Application of Content-Based Network for Sensor Data Distribution System

Hideya Ochiai,   Zhiyong Wang,   Ryo Oguchi,   Tetsuhiro Sugiyama,
Yusuke Sakamoto,   Shinichi Ishida,   Hiroshi Esaki
Graduate School of Information Science and Technology, The University of Tokyo, Japan

Abstract

This paper proposes and evaluates a new sensor data distribution system, based on the content based network technology. This system is an overlay XML-based content-based routing system over the plain unicast IP network. As the application of the prototype system, we have selected Live E! project, whose object is collecting the weather information from various weather stations and distributing these digital data to the appropriate receiver nodes using the XML-based content-based multicast service. We have designed (1) the representation format of sensor information with XML, (2) content-based multicast routing algorithm using XPath. We have implemented the proposed system for its technical and functional validation.
Keyword: Sensor Network, CBN.

1. Introduction

While the attention and the consideration on ubiquitous networking research has been progressed, it has been realized the importance of sensor networking systems, that must accommodate wide spectrum of sensor devices and information generated by wide variety of embedded sensor devices. It shall be expected that the sensor networks are going to play an important role, as the fundamental infrastructure for at large computer networks that include wide variety of embedded devices (e.g., actuator devices). The role of sensor networks are generating, collecting, integrating, aggregating, processing, delivering and sharing the sensor information for any users on the global Internet. Due to the diversity of sensor devices and the large number of sensor devices, the design and operation of (global-scale) sensor networking would need new architectural concepts. The sensor networks must manage the sensor devices, as well as the sensor information generated by these sensor devices.

This paper discusses and gives preliminary system evaluation for a sensor networking architecture, which

is based on Content-Based Networking paradigm [1]. The designed and implemented system is focusing on the applying as the Live E! project [2] platform. We implemented a prototype system over MANET (Mobile Ad hoc Network) IP network in the laboratory so as so validate the proposed architecture and its functionality.

2. Live E! Project

Live E! project is aiming to establish the global information platform, that achieves autonomous digital data circulation and sharing of any information related with the earth.   As a first exampled application, Live E! project picks up the climate (weather) information, such as temperature, pressure or humidity both for public and private use. In this project, these data are collected by sensors spread all over the globe. Various types of sensor nodes will establish a large scale sensor network system over the globe. Using this sensor network, any user on the Internet must be able to obtain any of the sensing information on the globe.   In order to improve the service usability, we adopt a web service to provide our climate data for the users. Users can include or introduce the deseeded functionalities into their own application using SOAP.
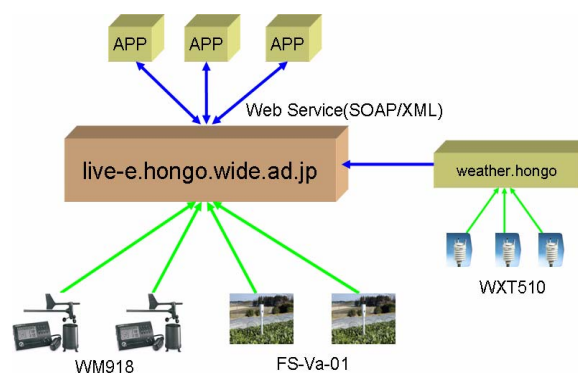

Figure 1, Live E! architecture

Existing Live E! Project network is basically adopting a server-central architecture. This is because; (1) we want to start the project as soon as possible, and (2) the current system has not yet accommodated huge

number of sensor nodes. Since the number of sensor node in the network is not huge, the required processing capability at the server node is acceptable, at this time. In the current system, a sensor periodically sends the data to the specific DB (Data Base) node, whenever the sensor node receives the data that should be sent to DB node. Since the total volume of information and the frequency of data reception event observed at the DB node are growing dramatically according to the increase of the number of sensor node, it is obvious for us that only a single DB node can not be applicable. Also, the data queries to obtain the data in the DB node issued by the user increase, according to the increase of users. The queries issued by the users would or would not synchronize. The queries issue the interruption, which is very expensive operation for the computer system, at the DB node. In order to solve these technical issues, we have proposed a system architecture, which can achieve the above functions with the distributed fashion. The proposed architecture in this paper is based on the adoption of Content-Based Networking paradigm.

