Extending Policy Languages for Expressing the Self-Adaptation of Web Services

Haithem Mezni
(University of Jendouba, SOIE, Jendouba, Tunisia
haithem.mezni@gmail.com)

Walid Chainbi
(National School of Engineers, SOIE, Sousse, Tunisia
walid.chainbi@gmail.com)

Khaled Ghedira
(Higher Institute of Management, SOIE, Tunis, Tunisia
khaled.ghedira@isg.rnu.tn)

Abstract: With the growing demand on Web Services, self-adaptation in the highly-dynamic environment is becoming a key capability of service-based systems. As a solution for Web services to provide added value and high QoS, combining self-* and policies allows reducing management complexity and effectively drives adaptation. Also, providers must participate in the self-adaptation process as they are aware of the capabilities of their offered services and exceptions that may occur. Despite the important role of service providers, existing approaches did not address this major issue. Thus, the description of self-adaptive Web services must not be limited to functional and QoS data. To address these issues, we extend the WS-Policy framework to represent capabilities and requirements of self-* Web services. We also extend UDDI in order to store and manage service policies, as the current UDDI model does not offer these capabilities. Finally, we propose an ECA-based planning mechanism to specify decision making in the self-adaptation process.

Keywords: Web services, self-adaptation, Autonomic computing, Quality of service, WS-Policy, UDDI, ECA rules
Categories: C.2.4, H.3, H.3.5

1 Introduction

Modern distributed systems require to dynamically take into account at runtime the changes in the users’ needs and the execution environment variations in order to improve QoS [Zouari et al. 2014]. In this context, and with the increasing adoption of SOA, service-based systems face an unprecedented level of change and dynamism. Therefore, there is a growing need for efficient mechanisms for service self-adaptation. Although extensive work has been done in this area, existing approaches usually use predefined adaptation strategies to deal with abnormal behavior and do not consider all properties of SOA actors that are subject to change [Kazhamiakin et al. 2010]. Moreover, abnormal behavior should be seen from the perspective of service providers as they are aware of the capabilities and requirements of their published services and the exceptions that may occur. In order for Web services to
provide high QoS, providers must participate in the self-adaptation by implementing their specific adaptation policies. Thereby, relieving adaptation system from the situation where no recovery alternative is applicable in a given problem situation.

To tackle these issues, we adopt the autonomic computing paradigm [Kephart and Chess 2003] and we consider Web services and UDDI registries as autonomic systems [Chainbi et al. 2012] i.e., systems endowed with self-* capabilities including self-configuration, self-healing, self-optimization, and self-protection. Additionally, to effectively drive adaptation, we propose a rich information model to allow describing autonomic Web services (AWS) based, not only on functional and QoS data, but also on providers’ recommendations. This new kind of information may be used at runtime and gives the adaptation system additional alternatives (i.e. adaptation plans) to manage the executing service and to allow a better precision in driving adaptation.

Since managing Web services requires a structured representation of their capabilities, we adopt a policy-based approach in order to endow services with self-* behavior. Policies have proven their popularity and acceptance in research and academia. They have long been employed in the management of traditional distributed systems [Boutaba and Aib 2007] and become a key for SOA management. In our work, policies have a core position. They are used at different levels including service description, QoS monitoring, service self-adaptation, etc. The adoption of policy specification languages has been suggested in several approaches as a solution to describe QoS. Specifications like WS-Policy [Bajaj et al. 2007] are widely deployed to address SOA issues, as they offer a means to express QoS requirements and capabilities in Web service systems [Phan et al. 2008]. However, existing extensions to WS-Policy [Tosic et al. 2007, Chhetri and Kowalczyk 2010, Badidi and Esmah 2011] did not express information necessary to service self-adaptation. Also, there is no detail about self-* policies and how policies are specified and processed.

The main objective of this paper is to extend WS-Policy in order to express requirements for the run time adaptation of Web services. However, knowing that WS-Policy provides only textual descriptions, this raises a question about how to use adaptation policies in the self-adaptation process. For this reason, a part of our proposed work uses the ECA rule-based approach to convert adaptation policies to executable ECA rules [Bassiliades and Vlahavas 1997]. ECA rules are extracted from the policy documents and can, then, be executed by autonomic service managers.

The rest of this paper is organized as follows. Section 2 introduces a rich AWS information model. Section 3 presents our extension called AWS-Policy. Section 4 presents an extension to UDDI information model. In section 5, we show how specific adaptation policies are used in the self-adaptation process. Section 6 validates the proposed approach through a set of experimental results. In section 7, we discuss the relevance of our approach compared to related work. The last section is devoted to the conclusion and the future work.

2 Autonomic Web Service Information Model

In our previous work, we have shown that an AWS involves two parts: a managed service and an autonomic manager, which endows the invoked service with self-* mechanisms (see [Fig. 1]). AWSs collaborate with autonomic registries (autonomic systems composed of the traditional UDDI registry and a controller named autonomic...
registry manager, which has the ability to manage the registry content) to ensure self-adaptation and to allow updating registry content [Chainbi et al. 2012].

However, to effectively ensure self-adaptation, an AWS must have a rich description of its capabilities. We believe that functional and non-functional data are not sufficient to effectively drive self-adaptation, and AWSs need to be specified with additional information, such as specific adaptation policies. Unlike other approaches that design their management systems based on predefined management rules, we propose two types of adaptation policies and we give service providers the possibility to define and register their recommended adaptation policies into UDDI registry, in order to effectively drive the autonomic managers in the self-adaptation process. In our work, a service behavior may be adjusted using what we call default and/or specific adaptation plans. While default plans are predefined actions included in the implementation of autonomic managers as internal knowledge and executed to adapt any failed service, specific plans are priority actions defined by providers at publication time (e.g. specific mediation solutions, invoking a trusted service).

