A QoS Policy Configuration Modeling Language for Publish/Subscribe Middleware Platforms

Joe Hoffert, Douglas Schmidt, and Aniruddha Gokhale
Institute for Software Integrated Systems
Vanderbilt University
2015 Terrace Place
Nashville, TN 37212
{jhoffert, schmidt, gokhale}@dre.vanderbilt.edu

ABSTRACT
Publish/subscriber (pub/sub) middleware platforms for event-based distributed systems often provide many configurable policies that affect end-to-end quality of service (QoS). Although the flexibility and functionality of pub/sub middleware platforms has matured, configuring their QoS policies in semantically compatible ways has become more complex. This paper makes two contributions to reducing the complexity of configuring QoS policies for event-based distributed systems. First, it evaluates various approaches for managing complex QoS policy configurations in pub/sub middleware platforms. Second, it describes a domain-specific modeling language (DSML) that automates the analysis and synthesis of semantically consistent QoS policy configurations.

Keywords
Pub/sub Middleware, Event-based Distributed Systems, Domain-Specific Modeling Languages, Data Distribution Service

1. INTRODUCTION
With increasing advantages of cost, performance, and scale over single computers, the proliferation of distributed systems in general and distributed event-based systems in particular have increased dramatically in recent years [7]. In contrast to distributed object computing middleware (such as CORBA and Java RMI) — where clients invoke point-to-point methods on distributed objects — pub/sub middleware platforms distribute data from suppliers to (potentially multiple) consumers. Examples of standardized pub/sub middleware include the Java Message Service (JMS) [14], Web Services Notification [12], CORBA Event Service [10], and OMG Data Distribution Service (DDS) [11]. These event-based services allow the propagation of data throughout a system using an anonymous publish/subscribe (pub/sub) model that decouples event suppliers from event consumers.

To support the requirements of a broad spectrum of application domains, pub/sub middleware for event-based distributed systems typically provides many policies that affect end-to-end system QoS properties, such as persistence, i.e., determining how much data to save for current subscribers, durability, i.e., determining whether to save data for late joining subscribers, and grouped data transfer, i.e., determining if a group of data needs to be transmitted and received as a unit. Each QoS policy may have multiple attributes associated with it, such as the data topic of interest, data filter criteria, and the maximum number of data messages to store when transmitting data. Moreover, each attribute can be assigned one of a range of values, such as the legal set of topics, a range of numbers for the maximum number of data messages stored for transmission, or the set of criteria used for filtering. Choosing the right set of values for these QoS policies and ensuring that the QoS policies are configured together in a semantically compatible way, i.e., so that they do not conflict with or contradict each other, is essential for proper system behavior.

This paper uses the DDS middleware (described in Sidebar 1) to illustrate the challenges of ensuring semantically compatible QoS policy configurations. Semantic compatibility is accomplished when the combination and interaction of the specified QoS policies produce the overall desired QoS for the system, i.e., when the system executes with the QoS that is intended. It is hard to achieve semantic compatibility using conventional programming techniques, i.e., manually specifying valid QoS policy configurations with particular QoS policies that are used to the attributes of the policies. Achieving semantic compatibility is particularly hard when QoS policies can conflict with each other, e.g., specifying unreliable data transmission for the data sender with reliable data reception for the data receiver. Moreover, this problem is exacerbated in distributed systems where the overall QoS policy configuration is not globally defined or controlled but is defined by the aggregation of the locally defined QoS policy configurations.

One way to address key QoS policy configuration concerns for event-based distributed systems involves the use of model-driven engineering (MDE) tools [13] based on domain-specific modeling languages (DSMLs) [15]. The type systems of DSMLs formalize the application structure, behavior, and requirements within particular domains, such as software-defined radios, avionics mission computing, online finan-
cial services, warehouse management, or even the domain of pub/sub middleware platforms. DSMs are described using metamodels, which define the relationships among concepts in a domain and precisely specify the key semantics and constraints associated with these domain concepts. Developers use DSMs to build applications using elements of the type system captured by metamodels and express design intent declaratively rather than imperatively.

