

## [MS-PCCRR]:

# Peer Content Caching and Retrieval: Retrieval Protocol

---

### Intellectual Property Rights Notice for Open Specifications Documentation

- **Technical Documentation.** Microsoft publishes Open Specifications documentation (“this documentation”) for protocols, file formats, data portability, computer languages, and standards support. Additionally, overview documents cover inter-protocol relationships and interactions.
- **Copyrights.** This documentation is covered by Microsoft copyrights. Regardless of any other terms that are contained in the terms of use for the Microsoft website that hosts this documentation, you can make copies of it in order to develop implementations of the technologies that are described in this documentation and can distribute portions of it in your implementations that use these technologies or in your documentation as necessary to properly document the implementation. You can also distribute in your implementation, with or without modification, any schemas, IDLs, or code samples that are included in the documentation. This permission also applies to any documents that are referenced in the Open Specifications documentation.
- **No Trade Secrets.** Microsoft does not claim any trade secret rights in this documentation.
- **Patents.** Microsoft has patents that might cover your implementations of the technologies described in the Open Specifications documentation. Neither this notice nor Microsoft's delivery of this documentation grants any licenses under those patents or any other Microsoft patents. However, a given Open Specifications document might be covered by the Microsoft [Open Specifications Promise](#) or the [Microsoft Community Promise](#). If you would prefer a written license, or if the technologies described in this documentation are not covered by the Open Specifications Promise or Community Promise, as applicable, patent licenses are available by contacting [iplg@microsoft.com](mailto:iplg@microsoft.com).
- **Trademarks.** The names of companies and products contained in this documentation might be covered by trademarks or similar intellectual property rights. This notice does not grant any licenses under those rights. For a list of Microsoft trademarks, visit [www.microsoft.com/trademarks](http://www.microsoft.com/trademarks).
- **Fictitious Names.** The example companies, organizations, products, domain names, email addresses, logos, people, places, and events that are depicted in this documentation are fictitious. No association with any real company, organization, product, domain name, email address, logo, person, place, or event is intended or should be inferred.

**Reservation of Rights.** All other rights are reserved, and this notice does not grant any rights other than as specifically described above, whether by implication, estoppel, or otherwise.

**Tools.** The Open Specifications documentation does not require the use of Microsoft programming tools or programming environments in order for you to develop an implementation. If you have access to Microsoft programming tools and environments, you are free to take advantage of them. Certain Open Specifications documents are intended for use in conjunction with publicly available standards specifications and network programming art and, as such, assume that the reader either is familiar with the aforementioned material or has immediate access to it.

## Revision Summary

Date	Revision History	Revision Class	Comments
12/5/2008	0.1	Major	Initial Availability
1/16/2009	0.1.1	Editorial	Changed language and formatting in the technical content.
2/27/2009	0.1.2	Editorial	Changed language and formatting in the technical content.
4/10/2009	0.2	Minor	Clarified the meaning of the technical content.
5/22/2009	1.0	Major	Updated and revised the technical content.
7/2/2009	1.1	Minor	Clarified the meaning of the technical content.
8/14/2009	2.0	Major	Updated and revised the technical content.
9/25/2009	2.1	Minor	Clarified the meaning of the technical content.
11/6/2009	2.2	Minor	Clarified the meaning of the technical content.
12/18/2009	2.2.1	Editorial	Changed language and formatting in the technical content.
1/29/2010	2.3	Minor	Clarified the meaning of the technical content.
3/12/2010	2.3.1	Editorial	Changed language and formatting in the technical content.
4/23/2010	2.4	Minor	Clarified the meaning of the technical content.
6/4/2010	3.0	Major	Updated and revised the technical content.
7/16/2010	3.0	None	No changes to the meaning, language, or formatting of the technical content.
8/27/2010	3.0	None	No changes to the meaning, language, or formatting of the technical content.
10/8/2010	3.0	None	No changes to the meaning, language, or formatting of the technical content.
11/19/2010	3.0	None	No changes to the meaning, language, or formatting of the technical content.
1/7/2011	3.0	None	No changes to the meaning, language, or formatting of the technical content.
2/11/2011	3.0	None	No changes to the meaning, language, or formatting of the technical content.
3/25/2011	3.0	None	No changes to the meaning, language, or formatting of the technical content.
5/6/2011	3.0	None	No changes to the meaning, language, or formatting of the technical content.
6/17/2011	3.1	Minor	Clarified the meaning of the technical content.
9/23/2011	3.1	None	No changes to the meaning, language, or formatting of the technical content.
12/16/2011	4.0	Major	Updated and revised the technical content.
3/30/2012	5.0	Major	Updated and revised the technical content.

