Priority-based QoS Extensions and IAM Improvements

Gyudong Park and Hyoek Jin Choi*

ADS&TR Institute – Command and Control Systems PMO, Agency for Defense Development, Seoul, Korea
E-mail: iobject@add.re.kr; mycult@add.re.kr
**Corresponding Author*

Received 16 June 2024; Accepted 26 June 2024

Abstract

The command and control system operates in a harsh and dynamic environment with limited resources and have a very high risk of failure or malfunction. In the case of military information systems, including the command and control system, the efficiency and effectiveness of system resource management are very important and required. Therefore, the application of a QoS-like approach is necessary to improve the operational effectiveness of all command and control system resources. However, supporting QoS at the entire command and control system level incurs additional costs and burdens for implementation and operation. This paper describes the necessity and possibility of collaboration with QoS and IAM (identity and access management) among the collaboration between core functions within the command and control system. This paper proposes an extended QoS approach to improve the operational effectiveness of the entire command and control system resources. As a result of this research, expanded concepts, structures, standards, and methods of collaboration between QoS and IAM are developed and presented, and their feasibility is demonstrated through prototype development and experiments.

Keywords: Access congestion, priority-based access control (PBAC).

Journal of Web Engineering, Vol. 23_6, 769–786.
doi: 10.13052/jwe1540-9589.2362

1 Introduction

A system that has limited resources, such as a command and control system, and is operated in a harsh and dynamic environment, has a very high risk of failure, malfunction, and load explosion, while rapid recovery, repair, and expansion are greatly limited. Thus, a situation in which a lot of users or applications compete for insufficient resources of the system may frequently occur, and thus, some resource requests may be delayed or rejected.

"The quality of the command and control system is measured by the recognition of environmental changes, the swiftness of recognition, and the appropriateness and timeliness of responses [1]." In other words, delay or refusal of resource requests for mission execution causes deterioration in the quality of command and control. When requests for resources, especially for mission-critical tasks, are delayed or denied by requests for resources that are less critical, the results can sometimes be catastrophic.

When resources are scarce, an approach that uses available resources more effectively is critical. To this end, insufficient resources must first be allocated to requests from users or applications for critical missions. In the communication network, QoS (quality of service) [2] is applied, which differentially allocates insufficient bandwidth according to the priority of traffic when traffic congestion occurs.

In the case of military information systems, including the command and control system, the efficiency and effectiveness of system resource management are essential because acquiring and maintaining sufficient resources are greatly limited due to SWaP (size, weight, and power) issues and poor operating environments. Therefore, the application of QoS, which has been proven over a long period in the network domain, is necessary to improve the operational effectiveness of the entire command and control system. In order to support QoS at the entire command and control system level, monitoring and control capabilities for all resource requests are required. However, this creates considerable additional costs and loads for implementation and operation.

This paper describes the necessity and possibility of the collaboration between QoS and IAM (identity and access management) among several possible collaborations between core functions within the command and control system [3]. IAM is a core function of system security and consists of identity management, authorization, authentication, and access control functions. The primary purpose of IAM is to guarantee access by authorized users and restrict access by unauthorized users [4]. To this end, IAM monitors and controls all traffic between users and the system.

This paper proposes an extended QoS approach to improve the operational effectiveness of the entire set of command and control system resources. As a result of this research, expanded concepts, structures, standards, and methods of the collaboration between QoS and IAM were developed and presented, and their feasibility was demonstrated through prototype development and experiments.

2 Related Works

QoS is already a mature concept and technology. Various QoS standards and technologies have been developed, such as IntServ, DiffServ, time-sensitive networking (TSN), deterministic networking (DetNet), and network slicing. However, QoS standards and technologies so far have limited scope of application to network bandwidth resources [5–7].

QoS differentiates traffic based on priority in situations of insufficient bandwidth. Therefore, it is crucial in QoS which priority criterion to apply and how. As a representative priority criterion for military communication networks, there are triple-metric priority criteria of performance, importance, and urgency [8]. This criterion is very suitable for military communication network QoS implementation that prioritizes traffic with higher importance and urgency and meets different performance requirements for each traffic type or service class. However, it is still mainly applied to military communication network QoS implementations [9, 10].