3. Content-Based Networking

The Content-Based Network (CBN) is a kind of overlay networking. The conventional IP network transfers data based on the destination IP address given in each IP packet. On the contrary, in the CBN, the transmission of message flow is driven by the content of the messages, rather than by the explicit IP addresses assigned by senders. The routing process on the CBN is of content oriented. The sender do not need to know the receiver's individual IP addresses when it sends message and the routers are responsible to delivering the received message toward the appropriate destination node(s).
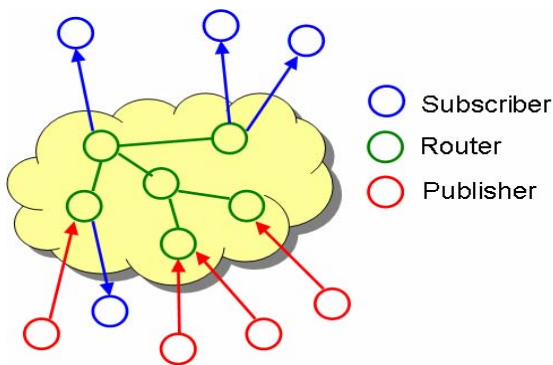

Figure2, Content-Based network

Since, in the CBN, the destination node does not need to resolve the destination node(s) for every message, it is suitable for a case where frequent configuration modification or changes occur in the network or at the destination nodes (i.e., sender or receiver). We realize that this functional feature of CBN is applicable to the nature and behavior of node in the Live E! project network. In order to apply the CBN system to the Live E! project network, we consider; (1)

data collection from the sensor node to DB nodes and (2) data delivery from the DB nodes to end user nodes.
(1) Data collection;
     Sensor node corresponds to information publisher node and DB node corresponds to a subscriber node.
(2) Data distribution;
     DB node corresponds to information publisher node and end user node corresponds to a subscriber node.

Here, the following discussion is for the case of data collection, i.e., sensor information is delivered to the appropriate DB node, based on the context of message.

The communication between these two kinds of nodes (i.e., publisher nodes and subscriber nodes) in our CBN network is of indirect. They are connected through a context-based router network. Each DB node (i.e., subscribers) declares the "context information" to the neighbor router(s). The context information is about forwarding conditions, such as "I need temperature or humidity data in Tokyo". The publisher node, i.e., sensor node, knows nothing about its final destination(s) (i.e., DB node(s)). What the publisher node has to do is to send the data it to the registered router(s). Routers, that are sitting between publisher node (sensor node) and subscriber nodes (DB node), analyze the content of messages arrived from publishers (sensors) or from the other routers, in order to determine which node(s) is(are) the next-hop interface(s) of them. Also, the routers receive the subscribing request from DB node at anytime, and distribute the corresponding context information to the other routers.

We realize that there are two benefits when adopting a Content Based Network paradigm in Live E! sensor network.
(1) We can establish an independent relationship between sensor nodes and DB nodes. End nodes, that are sensor nodes or DB nodes, can be independently and autonomously deployed. So the complexity of the operational management for the system can be relaxed. Especially, when a DB node declares a new subscribing request, such as a specific type of climate data, the routers registered by the DB node could forward these data by themselves and we do not need to inform this information to the corresponding sensor node(s).
(2) It is expected that, by the introduction of CBN paradigm, we could improve the system scalability, against the increase of sensor nodes and end user nodes. In the legacy client-server architecture, all the processing load must be laid on the server node. On the contrary, in the proposed system, such processing load can be distributed among the content-based routers. Any node can play a role of CBN router, whenever it installs specific protocol to play as CBN router.

