Collection Synchronization for WebDAV
draft-daboo-webdav-sync-00

Abstract

This specification defines an extension to WebDAV that allows efficient synchronization of the contents of a WebDAV collection.
# Table of Contents

1. Introduction ................................................. 3  
2. Conventions Used in This Document .......................... 3  
3. Open Issues ................................................ 3  
4. WebDAV Synchronization Report ............................. 4  
   4.1. Overview ................................................. 4  
   4.2. Report defined .......................................... 5  
   4.3. Example: Initial DAV:sync-collection REPORT .......... 6  
   4.4. Example: DAV:sync-collection Report with token ..... 8  
5. XML Element Definitions ................................. 10  
   5.1. DAV:sync-collection XML Element ........................ 10  
      5.1.1. DAV:sync-token XML Element ....................... 11  
      5.1.2. DAV:multistatus XML Element ....................... 11  
      5.1.3. DAV:sync-response XML Element .................... 12  
6. Security Considerations ................................. 12  
7. IANA Considerations ......................................... 12  
8. Acknowledgments .............................................. 12  
9. References .................................................... 12  
   9.1. Normative References .................................... 12  
   9.2. Informative References .................................. 13
1. Introduction

WebDAV [I-D.ietf-webdav-rfc2518bis] defines the concept of ‘collections’ which are hierarchical groupings of WebDAV resources on an HTTP [RFC2616] server. Collections can be of arbitrary size and depth (i.e. collections within collections). WebDAV clients that cache resource content need a way to synchronize that data with the server (i.e. detect what has changed and update their cache). This can currently be done using a WebDAV PROPFIND request on a collection to list all members of a collection along with their HTTP ETag values, which allows the client to determine which resources were changed, added or deleted. However this does not scale well to large collections as the XML response to the PROPFIND response will grow with the collection size.

This extension proposes a new WebDAV REPORT that results in the server returning to the client only information about those resources which have changed, are new or were deleted since a previous execution of that report on the collection.

2. Conventions Used in This Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

When XML element types in the namespace "DAV:" are referenced in this document outside of the context of an XML fragment, the string "DAV:" will be prefixed to the element type names.

3. Open Issues

1. Should we try and discriminate between changes to the body of a resource and changes to the properties?

2. If we indicate a property change, should we return the list of properties that changed on each resource (propnames NOT values)?

3. Should we provide a way to indicate that a ‘new’ resource is new in the collection as a result of a COPY or MOVE from another location, as opposed to being created?

4. Should we provide a way to indicate that a ‘deleted’ resource was removed from the collection as a result of a MOVE to another location, as opposed to being actually deleted?

5. How should ACLs be handled? e.g. a resource is visible to a user at one point in time, then its <read> privilege is removed.
Should the resource be marked as having been deleted when next synchronized?

6. Do we want a special indicator for a resource that was deleted and then re-created, as opposed to just indicating that the resource was ‘changed’?

4. WebDAV Synchronization Report

4.1. Overview

In order to synchronize data between two entities some form of synchronization token is required to define the state of the data to be synchronized at a particular point in time. That token can then be used to determine what has changed since that time and the current time.

HTTP already defines a synchronization token in the form of an entity tag which is attached to a resource. However, the entity tag is not always required to be ‘strong’ and thus cannot be relied on absolutely to as valid synchronization indicator. In addition, there is no concept of an entity tag for a collection’s contents.

The is specification defines a new WebDAV REPORT that is used to enable client-server collection synchronization.

In order to synchronize the contents of a collection between a server and client, the server provides the client with a synchronization token each time the synchronization report is executed. That token represents the state of the data being synchronized at that point in time. The client can then present that same token back to the server at some later time and the server will return only those items that are new, have changed or were deleted since that token was generated. The server also returns a new token representing the new state at the time the report was run.

Typically the first time a client connects to the server it will need to be informed of the entire state of the collection (i.e. a full list of all resources that are currently contained in the collection). That is done by the client sending an empty token value to the server. This indicates to the server that a full listing is done.