### Sidebar 1: Overview of DDS

The OMG Data Distribution Service (DDS) [11] is an Object Management Group (OMG) specification that defines a standard anonymous pub/sub architecture for exchanging data in event-based distributed systems. DDS provides a global data store in which publishers and subscribers write and read data, respectively. The DDS architecture consists of two layers: (1) the data-centric pub/sub (DCPS) layer that provides APIs to exchange topic data based on specified QoS policies and (2) the data local reconstruction layer (DLRL) that makes topic data appear local.

The DCPS entities in DDS include topics, which describe the type of data to be written or read; data readers, which subscribe to the values or instances of particular topics; and data writers, which publish values or instances for particular topics. Various properties of these entities can be configured using 22 QoS policies. Moreover, publishers manage groups of data writers and subscribers manage groups of data readers.

Table 1 summarizes the subset of the DDS QoS policies referenced in this paper. Each QoS policy has attributes with the majority of the attributes having a large number of possible values, e.g., an attribute of type long or character string.

<table>
<thead>
<tr>
<th>DDS QoS Policy</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>User data</td>
<td>Attaches application data to DDS entities.</td>
</tr>
<tr>
<td>Durability</td>
<td>Determines if data survives the time when written or read.</td>
</tr>
<tr>
<td>Durability</td>
<td>Details how durable data is stored.</td>
</tr>
<tr>
<td>Service</td>
<td>Delivers data as a group and/or in order.</td>
</tr>
<tr>
<td>Deadline</td>
<td>Determines rate at which periodic data is refreshed.</td>
</tr>
<tr>
<td>Latency</td>
<td>Sets guidelines for acceptable end-to-end delays.</td>
</tr>
<tr>
<td>Budget</td>
<td>Controls writer(s) of data.</td>
</tr>
<tr>
<td>Liveliness</td>
<td>Sets liveliness properties of topics, data readers, data writers.</td>
</tr>
<tr>
<td>Time Based</td>
<td>Modifies exchange between slow consumers and fast producers.</td>
</tr>
<tr>
<td>Filter</td>
<td>Controls logical partitioning of data.</td>
</tr>
<tr>
<td>Reliability</td>
<td>Controls reliability of data transmission.</td>
</tr>
<tr>
<td>Transport</td>
<td>Sets priority of data transport.</td>
</tr>
<tr>
<td>Priority</td>
<td>Sets whether data sender or receiver determines order.</td>
</tr>
<tr>
<td>Order</td>
<td>Sets how much data is kept to be read.</td>
</tr>
<tr>
<td>History</td>
<td>Controls resources used to meet requirements.</td>
</tr>
</tbody>
</table>

This paper describes how a particular MDE-based DSML, the DDS QoS Modeling Language (DQML), can help alleviate the complexities of QoS policy configuration for event-based distributed systems. QoS policy configurations can be defined as complex when, in general, a human being cannot manually manage the number of QoS policies, policy attributes, possible values for the attributes, and the interactions of the policies to ensure that the configurations are valid. DQML can be used to model QoS policy configurations within the domain context, i.e., using terminology and constructs within the domain of interest, while also checking for valid QoS policy configurations early in development at design time, when errors require less effort to fix. DQML can also be used to generate correct-by-construction QoS policy configuration artifacts, e.g., QoS policy configuration files that store particular attribute values for QoS policies which the system loads during execution.

The remainder of the paper is organized as follows: Section 2 uses the NASA Magnetospheric Multiscale (MMS) Mission [1] as an example of an event-based distributed system that can benefit from an MDE-based approach to QoS policy configuration. Section 3 highlights the challenges of QoS policy configuration in a DDS-based design for the MMS mission and analyzes various approaches to addressing these challenges; Section 4 describes our approach to resolving the challenges presented above using the DQML DSML; Section 5 compares DQML with related work; and Section 6 presents lessons learned guiding future work on DQML.

2. MOTIVATING EXAMPLE: NASA'S MAGNETOSPHERIC MULTISCALE MISSION

We chose NASA's Magnetospheric Multiscale (MMS) Mission to highlight the challenges of configuring QoS policies in event-based distributed systems. This mission is designed to study particular aspects of the earth's magnetosphere, such as magnetic reconnection, charged particle acceleration, and turbulence, in key boundary regions. MMS utilizes five co-orbiting coordinated spacecraft with identical instrumentation to perform the desired measurements. The spacecraft can be (re)positioned into different spatial/temporal relationships so that a three-dimensional view of the plasma field, and current structures can be constructed.

