

# QoS Quorum-Constrained Resource Management in Wireless Grid

Chan-Hyun Youn<sup>1</sup>, Byungsang Kim<sup>1</sup>, Dong Su Nam<sup>1</sup>, Eung-Suk An<sup>1</sup>,  
Bong-Hwan Lee<sup>2</sup>, Eun Bo Shim<sup>3</sup>, and Gari Clifford<sup>4</sup>

<sup>1</sup> School of Engineering, Information and Communications University  
119 MunjiRo, Yousong-gu, Daejeon 305-732, Korea  
{chyou, bskim, dsnam, sunyan}@icu.ac.kr

<sup>2</sup> Dept. of Information and Communications Engineering, Daejeon University  
96-3 Yongun-dong, Dong-gu, Daejeon 300-716, Korea  
blee@dju.ac.kr

<sup>3</sup> Dept. of Mechanical Engineering, Kwangwon National University  
192-1 Hyoja 2-dong, Chunchon 200-701, Kwangwon-do, Korea  
ebshim@kangwon.ac.kr

<sup>4</sup> Harvard-MIT Division of Health Science Technology, MIT  
Cambridge, MA 02139, USA  
gari@mit.edu

**Abstract.** In wireless Grid computing environment, end-to-end Quality of Service (QoS) could be very complex, and this highlights the increasing requirement for the management of QoS itself. Policy quorum-based management offers a more flexible, customizable management solution that allows controlled elements to be configured or scheduled on the fly, for a specific requirement tailored for a customer. This paper presents a QoS guaranteed management scheme, Policy Quorum Resource Management (PQRM), which facilitates the reliable scheduling of the wireless Grid system with dynamic resource management architecture aimed at fulfilling the QoS requirements. Experimental results show the proposed PQRM with resource reconfiguration scheme improves both performance and stability, which is suitable for a wireless Grid services.

## 1 Introduction

Grid computing provides widespread dynamic, flexible and coordinated sharing of geographically distributed heterogeneous networked resources among dynamic user groups. Wireless communications is a rapidly evolving and promising sector of the communications arena, and even in a challenging time for the telecommunications industry, represents a significant development opportunity for companies and organizations in creating a global market. The increasing reliance on wireless networks for information exchange makes it critical to maintain reliable and secure communications even in the instances of a component failure or security breach. In wireless networks, mobile application systems continuously join and leave the network and change locations with the resulting mobility impacting

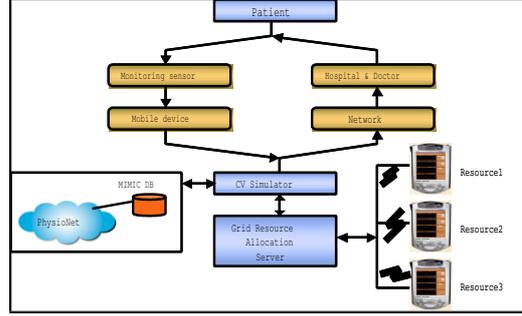
the degree of survivability, security and reliability of the communication. Our research focuses on policy based dynamic resource allocation and management for Quality of Service (QoS) guaranteed applications in a wireless Grid [1], [2].

Our resource management policy will need to effectively map the pre-defined QoS requirements to the actual resources on the wireless network. Ideally some of these QoS requirements will be defined at the local system level to minimize protocols utilized at the middleware system, however, we need to address the problem of combining different kinds of QoS from the multiple resources available in the Grid and among multiple Grids. We present a dynamic resource management architecture aimed at the fulfillment of the above requirements in wireless Grid services. We discuss the efficiency of the proposed reconfiguration policy on the experimental Grid testbed. We imagine that these kinds of reliable, secure and QoS guaranteed Grid services would be required for future mobile health care applications.