In some cases a server may only wish to maintain a limited amount of history about changes to a collection. In that situation it will return an error to the client when the client presents a token that is "out of date". At that point the client has to fall back to synchronizing the entire collection by re-running the report request
using an empty token value.

4.2. Report defined

The DAV:sync-collection REPORT is used to provide an overview of what has changed on the server since a previous execution of this report. The primary purpose of this is to aid the client in synchronizing its local cache of data with the server, in an efficient manner with as few round trips as possible.

If this report is implemented by a WebDAV server, then the server MUST include "collection-sync" in any DAV: response header to an OPTIONS request. The report MUST also be listed in the "DAV: supported-report-set" property on any collection supporting synchronization. If the report is not available, clients MUST NOT attempt to execute one.

To implement the behavior for this report a server needs to keep track of changes to any resources in a collection. This includes noting the addition of new resources, noting changes to resources and noting removal of resources (where "removal" could be the result of a DELETE or MOVE operation). The server will track each change and provide a synchronization "token" to the client that describes the state of the server at a specific point in time. This "token" is returned as part of the response to the "collection-sync" report. Clients include the last token they got from the server in the next "sync-collection" report that they execute and the server provides the changes from the previous state represented by the token to the current state, represented by the new token returned.

The synchronization token itself is an "opaque" string - i.e. the actual string data has no specific meaning or syntax. A simple implementation of such a token would be a numeric counter that counts each change as it occurs and relates that change to the specific object that changed.

Marshalling:

The request URI MUST be a collection. The "Depth" header MUST be ignored by the server and SHOULD NOT be sent by the client. The request body MUST be a DAV:sync-collection XML element (see Section 5.1, which MUST contain one DAV:sync-token XML element, and optionally a DAV:propstat XML element.

The response body for a successful request MUST be a DAV: multistatus XML element, which MUST contain one DAV:sync-token element in addition to any DAV:sync-response elements.
The response body for a successful DAV:sync-collection REPORT request MUST contain a DAV:sync-response element for each resource that was created, has changed or been deleted since the last synchronization operation as specified by the DAV:sync-token provided in the request.

The DAV:status element in each DAV:sync-response element is used to indicate how the resource may have changed:

- A status code of '201 Created' is used to indicate resources that are new.
- A status code of '200 OK' is used to indicate resources that have changed.
- A status code of '404 Not Found' is used to indicate resources that have been removed.

Preconditions:

- (DAV:valid-sync-token): The DAV:sync-token element value MUST map to a valid token previously returned by the server;

Postconditions:

- None.

4.3. Example: Initial DAV:sync-collection REPORT

In this example, the client is making its first synchronization request to the server, so the DAV:sync-token element in the request is empty, and it also asks for the DAV:getetag property. The server responds with the items currently in the targeted collection (indicating that they are ‘new’ via the ‘201 Created’ status code). The current synchronization token is also returned.
REPORT /home/cyrusdaboo/ HTTP/1.1
Host: webdav.example.com
Content-Type: text/xml; charset="utf-8"
Content-Length: xxxx

<?xml version="1.0" encoding="utf-8" ?>
<D:sync-collection xmlns:D="DAV:">
  <D:sync-token/>
  <D:prop>
    <D:getetag/>
  </D:prop>
</D:sync-collection>
HTTP/1.1 207 Multi-Status
Content-Type: text/xml; charset="utf-8"
Content-Length: xxxx

