patients. Artemis Cloud utilizes a Hospital Interface (top middle box) which performs the extraction, transformation and load of data (i.e., ETL capabilities) on the one hand, and on the other hand facilitates the management of hospitals' connections to the back end cloud. This interface improves the extendability and modularity of the cloud based Artemis.

The core of Artemis Cloud's real-time functionality is a stream computing middleware component (top right box), IBM InfoSphere Streams (Streams, hereafter), which provides scalable processing of multiple streams of high-volume, high-rate data. Streams provides Artemis Cloud with a very extendable real-time execution environment. An application in Streams, consists of a set of operator nodes interconnected in a graph. Each operator node inputs one or more streams and produces one or more output streams.

In addition to real-time analytics capabilities, Artemis Cloud is able to provide at-rest analytics (i.e., retrospective analysis) for stored data (bottom right box). Incorporating IBM InfoSphere BigInsights (IBM distribution of Hadoop ecosystem), offers great power of analysis as well as persistence storage. Researchers can apply data mining techniques [13], machine learning algorithms and statistical modeling, against vast amount of stored information and find new rules which may help provide earlier detection of diseases. Such new rules or modified parameters can be deployed to the real time analysis platform seamlessly.

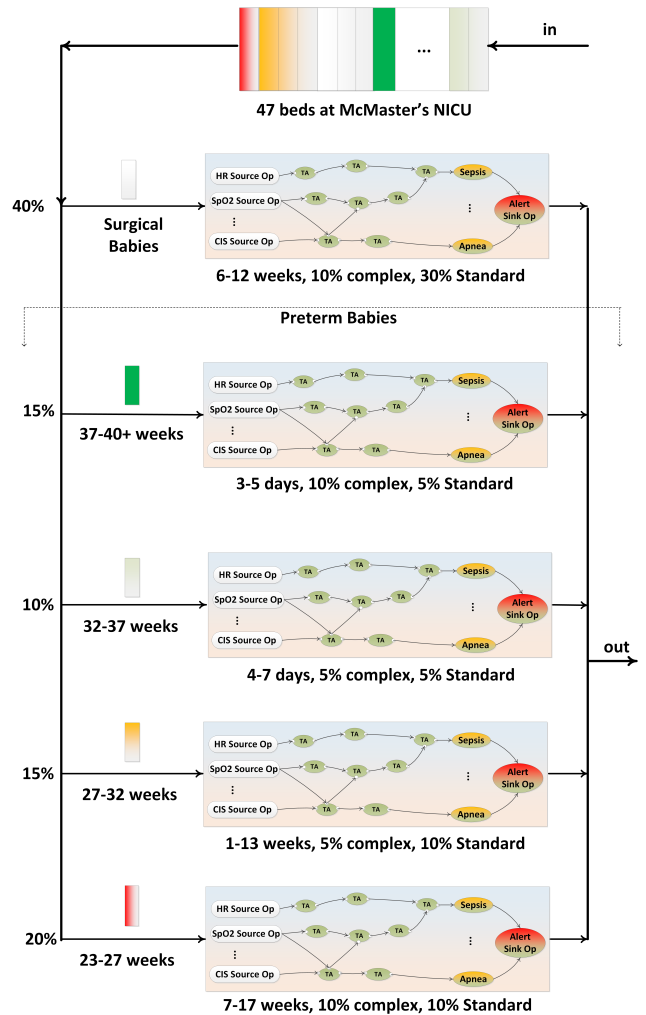


FIGURE 2. Types of patients (preterm and surgical term babies) at McMaster NICU. Neonates have been classified based on their Gestational age.

The deployment server is responsible for applying new rules and parameters to the real-time environment.

Artemis Cloud also utilizes a relational database in the data integration component which is interfaced with Streams to store the fresh data arriving from hospitals (Data Integration box). Research user interface and visualization components (bottom left box) visualize the real-time and historical data. Fig. 1 shows the architecture of the Artemis Cloud that is currently being deployed at McMaster. As IBM is one of the partners in this research, we have realized the architecture by IBM products.