Figure 1 shows an example MMS spacecraft deployment, which includes a ground station the spacecrafts can communicate with during a high-capacity window and a non-MMS satellite that can communicate with the MMS spacecraft. The figure also shows data flows between MMS systems and the QoS requirements applicable to the MMS mission outlined below. MMS spacecraft will be equipped with uplink and downlink capability to transport telemetry. Each spacecraft will need to gather, store, and transmit information regarding neighboring spacecraft to enable precise coordination for particular types of telemetry and positioning. The instruments aboard each spacecraft will also generate ~250 megabytes of data per day. To minimize ground station cost each spacecraft will store up to 2 weeks worth (i.e., 3.5 GB) of data so that the spacecraft can wait for high-rate transmission windows. With these data requirements, the DDS pub/sub middleware would be appropriate since its QoS policies support the following capabilities:
which is outlined in Table 1, which can be configured to meet the needs of event-based distributed systems with diverse QoS requirements. DDS’s flexible configurability, however, requires careful management of interactions between various QoS policies so that the system behaves as expected. This section uses our MMS example from Section 2 to present the challenges of configuring DDS QoS policies so the system executes as intended.

### 3.1 DDS QoS Policy Configuration Challenges

The following are three general types of challenges that arise when creating DDS QoS policy configurations.

**Challenge 1: QoS compatibility.** The DDS specification defines constraints for compatible QoS policies. When these constraints are violated data will not flow between data writers and data readers. For example, an incompatibility between reliability QoS policies will occur if an MMS ground station requests data be sent reliably but an MMS spacecraft offers only a best-effort policy. The data will not flow between the spacecraft and the ground station because the QoS policies are incompatible, as shown in Figure 2.

### 3.2 Potential Incompatible DDS QoS Policies

Table 2 lists the QoS policies that can be incompatible and the relevant types of entities for those policies. Incompatibility applies to multiple QoS policies of the same type across various entities. Section 4.3.1 presents our approach to addressing QoS compatibility challenges.

<table>
<thead>
<tr>
<th>QoS Policies</th>
<th>Affected DDS Entities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Durability</td>
<td>Topic, Data reader, Data writer</td>
</tr>
<tr>
<td>Presentation</td>
<td>Publisher, Subscriber</td>
</tr>
<tr>
<td>Deadline</td>
<td>Topic, Data reader, Data writer</td>
</tr>
<tr>
<td>Latency Budget</td>
<td>Topic, Data reader, Data writer</td>
</tr>
<tr>
<td>Ownership</td>
<td>Topic, Data reader, Data writer</td>
</tr>
<tr>
<td>Lifespan</td>
<td>Topic, Data reader, Data writer</td>
</tr>
<tr>
<td>Reliability</td>
<td>Topic, Data reader, Data writer</td>
</tr>
<tr>
<td>Destination Order</td>
<td>Topic, Data reader, Data writer</td>
</tr>
</tbody>
</table>

Table 2: Potential Incompatible DDS QoS Policies

**Challenge 2: QoS consistency.** The DDS specification also defines when QoS policies are inconsistent, indicating that the multiple QoS policies associated with a single DCPS entity are not allowed. For example, an inconsistency between the Deadline and Time-based Filter QoS policies will occur if an MMS ground station tries to set the Deadline QoS policy’s deadline period to 5 ms and the Time-based Filter QoS policy’s minimum separation between incoming pieces of data to 10 ms, as shown in Figure 3.

This policy configuration specifies that the ground station requires data at time interval 5 ms but that it also should only receive data at an interval of at least 10 ms. Unfortunately, this policy configuration is inconsistent with the DDS constraint that deadline period ≥ minimum separa-
tion. Table 3 describes cases where QoS policies associated with a single DDS entity can be inconsistent.

