

Siteaccess Matching

- Configuration
- Available matchers
- Compound siteaccess matcher
- Matching by request header
- Matching by environment variable
- URILexer and semanticPathinfo

Siteaccess matching is done through **eZ\Publish\MVC\SiteAccess\Matcher** objects. You can configure this matching and even develop custom matchers.

Configuration

You can configure siteaccess matching in your main **ezpublish/config/ezpublish.yml** :

```
ezpublish.yml
ezpublish:
  siteaccess:
    default_siteaccess: ezdemo_site
    list:
      - ezdemo_site
      - eng
      - fre
      - fr_eng
      - ezdemo_site_admin
    groups:
      ezdemo_site_group:
        - ezdemo_site
        - eng
        - fre
        - fr_eng
        - ezdemo_site_admin
    match:
      Map\URI:
        ezdemo_site: ezdemo_site
        fre: fre
        ezdemo_site_admin: ezdemo_site_admin
```

You need to set several parameters:

- **ezpublish.siteaccess.default_siteaccess**
- **ezpublish.siteaccess.list**
- *(optional)* **ezpublish.siteaccess.groups**
- **ezpublish.siteaccess.match**

ezpublish.siteaccess.default_siteaccess is the default siteaccess that will be used if matching was not successful. This ensures that a siteaccess is always defined.

ezpublish.siteaccess.list is the list of all available siteaccesses in your website.

(optional) **ezpublish.siteaccess.groups** defines which groups siteaccesses are member of. This is useful when you want to mutualize settings between several siteaccesses and avoid config duplication. Siteaccess groups are considered as regular siteaccesses as far as configuration is concerned.

A siteaccess can be part of several groups.

A siteaccess configuration has always precedence on the group configuration.

ezpublish.siteaccess.match holds the matching configuration. It consists in a hash where the key is the name of the matcher class. If the matcher class doesn't start with a \, it will be considered relative to eZ\Publish\MVC\SiteAccess\Matcher (e.g. Map\Host will refer to eZ\Publish\MVC\SiteAccess\Matcher\Map\Host)

Every **custom matcher** can be specified with a **fully qualified class name** (e.g. \My\SiteAccess\Matcher) or by a **service identifier prefixed by @** (e.g. @my_matcher_service).

- In the case of a fully qualified class name, the matching configuration will be passed in the constructor.
- In the case of a service, it must implement eZ\Bundle\EzPublishCoreBundle\SiteAccess\Matcher. The matching configuration will be passed to `setMatchingConfiguration()`.

Make sure to type matcher in correct case, if wrong case like "Uri" instead of "URI" it will happily work on systems like Mac OS X because of case in sensitive file system, while it will fail when you deploy it to a linux server. This is a known artifact of **PSR-0 autoloading** of PHP classes.

Available matchers

Name	Description	Configuration	Example
URIElement	Maps a URI element to a siteaccess. This is the default matcher used when choosing URI matching in setup wizard.	<p>The element number you want to match (starting from 1).</p> <pre> ezpublish: siteaccess: match: URIElement: 1 </pre> <p>Important: When using a value > 1, it will concatenate the elements with _</p>	<p>URI: /ezdemo_site/foo/bar</p> <p>Element number: 1 Matched siteaccess: ezdemo_site</p> <p>Element number: 2 Matched siteaccess: ezdemo_site_foo</p>
URIText	Matches URI using <i>pre</i> and/or <i>post</i> sub-strings in the first URI segment	<p>The prefix and/or suffix (none are required)</p> <pre> ezpublish: siteaccess: match: URIText: prefix: foo suffix: bar </pre>	<p>URI: /footestbar/my/content</p> <p>Prefix: foo Suffix: bar Matched siteaccess: test</p>
HostElement	Maps an element in the host name to a siteaccess.	<p>The element number you want to match (starting from 1).</p> <pre> ezpublish: siteaccess: match: HostElement: 2 </pre>	<p>Host name: www.example.com</p> <p>Element number: 2 Matched siteaccess: example</p>

HostText	Matches a siteaccess in the host name, using <i>pre</i> and/or <i>post</i> sub-strings.	<p>The prefix and/or suffix (none are required)</p> <pre> ezpublish: siteaccess: match: HostText: prefix: www. suffix: .com </pre>	<p>Host name: www.foo.com</p> <p><u>Prefix:</u> www. <u>Suffix:</u> .com <u>Matched siteaccess:</u> foo</p>
Map\Host	Maps a host name to a siteaccess.	<p>A hash map of host/siteaccess</p> <pre> ezpublish: siteaccess: match: Map\Host: www.foo.com: foo_front adm.foo.com: foo_admin www.bar-stuff.fr: bar_front adm.bar-stuff.fr: bar_admin </pre>	<p>Map:</p> <ul style="list-style-type: none"> • www.foo.com => foo_front • admin.foo.com => foo_admin <p>Host name: www.example.com</p> <p><u>Matched siteaccess:</u> foo_front</p>
Map\URI	Maps a URI to a siteaccess	<p>A hash map of URI/siteaccess</p> <pre> ezpublish: siteaccess: match: Map\URI: something: ezdemo_site ezdemo_site_admin foobar: </pre>	<p>URI: /something/my/content</p> <p><u>Map:</u></p> <ul style="list-style-type: none"> • something => ezdemo_site • foobar => ezdemo_site_admin <p><u>Matched siteaccess:</u> ezdemo_site</p>
Map\Port	Maps a port to a siteaccess	<p>A has map of Port/siteaccess</p> <pre> ezpublish: siteaccess: match: Match\Port: 80: foo 8080: bar </pre>	<p>URL: http://ezpublish.dev:8080/my/content</p> <p><u>Map:</u></p> <ul style="list-style-type: none"> • 80: foo • 8080: bar <p><u>Matched siteaccess:</u> bar</p>