One of the key benefits of Artemis Cloud is that it is automated and requires almost no staff input for gathering the data. The output of information from this system will require both research and systematic education however, with the graphical interfaces that are being designed to do this it should require no more education than a bedside medical device. As a result such a system when delivered through the cloud will require only one or two experts such as medical/computing engineers to maintain the local system

TABLE 2. Estimate of patients and their infusion pumps statistics in McMaster’s NICU. This estimates are based on expert opinion and data from the last 3 months of the unit.

Gestational Age (weeks)	Complex		Standard	
	Continues infusion (#pumps – #days)	Intermittent infusion (#pumps – #days – #hours)	Continues infusion (#pumps – #days)	Intermittent infusion (#pumps – #days – #hours)
37 – 40+	3 – 10, 1 – 5 days	2 – 14 days – 2 hrs/day	3 – 5 days	1 – 5 days – 1 hr/day
32 – 37	3 – 14 days	2 – 14 days – 2 hrs/day	3 – 5 days	1 – 3 days – 1 hrs/day
27 – 32	5 – 30 days	2 – 14 days – 2 hrs/day	3 – 10 days	1 – 3 days – 1 hr/day
23 – 27	5 – 30 days	2 – 14 days – 3 hrs/day	3 – 12 days	1 – 3 days – 2 hrs/day
Surgical	6 – 14, 3 – 60 days	2 – 14 days – 2 hrs/day	3 – 15 days	1 – 10 days – 1 hr/day

and this level of training is the same as any hospital computer system. By careful design of the output we mitigate the need for long term extensive on site training and improve acceptance.

IV. MODELING OF McMaster NICU

In order to model the NICU at McMaster hospital we tried to use real data as much as possible. In this section, we describe the patient journey and infusion pumps statistics obtained directly from McMaster’s NICU. All these inputs are required for characterizing the target infrastructure at McMaster’s NICU.

Fig. 2 and Table 2 show the patient journey and infusion pumps statistics for the NICU at McMaster Hospital respectively. These two describe the NICU as follows.

As can be seen (Fig. 2), McMaster Children’s Hospital has 47 NICU beds including different types of patients. A new patient will be accommodated to NICU once per hour on average. Patients are categorized based on their ages and conditions into preterm/term and complex/standard respectively. Forty percent of patients, including term and preterm babies, are referred to McMaster’s NICU for surgical purposes. Approximately, surgical babies stay at hospital for 6–12 weeks and 9 medical algorithms will be applied for after-surgery monitoring. The rest of patients, i.e. preterm babies, are classified into four categories; babies who are born in 37–40+, 32–37, 27–32 and 23–27 weeks of their gestation age. All these patients will be monitored by 8 or 9 real-time medical algorithms.

Table 2 shows that the complex surgical patients (i.e., 10%) requires 6 pumps for 14 days and then 3 pumps for 60 days; they also need 2 pumps for 14 days which are operating only 2 hours a day. If a surgical baby is indicated as standard patients (i.e., 30% of the whole patient population), they will be in need of 3 pumps for 15 days continuously, and one pump for 10 days for 1 hour a day intermittently. The first group of preterm babies (i.e., 37–40+) will be hospitalized for 3–5 days; in case of complex patients, 3 infusion pumps operates for 10 days and then 1 pump for 5 days continuously. Also 2 pumps for 14 days will complement the injection for only 2 hours a day. If they are standard patients, they require 3 pumps for 5 continuous days and one pump for 5 intermittent days, activated only for 1 hour per day. It worth noting that Artemis platform does not intend to control and/or automatically calculate drug dosing in infusion pumps.

As Fig.2 suggests, McMaster NICU can be modeled as a single heterogeneous finite queue with multiple service facilities which are working in saturation regime. Each type of patients has distinct characteristics in terms of service duration and number of algorithms. Algorithms are also different in terms of required computational resources. Each class of patients needs a different number of infusion pumps for various length of stay continuously or discreetly. The McMaster NICU receives more admission requests that it has space for so that the NICU is always full and there is rarely any vacant bed - and this assumption has been used for the model.

In order to characterize medical algorithms, we use the real data obtained from a previous pilot deployment at SickKids Hospital NICU. Provided that, we are able to model the infrastructure at McMaster and predict the amount of computational capacity, storage and memory that are required to support a reliable real-time monitoring followed by storing all relevant data for further analysis. (The clinical data provided in this paper is from expert opinion based from summarized data about the NICU stays and as a result research ethics board approval was not deemed required for this research.)

