Communication Method for Remote Device Control using the Internet and its Evaluation

SHINJI KITAGAMI1, YOSUKE KANEKO1, AKIHISA YASUDA1, HARUMI MINEMURA1 and JUN SAWAMOTO2

1 Mitsubishi Electric Corporation Information Technology R&D Center
Kamakura, Kanagawa 247-8501
2 Iwate Prefectural University
Takizawa, Iwate 020-0193
JAPAN
Kitagami.Shinji@bp.MitsubishiElectric.co.jp

Abstract: Machine-to-machine (M2M) service systems have been receiving widespread attention for connection of a number of devices to the internet and for provision of various services. For network access of the M2M service system, not only security and immediacy, but also multi-service connectivity is required, which makes it possible to connect devices with multiple services. As described in this paper, problems that occur when applying conventional communication protocols to the M2M access network are investigated to propose a “M2M proxy communication method” that simultaneously satisfies the three requirements: security, immediacy, and multi-service connectivity.

Key-Word: M2M Service System, Multi-Service Connectivity, Security, Immediacy, Proxy Server

1 Introduction
The use of so-called machine-to-machine (M2M) service systems have spread among people seeking to connect between devices or a device and a remote system via the internet to provide services such as remote maintenance and remote energy management without human intervention [1][2][3]. Such M2M service systems comprise a device that is connectable to the network, an access network for connecting the device with the internet service, and an application server for providing various services. For the access network (hereinafter, M2M access network), which collects data and performs remote control among numerous devices and the application server, three fundamental requirements is required; “security” -- prevents unauthorized access to the device, “immediacy” -- gives control from the application server to the device and “multi-service connectivity” -- connects the device with multiple different services. In particular, with the wide popularity M2M services, the number of cases is growing in which a single device is shared by multiple servers; multi-service connectivity has taken on increasing importance.

However, it remains problematic that conventional communication methods cannot achieve multi-service connectivity such as a direct access method, in which a server function is implemented on the devices, and a polling method, in which requests are sent by the devices to the application server on a regular basis.

In this paper, an “M2M proxy communication method” is proposed as a communication method that achieves multi-service connectivity while providing security and immediacy. In the M2M proxy communication method, a proxy server located between the devices and the application server relays a long-polling communication from the devices and a request–response communication from the application server. In relay processing, multi-service connectivity of the M2M access network is achieved by performing access control and exclusion control to the devices. By adopting the M2M proxy communication method, the M2M access network simultaneously provides security, immediacy, and multi-service connectivity at low cost without introducing any special scheme into the devices.

2 Fundamental requirement of an M2M access network
In this chapter, descriptions will be given of M2M service system characteristics and fundamental requirements for the communication method applied to the M2M access network.
2.1 Configuration and characteristics of an M2M service system

Fig. 1 portrays a common internet service system (hereinafter, an ASP model) such as electronic shopping and a common system configuration of the M2M service system (hereinafter, an M2M model).

The ASP model and the M2M model are in a form in which a user PC is replaced by a device topologically. However, the positioning of data and applications is different: in the ASP model, both data and the application exist on the application server side. In contrast, in the M2M model, the generated data exist on the device side and the application server collects them. Various actuator functions of the device are considered to be part of applications of services. Therefore, applications are distributed and arranged on both the server side and the device side.

When particularly addressing communication on the network between the user PC and device and the application server, as presented in Fig. 2, a request–response type communication with the user PC being an initiator is the basis for the ASP model. However, bi-directional communication is necessary for the M2M model for data transmission and for abnormality notices from the device to the server and control requests from the server to the device.

Consequently, the communication form of the M2M access network differs from that of the ASP model. A communication method that is suitable for the characteristics is necessary for adoption.

2.2 Fundamental requirements

When investigating the communication method of the M2M access network, fundamental requirements for adapting to the communication of the M2M model described in 2.1 are “security”, “immediacy”, and “multi-service connectivity.”