As we discussed above, deployment of sensor nodes and DB nodes in the proposed CBN network can also be very flexible, and we could easily modify the system topology. By considering such benefits, we implemented a prototype system.

4. Prototype System Implementation

There are three types of units in our prototype system. They are; (1) sensor node, (2) router and (3) DB node. We implemented each function by Java.
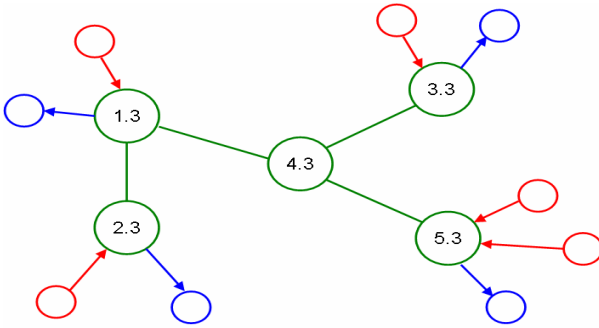

Figure 3, Network topology
Green: CBN Router
Red: Sensor node (Publisher)
Blue: DB (Subscriber)

Sensor node is called as "publisher" in our system, since it generates message to send it to the corresponding neighbor CBN router(s). These messages are in XML format containing the climate data and other information, such as the sensor ID, location and time-stamp of data. Here, we design the XML format so that the definition and instances of content can be expandable.

DB nodes are acting as a data receiver, so that one DB node is called as "subscriber" in our system. It should be configured to know which kind of data it wants to collect. By sending a subscription message to the corresponding neighbor routers, the DB node completes the registration to the CBN. The message sent by DB node is represented as an XPath format to inform the router(s) informing its forwarding condition. The following is an example of XPath information.

/sensor_data[sensor_type='Temperature']

This XPath indicates that the DB node needs any temperature data gathered by sensor nodes.

The router is the most complicated unit in the prototype system. In general, a router mainly provides the two functions. These are the forwarding (i.e., data-plane) and the routing (i.e., control-plane). Both functions must be of content-based. The forwarding function performs the transmission of the received messages from source node (i.e., router or sensor node) to the appropriate next hop node (router or DB node), according the routing table generated and managed by routing function. The routing function generates and manages the routing table, through the exchange of route message among the routers. Here, the route message is represented by XPath format.

The content-based network in our prototype system is of application layer networking, rather than of network layer networking, since the topology of routers,

publisher nodes and subscriber nodes in the network is defined as an overlay network over legacy layer 3 IP network.
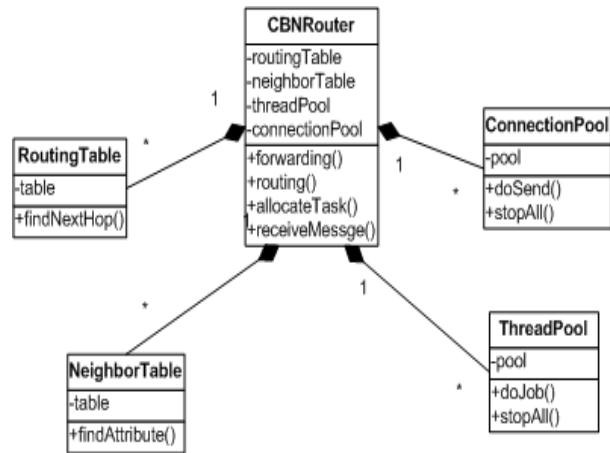

Figure 4, A CBN Router

When we were implementing the prototype system, we focused on how to improve the efficiency and performance of message forwarding in the system. Actually, we adopted the following two techniques;

(1) We developed a connection pool and a thread pool for a router node using JAVA thread support.
(2) Design sophisticated forwarding and routing algorithm in control-plane, in order to avoid unnecessary message transmission in the data-plane. Since, in CBN, the message transmission is performed by the routers, we need some routing algorithm running in routers. When we can not apply a proper routing algorithm, a message flooding would occur in the network. We applied an algorithm to establish broadcast tree in the network, to reduce the message transmission among routers. Router can only send messages to the next hop router(s) or DB node(s), through the path of the appropriate broadcast tree. This mechanism could avoid the message forwarding loop. In addition to this operation, by sharing the context information, issued by DB nodes among routers, we decrease the forwarding time and distance.