Algorithms for the Artemis platform are developed either using data mining techniques or identifying patterns described in the medical literature that have not previously been detectable using automated methods. These algorithms are validated in robust clinical trials before being used to provide decision support for clinicians. Examples of this include our research using Artemis for late onset neonatal sepsis [22], apnoea of prematurity [3], [23], premature infant pain [24], anemia of prematurity [25] and sleep wake cycling detection [26] in neonates. For example, the clinical rule “If a pause in breathing occurs for greater than 20 seconds, or a pause in breathing that is associated with a change in heart rate, or blood oxygen saturations happens”, then a reportable condition of apnea is present [3].

V. ANALYTICAL MODELING OF THE PROPOSED METHOD

We use Kendall notation to describe the characteristics of the queuing systems that we focus on. The notation is of the form “A/S/s/c/p/D” in which A stands for the description of the arrival process; S stands for the service time distribution, s stands for the number of servers in the system and can be any integer equal or larger than 1; c stands for the capacity of the queue. If this argument is missing, then, by default, the queue

capacity is infinity; p stands for the system population. If this argument is missing then, by default, the system population is infinity; D stands for the queuing discipline, which can be FIFO, LIFO, or any other queuing discipline. If this argument is missing, then, by default, the queuing discipline is FIFO.

We model the Artemis Cloud platform as an $M/G/m/m$ queuing system which (M stands for Markovian, corresponds to homogeneous point process) indicates that the inter-arrival time of patients arrival is exponentially distributed with the mean value of λ , while patients' resident time at NICU are independently and identically distributed random variables that follow a general distribution with mean value of μ . The system under consideration contains m servers (i.e., bed spaces) that renders service in order of patients arrivals (First-In-First-Serve, FIFS). The capacity of system is m which means there is no extra room for queuing patients. As the population size of newborns is relatively high while the probability that a given newborn baby to be preterm is relatively small, the arrival process can be modeled as a Markovian process [20]. By such a modeling we are not only able to characterize current McMaster NICU but also predict the required resources in case of expansion in the future.

Since there is no indication either in academia or industry to assume any well-known probability distribution for duration of patients' residence in NICU, we consider a generally distributed random variable for the patient resident time in NICU. One possible scenario is to consider the hospitalization of each type of patient at McMaster' NICU as an exponentially distributed random variable with distinct mean value. Therefore, the overall visiting time for all patients will be a five-stage hyper-exponentially (HE) distributed random variable. We assume this scenario for numerical validation. Assuming that, the mean value of hospitalization is [20]:

$$E[x] = \int_{-\infty}^{\infty} xf(x)dx = \sum_{i=1}^n p_i \int_0^{\infty} x\lambda_i e^{-\lambda_i x} dx = \sum_{i=1}^n \frac{q_i}{\lambda_i} \quad (1)$$

in which q_i and $1/\lambda_i$ are the probability of being patient type i and corresponding mean value of residence time in NICU, respectively. Thus, the queuing system that we need to solve and obtain the performance metrics is of type $M/HE/m/m$. Characterizing multiple queuing systems with non-exponential distributed service time is not exactly tractable [20], [27]; however, since $M/HE/m/m$ queuing system has no extra capacity than service facilities, it works exactly the same as $M/M/m/m$ queuing systems [20], for which the steady-state probabilities (i.e., the probability of having k patients in the NICU in equilibrium) is given by

$$p_k = \frac{\frac{\rho^k}{k!}}{\sum_{n=0}^m \frac{\rho^n}{n!}} \quad (2)$$

in which $\rho = \lambda/m\mu$ and m is the number of bed spaces. Traffic intensity (aka offered load) is defined as arrival rate over service rate. In this model, blocking refers to when an admission request to McMaster's results in the patient being

redirected to other hospitals. Blocking probability can be obtained as:

$$P_b = \frac{\frac{\rho^m}{m!}}{\sum_{n=0}^m \frac{\rho^n}{n!}} \quad (3)$$

The probability generating function (PGF) will be:

$$P(z) = \sum_{k=0}^m p_k z^k \quad (4)$$

And the effective patients arrival rate (i.e., the rate of patients who can get through the NICU) is

$$\lambda_e = \lambda(1 - P_b)$$

Now, we can obtain the desired performance metrics. The mean number of patients in NICU is the first derivative of $P(z)$ when $z = 1$.