<table>
<thead>
<tr>
<th>Inconsistent QoS Policies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource Limits max samples per instance &lt; History depth</td>
</tr>
<tr>
<td>Deadline period &lt; Time Based Filter minimum separation</td>
</tr>
</tbody>
</table>

Table 3: Consistency of DDS QoS Policies

Even with a relatively small system, the incompatibility and inconsistent QoS policies that exist between the reader and writer may not be obvious. This complexity is exacerbated in larger systems, which must consider the number of possible values for each attribute of a QoS policy, the number of attributes for any one QoS policy, the number of QoS policies used, the number of entities associated with QoS policies, the interactions of the QoS policies, and the number of computing nodes in a distributed system.

For example, in the MMS system it is hard to check all incompatible and inconsistent QoS policies manually. If there are any QoS settings that are semantically incompatible, the intended QoS behavior will not be realized and the system will not perform as needed, e.g., data will not be transmitted as required between the spacecraft and the ground stations, between the spacecraft themselves, or within a single spacecraft. Section 4.3.2 describes our approach to addressing QoS consistency challenges.

Challenge 3: Accurate QoS policy configuration transformation. After a valid QoS policy configuration has been designated it must be transformed from design into implementation. A conventional approach to this transformation is to (1) document the desired QoS policies, attributes, values, and associated entities designed often in an ad-hoc manner such as with handwritten notes or conversations between developers and then (2) transcribe this information into the source code. However, this process creates opportunities for accidental complexities as the QoS policies, attributes, values, and related entities can be misread, mistyped, or misunderstood so that the QoS policy configuration that is encoded in the system differs from the valid configuration originally intended.

3.2 Evaluating Common Alternative Solution Techniques

Below we evaluate three common alternatives for addressing the challenges outlined Section 3.1 in terms of their ability to document and realize proven QoS policy configurations robustly.

3.2.1 Point Solutions

In this approach, modifications are made to the existing system’s QoS policies, feedback is gathered, and further modifications are made based on the feedback. This iterative process can be done either at (1) run-time, i.e., while the system is executing, or (2) development time, i.e., while the system is being developed. In either case, developers must not only design a proper QoS policy configuration, but also ensure that the design is transferred faithfully from design to implementation.

A point solution approach works best when a configuration expert is available, the configuration is simple, and the configuration need not be maintained nor enhanced. Under these circumstances the problem is simplified and the overhead of training others, codifying the expertise, or otherwise developing for modifiability may not be needed.

Point solutions, however, make it hard to capture proven QoS policy configurations or leveraging from the expertise of others. Moreover, point solutions do not support automated transformation of configuration solutions from design to implementation. Developers must instead rely on human experts who have developed approaches to solving this transformation challenge or must reacquire this expertise.

3.2.2 Patterns-based Solutions

This approach to addressing the DDS QoS policy configuration challenges uses configuration patterns for DDS, which document the use of QoS policies that provide management, prioritization, and shaping of a data-flow in a network [8]. For example, system developers can limit access to certain data by utilizing the DDS Controlled Data Access pattern, which uses the DDS Partition and User Data QoS Policies along with other DDS elements to provide the desired QoS.

Configuration patterns enable the codification of configuration expertise so it is documented clearly and can be reused broadly. It addresses the problems of the availability of human experts by making the configuration policy expertise generally available. A drawback with a patterns-based approach, however, is that developers are still responsible for implementing the policies manually, which can be tedious and error-prone. Moreover, developers may choose to implement the patterns in different ways, which can impede reuse and integration.

3.2.3 DSM-based Solutions

This approach to addressing the complexity of managing QoS policy configurations involves the use of DSMLs. DSMLs not only codify configuration expertise in the meta-models that are developed for the particular domain but also use an executable form of that expertise to synthesize part or all of an implementation. For example, DSMLs can generate valid QoS configuration files from valid QoS policy configurations modeled in the DSMLs.

DSMLs can also ensure proper semantics for specifying QoS policies and enforce all attributes for a particular QoS policy to be specified and used correctly, as described in Section 1. They can therefore detect many types of QoS configuration problems at design time and can automate the generation of implementation artifacts (e.g., source code and configuration files) that reflect the design intent.