(1) Security

It is necessary to connect devices with the internet to make it possible to perform remote control from the application server in the M2M model. Therefore, to obtain security to prevent unauthorized acquisition of the device data and unauthorized operation of the devices becomes an important requirement [4]. In particular, extraordinarily large number of devices will be connected in the M2M model; security must be obtained on the device side without much cost.

(2) Immediacy

With the M2M model, in addition to security, immediacy must be satisfied when performing remote control from the application server to devices [5]. Requests and responses in the M2M model have reverse directions in the ASP model.

(3) Multi-service connectivity

In the M2M model, multiple services must use a single device [4]. For example, in the case of an air conditioning device, communications are necessary with application servers of both remote energy saving service and remote maintenance service: multiple servers provide services to when viewed from the devices.

3 M2M proxy communication method

In this chapter, after describing adaptability of fundamental requirements of the M2M access network explained in 2.2 for existing communication methods, an “M2M proxy communication method” will be developed as the communication method to solve problems of conventional communication methods.

3.1 Conventional communication method

As a communication method that has been applied traditionally to the M2M access network, a direct access method, a polling method, and long-polling method will be adopted from system examples presented in the past [2][5][10].
3.1.1 Direct access method
In the direct access method, by implementing a communication service function on the device side and making the application server an initiator, a request–response mode of communication is initiated (Fig. 3(a)). For use as a communication protocol based on the direct access method, BACnet/WS, oBIX [6], and other similar protocols are available.

In the direct access method, the global IP address must be open to ensure network accessibility to the device. Methods such as TURN and STUN [7] have been developed for solving the problem of NAT traversal, which achieves bi-directional communication on the internet. Nevertheless, each has a limit on the applicable router. To prevent unauthorized accesses to devices from outside and to control access from multiple application servers, an access control function must be implemented on the device side, entailing the problem of increased cost.

3.1.2 Polling method
Contrary to the direct access method, in the polling method, polling is performed to the application server by a request–response type communication with the device being the initiator (Fig. 3(b)). That is, from the device side to the application server, whether a request to the device exists or not is determined on a regular basis. Only when the request exists is it taken into the device. The form of this method is the same as that in the case in which a Web site on the internet is referred to in the ASP model. There is no need to open the global IP address of the device and the problem of NAT traversal can be avoided. Unauthorized access from outside is prevented in the router.

Immediacy of the polling method depends on a polling interval and whether it is equal to an update interval of the server side, the polling method is effective [8]. In the M2M model, however, because the timing of request generation for the device is usually uncertain, to ensure immediacy, the polling interval must be shortened. Consequently, unnecessary polling communications frequently occur to impose a large load on the entire system disadvantageously. Furthermore, the destination of polling becomes one server, failing to ensure multi-service connectivity.

3.1.3 Long-polling method
The long-polling method solves problems of polling. The communication form is the same as the polling method. If no request exists to the device on the server side, however, then a session is made to wait until a request occurs on the server side. (Fig. 3(c)) Bayeux [9] is a communication protocol that achieves a pseudo-direct access by adopting the long-polling method. With iopeNet [10], Bayeux is adopted as a communication protocol for an equipment network.

The merit of the long-polling method is that security and immediacy are obtainable simultaneously in the M2M access network. However, the problem exists that multi-service connectivity cannot be ensured because it is not possible to perform polling to multiple application servers as in the polling type.

3.2 Proposed method
As described in 3.1, a multi-service connection cannot be provided using conventional communication methods. Accordingly, an “M2M proxy communication method” is developed in this study as a communication method to solve the problem.

3.2.1 System configuration and communication sequence
In the M2M proxy communication method, as portrayed in Fig. 4, the M2M proxy server is located between the device and the application server belonging to neither of them. Between the device and
the M2M proxy server, the long-polling method is adopted. Between the M2M proxy server and the application server, a request–response type communication is adopted in which the application server is an initiator.

Fig. 5 presents the communication sequence of the M2M proxy communication method. The M2M proxy server relays the request from the application server using request–response type communication to the response for the request queue from a stand-by device using the long-polling method. When a response is made from the device, The M2M proxy server relays it to the response of the request from the application server.

