Towards a Reconfiguration Service for Distributed Real-Time Java

P. Basanta-Val, and M. García-Valls
Dpto. de Ingeniería de Telemática
Universidad Carlos III de Madrid
Avda. Universidad nº 30, Leganés- Madrid (Spain) 28911
{pbsanta, mvalls}@it.uc3m.es

Abstract—Ancient monolithic distributed systems were attached to well-known development practices and offline analysis. Current scenarios are more dynamic, and open, plenty of applications and services which appear and disappear dynamically at runtime. Likewise, these scenarios require taking into account actions that were traditionally addressed offline, this time in an online scenario. This paper contributes a reconfiguration service in the context of distributed real-time Java application as a means to include real-time reconfiguration into next generation real-time Java systems. The paper addresses the integration taking into account changes required in the API and the cost of some reconfiguration strategies.

Keywords- DRTSJ, real-time Java, real-time middleware, predictable reconfiguration.

I. INTRODUCTION

Many real-time infrastructures consider admission control as an offline feature [1, 2][3]. Using accurate analysis tools, (real-time) systems engineers analyze system requirements executing feasibility tests before deploying their applications. These algorithms check if all tasks of the system get satisfied their requirements and generate also a static configuration that can be deployed on top of real-time operating systems and/or real-time middleware. However, many applications are much more dynamic and may benefit from a on-line scheduling analysis integrated within its core ([4][5][6][7] [8]). New nodes and tasks may appear and disappear dynamically without deadline misses. Also, there is an increased need for adaptation and reconfiguration.

This paper addresses a real-time reconfiguration service for distributed real-time Java. Its goal is to include algorithms able to modify the behavior of a whole distributed application by adding and removing groups of tasks. The reconfiguration service proposed is also of interest for the DRTSJ (i.e. the Distributed Real-Time Specification for Java) community (e.g. [9][6][10][11][12][13]) that may include this new service within its core. Currently, DRSTJ is silent on real-time reconfiguration issues.

The reconfiguration service proposed in this paper is similar to the reconfiguration strategies proposed in [5] [8] for the iLAND middleware. Both approaches aim at producing online admission control for predictable end-to-end response-times. However, their contexts and goals are rather different. While iLAND [14] is described in the context of SOAs (Service Oriented Architectures) with multiple service implementations for C-applications; the proposed service is focused on distributed real-time Java applications.

Thus, the goal of the service in distributed real-time Java is to produce a common framework useful for real-time [re]configuration. The rest of the paper describes the approach and evaluates a simple [re]configuration strategy. Section 2 describes the context of the reconfiguration service. Section 3 relates its performance on a networked infrastructure. Section 4 connects the work to the state of the art. Finally, Section 5 introduces conclusions and outlines related work.

II. RECONFIGURATION SERVICE

The service proposed is described for DREQUIEMI a framework for distributed real-time Java, and its particular API ([11]). The extensions and techniques described in this section may be included without major changes into DRTSJ (The Distributed Specification for Real-time Java).

A. Distributed Real-Time Java middleware

DREQUIEMI (Figure 1 and Table I) divides the system distribution infrastructure into a set of layers that contribute to hide distribution issues. The model identifies three main resources that may be managed (namely: memory, CPU, and the network). These services are accessed via an infrastructure middleware (which could be supported via a real-time Java virtual machine) which accesses these resources. On top of the infrastructure middleware, two new layers (distribution and common services) are in charge of providing standard services for distribution. The list includes stub/skeleton in charge of real-time invocations ([15] [16]); DGC in charge of distributed garbage collection, naming for a white page service; and synch events for end-to-end event models ([17][18]). The common services use three main pools (connection, thread, and memory area-pools [19]) to parameterize the behavior of these services. Globally, resource allocation is managed by three managers: distributed memory manager, processor manager and distributed connection manager. On top of common services, there is application subsystem which divides applications into a set of reusable components (currently addressed with OSGi technology [20][21]).

In DREQUIEMI, there are three natural reconfiguration points. The first is placed at the infrastructure level, as a means to control locally memory processor and network. The second extends this control to a distributed scheme and it is placed at the distributed
manager which controls distributed applications performance. The last level of control is at component level performing adaptations taking into account the components of an application.

<table>
<thead>
<tr>
<th>Resources</th>
<th>Infrastructure middleware</th>
<th>Distribution middleware</th>
<th>Common services middleware</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>Network</td>
<td>Connection pool</td>
<td>Stub/Skeleton</td>
<td>Comp. manag.</td>
</tr>
<tr>
<td>Processor</td>
<td>Network</td>
<td>Processor pool</td>
<td>DGC</td>
<td>Comp. proc. manag.</td>
</tr>
<tr>
<td>Network</td>
<td>Network</td>
<td></td>
<td>Scheduling</td>
<td>Comp. comm. manag.</td>
</tr>
</tbody>
</table>

Figure 1. DREQUIEMI’s architecture

The service analyzed in this section is at distribution level, it manages adaptation at distribution level by controlling resource in all distributed nodes. Internally, it uses the resource manager included in each node. It also offers support to the component manager defined at the application level of DREQUIEMI.