On the other hand, IAM technology has been continuously developed for a long time for system security, supporting multiple users. A lot of systems in the past adopted a perimeter-based security approach, and all users could freely access all resources within the system after authentication [11]. However, owners or administrators of the system resources want to be able to authorize and control user access to resources in a more sophisticated way.

IBAC (identity-based access control) [12] was developed to control access to specific resources. IBAC utilizes access control lists or matrices to link permission to access specific resources with user identities. However, the complexity and load of authorization may increase excessively due to the number of resources and users.

RBAC (role-based access control) [13] has been proposed to reduce authorization complexity and load. Authorization of RBAC sets the user's resource access permission as a connection through a role. However, as the scale of systems and users increases and access control requests from resource owners diversify, RBAC's authorization complexity and load also increase a lot.

However, the authorization methods up to RBAC are performed by establishing an explicit relationship between individual users and individual resources. Therefore, when updating information according to policy and environmental changes, all relationships between users and resources must be re-examined, resulting in high renewal costs.

ABAC (attribute based access control) [14] was proposed to improve these disadvantages. In ABAC, the authorization method is performed by creating, modifying, and deleting rules based on Boolean algebra using attributes of users and resources. ABAC can significantly reduce authorization's difficulty, complexity, and load by setting the relationship between users and resources through abstracted rules.

Most of the properties of users and resources are static properties that do not change over a long time. The result of access control that relies on static attribute-based authorization is always the same at anytime, anywhere. However, the need and demand for different access controls are increasing according to changes in the situation recently.

Furthermore, the ABAC improvement that reflects this is CBAC (context-based access control) [15]. CBAC additionally exploits the dynamic attributes of the situation or context, as well as the attributes of users and resources.

Until now, QoS and IAM have been developed as independent technologies in the communication network and security fields, respectively. However, in this paper, the application target of QoS is extended to the entire system resources. The purpose of IAM extends to priority-based system efficiency improvement in situations where system resources are scarce, and QoS and IAM are modified to work together to improve system efficiency.

Moreover, for this purpose, new collaboration concepts, structures, and methods between QoS and IAM were presented. Finally, in this paper, PBAC (priority based access control) and priority criteria, which are suitable for new access control methods of this study are proposed.

3 Proposed Concept

3.1 Current Concept

Figure 1 shows the concept of system resource access and use for users. First, the user logs into the system through IAM's authentication. Also, system

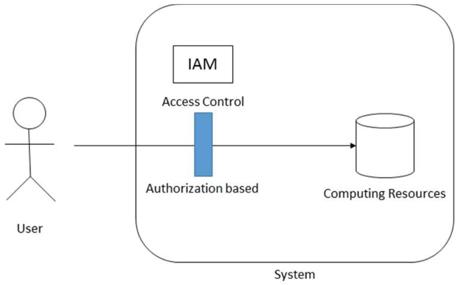


Figure 1 Resource access concept in the current framework.

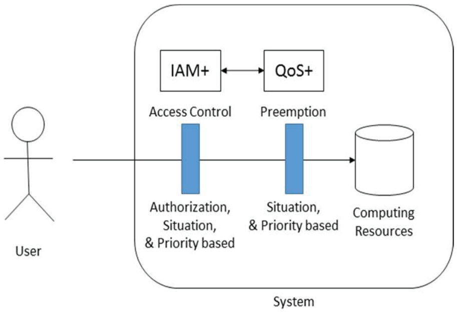

IAM+
QoS+
Access Control
Preemption
Authorization, Situation, & Priority based
Situation, & Priority based
Computing Resources
User
System

Figure 2 Extended access concept.

access requests from logged-in users are granted or denied based on authentication level and authorization through IAM access control. In addition, QoS is applied only to communication network bandwidth resources and not to other computing resources.