5. Discussion and Further Study Items

There are some features in CBN that differentiate the model from other legacy network architectures such as plain IP network. One of differentiated features can be that the message transmission is not performed in the form of "end to end". This means that an end node does not have idea about the final destination of the message before sending them. Though, in the legacy IP network, the source node must know the destination IP address for all the IP packets sent out from the source node. With this unique feature of the proposed CBN architecture, we want to perform the further study on extending our sensor network architecture toward the following three directions.

(1) Automatic configuration mechanism of sensor node
When the number of sensors is getting large number, the automatic system configuration at each sensor node is realized as the mandatory requirement for the sensor network. The initial system configuration may need some small and simple manual configuration. After the initial installation, the sensor node may need detailed configuration according to its physical location and according to historical system updating. When we implement the system where such an updating and detailed configuration and modification can be performed automatically, it is highly useful for huge scale of sensor node network operation. Especially, when we need replace the sensor node due to some reason (e.g., node failure), what we have to do is putting the sensor node with minimum configuration. We could implement this function via the inverse direction data transmission in our CBN model. A sensor node acts as a subscriber to obtain the appropriate system configuration from DB nodes. DB nodes publish the detailed system configuration information to the appropriate sensor node. In order to develop this system, we will need the architecture to synchronize the information among the corresponding DB nodes.
(2) Improve the flexibility and scalability of the sensor network.
In order to come up with the increase of number of sensor nodes, we must develop the scale-free CBN system. By nature, we may realize that the CBN system could be scale-free networking architecture. This is because we can define the recursive CBN system, when we deploy the hierarchical CBN network. We can define the DB nodes (i.e., subscriber node) in the lower level of network can act as publisher node in the higher level of network. This is the same system definition in the routing system of IP network.
(3) Performance improvement of message forwarding capability
Apparently, the message forwarding performance of the proposed XPath based routing is not good enough. In order to improve the performance of message forwarding, some message forwarding mechanism may have to be applies to. For example, defining the forwarding cache of hashed values for the XPath's of messages that are recently forwarded at the CBN routers would be effective. This technique would be same as the CAM table looking up in the recent routers, acting as the first cache of routing entries, or would be same as the label switching architecture.

6. Conclusion

In this paper, we describe the prototyping sensor network focusing on the application on "Live E!" project network, applying the Content-Based Networking paradigm. The benefits of applying the CBN paradigm in the sensor network is discussed and clarified. We have designed the following prototype system; (1) the representation format of sensor information with XML, (2) content-based multicast routing algorithm using XPath. We have implemented the proposed system to demonstrate it's technical and functional validation.

Acknowledgement

References

[1]. A. Carzaniga and A. Wolf, "Content-Based Networking", *White Paper*, 2001.

[2]. Live E! Project
http://www.live-e.org/

[3]. A. Carzaniga and A. Wolf, "Forwarding in a Content Based Network," *SIGCOMM* 2003.

[4]. A. Carzaniga, M.J. Rutherford, and A.L. Wolf, "A Routing Scheme for Content-Based Networking", *INFOCOM*, 2004.

[5]. Masashi YAMADA, Takashige HOSHIAI, and Shoji KIMURA,"Brokerless Trust and Reputation Implementation for Intelligent Community Spaces", Special Issue on P2P Networking and P2P Services, European Transactions on Telecommunications (ETT), Vol.15, No.6, pp.589-pp.598, Wiley Inter Science, 2004.11