B. Extension for Reconfiguration Service

Figure 2 shows the API changes introduced to accommodate the reconfiguration service in the DREQUIEMI ecosystem, which includes new APIs for centralized and distributed real-time Java.

![Diagram showing API changes](image)

Figure 2. New classes proposed for the reconfiguration service

In centralized Java, the main extension required is network support. Currently RTSJ [22], the main approach in real-time Java, does not take into account the existence of different types of networks. In the support proposed for reconfiguration two new elements are included to introduce the network into the system. The first is a characterization for networks similar to memory parameters and scheduling parameters currently included in RTSJ via a tagging interface (NetworkParameters). All network classes include a tagging interface that extends scheduling parameters list included in Table I with network parameters. Network parameters are domain dependant and may consist of a simple non preemptive priority for the router. (e.g., in [23] the authors used a non-preemptive model with network release times, and scheduling parameters to model a prioritized IP-based Switched-Ethernet.) The possibility of including additional networks is added via a new scheduler template named ResourcesScheduler.

Then, the distributed scheduler API extends the API scheduler interface of RTSJ to be accessible as a remote object via DRTSJ. This simple mechanism is enough to allow the allocation of schedulable entities from another remote node. The code for the extension is provided in Listing 1. In all this code all the parameters of the remote object have to be serializable in order to be able to be transferred via the network. (Note: a way to transfer this non-functional information among in distributed remote invocations is described addressed in [24] [25]).

From the point of view of the definition of a reconfiguration service (ReconfigurationService in Listing 2), the approach analyzed in this article is simple.
It is based on modeling the distributed scheduler (described in Listing 1) as an RTSJ’ schedulable object. This simple approach enables to define generic real-time parameters for the reconfiguration service.

Listing 1. API interface for the distributed scheduler

```java
01: public interface DistributedScheduler
02:     extends java.rmi.Remote {
03:         boolean addFeasible(Schedulable s,
04:             Schedulable rs,
05:             ResourceParameters rp,
06:             NetworkParameters np,
07:             NetworkParameters np )
08:         throws RemoteException;
09:         boolean removeFeasible(Schedulable s)
10:         throws RemoteException;
11: }
```

Listing 2. API interface for the reconfiguration service.

```java
01: public interface ReconfigurationService extends
02:     DistributedScheduler, Schedulable {
03:     MemoryParameters getMemoryParameters();
04:     throws RemoteException;
05:     ProcessingGroupParameters getProcessingGroupParameters();
06:     throws RemoteException;
07:     ReleaseParameters getReleaseParameters();
08:     throws RemoteException;
09:     Scheduler getScheduling();
10:     throws RemoteException;
11:     SchedulingParameters getSchedulingParameters();
12:     throws RemoteException;
13:     NetworkParameters getNetworkParameters();
14:     throws RemoteException;
15:     void setMemoryParameters(MemoryParameters mp)
16:     throws RemoteException;
17:     void setProcessingGroupParameters(ProcessingGroupParameters ppg)
18:     throws RemoteException;
19:     void setReleaseParameters(ReleaseParameters rp)
20:     throws RemoteException;
21:     void setScheduler(Scheduler sch)
22:     throws RemoteException;
23:     void setSchedulingParameters(SchedulingParameters sp)
24:     throws RemoteException;
25: }
```

By using this approach, the API of the reconfiguration could be bounded time. It offers mechanisms to limit declare the amount of CPU, memory, bandwidth in the network required for the reconfiguration process. Furthermore, the reconfiguration process could be globally modeled as a periodic, or sporadic process which may scheduled with the remaining tasks in each remote node.

Lastly, notice that by default there is not a dominant reconfiguration algorithm or technique. This is a distinctive feature of the scheduler of RTSJ too. In both cases the idea is that different algorithms may extend the basic tasking template to include system requirements. A developer may also extend the basic reconfiguration policy to take into account different application domains.

C. End-2-End Flow-Shop Reconfiguration Use-Case

The use case, which is empirically explored in the evaluation of Section III, is the simple end-to-end flow-shop model used by Sun [26], Tindell [27], and Palencia [28]. In this model, an end-to-end restriction is modeled (see Figure 3) as a sequence of subtasks that execute in order (i.e. with precedence constraints). Each end-to-end transaction ($T_i$ | $j \in 1..M$) is defined as follows:

- Global deadline ($D_i$)
- A global period ($T_i$)
- A set of $s_i$ schedulable segments ($S_i | \cdots | S_i$)

A local execution priority ($P_a$) for each end-to-end transaction segment ($S_a$).

A local worst-case execution time ($C_a$) for each segment ($S_a$).

In addition, each node runs a periodic enforcer [29] that allows analyzing each flux in a node as it was local.

![Figure 3. Reconfigurable Priority-Based End-To-End transactions classes for the reconfiguration service. It consists of five schedulable segments.](image)

III. EMPIRICAL EVIDENCES

An empirical evaluation of the mechanism described in Section II was carried out. The goal pursued in this
evaluation was to obtain a reference performance for the system. This performance has been evaluated on a networked infrastructure. The evaluated application consists of N Java nodes (Figure 4) that host distributed applications. There is a single local area network (LAN) supported with an IP-router that connects nodes one each other. The router enforces a priority driven policy via end-to-end communications.