For this purpose, we characterize self-* Web services by three types of information: Functional properties, QoS properties, and Specific adaptation plans. Since managing Web services requires a structured representation of their capabilities, we adopt the WS-Policy framework to treat them as policies [Bajaj et al. 2007]. WS-Policy is a specification that allows Web Services to advertise their characteristics in a flexible and extensible grammar using XML format. A policy is defined as a collection of alternatives which is, itself, defined as a collection of assertions. Assertions are the basic building blocks of policies and are used to represent a requirement, capability or a behavior of a Web service.

However, WS-Policy does not allow defining capabilities and requirements of
self-adaptive Web services, and still lacks mechanisms for semantic description and matching of service properties. We believe that the more Web services are attached with QoS and adaptation policies, the more important is the automation of the service execution and adaptation. Indeed, by taking these policies into account, a Web service behavior is always adjusted based on self-* capabilities (e.g. self-healing policies) allowing, thus, a constant execution, even in case of QoS constraints’ violation, and a high availability while meeting QoS constraints (e.g. response time and reliability).

To remedy these restrictions, we propose to extend WS-Policy in order to represent AWS capabilities and requirements necessary to trigger self-adaptation and to adjust the behavior of a managed Web service.

3 AWS-Policy: An Extension for Autonomic Web Services’ Description

The AWS-Policy specification allows providers to define the behavior of their offered services. It makes Web services expose autonomic behaviors in response to changes detected in the execution environment and predicted by a service provider (e.g. self-healing and self-configuration behavior, in case of substitution and mediation actions). Each construct in AWS-Policy is considered as a service behavior (e.g. unavailability QoS attribute, timeout event, SLA Negotiation action, bandwidth context). Separation of concern and flexibility of AWS-Policy allows the services’ implementations to focus only on the functional capabilities, whereas policies are responsible for providing and guaranteeing a certain behavior and a constant service execution, while taking into account the changing conditions and the required QoS.

AWS-Policy offers an extensible grammar based on XML format. The building blocks in AWS-Policy are assertions, which allow defining a QoS preference, a QoS capability, an event or a self-adaptation action. Assertions are grouped using a set of policy operators which can be recursively nested to express complex policies.

3.1 QoS Policies

This section presents the AWS-Policy document model that may be used by service providers to describe the QoS capabilities of their offered Web services, or by users to specify their QoS requirements. The QoS policies XML schema is shown in [Fig. 2].

![Figure 2: QoS policies XML schema.](image-url)
As shown in [Fig. 2], an AWS must be described by a set of QoS attributes. Each one is described in a \textit{QoSAssertion} element and is characterized by three properties (\textit{name}, \textit{value}, and \textit{unit}) containing, respectively, the name, the value, and the measure unit of the QoS attribute. Since providers may describe their services using complex QoS attributes, we define the \textit{QoSAll} alternative element to group the QoS attributes into a composite one. Fig. 3 shows an example of AWS-Policy QoS document.

![Image](http://example.com/qos-policy-example.png)

\textbf{Figure 3: QoS policies of the FileSendWS Web service.}

An overall view shows that the offered \textit{FileSendWS} service is a free Web service characterized by a high degree of performance and availability.

\section*{3.2 Specific Adaptation Policies}

As shown in [Fig. 4], a specific adaptation policy is defined as a set of plan alternatives. Each one contains two policy alternatives: \textit{ExactlyOneEvent} and \textit{ExactlyOnePlan}. The first element specifies the events that may affect service execution, whereas the second element describes specific adaptation plans that may be triggered by the events defined in the \textit{ExactlyOneEvent} alternative.

AWS-Policy allows specifying execution events in a flexible manner. Indeed, self-adaptation may be triggered by a primitive or a composite event. For primitive events, we define the \textit{EventAssertion} element which may contain a set of attributes depending on the mentioned event. Composite events are of type “AND” ($e_1 \land e_2 \land \ldots e_n$). They require that all of the sub-events must occur during service execution so that they may trigger some potential adaptation actions. A composite event is expressed using the \textit{AllEvent} policy alternative. Since different events may trigger the same adaptation actions, we also consider the events of type “OR” ($e_1 \lor e_2 \lor \ldots e_n$) which means that at least one of these events must occur during invocation so that it may trigger adaptation. Such events are expressed using the \textit{ExactlyOneEvent} alternative.
The second part of the self-adaptation policy is an ExactlyOnePlan alternative defining self-adaptation actions that may be executed when one of the events described in the ExactlyOneEvent alternative occurs. The ExactlyOnePlan element indicates that only one specific plan must be executed by the autonomic manager. We define the ActionAssertion assertion to describe an atomic self-adaptation action. Since an adaptation plan may be composed by a set of adaptation actions, it is expressed by grouping ActionAssertion elements in the AllPlan alternative. The AllPlan alternative requires all primitive actions to be executed in the specified order.

The priority attribute in the XML schema is used by a provider to indicate that a specific adaptation action has to be firstly executed by the autonomic manager, or executed when the default adaptation plans fail to resolve a service failure. This allows a better precision in choosing adaptation actions. Another important attribute is the planCombiningAlgo which is used by planning agents to select between more than one suitable plan and to resolve conflicts between self-adaptation plans. Details on plan combining mechanisms will be presented in section 5.2.