$$\bar{p} = P'(z)|_{z=1} \quad (5)$$

By Little's law [20], the mean patient residence time is

$$\bar{t} = \frac{\bar{p}}{\lambda_e} \quad (6)$$

VI. NUMERICAL RESULTS

The analytical model presented above has been implemented in Maple 18 [28] in order to obtain the numerical results. First we characterize the performance metrics for the current configuration of Artemis Cloud at McMaster which was described in section IV. Table 3 shows the amount of data collected from each NICU bed during 24 hours in Megabytes [9]. Artemis collects all these physiological data regardless of patients' type for the sake of completeness of archival. Table 4 shows the obtained performance metrics and important exogenous parameters for the current capacity.

The average length of stay for patients is 42 days and each patient requires 9 algorithms on average in Streams.

TABLE 3. Type and amount of physiological data acquired by Artemis Cloud at SickKids: for one patient during 24 hours.

Type	Description	Amount (MB)
HR	Heart Rate	5
SpO ₂	Blood Oxygen Saturation	5
SBP	Systolic Blood Pressure	5
DBP	Diastolic Blood Pressure	5
MBP	Mean Blood Pressure	5
ADBP	Arterial DPB	5
AMBP	Arterial MBP	5
ASBP	Arterial SBP	5
RR	Respiratory Rate	5
IRW	Impedance Respiratory Waveform	50
CO ₂	CO ₂ Waveform	50
PLETH	Plethysmography Waveform	50
ECG	Electrocardiography	500
TOTAL	Size of Raw Data	≈ 700

TABLE 4. Configuration parameters and performance metrics for current capacity of McMaster NICU.

Parameter	Value (unit)
No. of beds in NICU	47
Mean patient arrival rate (patient/day)	2
Mean no. of patients in NICU	45.8
Mean length of stay for patients (days)	41.5
Mean no of algorithms for 1 patient (per day) – [9]	9
No. of all running algorithms on Streams (per day)	423
NICU’s Service rate (patient/day)	0.024
Blocking probability	0.448
Mean value of memory per algorithm – [9]	110 (MB)
Req. memory on Streams cluster	41.3 (GB)
Req. storage for a patient’s data (per day) – [9]	700 (MB)
Req. storage for a pump data (per hour)	360 (MB)
Req. storage on BigInsights cluster (per year)	40 (TB)

Each algorithm is approximately consuming 110 MB of memory which indicates the requirement of at least 41 GB of memory for the stream computing cluster. Note that this amount of memory is just for application hosts and the management hosts require at least 2 more GBs of memory. As can be seen, the amount of minimum storage for Hadoop cluster in order to only support the accommodation of raw physiological and infusion pump data for one year is 40 Terabytes. Depending on data schema design on BigInsights cluster, additional storage might be required for the metadata. Moreover, the storage required for non-physiological data such as patient information, lab results and other related medical data should be added on top of these calculations.

Fig. 3 shows the amount of storage for BigInsights cluster, for 20, 47, 80, 110, 140 and 170 beds in the NICU.

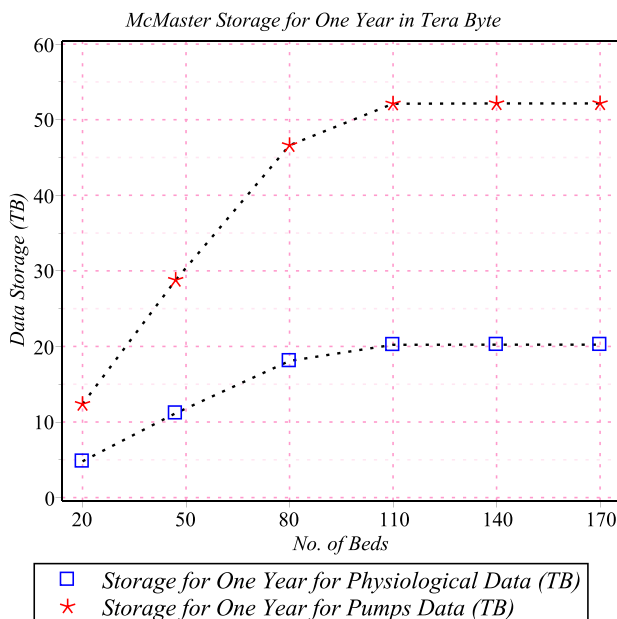


FIGURE 3. Required storage for BigInsights cluster for different configurations.