3.2 Extended Concept

This paper proposes an improved concept, as shown in Figure 2. This concept allows QoS to be applied across system resources, not just bandwidth resources. Even when computing resources are insufficient, the system differentiates requests for corresponding resources according to priority. Furthermore, this processing is performed by IAM, which can monitor and control all system resource access requests.

As shown in Figure 2, the extended concept expands and improves the existing QoS and IAM, and the expanded and improved QoS and IAM are named QoS+ and IAM+, respectively.

3.2.1 QoS+

In QoS+, existing QoS is utilized for network resources, and a priority-based preemption method is additionally applied to system resources. When a new resource request arrives in a state where system resources are insufficient, resources in use from an existing request having a lower priority are retrieved and allocated to the new request. If existing requests with lower priority perform transactions that cannot be stopped or stopped immediately, a new request is queued. Additionally, existing requests that were interrupted by preemption will wait until resources become available before being resumed. Therefore, QoS+ should minimize the additional load generated in this process. Depending on policies, QoS+ can be applied to different types and levels of physical or logical resources. Examples include CPU, memory, and VM (virtual machine).

3.2.2 IAM+

IAM's ability to monitor and control all traffic between users and systems is too powerful and valuable to be used only for system security. Accordingly, IAM+ increased system effectiveness by expanding and improving the existing system security-focused permission-based access control. Also, for this purpose, a new access control method, PBAC, was proposed. PBAC approves or denies resource requests based on priority according to system resource utilization. That is, when system resources are insufficient, a login of a new user or a request for new system resources may be restricted according to priority.

3.2.3 Priority Criteria

This paper applies the triple-metric priority criterion widely used in communication network QoS implementation to QoS+ and IAM+ implementation. This is because performance requirements for computing resources are different for different applications or types of applications, and may be of different importance to different users or tasks.

PBAC classifies applications into three types: real-time, interactive, and batch. Each type has different performance requirements and requires differentiated processing. In this paper, similarly to traffic engineering in communication network QoS, we proposed a method of processing resource requests by application by appropriately allocating computing resources by application type in advance. In addition, there may be non-pre-emptive applications in which resource allocation must always be guaranteed, such as control class traffic of communication network QoS. In the communication network QoS, the control traffic is allocated to EF (expedited forward), to which sufficient resources are allocated and processed so that traffic congestion does not occur. QoS+ also handles non-pre-emptive applications similarly.

In addition, importance is assigned to each user and application for importance-based differentiation processing. The importance of each user can be assigned through a request and approval process, and the importance of each application can be set in advance according to the policy of each system. Each can be managed through identity management. In addition, when system resources are insufficient, users and applications with higher importance are allocated system resources first.

Urgency relates to the time limit for processing or responding to an application. A response time limit of all real-time applications is "immediate," therefore, it's impossible to differentiate according to the urgency. However, interactive applications and batch applications may have different response or processing time limits for each request. Expired requests can be cancelled to avoid wasting system resources. Also, it is possible for a request that has a margin in response or processing time to give way to a request that does not.

4 Design and Implementation

To prove the feasibility of the proposed concept, QoS+, and IAM+ prototypes were designed and developed, and experiments were conducted. Additionally, the implementation scope of the prototype is limited to include real-time or interactive applications that require an immediate response and only consider the priority of importance.

4.1 Major Functions

Figure 3 shows the QoS+ and IAM+ prototypes' main functions and interfaces.

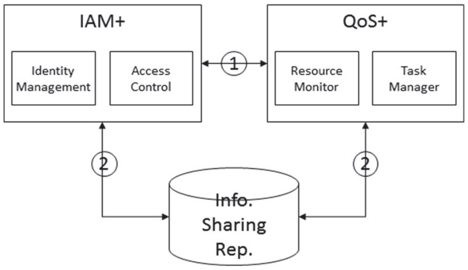


Figure 3 QoS+ and IAM+ prototypes' main functions and interfaces.

