

Understanding Heuristics for Grouping by Operation - Type

This topic explains how Virtualize groups messages in the "Generate Parameterized Messages from Traffic" wizard. The way that messages are grouped determines what responders are generated (one responder is generated for each message group).

Sections include:

- [Understanding Grouping](#)
- [Grouping by HTTP Method](#)
- [Grouping by URL Paths](#)
- [Grouping by Presence of URL Parameters](#)
- [Grouping by URL Parameter Values](#)
- [Grouping Based on Request Body Content](#)
- [Grouping By a Combination of Criteria](#)

Understanding Grouping

The way that Virtualize groups messages by operation/type is influenced by both a) Virtualize's analysis of the given traffic and b) the grouping criteria that you choose to apply. Recommended grouping strategies are automatically selected; these recommendations can be adjusted and overridden as you see fit.

Virtualize takes the approach of analyzing the traffic in two stages: first it recommends what grouping strategies should be enabled, and then it uses the selected strategies to divide the messages into groups. When creating PVAs, one responder is generated for each group.

The recommended grouping strategies are selected based on the following heuristics:

- **Grouping based on Request Body** is selected in the following cases:
 - SOAP Traffic
 - Non-XML message formats, including custom message formats
 - Non-HTTP traffic
- **Grouping based on HTTP Method and URL Paths** is selected in other cases—except that HTTP method is only selected if there are variations in the methods within the provided traffic.

When grouping, Virtualize:

- **Favors specific over general.** Virtualize generates as specific of an expression as possible matching all messages within that group only (not matching messages in other groups). This ensures that generated responders correlate correctly for the same traffic.
- **Follows the grouping order (HTTP method, URL path, URL parameter and then request body content).** For example, if URL paths are enough to differentiate between groups, Virtualize does not create Request Body correlations and so on for each of these grouping criteria steps.

Grouping by HTTP Method

Split the generated Message Responders based on the HTTP method (POST, GET, PUT, etc.). When this strategy is enabled, each unique HTTP method results in a Message Responder getting generated.

For example, if the traffic contains 10 POST messages and 15 GET messages (regardless of their order), Virtualize will create 2 responders in the generated PVA: one for the POST messages and one for the GET messages (assuming that no other grouping strategies have triggered further splits in the groups).

The resulting groups can be customized so that a group can capture messages of more than one HTTP method. For example, if both POST and PUT are enabled within a group, then messages using the POST method and messages using the PUT method will match the group and get assigned to it. In other words, the responder that is generated based on that group will correlate with both POST and PUT messages (it will accept either one).

Grouping by URL Paths

Split the generated Message Responders based on the uniqueness of URL paths. Each unique HTTP URL path results in a Message Responder getting generated. The configuration for grouping by HTTP URL Paths essentially defines that uniqueness.

When groups are first generated based on this strategy, the paths are analyzed to produce a set of path expressions that would best characterize the type of messages. That characterization uses a path analysis algorithm that is inherently subjective. It is designed to account for the most common patterns in REST services, but it is possible that the generated path criteria need to be adjusted for optimal results.

For example, if the traffic contains the following HTTP URL paths (omitting HTTP methods, URL parameters and other factors for brevity)

```
/service/order/v1/1/summary
/service/order/v1/2/summary
/service/order/v1/2/summary
/service/customer/v1/a/contact
/service/customer/v1/b/contact
/Foo
/Foo
```

Virtualize will create three responders in the generated PVA:

- Responder 1: for the `summary` paths (3 messages)
- Responder 2: the `contact` path (2 messages)
- Responder 3: for the `Foo` paths (2 messages)

Note that these responders will be generated using Ant-style wildcards, where `*` matches zero or more characters and `**` matches zero or more directories. See [URL Paths Grouping Configuration](#) for details and examples.

Configuration for Path-Based Data Source Correlation

Virtualize configures URL path segment indexes when creating a new PVA so that each group (responder) can respond to the same requests with varying URL Path values. It configures URL path parameters by providing a segment index. For example, given the following URLs

```
/rest/api/2/version/1234/relatedIssueCounts
/rest/api/2/version/4568/relatedIssueCounts
/rest/api/2/version/4567/relatedIssueCounts
```

Virtualize will designate the fifth path segment (the one with the 4 digit number) to be parameterized with a data source. Response data for each of these IDs can then be data-driven without the need to edit the PVA.

Grouping by Presence of URL Parameters

Split the generated Message Responders based on presence of HTTP URL parameters. Each set of one or more URL parameters results in a Message Responder getting generated.

For example, if the traffic contains the following URL parameters

```
?oid=1&category=women
?oid=2&category=women
?oid=1&category=babies
?ssn=1234567890&state=CA&category=silver
?ssn=1234567891&state=CA&category=gold
```

Virtualize will split the responders in the generated PVA based on the presence of `oid`, `ssn` and `state` parameters as follows (each responder to have the following responder correlation criteria)

- Responder 1: the URL parameter `oid` is present (3 messages)
- Responder 2: the URL parameters `ssn` and `state` are present (2 messages)

Note that this heuristic does not NOT generate groups based on URL parameter values. If you want to group based on URL parameter values, you can configure this manually. The grouping based on the presence of URL parameters is configured using the **Correlate only when the list of parameters matches exactly and Correlate whenever the parameter name is present regardless of value** options. See [URL Parameters Grouping Configuration](#) for details.

Grouping by URL Parameter Values

Split the generated Message Responders based on a list of values of specific HTTP URL parameters. Each list of one or more URL parameter values results in a Message Responder getting generated.

For example, if the traffic contains the following URL parameters

```
?oid=1&category=women  
?oid=2&category=women  
?oid=1&category=babies  
?ssn=1234567890&state=CA&category=silver  
?ssn=1234567891&state=CA&category=gold
```

Virtualize will split the responders in the generated PVA based on the following responder correlation criteria:

- Responder 1: category=women (2 messages)
- Responder 2: category=babies (1 message)
- Responder 3: state=CA, category=silver (1 message)
- Responder 4: state=CA, category=gold (1 message)

Grouping Based on Request Body Content

Split the generated responders based on a list of XPath expressions that are evaluated on the request content. Each XPath expression represents a group and XML Message correlation criteria.

When this strategy is enabled, SOAP messages are (by default) grouped based on first element name under Body. For non-SOAP XML content, messages are grouped based on root element name.

Note that when a custom message format (or a built-in non-XML one) is used, the XPath expressions are based on the converted XML format of the requests

Grouping By a Combination of Criteria

Generate a responder for each group of messages matching more than one grouping criteria from the use cases described above.

For example, given the following messages in the traffic file

```
GET /service/order/v1/1/summary?oid=1&category=women  
GET /service/order/v1/2/summary?oid=1&category=women  
GET /service/order/v1/2/summary?oid=1&category=babies  
POST /service/customer/v1/a/contact?ssn=1234567890&state=CA&category=silver  
POST /service/customer/v1/b/contact?ssn=1234567891&state=CA&category=gold  
POST /Foo  
POST /Foo
```

Virtualize will group them into responders based on the following criteria for each responder/group:

Group/Responder	HTTP Methods	URL Paths	URL Parameter Presence	Request Body Content
1	GET	summary	oid	
2	POST	contact	ssn, state	
3	POST	Foo		child 1 of Body = myOperation