The AWS-Policy document in Fig. 5 describes specific adaptation policies of a published service. An overall view shows that the service provider defines a set of self-adaptation plans for two possible events. The first set of plans may be triggered when a “binding fault” (lines 8 to 13) is detected by the autonomic manager or when the executing service is not available. The second plan may be triggered in case of an “execution fault” (see “response error” or “timeout” in lines 30 and 31).
Fig. 5 also shows that, in case of a detected binding fault (binding denied may occur if the authorization component denies the access, that is, either authorization denied, authentication failed, or accounting problems occur), the autonomic manager features two possible priority self-adaptation actions: reselect a new composition for the executing composite service or, in case of restarting failure, substitute the service with one of two preferred and trusted services (a mediation action may be performed in case of incompatibility in services’ interfaces). Fig. 5 also shows that the provider recommends re-executing the service operations in case of a delivered incorrect result.

The next section reveals how AWS policies are published in the service registry.

4 Enhancing UDDI Information Model with Policies

In our previous work, we have shown the way autonomic registries interact with each other in the execution environment, to allow exchanging and processing different kinds of management tasks, such as QoS updating and service discovery [Chainbi et al. 2012].

However, representing and storing providers’ policies is a main challenge, as the current UDDI model does not offer these capabilities and as service metadata are basic resources in driving the adaptation process. The problems of storing QoS and specific adaptation policies in UDDI may be resolved using tModels concept. tModels are core components of UDDI. They represent unique concepts or constructs and they are used to
describe compliance with a specification, a concept, a category, or an identifier system. Each tModel should contain an overviewURL, which references a document that describes the tModel and its usage in more details [OASIS 2004]. We use tModels as a categorization solution to facilitate service discovery by the autonomic registry managers (also called ARMs). Since tModels are considered as containers for references to service descriptions, they may be used to represent QoS policies and adaptation policies in UDDI.

Fig. 6 shows an extension of the original UDDI information model with two new data structures for describing AWS related information. Unshaded boxes and solid lines represent data structures in the original model and the association between them, whereas shaded boxes and dashed lines illustrate our extension. QoS policies data structure represents QoS information of a managed Web service, whereas specific adaptation policies data structure represents actions defined by a service provider that will eventually be used in priority to manage service deviations.

![Diagram of Extended UDDI Information Model](image)

*Figure 6: Extended UDDI information model.*

To these ends, when a Web service is published in the UDDI registry, a set of tModels is created to represent the QoS and specific adaptation policies of this service. To include Web service policies in UDDI, different ways may be envisioned, such as defining a tModel that points to an external resource or defining a tModel containing multiple categories, each reflecting a different policy. The first method has the disadvantage of requiring navigation to external resources to retrieve information. In addition, service data are not stored in UDDI in this case, which makes their management a difficult task. This is the case, for example, when services’ providers may restrict access to their servers, prohibiting ARMs to update QoS values.

We adopt the first method to store specific adaptation policies, and we integrate QoS policies using the second method. This choice is justified by the fact that specific adaptation policies do not interest the user in the discovery process and are only processed by autonomic managers. Furthermore, they do not require a representation
or a classification scheme like QoS policies, which need to be stored in the registry in order to facilitate service discovery. For this purpose, specific adaptation policies are considered as external resources and referenced in the `overviewURL` tag of a tModel. The UDDI `bindingTemplate` may, then, contain a reference to these policies, which is added to its `tModelInstanceDetails` collection.

Regarding QoS, we make a mapping between QoS policies and UDDI, in order to give service providers a way to store their QoS descriptions into UDDI registry. The QoS policies description is referenced in a tModel and each QoS assertion in the policy document is mapped to a category in the tModel structure (see [Fig. 7]). In UDDI, `categoryBag` tag allows to add categorization information into UDDI data structures to make a given entity a member of one or more categories. Therefore, ARMs can find the desired entity based on some classification scheme. In our work, categorization may be ensured on the basis of QoS policies. The `categoryBag` element in the QoS policies tModel acts as a collection of `keyedReference` structures (line 6), each containing a single QoS attribute-based categorization and characterized by a set of properties (keyName, keyValue, and keyUnit) describing respectively the name, value and measure unit of a QoS attribute (lines 11, 12, and 13).

![Figure 7: tModel with QoS policies.](image)

```xml
1  <tModel tModelKey="uddi:anyCompany:FileSend PrimaryBinding QoS Policies"/>
2    <name>QoS policies for FileSend</name>
3          <overviewDoc>
4            <overviewURL>http://localhost:3080/policies/FileSendQoS.cpdx</overviewURL>
5          </overviewDoc>
6    </tModel>
7      <categoryBag>
8        <keyedReferenceGroup tModelKey="uddi:uddi.org.QoSGroup performance">
9          <keyedReference>
10             tModelKey="uddi:uddi.org.QoSPolicy.execution"
11             keyName="execution"
12             keyValue="1830"
13             keyUnit="ms"/>
14          <keyedReference>
15             tModelKey="uddi:uddi.org.QoSPolicy.response-time"
16             keyName="response-time"
17             keyValue="1000"
18             keyUnit="ms"/>
19          </keyedReferenceGroup>

20      .......
21      .......
22          <keyedReference>
23             tModelKey="uddi:uddi.org.QoSPolicy.cost"
24             keyName="cost"
25             keyValue="True"/>
26      </categoryBag>
27  </tModel>
```

To categorize a managed service as being described by a composite QoS attribute, we propose to add a `keyedReferenceGroup` element (lines 7 and 8) to the `categoryBag`. This element is a collection of `keyedReference` structures that logically belong...
together and correspond to the component QoS attributes for the categorization.

