A Low-Delay, Light-Weight Publish/Subscribe Architecture for Delay-Sensitive IOT Services

Yunlei Sun, State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China & College of Computer and Communication Engineering, China University of Petroleum, Qingdao, China

Xiuquan Qiao, State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China

Wei Tan, IBM T. J. Watson Research Center, Yorktown Heights, NY, USA

Bo Cheng, State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China

Ruisheng Shi, State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China

Junliang Chen, State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China

ABSTRACT

In order to build a low-latency and light-weight publish/subscribe (pub/sub) system for delay-sensitive IOT services, the authors propose an efficient and scalable broker architecture named Grid Quorum-based pub/sub (GQPS). As a key component in Event-Driven Service-Oriented Architecture (EDSOA) for IOT services, this architecture organizes multiple pub/sub brokers into a Quorum-based peer-to-peer topology for efficient topic searching. It also leverages a topic searching algorithm and a one-hop caching strategy to minimize the search latency. Light-weight RESTful interfaces make the authors’ GQPS more suitable for IOT services. Cost analysis and experimental study demonstrate that the GQPS achieves a significant performance gain in search satisfaction without compromising search cost. The authors apply the proposed GQPS in the District Heating Control and Information Service System in Beijing, China. This system validates the effectiveness of GQPS.

Keywords: Event-Driven Service-Oriented Architecture (EDSOA), Grid Quorum-Based Pub/Sub (GQPS), Internet of Things, Publish/Subscribe (pub/sub), RESTful Interfaces

DOI: 10.4018/ijwsr.2013070104
INTRODUCTION

Scalability and low-latency is crucial for large-scale distributed delay-sensitive IOT applications. For instance, an earthquake early warning system needs a very short delay between the occurrence of an earthquake event and the disaster relief actions. The main challenge is to minimize the latency between the detection of an occurred event and the delivery of alarm messages in order to maximize the time available for preventing possible damages (Fischer, 2012). Let’s review an emerging IOT application scenario, i.e. the district heating control. This system often needs to automatically process a large volume of real-time streaming data, and make prompt and intelligent decisions in response to the changing situations reflected by those events. For example, when a leakage of heating water supply pipe occurs, the pipeline pressure, backwater volume, backwater temperature and household temperature will all have some abnormal fluctuations. The heating system needs to fuse the multi-source raw sensor data to derive problem diagnosis. Then, based on this implicit knowledge, the system will display the warning information in the real-time monitoring system, with maintenance notifications sent to the maintenance and rescue system. When the maintenance work has to influence the household heating, the customer management system will need to send the related notifications to the influenced customers by short messages. Once the fault problem becomes more serious, the emergency information also needs to be reported to the government heating administration department.

From the scenario above we can see that typically IOT applications involve the processing of complex multi-source real-time data streaming and the event-driven service dynamic coordination cross multiple service domains or organizations. The pub/sub system can facilitate the asynchronous and on-demand sensor information dissemination in a largely distributed and loosely-coupled IOT environment. However, existing pub/sub systems suffer from poor scalability and high delay (Paridel, 2010). P2P-based pub/sub techniques are usually based on the Distributed Hash Table (DHT) overlay network with a certain logical structure, making the length of the path between any two nodes generally not more than $O\left(\log_2 n\right)$, where $n$ is the number of nodes in the network (Stoica, 2003). In order to build a scalable and low-latency pub/sub system for large-scale delay-sensitive IOT services, we propose to use the Grid Quorum (Cheung, 1992) mechanism to construct an overlay network and tightly integrate the pub/sub application. In the proposed GQPS (Grid Quorum based Pub/Sub) architecture, each node maintains a $O\left(\sqrt{n}\right)$ length neighbor list, so that the path between any two nodes in the overlay does not exceed two hops.

Our GQPS architecture is motivated by the concept of quorum and is part of the EDSOA described in Section “Background and Motivation”. The rationale behind the design of GQPS is to distribute the sensor information of millions of IOT objects. To avoid the single point of failure, no single pub/sub broker maintains the system-wide global service information. The primary abstraction exported by the proposed GQPS is used to construct a scalable broker architecture for pub/sub systems. IOT objects publish sensor information to GQPS, which is responsible for event routing, topic searching and topic matching. When topic match is identified, notification messages are sent to IOT services that have subscribed the corresponding topic. In GQPS, we present a specific design for pub/sub systems. The caching strategy reduces the number of transmitted messages to accelerate topic search. Each broker maintains a cache of the subscriber lists of their immediate neighbors. The single-hop topic search algorithm helps to achieve a small and constant average search latency. A group of light-weight RESTful (Fielding, 2000) interfaces are also designed for IOT environment, taking into account the low processing power and energy constraints of small embedded devices, such as sensors.
Discovery of Web Services in a Multi-Ontology and Federated Registry Environment
[www.igi-global.com/article/discovery-web-services-multi-ontology/3062?camid=4v1a](www.igi-global.com/article/discovery-web-services-multi-ontology/3062?camid=4v1a)

Auction-Based Pricing Model for Web Service Providers
[www.igi-global.com/article/auction-based-pricing-model-web/3085?camid=4v1a](www.igi-global.com/article/auction-based-pricing-model-web/3085?camid=4v1a)