3.2.2 Correspondence to multi-service connectivity
In the M2M proxy communication method, communication between the M2M proxy server and the application server employs a request–response type communication, which makes the application server an initiator. Because a scheme is adopted that the M2M proxy server becomes a relay point to produce a loose coupling with the device side, a request can be transmitted from multiple application servers to a single device, enabling provision of multi-service connectivity.

Using the M2M proxy communication method, multi-service connectivity can be provided at low cost with no special scheme for the device and the application server. With this method, because a long-polling method is used for the communication between the device and the M2M proxy server, security and immediacy, which are the fundamental requirements of the M2M access network, are suitable in principle.

4 Experiments and evaluation of developed method
Chapter 3 explained that the M2M proxy communication method can simultaneously provide the fundamental requirements of the M2M access network: security, immediacy, and multi-service connectivity. However, because the developed method employs a system configuration in which an M2M proxy server is arranged that performs a communication relay, the communication overhead and a system load at the time of increase in the number of devices to be connected in the M2M proxy server might cause deterioration of its immediacy. To evaluate the influence, the M2M proxy server implementing the M2M proxy communication method is produced experimentally and an experiment is performed for evaluating the communication overhead and the load of the M2M proxy server.

4.1 Configuration of the experiment system
Table 1 and Fig. 6 present the configuration of the experiment system and specifications of the M2M proxy server. In the communication overhead measurement, a commercial FTTH network (download 40 Mbps, upload 5 Mbps) is used as the M2M access network. In the load measurement of the M2M proxy server, because the internet environment is not indispensable, the experiment was performed using a private network. The communication protocol was HTTP.

4.2 Experiment procedures
Measurements of the communication overhead and the server load of the M2M proxy server are performed respectively according to the following procedures.

(1) Measurement of communication overhead
By comparing the response time of the M2M proxy communication method and the direct access method, the degree of influence on the overhead is evaluated.

(a) A port of a PC side router simulating the device is opened. Then the necessary data are transmitted directly from the application server to the PC; the
time $Ta$ is measured until the response is returned. It is the response time of the direct access method. (b) The same required data as those described above (a) are transmitted to the PC via the M2M proxy server, and the time $Tb$ is measured until the response is returned. This is the response time of the present developed method. (c) Obtaining the difference between the $Ta$ and $Tb$ described above shows the communication overhead of the M2M proxy communication method. However, the processing time $Tc$ of the device depends on the applications running on the system. Therefore, it is omitted from each response time.

(2) **Measurement of the load of the M2M proxy server**

This server load measurement experiment simulates several connected devices to measure the M2M proxy server’s tendency of resource (memory and CPU utilization) consumption. (a) Up to 4,000 long-polling connections are performed simultaneously from a PC simulating multiple devices to measure the change of the memory utilization and CPU utilization. (b) Data transmission and reception are performed from the application server simultaneously as described above (a) to measure the response performance.

### 4.3 Experiment results

Measurement results of communication overhead and the server load of the M2M proxy server are shown.

#### (1) Measurement results of communication overhead

Table 2 presents measurement results that indicate the communication overhead of the M2M proxy communication method. Results of the experiment show that, although the response time of the direct access method was 1.46 sec, which of the M2M proxy communication method was 1.50 sec. That is, the overhead of the M2M proxy communication method was 0.04 sec, being 2.7% of all.