5 Self-adaptation Planning: from AWS-Policy to ECA Rules

In this section, we propose a planning technique that allows driving Web service self-adaptation based on predefined and specific self-adaptation policies. Despite the importance of this phase in the autonomic cycle, few approaches have been proposed to deal with the problem of autonomic decision making [Lu 2011].

As we mentioned above, providers can define their specific self-adaptation plans in the form of policies. This raises a question about how to use these policies by the autonomic manager to adapt a service. The effective use of adaptation policies in the autonomic cycle requires that the policies be captured and translated into actions within the autonomic manager. Since adaptation policies are textual descriptions, we propose to convert them to executable rules that may be executed by autonomic managers on their rules’ engines. Rules are increasingly being used to specify a variety of situations, such as business needs, conventional behavior, and policies [Sing and Huns 2005]. Rules are also desirable because they are executable, unlike textual descriptions or even some formal specifications. There is no additional step of converting specifications into formally executable implementations. We adopt the classical ECA (Event-Condition-Action) form [Bassiliades and Vlahavas 1997], where the event is a triggering condition. Such rules can be read as:

\[
\text{on event if condition then (perform) action}
\]

As an example of policies’ transformation, the policy elements contained in the AWS-Policy document in [Fig. 5] will be converted as follows:

| If the binding is denied or the authentication fails, then composition should be reselected instead of substituted whenever possible. |

Here, binding fault or authentication failure are the events. The condition is true (\textit{priority}="true"), and the action is re-selection of the composition (line 16) or substituting the failed service with specific one (lines 19 and 22).

Next, we present a transformation algorithm that allows generating self-adaptation ECA rules from the AWS-Policy documents.

5.1 Specific Adaptation Policies Transformation

In order to effectively monitor and adapt the executing Web service, the autonomic service manager starts by interacting with the autonomic registry (see [Fig. 1]) to get the AWS-Policy documents describing QoS and specific self-adaptation policies. This step is important as the autonomic manager must be aware of the capabilities of its managed service. Self-adaptation plans defined by the provider are then extracted and stored in the form of ECA rules. Connection between possible events (AWS-Policy constructs) and the corresponding adaptation actions is established when generating the different parts of each new rule. First, a specific self-adaptation rule is defined as follows:
Definition 1 (Self-adaptation rule set). A self-adaptation rule set for a managed Web service, SARS\textsuperscript{WS}, is a finite set of ECA rules, \( R = (E, C, A, P, RCA) \) such as:

1. \( E = \{ e | e = e_1; \ldots; e_n \} \). Event set \( E \) consists of one or more serializations of events, where the serialization is expressed as \( e_1; \ldots; e_n \).
2. Condition set \( C \) consists of the disjunction of zero or more conjunctions of condition predicates. Each conjunction is expressed as \( c_1 \land \ldots \land c_m \), and condition predicate \( c_i \) is expressed as \( r_1 \theta r_2 \), where \( \theta \) is an operator from the set \{=, \neq, <, \leq, >, \geq \}, and \( r_i \), \( i = 1, 2 \), is a constant, an event variable, or a QoS attribute.
3. Action set \( A \) consists of a serialization of primitive self-adaptation actions, where the serialization is expressed as \( a_1; \ldots; a_n \).
4. Priority \( P \) is a Boolean variable indicating whether or not the primitive actions in \( A \) are executed in priority instead of triggering default self-adaptation rules.
5. Rule combing algorithm \( RCA \) defines a strategy of combing a set of concurable rules in a single decision. \( RCA \) indicates if the self-adaptation rule overrides the other rules in the SARS\textsuperscript{WS} rule set.

The SARS\textsuperscript{WS} rule set contains two self-adaptation sub-sets: the SARS\textsuperscript{WS}\textsuperscript{specific} rule set which contains specific self-adaptation rules transformed from an AWS-Policy document defined by the provider, and the SARS\textsuperscript{WS}\textsuperscript{default} rule set containing self-adaptation actions predefined for each autonomic manager instance.

Definition 2 (Self-adaptation constraints). Self-adaptation constraints SAC\textsuperscript{WS} for a managed Web service is a power set of the self-adaptation actions’ set such that \( AC = \{ a_1, \ldots, a_n | a_1 \land \ldots \land a_n \text{ are not allowed to be executed simultaneously} \} \in SAC\textsuperscript{WS} \).

Definition 3 (Rule conflict). Let SARS\textsuperscript{WS}\textsuperscript{rig} be a rule set containing the rules that are triggered on the occurrence of a primitive event. It is said that SARS\textsuperscript{WS}\textsuperscript{rig} has a rule conflict if two or more service actions of SARS\textsuperscript{WS}\textsuperscript{rig} are contained in one of the element of SAC\textsuperscript{WS}.

Our algorithm for specific self-adaptation policies transformation (see [Fig. 8]) starts with an event alternative set EAS\textsuperscript{WS} and a plan alternative set PAS\textsuperscript{WS}, as inputs. The output of the algorithm is a specific self-adaptation ECA rule set SARS\textsuperscript{WS}\textsuperscript{specific}.