## 2 The Management for Cardiovascular Simulator in Wireless Grid

A general architecture for future wireless access networks involves a hybrid of technologies. Advances in patient care monitoring have allowed physicians to track an Intensive Care Unit patient's physiological state more closely and with greater accuracy. Modern computerized clinical information systems can take information from multiple data sources and store it in a central location. The Multi-parameter Intelligent Monitoring for Intensive Care (MIMIC) II database [3] takes advantage of the data collection systems installed at a partner hospital to collect large numbers of real patient records. The MIMIC II Relational Database Management System (RDBMS) is used to administer this database and create table definitions. Data such as nursing notes, medications, fluid input and output, updates to patient charts, lab results, CareVue data etc., can be downloaded from the hospital's Infrastructure Support Mart (ISM) and entered into the MIMIC II database. Waveforms such as electrocardiograms (ECGs), blood pressure, and pulse oximetry are stored in a separate waveform collection system in a separate database at the hospital, which is the waveform counterpart to MIMIC II.

Fig. 1 represents a computational physiological scenario using a wireless Grid. Cardiovascular (CV) data measured at home are transferred to a "patient monitoring manager" to assess those data. In this manager server, the monitored data is compared with the pathological data set stored in the MIMIC II database of PhysioNet [4]. If the signal is determined to be indicative of a pathological state, the cardiovascular information is used to determine the parameters of a CV model. CV hemodynamic variables under a specific physiological hypothesis are simulated and compared with the abnormal signal. In this simulation stage, Grid computing is utilized to speed up the analysis.



**Fig. 1.** Data flow for a mobile patient monitoring application using a CV simulation distributed on the wireless Grid interacting with a CV database

### 3 Quorum-based resource management model

#### 3.1 Definition of Quorum model

Most Grid applications specify the type of required resources such as various types of processing, storage, and communication requirements among such elements. To guarantee the QoS in user's application, the Grid manager should assign the resources to each application. The Quorum based management system will define the QoS vector to each application from the profile at input. We define two items; the one is system QoS vector and the other is network QoS vector. Each service requester must specify its QoS requirements for the Quorum Manager (QM) in terms of the minimum QoS properties for both the node system and network resources. We define the resource Quorum QR that represents current status of the resource. Each resource has its resource status vector which is represented both invariable and variable description [5]. System resource can take the processor specification or memory size as invariable description and processor utilization or available memory space as variable description and end-to-end available bandwidth, delay, data loss rate as variable specification.

$$Q_R = \left\{ \langle \vec{\theta}_i, \vec{\theta}_{jk} \rangle \mid i, j, k = 1, \dots, n \right\}, \quad (1)$$

where  $\vec{\theta}_i$  denotes the current available resource level of the system resource  $i$  and  $\vec{\theta}_{jk}$  represents the current available resource level of the network between system resource  $j$  and  $k$ .

We assume the system has to admit and allocate resources for the set  $A = \{a^1, \dots, a^k, \dots, a^m\}$  of applications. An application also represented by undirected graph with tasks and their communication relation. A required QoS level represents the vector of the resource description and has the range between minimum and maximum requirement. We denote them with  $q$  and  $Q$ , respectively.

We define the QoS-Quorum,  $\mathbf{Q}_A$ , which represents the required quality level for an application set  $A$ .

$$Q_A = \left\{ \left\langle \left[ \vec{q}_i^k, \vec{Q}_i^k \right], \left[ \vec{q}_{ij}^{kl}, \vec{Q}_{ij}^{kl} \right] \right\rangle \mid i \neq j, i, j = 1, \dots, \mu, k \neq l, k, l = 1, \dots, m \right\}, \quad (2)$$

where  $\vec{q}_i^k$  and  $\vec{Q}_i^k$  denote the minimum and maximum QoS level required for task  $k$  on the system resource  $i$ , respectively.  $\vec{q}_{ij}^{kl}$  and  $\vec{Q}_{ij}^{kl}$  represent the minimum and maximum QoS level required for communicating the task  $k$  in system resource  $i$  and task  $l$  in system resource  $j$ , respectively.