4.1.1 QoS+

QoS+ consists of two core functions. First, the resource monitor measures and manages the usage rate of computing resources in real-time and shares them with other functions. The task manager manages tasks using computing resources, requests information and performs a pre-emptive function according to priority when a new request arrives in a resource shortage situation.

4.1.2 IAM+

The IAM+ prototype was implemented based on KeyCloak, a leading open-source IAM solution. The KeyCloak's identity management and access control were improved, and the rest of its functionality was reused without modification. Identity management improved the function of [16] to additionally manage the importance and urgency of users and applications, and improved access control by applying PBAC.

4.1.3 Interface between IAM+ and QoS+

As shown in Figure 3, two interfaces are provided to share information between QoS+ and IAM+. One is an interface for real-time or synchronous information sharing, which is implemented as a function call between functions (1). The other uses an information-sharing repository as an interface for asynchronous information sharing (2). Identity management information is also shared with other functions, including QoS+, via the information-sharing repository (2). Additional functions or applications can collaborate with existing ones through information-sharing repositories.

QoS+ monitors system resource operating status in real-time and shares it with IAM+. IAM+ utilizes information about the operating state of system resources to determine and enforce priority-based access control levels. QoS+ also shares the user identification information with IAM when reclaiming system resources for low-priority users. If the system resources in use are retrieved, the user's system operation is also suspended. A common behavior for users in this case is to constantly retry, which can exacerbate the low system resource situation. To prevent this, IAM+ sends an appropriate message to users and restricts their access requests for some time.

4.2 Major Procedure

Figure 4 shows the system resource access procedure. A user's access to system resources proceeds according to the user's authentication and authority.

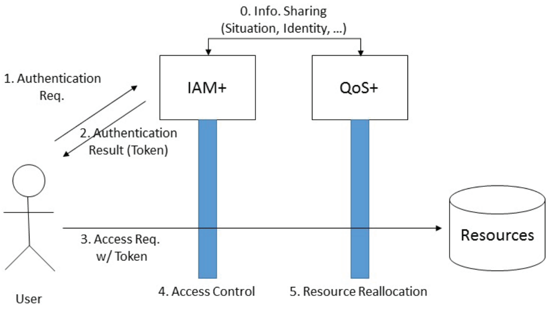


Figure 4 System resource access procedure.

```
Whether to proceed with resource preemption = FALSE
FOR # of resource free(for CPU, Memory)
    IF (current resource utilization + expected resource utilization > User resource threshold)
        Whether to proceed with resource preemption = TRUE
        Store resource threshold exceeded amount value.
    END IF
END FOR

RETURN Resource preemption progress, resource threshold exceeded amount
```

Figure 5 Resource preemption decision algorithm.

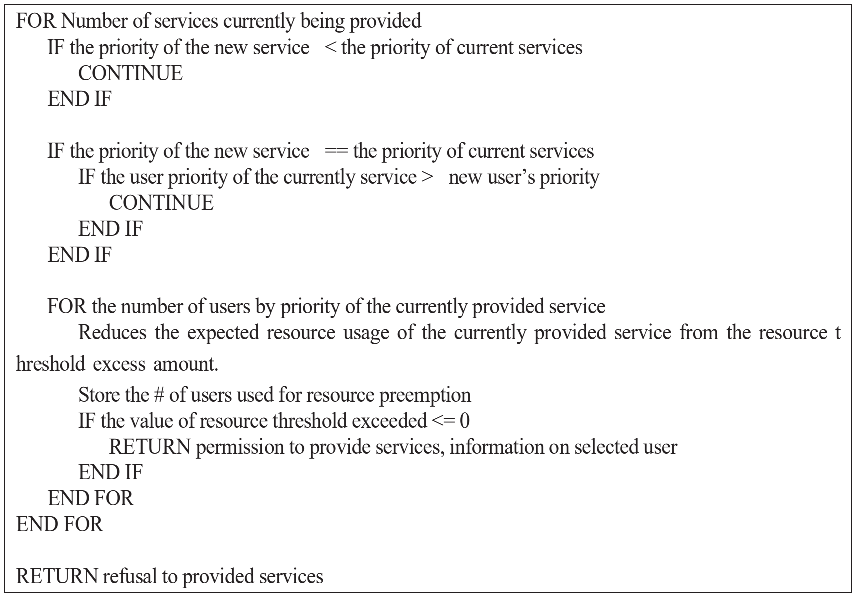
Most IAM solutions return results via tokens when user authentication is successful. One of the key ideas in this paper is to extend these tokens to include user-importance information and share it with IAM+ and QoS+.