Regex\Host	Matches against a regexp and extract a portion of it	<p>The regexp to match against and the captured element to use</p> <pre> ezpublish: siteaccess: match: Regex\Host: regex: "^(\w+_sa)\$" # Default is 1 itemNumber: 1 </pre>	<p>Host name: example_sa</p> <p><u>regex:</u> ^(\w+_sa)\$ <u>itemNumber:</u> 1</p> <p><u>Matched siteaccess:</u> example</p>
Regex\URI	Matches against a regexp and extract a portion of it	<p>The regexp to match against and the captured element to use</p> <pre> ezpublish: siteaccess: match: Regex\URI: regex: "^/foo(\w+)bar" # Default is 1 itemNumber: 1 </pre>	<p>URI: /footestbar/something</p> <p><u>regex:</u> ^/foo(\w+)bar <u>itemNumber:</u> 1</p> <p><u>Matched siteaccess:</u> test</p>

Compound siteaccess matcher

The Compound siteaccess matcher allows to combine several matchers together:

- <http://example.com/en> matches site_en (match on host=example.com *and* URIElement(1)=en)
- <http://example.com/fr> matches site_fr (match on host=example.com *and* URIElement(1)=fr)
- <http://admin.example.com> matches site_admin (match on host=admin.example.com)

Compound matchers cover the legacy **host_uri** matching feature.

They are based on logical combinations, or/and, using logical compound matchers:

- Compound\LogicalAnd
- Compound\LogicalOr

Each compound matcher will specify two or more sub-matchers. A rule will match if all the matchers, combined with the logical matcher, are positive. The example above would have used Map\Host and Map\Uri., combined with a LogicalAnd. When both the URI and host match, the siteaccess configured with "match" is used.

ezpublish.yml

```
ezpublish:
  siteaccess:
    match:
      Compound\LogicalAnd:
        # Nested matchers, with their configuration.
        # No need to precise their matching values (true will suffice).
        site_en:
          matchers:
            Map\URI:
              en: true
            Map\Host:
              example.com: true
          match: site_en
        site_fr:
          matchers:
            Map\URI:
              fr: true
            Map\Host:
              example.com: true
          match: site_fr
      Map\Host:
        admin.example.com: site_admin
```

Matching by request header

It is possible to define which siteaccess to use by setting a **X-Siteaccess** header in your request. This can be useful for REST requests.

In such case, **X-Siteaccess** must be the **siteaccess name** (e.g. *ezdemo_site*).

Matching by environment variable

It is also possible to define which siteaccess to use directly via an **EZPUBLISH_SITEACCESS** environment variable.

This is recommended if you want to get **performance gain** since no matching logic is done in this case.

You can define this environment variable directly from your web server configuration:

Apache VirtualHost example

```
# This configuration assumes that mod_env is activated
<VirtualHost *:80>
  DocumentRoot "/path/to/ezpublish5/web/folder"
  ServerName example.com
  ServerAlias www.example.com
  SetEnv EZPUBLISH_SITEACCESS ezdemo_site
</VirtualHost>
```

This can also be done via PHP-FPM configuration file, if you use it. See [PHP-FPM documentation](#) for more information.

Note about precedence

The precedence order for siteaccess matching is the following (the first matched wins):

1. Request header
2. Environment variable
3. Configured matchers

URILexer and semanticPathinfo

In some cases, after matching a siteaccess, it is necessary to modify the original request URI. This is for example needed with URI-based matchers since the siteaccess is contained in the original URI and it is not part of the route itself.

The problem is addressed by *analyzing* this URI and by modifying it when needed through the **URILexer** interface.

URILexer interface

```
/**
 * Interface for SiteAccess matchers that need to alter the URI after matching.
 * This is useful when you have the siteaccess in the URI like
 */
"/<siteaccessName>/my/awesome/uri"
interface URILexer
{
    /**
     * Analyses $uri and removes the siteaccess part, if needed.
     *
     * @param string $uri The original URI
     * @return string The modified URI
     */
    public function analyseURI( $uri );
    /**
     * Analyses $linkUri when generating a link to a route, in order to have the
     * siteaccess part back in the URI.
     *
     * @param string $linkUri
     * @return string The modified link URI
     */
    public function analyseLink( $linkUri );
}
```

Once modified, the URI is stored in the **semanticPathinfo** request attribute, and the original pathinfo is not modified.