<?xml version="1.0" encoding="utf-8" ?>
<D:multistatus xmlns:D="DAV:"
    <D:sync-response>
      <D:status>HTTP/1.1 201 Created</D:status>
      <D:propstat>
        <D:prop>
          <D:getetag>"00001-abcd1"</D:getetag>
        </D:prop>
      </D:propstat>
    </D:sync-response>
    <D:sync-response>
      <D:href>http://webdav.example.com/home/cyrusdaboo/vcard.vcf</D:href>
      <D:status>HTTP/1.1 201 Created</D:status>
      <D:propstat>
        <D:prop>
          <D:getetag>"00002-abcd1"</D:getetag>
        </D:prop>
      </D:propstat>
    </D:sync-response>
    <D:sync-response>
      <D:href>http://webdav.example.com/home/cyrusdaboo/calendar.ics</D:href>
      <D:status>HTTP/1.1 201 Created</D:status>
      <D:propstat>
        <D:prop>
          <D:getetag>"00003-abcd1"</D:getetag>
        </D:prop>
      </D:propstat>
    </D:sync-response>
    <D:sync-token>1234</D:sync-token>
  </D:multistatus>

4.4. Example: DAV:sync-collection Report with token

In this example, the client is making a synchronization request to
the server and is using the DAV:sync-token element returned from the
last report it ran on this collection. The server responds listing
the items that have been added, changed or removed. The (new) current synchronization token is also returned.

>> Request <<

REPORT /home/cyrusdaboo/ HTTP/1.1  
Host: webdav.example.com  
Content-Type: text/xml; charset="utf-8"  
Content-Length: xxxx  

<?xml version="1.0" encoding="utf-8"?>  
<D:sync-collection xmlns:D="DAV:">  
  <D:sync-token>1234</D:sync-token>  
  <D:prop>  
    <D:getetag/>  
  </D:prop>  
</D:sync-collection>
HTTP/1.1 207 Multi-Status
Content-Type: text/xml; charset="utf-8"
Content-Length: xxxx

<?xml version="1.0" encoding="utf-8" ?>
<D:multistatus xmlns:D="DAV:"

<D:sync-response>
<D:status>HTTP/1.1 201 Created</D:status>
<D:propstat>
<D:prop>
<D:getetag>"00004-abcd1"</D:getetag>
</D:prop>
</D:propstat>
</D:sync-response>

<D:sync-response>
<D:href>http://webdav.example.com/home/cyrusdaboo/vcard.vcf</D:href>
<D:status>HTTP/1.1 200 OK</D:status>
<D:propstat>
<D:prop>
<D:getetag>"00002-abcd2"</D:getetag>
</D:prop>
</D:propstat>
</D:sync-response>

<D:sync-response>
<D:status>HTTP/1.1 404 Not Found</D:status>
</D:sync-response>

<D:sync-token>1238</D:sync-token>
</D:multistatus>

5. XML Element Definitions

5.1. DAV:sync-collection XML Element

Name: sync-collection

Namespace: DAV:
Purpose: WebDAV report used to synchronize data between client and server.

Description: See Section 4.

<!ELEMENT sync-collection (sync-token, DAV:prop?)>

5.1.1. DAV:sync-token XML Element

Name: sync-token

Namespace: DAV:

Purpose: The synchronization token provided by the server and returned by the client.

Description: See Section 4.

<!ELEMENT sync-token CDATA>

5.1.2. DAV:multistatus XML Element

Name: multistatus

Namespace: DAV:

Purpose: Extends the DAV:multistatus element to include synchronization details.

Description: See Section 4.

5.1.3. DAV:sync-response XML Element

Name: sync-response

Namespace: DAV:

Purpose: Contains the synchronization results returned by the server.

Description: See Section 4.

<!ELEMENT sync-response (DAV:href, DAV:status, DAV:propstat?)>

6. Security Considerations

This extension does not introduce any new security concerns than those already described in HTTP and WebDAV.

7. IANA Considerations

This document does not require any actions on the part of IANA.

8. Acknowledgments

9. References

9.1. Normative References


9.2. Informative References

Author’s Address

Cyrus Daboo
Apple Computer, Inc.
1 Infinite Loop
Cupertino, CA 95014
USA

EMail: cyrus@daboo.name
URI: http://www.apple.com/