4.3 Resource Retrieval Algorithm

Figure 5 describes the resource retrieval algorithm. In order to retrieve resources, the estimated resource usage information required for service operation is used as additional information, and this value is stored in the information sharing repository.

To determine whether resource preemption is necessary, the sum of the amount of resources used in the current system and the expected resource usage of the new service is calculated. If this sum does not exceed the user priority threshold, the service is provided without resource preemption. If the threshold value is exceeded, resource preemption is performed.

Figure 6 shows an algorithm for selecting a target for retrieving existing service resources when resource preemption is required. Verification is



```
FOR Number of services currently being provided
    IF the priority of the new service   < the priority of current services
        CONTINUE
    END IF

    IF the priority of the new service   == the priority of current services
        IF the user priority of the currently service >   new user's priority
            CONTINUE
        END IF
    END IF

    FOR the number of users by priority of the currently provided service
        Reduces the expected resource usage of the currently provided service from the resource t
hreshold excess amount.
        Store the # of users used for resource preemption
        IF the value of resource threshold exceeded <= 0
            RETURN permission to provide services, information on selected user
        END IF
    END FOR
END FOR

RETURN refusal to provided services
```

Figure 6 Screening algorithm to retrieve resources.

performed for all services in progress. All priorities are subject to resource retrieval, except where the existing service has higher priority than the new service and the case where the priority of the existing service user is higher than that of the new service user. Resource retrieval is based on the expected resource usage of the service and proceeds so that the resource excess is reduced by the expected resource usage of the service selected as the target for resource retrieval. If there are more retrievable resources, the new service is accepted.

5 Experiment

Figure 7 shows the experimental environment and primary procedures.

When a user logs in through the service, the service requests IAM+ to issue a token. At this time, PBAC, a new access control method, is performed in IAM+ to determine service permission or denial. If login is authorized, the user receives a token and requests a service the system provides. The service receiving the service request performs the request for PBAC through IAM+. If the corresponding request is granted, service permission is requested through QoS+, and resources are allocated according to priority.

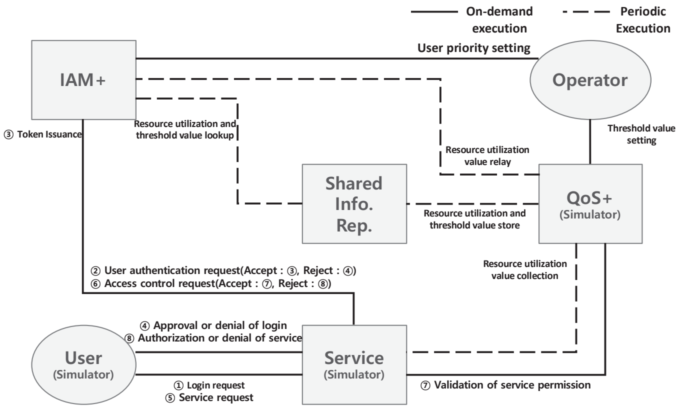

On-demand execution
Periodic Execution
IAM+
User priority setting
Operator
③ Token Issuance
Resource utilization and threshold value lookup
Resource utilization value relay
Threshold value setting
Shared Info. Rep.
Resource utilization and threshold value store
QoS+ (Simulator)
② User authentication request(Accept : ③, Reject : ④)
⑥ Access control request(Accept : ⑦, Reject : ⑧)
Resource utilization value collection
④ Approval or denial of login
⑧ Authorization or denial of service
User (Simulator)
Service (Simulator)
① Login request
⑤ Service request
⑦ Validation of service permission

Figure 7 Experiment environment.