The algorithm starts by parsing the set of events that may trigger self-adaptation plans recommended by the service provider (line 2). For each event, the algorithm creates the event part of the new ECA rule, by transforming event assertions and alternatives to simple (line 12) or composite events (lines 4-7). Next, the algorithm extracts, for each generated event, the condition part, according to the constraints of the specified events (e.g. “response time” subject for the “timeout” event) (lines 8-9). Note that the priority attribute in the plan alternatives is considered as a condition when executing the policies transformation process (line 25). Then, for each plan in PAS\textsuperscript{WS}, the action part of the new rule \( R_{\text{new}} \) is created according to the nature of the plan alternative, by taking the union of the simple actions specified in each plan in PAS\textsuperscript{WS} (lines 20-23). Finally, the algorithm, based on the set of self-adaptation constraints (i.e. adaptation actions that are not allowed to occur at the same time such as re-execute operation and skip operation) checks if the new rule has any conflict and eventually creates a conflict resolution rule for the detected conflict (line 27).

Performance of the policy transformation algorithm depends on the size of the events and actions’ sets, and on the self-adaptation constraints’ set. Time complexity
of the algorithm is \(O(n_e n_a n_c)\), where \(n_e\), \(n_a\) and \(n_c\) denote the number of events, the number of adaptation actions and the number of adaptation constraints, respectively. Function checkConflict (\(R_{\text{new}}, SARS^{WS}\)) requires \(O(n_a n_c)\) since for each adaptation action, the function checks the conflict with each action in the constraint set.

Algorithm 1 Specific_SelfAdaptation_Policies_Transformation (in \((EASWS, PASWS)\), out \((SARS^{WSpecific})\))

<table>
<thead>
<tr>
<th>Input: An event alternative set (EASWS) and a plan alternative set (PASWS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output: A provider defined self-adaptation rule set (SARS^{WSpecific}), where (SARS^{WSpecific} := { (Evt, Cond, Act, Pr, RCA)</td>
</tr>
</tbody>
</table>

1: \(SARS^{WSpecific} := \{\}\);
2: for each \(E \in EASWS\) do
3:   Cond := \{\};
4:   if \(E\.Type = \text{ALL}\) then
5:     Evt := \{\};
6:     for each \(EA \in E\) do
7:       Evt := Evt \(\cup EA\);
8:       \(C := \text{new Condition}(EA\.Subject, EA\.SubjectValue)\);
9:       Cond := Cond \(\cup C\);
10: end for
11: else
12:   Evt := \(E\);
13:   Cond := \text{new Condition}(Evt\.Subject, Evt\.SubjectValue);
14: end if
15: for each \(P \in PASWS\) do
16:   \(R_{\text{new}} := \text{new ECARule}()\);
17:   \(R_{\text{new}}\.setEvent(Evt)\);
18:   \(R_{\text{new}}\.setCondition(Cond)\);
19:   \(Act := \{\}\);
20:   for each \(A \in P\) do
21:     \(Action\ Ac := \text{new Action}(A\.Name, A\.Subject, A\.SubjectValue)\);
22:     Act.addAction(Ac);
23:     end for
24:   \(R_{\text{new}}\.setAction(Act)\);
25:   \(R_{\text{new}}\.setPriority(P\.Priority())\);
26:   \(R_{\text{new}}\.setCombiningAlgo(P\.CombiningAlgo())\);
27:   \(checkConflict(R_{\text{new}}, SARS^{WSpecific})\);
28:   \(SARS^{WSpecific} := SARS^{WSpecific} \cup R_{\text{new}}\);
29: end for
30: end for
31: return \(SARS^{WSpecific}\);

Figure 8: Algorithm Specific_Adaptation_Policies_Transformation.

5.2 Plan Selection and Conflict Resolution of Self-adaptation ECA Rules

As shown in the previous section, specific self-adaptation policies are transformed to ECA rules to be used by autonomic managers in the self-adaptation process. As a result, some rules may have inconsistencies with each other or with the predefined self-adaptation plans and may cause conflicts when triggered and executed.

To select the suitable self-adaptation plan, we propose a solution that allows reconciling self-adaptation policies when their evaluation is contradictory. The proposed AWS-Policy specification supports what we call plan combining algorithms, which allow resolving self-adaptation rule conflicts by representing a way of combining multiple adaptation plans into a single plan. The proposed method is based
on “event dominance” and “self-adaptation rule (plan) overriding”. Event dominance is defined to analyze the relationship between the events that occur during service execution. It is used to identify the rules that can be triggered together by a dominant event. Rule overriding consists of combining a set of rules that can be triggered by the same event into a single rule. These basic concepts are defined as follows:

**Definition 4 (Event dominance).** Let $e_i$ and $e_j$ be two serializations of events. $e_i$ is said to be dominant over $e_j$ (represented as $e_i \rightarrow e_j$) if the occurrence of $e_i$ implies the occurrence of $e_j$. That is, if $e_i = e_i^1; \ldots; e_i^n$ and $e_j = e_j^s; \ldots; e_j^t$ ($1 \leq s \leq t \leq n$).

**Definition 5 (Rule overriding).** Let $SARS^{WS}_{aov}$ be a self-adaptation rule set such that $SARS^{WS}_{aov} \subseteq SARS^{WS}$ and contains the rules triggered by occurred events, $SARS^{WS}_{aov} = \{R \mid R = R_1; \ldots; R_n\}$ and let $R_i$ and $R_j$ be two triggered rules. $R_i$ is said to override $R_j$ (represented as $R_i \rightarrow^{\forall} R_j$) if $R_i$ belongs to $SARS^{WS}_{aov}$, $R_i$ is triggered and each rule $R_j$ ($1 \leq j \leq n$ and $j \neq i$) contained in $SARS^{WS}_{aov}$ is rejected. $\forall R_j \in SARS^{WS}_{aov}, \forall R_i \in SARS^{WS}_{aov}, \exists a, a_i, a_j \in A_i \land a \in A_j$. That is, if $A_j = a^j_s; \ldots; a^j_t$, $A_i = a^i_1; \ldots; a^i_n$.