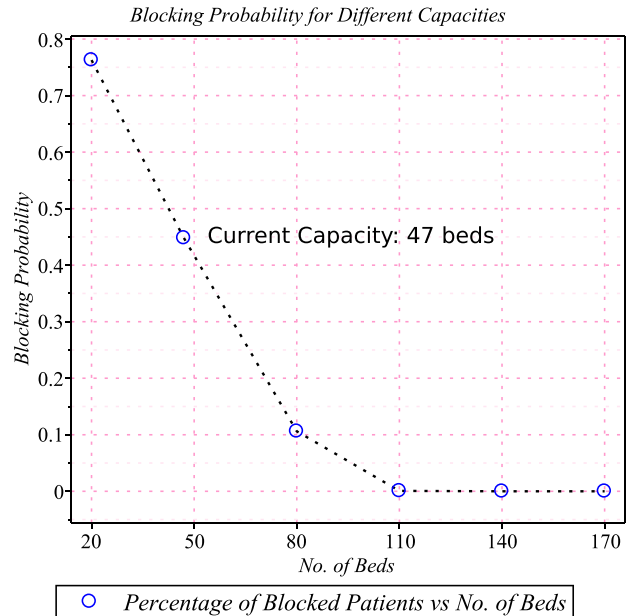


FIGURE 4. Blocking probability for different configurations.

These capacities are as sample extensions of the NICU in the future. Note that this amount is only for raw physiological and infusion pump data acquired from NICU. As can be seen, the amount of storage increases linearly with respect to NICU capacity up to 80 beds. However, afterward, since the traffic intensity, which is the ratio of arrival rate to service rate, gets decreased, the amount of required storage does not increase sharply. Also, as can be noticed for more than 110 beds the amount of required storage remains unchanged which indicates the departure rate of patients is more than patients arrival rate. In other words, for one year, 40 TB of storage is sufficient for the McMasters NICU regardless of NICU’s capacity (i.e., the number of bed spaces).

We are also interested in studying the number of patients that get blocked, i.e. redirected to another NICU, due to the capacity limitations of the NICU of interest. To this end, we characterize the blocking probability for the NICU with the capacity of 30 to 170 beds. As can be seen (Fig. 4), for the current capacity of McMaster’s NICU (i.e., 47 beds) %44.8 of patients get blocked. However, by increasing the capacity to 110 beds, the blocking will be less than a percent.

We also investigated the amount of memory and computation power for the stream computing cluster for different configurations. Fig. 5 shows the trend of required memory and number of CPU cores with respect to number of beds. Up to 110 beds there is a linear dependency between the required memory and capacity; however, as can be seen, results show that 75 GB of memory will suffice for the Streams cluster based on these arrival and departure rates. Our calculation for computation power is based on the standard CPU cores, i.e., 2.00GHz core, on IBM Softlayer [29] and our experiments that revealed that for each group of 20 algorithms we need a dedicate standard CPU core. As can be seen, the

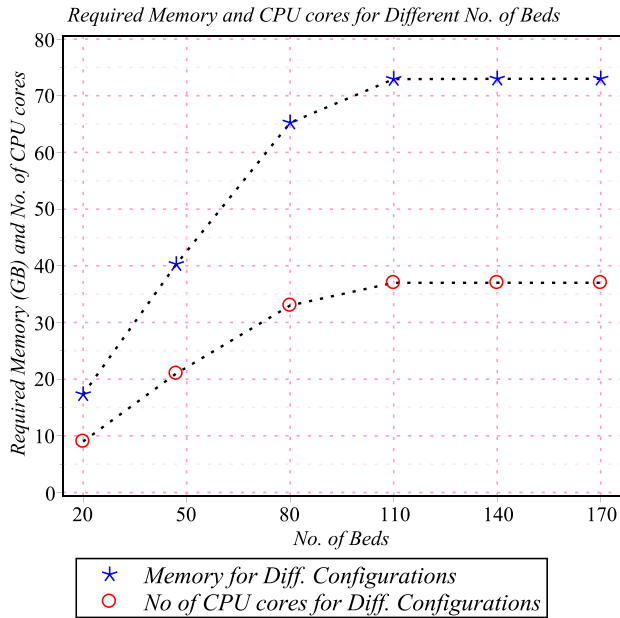


FIGURE 5. Required memory and cpu cores for Streams cluster.

trend for computation power is almost similar to memory which explained above. Therefore a cluster comprising of 40 cpu cores can handle the required computation for McMaster NICU. We shall repeat the fact that this amount of memory and computation power are just for application hosts and depending on the deployment of management servers, extra resources might be needed.