Table 1 Priority settings for each user

User ID	Priority
testuser1–testuser3	0
testuser4–testuser6	1
testuser7–testuser9	2
testuser10	3

Table 2 User and service settings

Item name	Configuration value
User ID	testuser1–testuser10
Service name	ServiceB

5.1 Identity DB

Table 1 shows the information of 10 users used in the experiment. The priority ranges from 0 to 3, and the higher the value, the higher the priority.

5.2 User Simulator

Table 2 shows the information of 10 users used in the experiment. The priority ranges from 0 to 3, and the higher the value, the higher the priority. For the operation of the user simulator, 10 users set in the identity DB were set, as shown in Table 1, to use the same service. The user simulator logs in from testuser1 to testuser10 every second to check the operation of the entire function of the prototype and requests service if the login is successful. If a login or service request is rejected, log in again.

Table 3 Service estimated resource usage and threshold settings

Item name	Configuration value
Service name	ServiceB
Expected resource utilization (CPU/memory)	20/20
CPU/memory threshold value	If 70, restrict users with priority 0
	If 80, restrict users with priority 1
	If 90, restrict users with priority 2

5.3 System Simulator

The system simulator makes the following assumptions. The services provided by the system simulator are real-time applications such as voice and video. The resource usage increases from the point of providing the service to the user, and the resource usage decreases when the service is denied, or the connection with the user is disconnected. There is a request limit of 3 seconds when IAM+ or QoS+ denies a user's login or service.

5.4 Shared Information DB

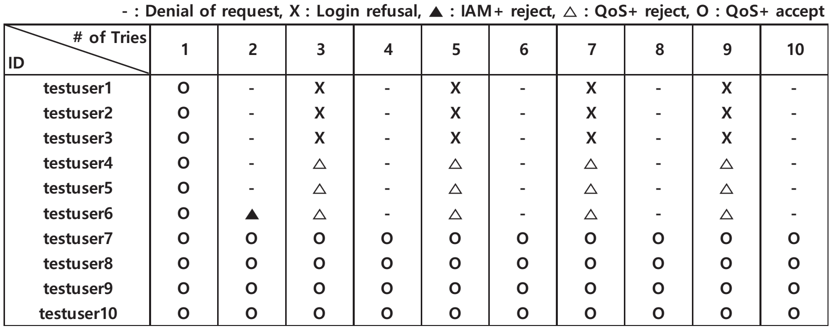
Table 3 shows the service to be used by users, the priority of the service, the expected resource usage for the service, and setting information for the threshold value. Expected resource usage is the expected usage of resources when using the service and is expressed as a percentage (%) of the total resources.

5.5 Experimental Results

The user simulator was run for 1 minute while IAM+, QoS+, shared information repository, and system-related processes were running.

First, Figure 8 shows the results when PBAC is applied, and the results are repeated 3 to 10 times even if the number of trials increases. When users request a service for the first time, the service is provided to each user, but as the number of users increases, resource usage increases and exceeds a set threshold. At this time, the user with the lower priority is retrieving the service resources used by QoS+. If the user tries to request the service again, they will be restricted from logging in via PBAC and using the service from IAM+. As a result, services for users with high priority are continuously provided.

Figure 9 shows the resources used when a user requests a service. It can be confirmed that resource usage is maintained within the threshold by IAM+ and QoS+.



- : Denial of request, X : Login refusal, ▲ : IAM+ reject, △ : QoS+ reject, O : QoS+ accept

ID \ # of Tries	1	2	3	4	5	6	7	8	9	10
testuser1	O	-	X	-	X	-	X	-	X	-
testuser2	O	-	X	-	X	-	X	-	X	-
testuser3	O	-	X	-	X	-	X	-	X	-
testuser4	O	-	△	-	△	-	△	-	△	-
testuser5	O	-	△	-	△	-	△	-	△	-
testuser6	O	▲	△	-	△	-	△	-	△	-
testuser7	O	O	O	O	O	O	O	O	O	O
testuser8	O	O	O	O	O	O	O	O	O	O
testuser9	O	O	O	O	O	O	O	O	O	O
testuser10	O	O	O	O	O	O	O	O	O	O

Figure 8 Service request results when PBAC is applied.