A drawback of using DSMLs is the learning curve required to use them effectively. Moreover, DSMLs typically depend on a particular modeling tool, so users must be familiar with the modeling tool, in addition to being familiar with the particular DSML.

4. THE DDS QOS MODELING LANGUAGE

This section describes the DDS QoS Modeling Language (DQML), which is a DSML we created using the Generic
Modeling Environment (GME)\cite{2} to automate the analysis and synthesis of semantically consistent DDS QoS policy configurations. We summarize the structure and functionality of DQML and then present how DQML helps address the challenges described in Section 3.1.

4.1 Structure of DQML

The DDS specification defines 22 QoS policies a subset of which is shown in Table 1 that control the behavior of DDS applications. DQML models all of these DDS QoS policies, as well as the seven entities that can have QoS policies, i.e., Data Reader, Data Writer, Topic, Publisher, Subscriber, Domain Participant, and Domain Participant Factory. Associations between the seven entities themselves and also between the entities and the 22 QoS policies can be modeled, taking into account which and how many QoS policies can be associated with any one entity as defined by the DDS specification.

There have been several decisions made regarding the design of the DQML metamodel that affects DQML's functionality and how it is used. These involve scope, constraints, and the relationship of QoS policies to DDS entities. To focus specifically on the QoS policy configuration challenges outlined above the scope of DQML is limited to only those DDS entities that can have QoS policies associated with them. In addition to data reader, data writer, and topic already mentioned and briefly defined, these include (1) publisher, which manages one or more data writers, (2) subscriber, which manages one or more data readers, (3) domain participant, which is a factory for DDS entities for a particular domain or logical network, and (4) domain participant factory, which is a factory for generating domain participants. While there are other entities and constructs specified by DDS none of them directly use QoS policies and are therefore not included within the scope of DQML.

The constraints placed on QoS policies for compatibility and consistency are defined in the DDS specification. DQML uses the Object Constraint Language (OCL)\cite{16} provided by GME for defining these constraints.

As noted in Section 3.1, Challenge 1, compatibility constraints involve a single type of QoS policy associated with more than one DDS entity whereas consistency constraints involve a single DDS entity with more than one QoS policy. Both types of constraints are included in DQML. The constraints are checked when the user explicitly invokes constraint checking.

4.2 Functionality of DQML

DQML allows DDS application developers to specify and control the following aspects of QoS policy configuration:

**Creation of DDS entities.** DQML allows developers to create the DDS entities involved with QoS policy configuration. DQML supports the seven DDS entities that can be associated with QoS policies.

**Creation of DDS QoS policies.** DQML also allows developers to create the DDS QoS policies involved with QoS policy configuration. DQML supports the 22 DDS policies that can be associated with entities to provide the required QoS along with the attributes, the appropriate ranges of values, and defaults.

**Creation of associations between DDS entities and QoS policies.** DQML supports the generation of associations between the entities and the QoS policies and ensures that the associations are valid.

**Checking compatibility and consistency constraints.** DQML supports checking for compatible and consistent QoS policy configurations. The user initiates this checking and DQML reports if there are any violations.

4.3 Resolving Design Challenges when Developing DQML

This section outlines how DQML is designed to address the challenges that arise when creating valid QoS policy configurations as outlined in Section 3.1.

4.3.1 Resolving QoS Compatibility

Challenge 1 in Section 3.1 describes the difficulties of ensuring compatible QoS policies between DDS entities that need to exchange data. DQML is designed to address this challenge by including compatibility checking in the modeling language itself. DQML users can invoke compatibility checking to make sure that the QoS policy configuration specified is valid. If QoS policies are found to be incompatible then the user is notified of this along with details of the incompatibility.

DQML is designed to allow for easy resolution of QoS incompatibilities. QoS policies are associated with DDS entities via connections made between them. When an incompatibility is found it can be quickly resolved in most cases by associating the incompatible DDS entities to the same QoS policy ensuring compatibility. DQML also decouples the QoS policy configuration from the application code through the use of interpreters that generate QoS settings for a given application in the format which the application expects.