To avoid conflict or violation of service constraints, the autonomic manager must take into account only the rules triggered by the dominant events. If none of the occurred events dominates each other, a technique called “self-adaptation rule overriding” is used to process the plans in order to choose a single plan.

We propose a dynamic conflict resolution algorithm which is based on event dominance and self-adaptation rule overriding. The algorithm is inspired from the combining algorithms defined for XACML access control policy language [XACML Technical Committee 2008]. To manage conflicts between access control rules, XACML supports different combining algorithms, each representing a way of combining multiple decisions, which are often conflicting, into a single decision. Possible decisions are: permit, deny, indeterminate and not applicable.

To this purpose, the planning technique uses the **Priority** and **PlanCombiningAlgo** attributes [section 3.2] to indicate an algorithm for combining the possible self-adaptation plans from the evaluation of a set of rules. Such algorithms can be:

- **Retry-overrides**: If any rule evaluates to retry, then the final generated plan is also re-executing the operations of the managed service. For example, when a detected event triggers three possible plans (e.g. retry, substitute and skip), and if the **planCombiningAlgo** attribute is set to “retry-overrides”, then the autonomic manager decides to re-execute the Web service operation.

- **Substitute-overrides**: If any rule evaluates to Invoke, then the final decision is also substitute Web service.

- **FirstPlan-applicable**: In this case, the self-adaptation rule is picked among the rules; according to the order given in the AWS-Policy document.

Our proposed algorithm for the selection and conflict resolution of self-adaptation ECA rules (see [Fig. 9]) starts with an event set $Evt^{WS}$, a self-adaptation rule set $SARS^{WS}$ and the plan combining algorithm RCA, as inputs. The output of the algorithm is a single decision $R_{select}$ from the multiple self-adaptation rules.

Knowing that the self-adaptation rules’ set $SARS^{WS}$ is triggered by a set of events $Evt^{WS}$, a rule set $SARS^{edom}$, which is included in $SARS^{WS}$, is only triggered by the most dominant events $Evtdom$ (line 1). The rule set $SARS^{edom}$ can be refined to a single self-adaptation rule $R_{select}$, which is the only rule that meets two criteria: “event
dominance” and “plan overriding”. In other words, $R_{select}$ is triggered by the most dominant events and is the result of combining the rules, in $SARS_{dom}$, based on the given combining algorithm (lines 15 or 17).

Algorithm 2 SelfAdaptation_RuleSelection_ConflictResolution (in $(Evt^{WS}, SARS_{dom}$, RCA), out ($R_{select}$))

- **Input:** An event set $Evt^{WS}$, a self-adaptation rule set $SARS_{dom}$ and a rule combining algorithm RCA.
- **Output:** a single self-adaptation ECA rule $R_{select}$ representing a decision from multiple adaptation rules.

1: $Evt_{dom}$ := findDominantEvents ($Evt^{WS}$);
2: if $Evt_{dom}$ ≠ $\emptyset$ then
3: $SARS_{dom}$ := $\{\}$;
4: for each $R$ $\in$ $SARS_{dom}$ do
5: exists := true;
6: for each $e$ $\in$ $Evt_{dom}$ do
7: if not (Trigger ($e$, $R$)) then
8: exists := false;
9: break;
10: end if
11: end for
12: end for
13: if exists then $SARS_{dom}$ := $SARS_{dom}$ $\cup$ $R$;
14: end for
15: $R_{select}$ := PlanCombiningAlgo ($SARS_{dom}$, RCA);
16: else
17: $R_{select}$ := PlanCombiningAlgo ($SARS_{White}$, “FirstPlan-Applicable”);
18: end if
19: return $R_{select}$;

Function PlanCombiningAlgo(SARS, RCA)

1: for each Rule $R$ $\in$ SARS do
2: Act := $R$.Actions();
3: for each $A$ $\in$ Act do
4: if $A$.CombiningAlgo = “FirstPlan-Applicable” OR $A$.CombiningAlgo = RCA then
5: return $R$;
6: end if
7: end for
8: if $R$.Priority = true and $R_{priority}$ = null then
9: $R_{priority}$ := $R$;
10: else if $R_{default}$ = null then $R_{default}$ := $R$;
11: end if
12: end for
13: if $R_{priority}$ ≠ null then return $R_{priority}$;
14: if $R_{default}$ ≠ null then return $R_{default}$;
15: return “Indeterminate”;}

Figure 9: Algorithm SelfAdaptation_RulesSelection_ConflictResolution

The algorithm starts by calling the function findDominantEvents ($Evt^{WS}$) in order to get the most dominant events among the triggering events in the $Evt^{WS}$ event set (line 1). For instance, let $e_1$, $e_2$ and $e_3$ be the occurred events and suppose that a set of rules $\{R_1, R_2, R_3, R_4, R_5\}$ are trigged as follows: $R_1$ by $\{e_1, e_2\}$, $R_2$ by $\{e_1\}$, $R_3$ by $\{e_1, e_3\}$, $R_4$ and $R_5$ by $\{e_1, e_2, e_3\}$. The most dominant events in this case are $\{e_1; e_2; e_3\}$. Consequently, $SARS_{dom}$ includes only the self-adaptation rules $R_4$ and $R_5$ (lines 6-13).