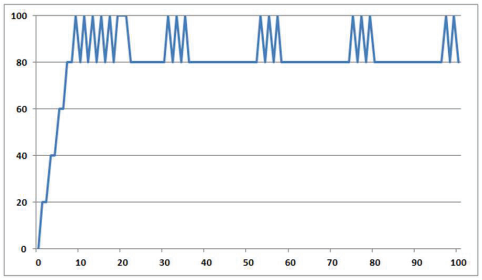
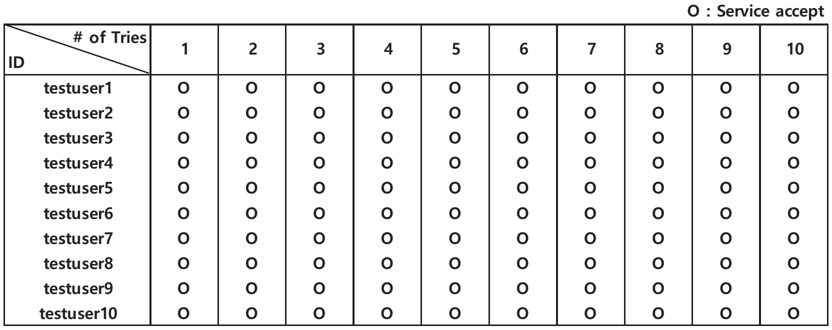


Figure 9 Changes in resource usage when applying PBAC.



O : Service accept

ID \ # of Tries	1	2	3	4	5	6	7	8	9	10
testuser1	O	O	O	O	O	O	O	O	O	O
testuser2	O	O	O	O	O	O	O	O	O	O
testuser3	O	O	O	O	O	O	O	O	O	O
testuser4	O	O	O	O	O	O	O	O	O	O
testuser5	O	O	O	O	O	O	O	O	O	O
testuser6	O	O	O	O	O	O	O	O	O	O
testuser7	O	O	O	O	O	O	O	O	O	O
testuser8	O	O	O	O	O	O	O	O	O	O
testuser9	O	O	O	O	O	O	O	O	O	O
testuser10	O	O	O	O	O	O	O	O	O	O

Figure 10 Service request results when PBAC is not applied.

Figure 10 shows the results when PBAC is not applied. Since there is no separate restriction, all services are permitted according to the user's request.

In Figure 11, it can be seen that the resource usage continues to increase according to the user's request, reaching 200%. In this case, it is difficult for users to receive normal services any longer.

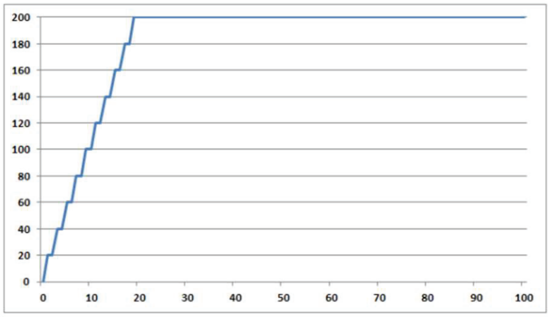


Figure 11 Changes in resource usage when PBAC is not applied.

6 Conclusion

This paper describes the necessity and possibility of differential allocation and processing of the entire system resources according to user or application priority. Based on this, solutions covering related concepts, structures, and methods were specifically presented and tested. In addition, one of the biggest obstacles to introducing new solutions is the introduction cost. However, a plan to significantly lower the cost by utilizing the existing IAM function was proposed.

Furthermore, in the process, ideas and technologies such as new access control methods and token expansion were devised and developed, and the feasibility of implementation were demonstrated through prototypes.

This paper is the first proposal and development attempt for QoS+ and IAM+, and further research and development are needed to raise the technology level to the level applicable to the actual system.