A resource universe  $R = \{r_1, \dots, r_n\}$  assumes a collection of resources that can be taken in the administration domain. A resource,  $r = \langle S, N \rangle$ , can be represented as undirected graph with system,  $S$  and their communication networks,  $N$ . To get the reliability in resource management, we define available resource Quorum,  $\mathbf{Q}_{AR}$ , which is a selected resource set which satisfying the QoS requirement from SLAs.

$$Q_{AR} = \{ \langle S_i, N_{ij} \rangle \subseteq R \mid \vec{q}_i^k \leq \vec{\theta}_i \leq \vec{Q}_i^k, \vec{q}_{ij}^{kl} \leq \vec{\theta}_{ij} \leq \vec{Q}_{ij}^{kl} \}, \quad (3)$$

where  $i, j = 1, \dots, n' \leq \mu$  and  $k, l = 1, \dots, m$ .  $\mathbf{Q}_{AR}$  is a set that satisfies a desired minimum QoS level of application. An available resource Quorum set is obtained from the policy module in the resource allocation server.

### 3.2 Resource configuration for resource scheduling

The Resource configuration for scheduling,  $\mathbf{F}(A, \mathbf{Q}_{AR})$ , is the mapping function;

$$F(A, Q_{AR}) = \{ \langle S_i^k, N_{ij}^{kl} \rangle \} = \{ \langle V^k, E^{kl} \rangle \xrightarrow{Q} \langle S_i, N_{ij} \rangle \}, \quad (4)$$

where  $i, j = 1, \dots, n' \leq u$  and  $k, l = 1, \dots, m$ .

$S_i^k(\vec{\theta}_s)$  or  $S_i^k(\vec{\theta}_n)$  is identified the resource topology from status of reflecting the resource quality. We intend to estimate current resource utilization of each application activity. Assuming that application represents its previous execution progress as an event, we can predict the current resource allocation sensitivity for the target application. On the network performance vector, end-to-end resource performance, such as bandwidth, delay, or data loss rate is peer-to-peer value. If a Grid resource management system should measure all the peer systems, it surely faces scalability problem and unrealistic. Thus, usually, in the initial step of the resource allocation, resource allocation and policy server generate the available resource Quorum in the network established by the resource brokers and each peer.

## 4 Policy Quorum based resource management in wireless Grids

In order to address the resource allocation problem, we propose a Policy Quorum-based Resource Management (PQRM) System that provides a reliable resource scheduling scheme in wireless Grid. A PQRM shown in Fig. 2 has the layered architecture involving a session admission controller with an application profile, and requested applications as inputs [6], [7].

The Session Admission Controller (SAC) gives inform to Quorum Manager (QM). The QM inspects available resources that satisfy minimum QoS requirements through MDS(Monitoring and Discovery Service). The Monitoring Manager (MM) obtains the information of Grid resources from the agent that queries the resources. The SLA is an important work for the SAC, which interacts with users and their requirement such as time or cost deadlines.