The next step of the algorithm is to check the self-adaptation rules in $SARS_{dom}$. This is by evaluating each resulting rule against the plan combining algorithm. To do that, the algorithm calls the planCombiningAlgo function (line 15) in order to combine the $SARS_{dom}$ rule set ($R_4$ and $R_5$) into a single decision.

Performance of the algorithm depends on the size of the triggered rule set and the
occurred event set. Time complexity of the algorithm is $O(n_r n_e n_a^2)$, where $n_r$, $n_e$ and $n_a$ denote the number of triggered self-adaptation rules, the number of events and the number of adaptation actions, respectively. Creating the $SARS_{dom}$ rule set takes the time $O(n_r n_e^2)$ as the size of $Evtdom$ set is smaller than $n_r$. Function planCombiningAlgo $\text{(SARS, RCA)}$ takes the time $O(n_r n_a)$ as the size of $SARS_{dom}$ is smaller or equal to $n_r$.

6 Implementation and Experimental Results

To validate our goal of making Web services self-*, we have used JADE platform to implement autonomic managers. The autonomic registry is implemented by extending Apache jUDDI which is an open source UDDI implementation compliant with version 2.0 specification. We have used JADE platform and JAX-WS to implement autonomic registry managers as agent-based Web service clients to the UDDI registry. To support AWS-Policy descriptions, we have implemented the main policy management aspects by extending Apache Neethi 2.0, which provides a general framework for developers to use WS-Policy.

We have evaluated the availability, reliability and response time of a set of Web services. We started the invocation without injecting any autonomic capability. Then we performed the same steps by associating an autonomic manager to each service.

Compared with normal execution, self-healing actions (retry or substitute in our case) have reduced the execution failures and consequently have improved the availability of the managed Web service. As an example, the EmailValidator Web service is executed 1000 times and fails to return a result 143 times (see [Fig. 10]). Using an autonomic manager to invoke Web service functionalities has allowed resolving 137 detected deviations from the 143 occurring events. However, the contract violation still happens (6 invocation failures in these experiments). This is explained by the failure of some generated self-healing plans.

![Figure 10: Evaluation of Web service execution with autonomic behavior.](image-url)
Regarding availability and response time (see [Fig. 11]); we can see that, without autonomic behavior, the response time, in most cases, is higher than the required response time (increased by 104 milliseconds). In the same way, availability is not constant when we vary the number of invocations. Self-healing actions have improved the availability of EmailValidator (between 82% and 97%). As shown in [Fig. 11 (b)], during the trials, the availability is increased by 14% and more.

We can also see that self-optimization improves the response time as autonomic manager continuously looks for service opportunities. However, when executing a self-healing action, the autonomic service can exceed the response time fixed in the original contract (e.g. case of 50 and 60 queries). This is explained by the cost of performing such action (i.e. time taken to select and execute self-healing actions).

Figure 11: (a) Web service response time (b) Web service availability.

We have compared our work to the approaches proposed in [Ben Halima et al. 2008, Cardellini and Iannucci. 2010]. In the case of the QOSH middleware [Ben Halima et al. 2008], the execution time with a “service monitor” is higher than the response time returned without a service monitor (increase between 7.5% and 37.22% and in some cases 43.6%). In our work, self-healing actions have increased the response time by 7.2%, only in few cases. Regarding the MOSES system [Cardellini and Iannucci 2010], 75% of the time is wasted in analyzing data, unlike our approach, which transforms self-adaptation policies into executable rules, before starting invocation, to reduce the processing time during autonomic execution.

We have also evaluated the results of using default and specific adaptation plans in the autonomic execution (see [Fig. 12]). We have measured the time taken by autonomic managers to decide about an adaptation and to execute the generated plan (in these experiments, retry execution, skip execution, and substitute a failed service).
We can see that a specific plan–driven adaptation requires a time lower than the
time spent by the ASM to adapt a Web service using predefined plans. For example,
in case of a predefined substitution plan, the ASM has to start a new Web service
discovery process that consists of searching in the whole registry. However, using a
specific adaptation plan in which the provider has defined a set of trusted substituting
services, the ASM can rapidly locate the WSDL file and thus, perform substitution
without starting a new expensive discovery process. For the five Web services, we
note that the gain in self-adaptation time varies between 14% and 44.22%, except for
the EmailValidator Web service (see [Tab. 1]).

<table>
<thead>
<tr>
<th>Web services</th>
<th>Self-adaptation failures</th>
<th>Self-adaptation time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>default</td>
<td>specific</td>
</tr>
<tr>
<td>MyCalculator</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Mathematica</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CurConverter</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>EmailValidator</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>EmailVerifier</td>
<td>17</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 1: Evaluation of Default and Specific Self-adaptation plans

Tab. 1 gives an idea about the nature of plans that caused adaptation failure and
shows the importance of involving providers in the self-adaptation process. Unlike
specific self-adaptation plans, default plans are frequently prone to failures. This can
be explained by the fact that providers are aware of the capabilities and requirements
of their published services and exceptions that may occur. Consequently, they may
predict the possible deviation and fix the suitable self-adaptation actions. For instance,
the effectiveness of specific adaptation plans for EmailValidator and EmailVerifier
Web services is about 100% and 96.72%, respectively.