VII. CONCLUSION

We have described and modeled the Artemis cloud deployment at McMaster’s hospital. In light of the proposed architecture and patient journey, the corresponding analytical model has been designed and developed. Using the performance model, important performance metrics such as mean number of patients in NICU, mean patient residence time, mean number of required medical algorithms and blocking probability were characterized and discussed. Based on our pilot project at SickKids, we identified the amount of required storage, memory and computation power for analytics and real time components respectively. We obtained interested performance indicators and design parameters for different configurations. Provided that, capacity planning and what-if analysis can be attainable for big data growth introduced by extension of the NICU at McMaster hospital.

Analytical modeling described in this paper is generalizable to other NICUs. We only need to identify the capacity (i.e., no. of beds), types of patients, statistics of infusion pumps and arrival rate of patients. This approach can be applied to other ICUs beyond the NICU such as the pediatric and adult ICUs with the new profiling information for patient populations within those ICUs. In the long run, this work supports our implementation of the expanded Artemis Cloud as a commercial offering to healthcare facilities in Canada

and worldwide to provide a cloud computing service to critical care.

Please note that medico-legal aspects of the use Big Data techniques to acquire, process and store high frequency physiological data and its impact on assessment of liability are beyond the scope of this paper. Our current deployments are all researched based and as such do not impact liability of clinicians as the data is not available to the public and is provided by parents for the sole use of research. In the future our architecture enables individual healthcare facilities to determine the data retention and use policies to support their needs.

REFERENCES

- [1] C. McGregor, “Big data in neonatal intensive care,” *Computer*, vol. 46, no. 6, pp. 54–59, Jun. 2013.
- [2] M. S. Kramer *et al.*, “Secular trends in preterm birth: A hospital-based cohort study,” *J. Amer. Med. Assoc.*, vol. 280, no. 21, pp. 1849–1854, Dec. 1998.
- [3] A. Thommandram, J. M. Eklund, C. McGregor, J. E. Pugh, and A. G. James, “A rule-based temporal analysis method for online health analytics and its application for real-time detection of neonatal spells,” in *Proc. IEEE Int. Congr. Big Data (BIGDATA CONGRESS)*, Jun. 2014, pp. 470–477.
- [4] T. De Smet, M. M. Struys, M. M. Neckebroek, K. Van den Hauwe, S. Bonte, and E. P. Mortier, “The accuracy and clinical feasibility of a new Bayesian-based closed-loop control system for propofol administration using the bispectral index as a controlled variable,” *Anesthesia Analgesia*, vol. 107, no. 4, pp. 1200–1210, 2008.
- [5] T. M. Hemmerling, S. Charabati, C. Zaouter, C. Minardi, and P. A. Mathieu, “A randomized controlled trial demonstrates that a novel closed-loop propofol system performs better hypnosis control than manual administration,” *Can. J. Anesthesia*, vol. 57, no. 8, pp. 725–735, 2010.
- [6] A. A. Flower, J. R. Moorman, D. E. Lake, and J. B. Delos, “Periodic heart rate decelerations in premature infants,” *Experim. Biol. Med.*, vol. 235, no. 4, pp. 531–538, 2010.
- [7] M. Blount *et al.*, “Real-time analysis for intensive care: Development and deployment of the artemis analytic system,” *IEEE Eng. Med. Biol. Mag.*, vol. 29, no. 2, pp. 110–118, Mar./Apr. 2010.
- [8] J. Cirelli, C. McGregor, B. Graydon, and A. James, “Analysis of continuous oxygen saturation data for accurate representation of retinal exposure to oxygen in the preterm infant,” *Stud. Health Technol. Informat.*, vol. 183, pp. 126–131, Feb. 2013.
- [9] H. Khazaei, C. McGregor, M. Eklund, K. El-Khatib, and A. Thommandram, “Toward a big data healthcare analytics system: A mathematical modeling perspective,” in *Proc. IEEE 10th World Congr. Services (DSS)*, Jun. 2014, pp. 208–215.
- [10] Q. Zhang, C. Pang, S. McBride, D. Hansen, C. Cheung, and M. Steyn, “Towards health data stream analytics,” in *Proc. IEEE/ICME Int. Conf. Complex Med. Eng. (CME)*, Jul. 2010, pp. 282–287.
- [11] N. Bressan, C. McGregor, K. Smith, L. Lecce, and A. James, “Heart rate variability as an indicator for morphine pharmacokinetics and pharmacodynamics in critically ill newborn infants,” in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2014, pp. 5719–5722.
- [12] N. Bressan, C. McGregor, and A. James, “Contextualizing complex high volume physiological and drug data in the neonatal intensive care unit,” in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBS)*, Aug. 2015.
- [13] C. McGregor, “System, method and computer program for multi-dimensional temporal data mining,” U.S. Patent 8583686, Nov. 12, 2013.
- [14] A. Iosup, S. Ostermann, M. N. Yigitbasi, R. Prodan, T. Fahringer, and D. H. Epema, “Performance analysis of cloud computing services for many-tasks scientific computing,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 6, pp. 931–945, Jun. 2011.
- [15] E. Deelman, G. Singh, M. Livny, B. Berriman, and J. Good, “The cost of doing science on the cloud: The montage example,” in *Proc. Int. Conf. High Perform. Comput. Netw. Storage Anal.*, Nov./Feb. 2008, pp. 1–12.