In the example of Figure 4, the application consists of two end-to-end transactions ($A_1$ and $A_2$). $A_1$ is made of five consecutive schedulable segments (one local in node 1, another two for the network and another at node 3). $A_2$ is allocated in Node1. All updates are done via the manager node which adds locally the two applications and transfers this information to the required nodes.

These results are complemented considering the cost of sending setup information from the manager node to the subscribed nodes. Figure 7 shows the total cost of creating a remote entity in a local node considering a different number of schedulable entities.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Evaluated scenario (All 796Mhz DREQUIEMI- 100 Mbps Ethernet)}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Minimum Cost of partitioning the system dependence using DREQUIEMI (796Mhz-100Mbps at manager node)}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{Task instantiation (796Mhz-100Mbps) cost, locally in a JVM node}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure7.png}
\caption{Remote allocation costs (796Mhz-100Mbps) using DREQUIEMI}
\end{figure}

\subsection*{A. Cost Tributaries}

The results (Figure 5) show a strong dependence among the number of schedulable entities (tasks) and the time required to carry out their deployments. In the experiments 4096 schedulable tasks keep the cost over 1 second. The minimum cost, which happens with one task and one processor, keeps it in below 100 µs. There is also a linear dependence among the number of processors and the time required to run the feasibility algorithm. This dependence is more moderate than the existing one in the number of JVMs.

The previous results (Figure 5) do not take into consideration the cost of allocating tasks in a remote node. Figure 6 shows the cost of allocating a number of schedulable objects in a JVM node. It considers three different schedulable segments: threads ($R_{thread}$), event handlers ($AsyncEventHandler$), and remote objects ($R_{tRO}$).

\subsection*{B. Total Configuration Time}

The total (re)configuration time is calculated as the time consumed deciding in which node is allocated each schedulable part of an application and the cost of deploying this task. Figure 8 shows this time for the given configuration (i.e. considering [4-32] JVMs (Java Virtual Machines) and [1-4096] $AsyncEventHandler$ schedulable segments). The results showed that the service is not able to perform 1024 assignments and deployments per second, even when there is a single JVM that has to host all tasks.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure8.png}
\caption{1,000-10,000 tasks costs (796Mhz-100Mbps) using DREQUIEMI}
\end{figure}
IV. RELATED WORK

This section focuses on middleware aspects. It goes through the real-time middleware reviewing the way different architectures have dealt with similar problems.

DRTSJ [9, 30], the leading effort towards a distributed real-time Java, technology is focused on a scheduling framework to accommodate the distributed real-time thread. The proposed real-time reconfiguration service may be adapted to the distributable threads used by DRTSJ to configure a global system dynamically. York’s approach [31] may also use the real-time reconfiguration service to carry out on-line configurations. From the point of view of the three profiles defined by UPM [32], the described reconfiguration service is useful for its flexible profile.

Recently, researchers working in RT-CORBA have described a distributed real-time manager [33, 34]. This distributed real-time manager connects different ORBs with a task manager in charge of allocating components to ORBs and performing load balancing and admission control. The described real-time reconfiguration service is closer to their model, i.e. both carry out admission control dynamically.

In RT-SOA (Service Oriented Architecture), it stands out the work carried out in real-time service composition ([4,35][5][8][14]), which deals with the problem of dynamic allocation of services with multiple available implementations. The authors addressed the problem of finding an optimum system configuration with composition algorithms, which have heuristics and figures of merit. The use of these techniques allows reducing the overhead of executing these algorithms several orders in magnitude. The proposed real-time reconfiguration service could be extended with these parameters.

An eventual piece of work is the distributed transaction manager (DTM) [36]. This service extends the contract model of FRESCOR [37] from a local scenario to a distributed system. The main difference between DTM and the service designed in this paper is in terms of entities negotiated and negotiation algorithms. Nevertheless, the algorithms described for the configuration service could be also included in DTM and vice versa.

V. CONCLUSIONS AND FUTURE WORK

Next generation real-time systems will benefit from having an enhanced infrastructure able to cope with predictable reconfiguration. The proposed reconfiguration service provides a template for a reconfiguration service in distributed real-time Java. The template may be extended to include different strategies and different kinds of applications. The empirical evaluation of a simple configuration strategy showed total reconfiguration costs that may be in hundreds milliseconds for a small number of nodes and schedulable segments.

Our ongoing work is focused on mode change protocols ([38, 39]). Implicitly, the addressed configuration service assumes that applications may be stopped and started at runtime without requiring configuration from the user. However, this is not realistic because reconfiguration is application-dependent in general. One solution to this problem is mode change protocols, which carry out this activity taking into consideration application nature. Additionally, the authors are exploring reconfiguration for real-time OSGi platforms (like [20, 40][41][42]).

ACKNOWLEDGEMENTS

This research was partially supported by the European Commission (ARTIST2 NoE, IST-2004-004527; iLAND ARTEMIS-JU Call 1) and by the Spanish national project REM4VSS (TIN-2011-28339).

REFERENCES