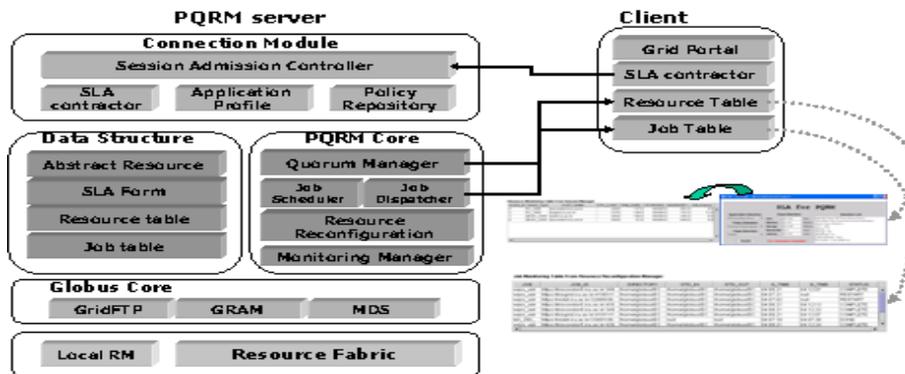


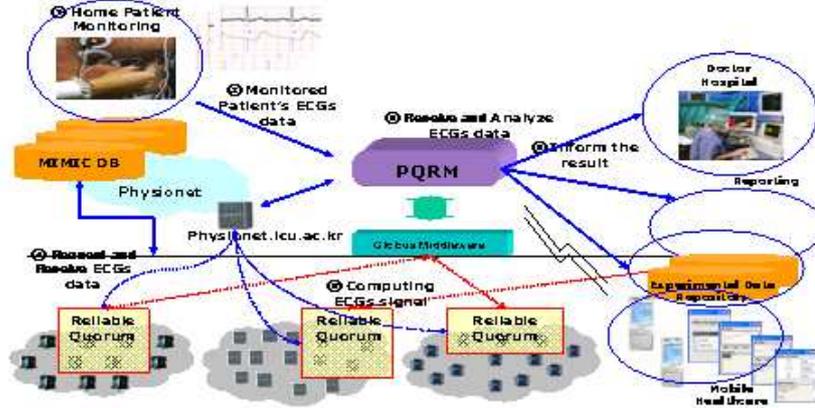
Fig. 2. Operational architecture of the PQRM system

We can obtain the initial resource configuration set as mapping from sub-jobs to available resource Quorum set: the initial resource configuration will satisfy the QoS requirement because it is subset of the  $Q_{AR}$  which is the set of establishment from the user's SLA. Failure of system or admission of the other application in the resource has influence on the variable QoS vector. Temporal instants  $T_i$ 's are event generation points of each application. In general, in order to measure the application performance, instrumentation technique is a useful method that recognizes the application behavior.

## 5 Experimental results and discussions

Fig. 3 represents a computational physiological scenario using Grid computing. In this figure, the human ECG is monitored and transmitted to the PhysioNet

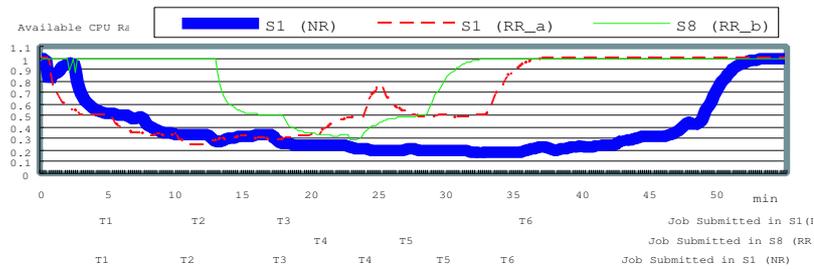
server. The monitored data is then delivered to a simulator via a mobile computer, where the system compares it with the pathological data set stored in MIMIC [3]. In case of an abnormal signal, identified through a nodal fit to data in MIMIC, the algorithm may provide pertinent advice to the monitored subject. The information can then be delivered to a medical doctor. Furthermore, personally optimized medication information can also be delivered to the patient client.



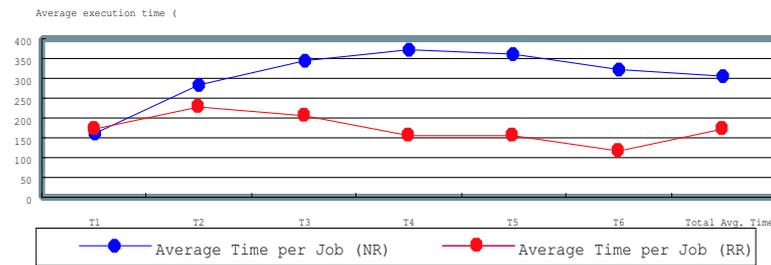
**Fig. 3.** Data flow for a patient monitoring application using an ECG computation distributed on the Grid interacting with an ECG database (MIMIC DB) via PhysioNet (a traditional client-server model)

In a wireless Grid infrastructure, all computation needed to simulate the monitored ECG signal will be executed on the available resources provided by a PQR system. For real-time processing, it is necessary to identify the resources available in the Grid infrastructure. In particular, in order to ascertain the reliability of the wireless environment, policy based Quorum will be considered in the context of resource management. If a user wishes to access his/her daily biometrics from a Grid server, it is necessary for some computational resources to remain online regardless of its location with respect to the VOs. The systems used in this experiment are eight in total, which consist of six systems in Korea and a PhysioNet server at MIT in the US. We assume that each user jobs J1, J2, J3, J4, J5 and J6 starts at time instants, T1, T2, T3, T4, T5, and T6, respectively. Each job is executed independently at each node. Each job is composed of 4 sub-jobs. Each sub-job calculates the entropy of the ECG signal that is transmitted from users. To identify the mobility-based resource management, we identified two types of experiments, e.g. a case of no resource reconfiguration after initial configuration and a case of resource reconfiguration. In case of no reconfiguration approach, the jobs were executed with resource topology determined at S1. The initial Quorum Set is {S1, S2, S3, S4, S5, S6}.