<table>
<thead>
<tr>
<th>Test Case</th>
<th>Mean Response Time [sec] (extract device process time $Tc$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Access</td>
<td>$Ta - Tc = 1.46$</td>
</tr>
<tr>
<td>via M2M Proxy</td>
<td>$Tb - Tc = 1.50$</td>
</tr>
</tbody>
</table>

( Number of Trials $= 10,000$)

![Fig. 6 Configuration of experiment system.](image)

![Fig. 7 Change of memory utilization and CPU utilization of M2M proxy by number of device accesses.](image)

![Fig. 8 Change of response performance by number of device accesses.](image)

Table 2 Measurement results of communication Overhead

<table>
<thead>
<tr>
<th>Test Case</th>
<th>Mean Response Time [sec] (extract device process time $Tc$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Access</td>
<td>$Ta - Tc = 1.46$</td>
</tr>
<tr>
<td>via M2M Proxy</td>
<td>$Tb - Tc = 1.50$</td>
</tr>
</tbody>
</table>

( Number of Trials $= 10,000$)

![Fig. 6 Configuration of experiment system.](image)

![Fig. 7 Change of memory utilization and CPU utilization of M2M proxy by number of device accesses.](image)

![Fig. 8 Change of response performance by number of device accesses.](image)

Table 1 Configuration of M2M proxy server.

<table>
<thead>
<tr>
<th>Software</th>
<th>Overhead Measurement</th>
<th>System Load Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU (Intel Xeon) $2$GHz</td>
<td>4-core $2.95$GHz</td>
<td></td>
</tr>
<tr>
<td>Memory $6$GB</td>
<td>$4$GB</td>
<td></td>
</tr>
<tr>
<td>OS (Linux) $RHEL$ $4.4$</td>
<td>$CentOS$ $5.4$</td>
<td></td>
</tr>
<tr>
<td>Java SE $1.6$</td>
<td>$1.0$</td>
<td>$1.6$</td>
</tr>
<tr>
<td>Tomcat $6.0$</td>
<td>$18$</td>
<td>$6.0$</td>
</tr>
</tbody>
</table>
Measurement results of the M2M proxy server load

Figs. 7 and 8 present measurement results of the M2M proxy server load. Memory utilization increases in proportion to the number of connected devices. Although the CPU utilization is 2% or less in the case in which the number of connected devices is 1,000 or less, it increases to 5% in the case of 1,000–3,000 connections. The CPU utilization abruptly increases when the number of connected devices exceeds 3,000. Immediacy is stable until the number of connected devices is 2,000; it later deteriorates drastically when exceeding 3,000.

5 Considerations

5.1 communications overhead

Effects of the M2M proxy communication method on communication overhead are limited to 3% according to measurement results of the communication overhead. Therefore, the service assuming the best effort communication performance in the internet environment has come to be useful on a practical level.

The communication overhead can be kept low because the long-polling method used in the request queue by the device establishes a communication session with the device before requests occur in principle. Therefore, the overhead for establishing a session for each request can be avoided.

5.2 M2M proxy server load

Memory utilization of the M2M proxy server increases in direct relation to the number of connected devices because the response wait of the long-polling increases in direct relation to the number of connected devices. The CPU utilization increases abruptly when the number of connected devices is greater than 3,000. The thread dispatch for communication standby of the long-polling becomes complicated and its overhead seems to have some influence. Accordingly, if the number of connected devices increases, then immediacy is degraded.

When applied to a large-scale system, the load balancing of the M2M proxy server is required to ensure immediacy. The M2M proxy server is in charge of the communication relay between the devices and the application server. Therefore, a study of load balancing method is necessary to identify a suitable M2M proxy communication method because no simple server load balancing method such as DNS round-robin is applicable.

6 Conclusion

As described herein, we presented M2M access network characteristics and problems. To solve the problems, this M2M proxy communication method was developed. The developed method can satisfy needs for security, immediacy, and multi-service connectivity simultaneously at low cost on the device side. We have presented and examined experimental results related to the communication overhead and the server load of the M2M proxy server when the number of connected devices increased, demonstrating the core capabilities of the developed method.

Future research should be undertaken to address establishment of a load balancing method suitable for the M2M proxy communication method to obtain scalability of the M2M proxy server.

References:

[4] ETSI, Machine-to-Machine communications (M2M); M2M service requirements, ETSI TS 102 609 v0.4.1, 2009
[6] OASIS, oBIX 1.0 (obix-1.0-cs01) Committee Specification 01, 2006