4.3.2 Resolving QoS Consistency

Challenge 2 in Section 3.1 describes the difficulties of ensuring consistent QoS policies so that the policies will not conflict with each other when associated with the same DDS entity. DQML is designed to address this challenge by including consistency checking in the modeling language itself. Just as with compatibility checking, the user of DQML can invoke consistency checking to make sure that the QoS policy configuration specified is valid. If any inconsistent QoS policies are found then the user is notified along with detailed information about the inconsistency to aid in correcting the problem.

4.3.3 QoS policy Configuration transformation

Challenge 3 in Section 3.1 outlines the problems of propagating a valid QoS policy configuration from design to implementation and then deployment. DQML addresses this challenge by creating interpreters that will iterate over the QoS policy configuration model designed in DQML to create appropriate implementation artifacts (e.g., source code, configuration files) that will faithfully recreate the QoS policy configuration as designed.

5. RELATED WORK

This section compares our work on DQML with related R&D efforts.

**DSMLs for configuring QoS.** There are currently several DSMLs developed to model QoS requirements for distributed real-time embedded (DRE) systems. The Distributed QoS Modeling Environment (DQME)\cite{5} is a modeling tool that targets essential elements of dynamic QoS adaption.
DQME is a hybrid of domain-specific modeling and run-time QoS adaptation methods, with emphasis on adapting QoS to changing conditions with limited system resources. In contrast, DQML focuses on designing and generating correct QoS policy configurations for data-centric pub/sub middleware, such as DDS.

The Options Configuration Modeling Language (OCML) is a DSL that aids developers in setting compatible component configuration options for the system being developed whereas DQML models QoS policy configuration for data-centric middleware that can be applicable across various endpoints such as processes, objects, or components.

Runtime monitoring. Real-time Innovations Inc. (RTI) (www.rti.com) and Prism Technologies (PrismTech) (www.prismtechnologies.com) have developed DDS products along with MDE tools that monitor and analyze the flow of data within a system using DDS. These tools help verify that a system is functioning as designed for a particular QoS policy configuration and for a particular point of execution. Discovering configuration problems at run-time is very late in the development cycle, when problems are more expensive and obtrusive to fix. In contrast, DQML allows QoS policy designers to create semantically compatible QoS policies at design time. DQML can also use the QoS policy configuration model to construct semantically compatible and syntactically correct implementation artifacts, e.g., source code, configuration files, that can be incorporated into the system implementation without human intervention.

6. CONCLUDING REMARKS

Highly configurable pub/sub middleware platforms are increasingly being used as the basis for event-based distributed systems. This adoption has also increased demand for powerful QoS policies. We developed DQML to address the challenges of managing the complexity of the wide variety of possible QoS policies, attributes, and values available with DDS, though its DSL techniques can also be generalized to other pub/sub middleware, such as JMS.

The following is a summary of lessons learned from our experience using DQML to model QoS policy configurations for DDS that are guiding our future work.

- **Ensuring semantic compatibility of QoS policies is crucial to proper deployment of event-based pub/sub systems.** When incompatibilities exist the system will not perform as designed with potentially cascading effects of non-performance and unpredictability.
- **Ensuring that the designed QoS policy configuration is the deployed one is crucial to the integrity of the system.** Even with the valid and proper design of a QoS policy configuration, a system will not perform as intended if the configuration is not faithfully transferred into the implementation. In future work we plan to incorporate DQML into model-based deployment tools [4] to provide “correct-by-construction” deployment for DDS.
- **Becoming proficient with OCL is hard and it can be easy to misuse.** Most systems developers are not familiar with rule-based languages such as OCL so there is the added overhead of training. Development tool support for OCL is often rudimentary, e.g., little debugging support, which lowers productivity. In future work we therefore plan to address enforcing constraints by looking at other constraint solving technologies, such as the Constraint Logic Programming Finite Domain (CLP(FD)) solver [6], [9].

- **Real-world use and feedback are crucial elements to developing a robust DSML.** As with any development tool, DSMLs can only be as good as the particular environments and systems with which they are used. DSL designers cannot foresee all the potential environments and scenarios a priori. In future work, we therefore plan to use DQML to generate QoS policy configurations for benchmarking of DDS implementations and the various DDS QoS policies.

7. REFERENCES