[16] L. Wang, J. Zhan, W. Shi, Y. Liang, and L. Yuan, "In cloud, do MTC or HTC service providers benefit from the economies of scale?" in *Proc. 2nd Workshop Many-Task Comput. Grids Supercomput. (MTAGS)*, vol. 2, Dec. 2010, Art. ID 7.

[17] S. Alam et al., "Early evaluation of IBM BlueGene/P," in *Proc. ACM/IEEE SC*, May 2008, Art. ID 23.

[18] E. Walker, "Benchmarking Amazon EC2 for high-performance scientific computing," *LOGIN*, vol. 33, no. 5, pp. 18–23, Oct. 2008.

[19] S. Saini, D. Talcott, D. Jespersen, J. Djomehri, H. Jin, and R. Biswas, "Scientific application-based performance comparison of SGI Altix 4700, IBM POWER5+, and SGI ICE 8200 supercomputers," in *Proc. ACM/IEEE Conf. SC*, Dec. 2008, Art. ID 7.

[20] H. Khazaei, J. Mišić, and V. B. Mišić, "Performance analysis of cloud computing centers using $M/G/m/m+r$ queueing systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 5, pp. 936–943, May 2012.

[21] G. Hayes, H. Khazaei, K. El-Khatib, C. McGregor, and M. Eklund, "Design and analytical model of a platform-as-a-service cloud for health-care," *J. Internet Technol.*, vol. 16, no. 1, pp. 139–150, 2014.

[22] C. McGregor, C. Catley, J. Padbury, and A. James, "Late onset neonatal sepsis detection in newborn infants via multiple physiological streams," *J. Critical Care*, vol. 28, no. 1, pp. e11–e12, Feb. 2013.

[23] A. Thommandram, J. M. Eklund, and C. McGregor, "Detection of apnoea from respiratory time series data using clinically recognizable features and KNN classification," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 5013–5016.

[24] T. Naik, N. Bressan, A. James, and C. McGregor, "Design of temporal analysis for a novel premature infant pain profile using artemis," *J. Critical Care*, vol. 28, no. 1, p. e4, 2013.

[25] J. E. Pugh, A. Keir, C. McGregor, and A. G. James, "The impact of routine blood transfusion on heart rate variability in premature infants," in *Proc. AMIA*, 2013, p. 1.

[26] J. M. Eklund et al., "Automated sleep-wake detection in neonates from cerebral function monitor signals," in *Proc. IEEE 27th Int. Symp. Comput.-Based Med. Syst. (CBMS)*, May 2014, pp. 22–27.