Acknowledgments

This work was supported by the Agency For Defense Development Grant funded by the Korean Government (UE201115ED).

References

[1] David S. Alberts, Richard E. Hayes, Understanding Command and Control, CCRP, 2006

[2] Cristina Aurrecoechea, Andrew T. Campbell & Linda Hauw, A survey of QoS architectures, Vol. 6, pp. 138–151, 1998

[3] Gyudong Park, Hocheol Jeon, AIQIA : An Integration Architecture between System Layers to Improve Interoperability, KIMST Annual Conference Proceedings, pp. 1116–1117, 2022

[4] Ishaq Azhar Mohammed, Intelligent authentication for identity and access management: a review paper, IJMIE, Vol. 3, Issue 1, pp. 696–705, 2013

[5] Lihao Chen, Jiayi Zhang, Tao Gao, and Tongtong Wang, Analysis of QoS Schemes and Shaping Strategies for Large Scale IP Networks based on Network Calculus, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Vol. 355, Springer,

[6] Toerless Eckert & Stewart Bryant, Quality of Service (QoS), Future Networks, Services and Management pp. 309–344, 2021

[7] J. M. Ppallan, K. Arunachalam, S. S. Gantha, S. Jaiswal, S. Song and A. Nigam, A Method for Enabling Context-Awareness at Transport Layer for Improved Quality-of-Service Control, IEEE Access, Vol. 9, pp. 123987–123998, 2021,

[8] Y. Xue, C. Gedo, C. Christou, B. Liebowitz, A Framework for Military Precedence-Based Assured Services in GIG IP Networks, IEEE MILCOM, 2007.

[9] Sunghwa Son, Gwangjin Wi and Kyung-Joon Park, Situation-Aware Survivable Network Design for MDPI, Applied sciences, 2022.

[10] Gyudong Park, Hocheol Jeon, Gyu Myoung Lee, Byungchun Jeon, A Study on Implementation and Improvement of Triple-Metric Based QoS for Military Networks, The Journal of Korean Institute of Communications and Information Sciences, 2022.

[11] Bharatha Sreeja G Mubeen Begum Saleem Venkata Sravya Divya K Jayashree R, Issues with perimeter based network security and a better model to resolve them, European Journal of Molecular & Clinical Medicine, Vol. 7, Issue 9, pp. 2437–2444, 2020.

[12] Shabnam Mohammad Hasani, Nasser Modiri, Criteria Specifications for the Comparison and Evaluation of Access Control Models, I. J. Computer Network and Information Security, Vol. 5, pp. 19–29, 2013.

[13] David F. Ferraiolo, John F. Barkley, D. Richard Kuhn, Authors Info & Claims, A role-based access control model and reference implementation within a corporate intranet, ACM Transactions on Information and System Security. Vol. 2, Issue 1, pp. 34–64, 1999.

[14] Vincent C. Hu; D. Richard Kuhn; David F. Ferraiolo; Jeffrey Voas, Attribute-Based Access Control, IEEE Computer, Vol. 48, Issue 2, pp. 85–88, 2015.

[15] A. Corradi; R. Montanari; D. Tibaldi, Context-based access control for ubiquitous service provisioning, Proceedings of the 28th Annual International Computer Software and Applications Conference, COMPSAC 2004.

[16] Gyudong Park, Gi-Yoon Jeon, Jong-Oh Kim, A Study on Improvement of the Military IdAM Using Edge-Sovereign Identity (ESI), Journal of Web Engineering, pp. 1435–1448, 2022.

Biographies

Gyudong Park received his Ph.D. in computer engineering from Hongik University, Korea, in 2014. He has been working for the Agency for Defense Development (ADD), Seoul, Korea as a researcher since 1998. His research areas include command and control, interoperability, network, information exchange, and security.

Hyoek Jin Choi received his MD in Dongguk University in 1995. From 1995 to now, he has been a researcher for the Agency for Defense Development (ADD), Seoul, Korea. His research areas are the command and control system, cyber security and artificial intelligence.