A non-reconfiguration method executes jobs sequentially at the system S1 and resource reconfiguration method performs at each temporal instant  $T_i$ , such as Quorum  $S1 = \{(J1, T1), (J2, T2), (J3, T3), (J4, T4), (J5, T5), (J6, T6)\}$ . On the other hand, user's mobility will invoke a policy-Quorum management module in PQRM to reconfigure the resource to guarantee the initial SLAs. A reconfiguration algorithm may change the current Quorum set through reshuffling the resources topology. Fig. 5 shows that the resource reconfiguration was triggered at  $T4$ . After triggering the resource reconfiguration, newly generated resource Quorum and resource re-allocation will be determined with  $S1 = \{J1, J2, J3, J6\}$ ,  $S7 = \{J4, J5\}$  at  $T4$ . The S8 is added in new resource quorum and J4 and J5 are moved from system S1 to S8. Fig. 4 represents resource status after reconfiguration. Fig. 5 depicts the performance comparison of the average execution time with and without a Resource Reconfiguration (RR) policy. Quorum management policy with resource configuration (in the bottom two bars) shows more stable execution time than with no reconfiguration management scheme.



**Fig. 4.** Comparison of resource status (available CPU ratio) in S1 and S8 after reconfiguration at  $T4$  (NR: No reconfiguration policy, RR: Resource Reconfiguration policy)



**Fig. 5.** Performance comparison of execution time with and without policy of resource reconfiguration (NR: No reconfiguration policy, RR: Resource Reconfiguration policy)

## 6 Conclusions

End-to-end QoS can be very complex in the wireless Grid computing environment. Policy based management offers a more flexible, customizable management solution that allows controlled elements to be configured or scheduled on the fly, for a specific requirement tailored for a customer. The proposed Quorum based resource management policy is suitable to effectively map the pre-defined QoS requirements to the actual resources on the wireless network. This paper has presented a QoS guaranteed management scheme that facilitates the reliable scheduling of the wireless Grid system with dynamic resource management architecture aimed at fulfilling the QoS requirements. We have discussed the performance and stability of the PQRM with resource reconfiguration scheme. An ECG analysis for mobile medical care has been presented, which would benefit from such a scheme.

**Acknowledgements.** This paper was supported in part by IITA (Institute of Information Technology Assessment) through ITRC project.

## References

1. L.W. McKnight, D. Anius and O. Uzuner, "Virtual Market in wireless Grid: Peering policy prospects", *TPRC2002 Research conference on Communication, Information and Internet Policy*, 2002.
2. F. Berman, G.C. Fox, A.J.G. Hey, *Grid Computing*, Wiley, 2003.
3. M. Saeed, C. Lieu, G. Raber, R.G. Mark, "MIMIC II: A massive temporal ICU patient database to support research in intelligent patient monitoring", *Computers in Cardiology*, pp641-644, 2002.
4. A. Golberger, L. Amaral, L. Glass, J.M. Hausdorff et al, "PhysioBank, PhysioToolkit, and PhysioNet :Component of a New Research Resource for Complex Physiologic Signals," *Circulation* 101 (23), June,2000.
5. L.J. Franken, B.R. Haverkort, The performability manager, *IEEE Network*, pp.24-32, Jan/Feb. 1994.
6. I. Cardei, S. Varadarajan, M. Pavan, M. Cardei, M. Min, "Resource Management for Ad-hoc wireless networks with cluster organization", *Journal of Cluster Computing in the Internet*, Kluwer Academic Publishers, Jan. 2004.
7. R.J. Alali, O.F. Rana, D.W. Walker, "G-QoS: Grid Service Discovery Using QoS Properties", *Computing and Informatics Journal*, Vo. 21, pp.363-382, 2002.