[27] H. Khazaei, J. Mišić, and V. B. Mišić, "Modeling of cloud computing centers using $M/G/m$ queues," in *Proc. 1st Int. Workshop Data Center Perform.*, Mar. 2011, pp. 87–92.

[28] Maplesoft, Inc. *Maple 18*. [Online]. Available: <http://www.maplesoft.com>, accessed Mar. 2014.

[29] IBM. *Soflayer, an IBM Company*. [Online]. Available: <http://www.soflayer.com>, accessed Sep. 2014.



NADJA MENCH-BRESSAN was born in Brazil in 1978. She received the B.A. degree in automation from the University of Caxias do Sul, RS, Brazil, in 2005, and the M.Sc. degree in automation, instrumentation and control and the Ph.D. degree in biomedical engineering from the Faculty of Engineering, University of Porto, Portugal, in 2007 and 2011, respectively.

She joined the Department of Business and Information Technology, University of Ontario

Institute of Technology, as a Post-Doctoral Fellow at the Health Informatics Research (HIR) laboratory co-ordinated by Prof. Carolyn McGregor, in 2011. She was also involved in nosocomial infection, apnoea, and intraventricular haemorrhage studies. Her research topic at HIR was physiological stream processing in real-time and retrospective analysis to support drug delivery control in Neonatal Intensive Care Units (NICUs) and the critical parameters of its application.

She is currently a Research Fellow with the Department of Neonatology, Hospital for Sick Children, Toronto, Canada. She continues her research with physiological stream processing in real-time and the impact of drug titration in the patients of the NICU.



CAROLYN MCGREGOR received the B.A.S. (Hons.) degree in computer science, and the Ph.D. degree in computer science from the University of Technology Sydney.

She is the Canada Research Chair in Health Informatics and Professor with the University of Ontario Institute of Technology. She has led pioneering research on big data analytics, stream computing, event stream processing, temporal data stream data mining, business process modeling,

and cloud computing. She now progresses this research within the context of critical care medicine, mental health, astronaut health, and military and civilian tactical training.

Prof. McGregor has a track record of leadership in Health Informatics across, research, teaching, university governance, and service to the profession. She received the B.A.S. (Hons.) degree in computer science, and the Ph.D. degree in computer science from the University of Technology Sydney. She is an International Leading Researcher in the area of critical care health informatics and in particular, neonatal health informatics for which she has specialized for over 15 years. She has authored over 100-refereed publications, holds three patents, and has established two startup companies.



JAMES EDWARD PUGH received the degree from the St Marys Hospital Medical School, London, U.K., in 2003. He completed six years of post-graduate medical education in U.K., before immigrating to Canada in 2009 to complete his pediatric residency training with The Hospital for Sick Children. Upon finishing his residency, he continued with the University of Toronto to pursue a fellowship in neonatal and perinatal medicine, which he completed in 2012. He is currently an

Assistant Clinical Professor with the Department of Pediatrics' Division of Neonatology.

His research interests include the development of information/communication technology in medicine. He is currently enrolled in a master's program in health informatics with the University of Ontario Institute of Technology. Specifically, he is interested in real-time computer assisted decision support and early warning systems. For his master's thesis, he is developing and validating a novel computer algorithm to detect and classify neonatal spells in real-time.

In addition to his passion for technology, he is reputed for his patient care. During his neonatology fellowship, he received the University of Toronto's 2011 Audrey-Tan Day Humanitarian Award and the 2012 SickKids Student Humanitarian Award in recognition of his compassion and humanitarianism in providing care to children and their families.



HAMZEH KHAZAEI was a Research Scientist with the IBM Canada Research and Development Centre for two years. He was working on Artemis project, in which real-time and retrospective analytics, infrastructure scalability, capacity modeling, and performance analysis were among his focus of research. He is currently a Post-Doctoral Researcher with the Center of Excellence for Research in Adaptive Systems, York University. His current research includes big data management

for smart cities, adaptive systems, and NoSQL data stores.

He received the bachelor's and master's degrees in computer science from Amirkabir University of Technology, a.k.a Tehran Polytechnic, Iran, and the Ph.D. degree in computer science from the University of Manitoba. He was involved in research on performance and availability modeling of cloud computing centers.