However, although our planning technique has significantly improved the
availability of the managed services (see [Fig. 10]) and seems to be effective in terms
of decision and adaptation quality, it presents some limitations regarding the time

Figure 12: Evaluation of default and specific self-adaptation plans.
spent in plan selection (especially for services with short execution time) and requires a considerable processing time when the decision making complexity increases.

In the future, we plan to improve the planning algorithm in order to optimize the selection process of the self-adaptation rules and to reduce the adaptation time.

7 Related Work

Several approaches have been proposed to deal with Web service self-adaptation. These approaches discuss various issues, such as service interactions [Kongdenfha et al. 2009, Taher et al. 2011], monitoring and prevention of SLA violations [Fugini and Siadat 2010, Mahhub and Spanoudakis 2010, Schmieders et al. 2011], proactive adaptation based on online testing [Hielscher et al. 2008, Sammodi et al. 2011], multi-layered monitoring [Mos et al. 2009, Guinea et al. 2011], etc.

Regarding policy-based management, WS-Policy has been extended in several works to describe QoS [Tosic et al. 2007, Chhetri and Kowalczyk 2010, Badidi and Esmah 2011]. However, WS-Policy is not used to specify adaptation policies and existing approaches lack mechanisms for using and managing the specified policies during the execution.

A recent trend is to complement approaches for reactive adaptation with proactive capabilities [Baresi et al. 2012]. Different solutions have been proposed such as equipping particular execution points in a composition with a set of alternative behaviors [Leitner et al. 2010, Aschoff and Zisman 2011], the use of testing to anticipate problems in service compositions and trigger adaptation requests [Hielscher et al. 2008, Tosi et al. 2009, Metzger et al. 2010], the combination of reactive and proactive adaptation to support self-protection and self-healing [Na et al. 2010], the use of semi-Markov models for performance predictions [Dai et al. 2009], the use of a composition tree to determine the impact of each service in a composition on its overall performance [Bodenstaff et al. 2009], risk-driven management [Ma et al. 2014]. However, these approaches lack reactive behavior and do not support changes that may occur in any part of the composition. In addition, they lack dynamism regarding test cases and adaptation strategies, when a composition is modified, in order to take into account changes and new requirements for the prediction process.

Recently, some approaches have been proposed to allow a monitoring and adaptation across multiple layers [Gjørven et al. 2008, Mos et al. 2009, Guinea et al. 2011, Zengin et al. 2011]. Their aim is to combine and correlate monitoring data from different sources and to avoid potential conflicts that may arise due to uncoordinated adaptations in different layers. These approaches use different techniques such as event monitoring and logging (e.g. [Zeginis et al. 2011]), detection of event patterns, use of dependency graphs, correlation and mapping between events and adaptation strategies (e.g. [Popescu et al. 2012]), etc. The major limitation of these approaches is that the control of execution is performed at each layer in isolation, and the events are processed independently of each others. Consequently, the monitored data are not correctly analyzed, which may lead to the wrong identification of the source of problem. Also, monitoring is often based on static event correlation rules. Finally, cross-layer approaches do not take into account properties and needs of other layers. This leads, in most cases, to incompatibility and conflict problems.
Our approach is different as it encompasses the main requirements for adaptive service-oriented systems [Metzger and Marquezan 2011], in particular autonomic decision making, the way Web service capabilities and adaptation requirements are specified, the involvement of different actors (autonomic registries, service providers and autonomic services themselves) in the self-adaptation process, self-adjustment of the adaptation system through the dynamic integration of providers’ policies.

8 Conclusion

In this paper, we have proposed an approach for policy-based self-adaptation of Web services. As a main contribution, we have proposed AWS-Policy, an extension for the description of autonomic Web services. The extension, based on a rich Web service information model, allows not only specifying QoS capabilities, but also provider’s recommendations. Indeed, as part of our work, service providers have the ability to participate in the self-adaptation process by publishing their specific adaptation policies, which will be used by autonomic managers during service execution.

To allow representing and storing AWS policies as well as achieving their self-management, we have extended the original UDDI information model with new data structures. We have also proposed a planning technique, which allows transforming specific adaptation policies to executable ECA rules, in order to enable the autonomic managers to determine an order of the actions that achieve the specified goal.

In the future, we intend to eliminate the limitations of this work through the introduction of semantics in all self-management steps. Indeed, semantics play an important role in the autonomic service behavior, and self-adaptation requires a semantic understanding of QoS and autonomic capabilities. AWS-Policy allows only for a syntactic description which may lead to inefficient matching of service policies. For this reason, we will integrate our AWS-Ont ontology [Mezni 2014], which is a complete ontology that covers all the behavioral aspects of self-* Web services. Using this ontology brings autonomic managers to a common conceptual space and helps to apply reasoning mechanisms to find a better match. AWS-Ont allows to reason about Web services’ policies and to carry out more intelligent tasks on behalf of service providers. We suggest adding semantics to AWS-Policy, using extensibility in policy assertions and alternatives. Also, a well-defined semantic matching algorithm is being implemented to check the compatibility between QoS policies. In this case, comparison between AWS-Policy constructs is performed based on assertions’ QNames and semantic extensions, i.e. domain concepts of our AWS-Ont ontology.

We also look for enhancing the process of creating specific self-adaptation policies by giving providers the possibility to define their recommended plans based on existing plans published by other providers. The idea of discovering autonomic Web services based on their self-* capabilities is underway.

Finally, for the future Internet, cross-layer self-adaptation is an emerging research issue in service-based systems. We intend to inject self-* properties in multiple layers, in order to support autonomic execution across the whole service technology stack.
References