<b>Date</b>	<b>Revision History</b>	<b>Revision Class</b>	<b>Comments</b>
7/12/2012	6.0	Major	Updated and revised the technical content.
10/25/2012	7.0	Major	Updated and revised the technical content.
1/31/2013	7.0	None	No changes to the meaning, language, or formatting of the technical content.
8/8/2013	8.0	Major	Updated and revised the technical content.
11/14/2013	8.0	None	No changes to the meaning, language, or formatting of the technical content.
2/13/2014	9.0	Major	Updated and revised the technical content.
5/15/2014	9.0	None	No changes to the meaning, language, or formatting of the technical content.
6/30/2015	10.0	Major	Significantly changed the technical content.
10/16/2015	10.0	None	No changes to the meaning, language, or formatting of the technical content.
7/14/2016	10.0	None	No changes to the meaning, language, or formatting of the technical content.

# Table of Contents

<b>1</b>	<b>Introduction .....</b>	<b>6</b>
1.1	Glossary .....	6
1.2	References .....	7
1.2.1	Normative References .....	8
1.2.2	Informative References .....	8
1.3	Overview .....	8
1.4	Relationship to Other Protocols .....	9
1.5	Prerequisites/Preconditions .....	9
1.6	Applicability Statement .....	9
1.7	Versioning and Capability Negotiation .....	10
1.8	Vendor-Extensible Fields .....	10
1.9	Standards Assignments.....	10
<b>2</b>	<b>Messages.....</b>	<b>11</b>
2.1	Transport.....	11
2.1.1	Peer Download Transport.....	11
2.1.2	Transport Security .....	11
2.2	Message Syntax.....	11
2.2.1	Common Data Types.....	11
2.2.1.1	BLOCK_RANGE .....	12
2.2.1.2	SEGMENT_RANGE .....	12
2.2.1.3	BLOCK_RANGE_ARRAY .....	12
2.2.1.4	SEGMENT_RANGE_ARRAY .....	12
2.2.1.5	ENCODED_SEGMENT_AGE .....	13
2.2.2	TRANSPORT_RESPONSE_HEADER.....	13
2.2.3	MESSAGE_HEADER .....	13
2.2.4	Request Message.....	15
2.2.4.1	MSG_NEGO_REQ .....	16
2.2.4.2	MSG_GETBLKLIST .....	16
2.2.4.3	MSG_GETBLKS .....	17
2.2.4.4	MSG_GETSEGLIST .....	18
2.2.5	Response Message.....	19
2.2.5.1	MSG_NEGO_RESP .....	20
2.2.5.2	MSG_BLKLIST .....	20
2.2.5.3	MSG_BLK.....	21
2.2.5.4	MSG_SEGLIST .....	22
2.2.6	Extensible BLOB .....	23
2.2.6.1	Extensible Blob Version 1 .....	24
2.2.6.1.1	Extensible Blob Version 1 Restrictions and Validation .....	24
<b>3</b>	<b>Protocol Details.....</b>	<b>26</b>
3.1	Client Details.....	26
3.1.1	Abstract Data Model.....	26
3.1.2	Timers .....	27
3.1.3	Initialization .....	27
3.1.4	Higher-Layer Triggered Events .....	27
3.1.4.1	MSG_NEGO_REQ Request .....	27
3.1.4.2	MSG_GETBLKLIST Initiation .....	27
3.1.4.3	MSG_GETBLKS Initiation.....	27
3.1.4.4	MSG_GETSEGLIST Initiation .....	28
3.1.5	Message Processing Events and Sequencing Rules .....	28
3.1.5.1	MSG_NEGO_RESP Received .....	28
3.1.5.2	MSG_BLKLIST Response Received.....	29
3.1.5.3	MSG_BLK Response Received .....	29
3.1.5.4	MSG_SEGLIST Response Received .....	29

3.1.5.5	Other Messages Received.....	30
3.1.6	Timer Events.....	30
3.1.6.1	Request Timer Expiration .....	30
3.1.7	Other Local Events.....	30
3.2	Server Details.....	30
3.2.1	Abstract Data Model.....	30
3.2.2	Timers .....	30
3.2.3	Initialization.....	31
3.2.4	Higher-Layer Triggered Events .....	31
3.2.5	Message Processing Events and Sequencing Rules .....	31
3.2.5.1	MSG_NEGO_REQ Received .....	31
3.2.5.2	MSG_GETBLKLIST Request Received .....	31
3.2.5.3	MSG_GETBLKS Request Received.....	31
3.2.5.4	MSG_GETSEGLIST Request Received.....	32
3.2.5.5	Other Messages Received.....	33
3.2.6	Timer Events.....	33
3.2.6.1	Upload Timer Expiration.....	33
3.2.7	Other Local Events.....	33
<b>4</b>	<b>Protocol Examples.....</b>	<b>34</b>
4.1	Download with GetBlockList and GetBlocks Exchanges .....	34
4.2	Simple Download with GetBlocks Download Sub-Sessions only .....	35
<b>5</b>	<b>Security.....</b>	<b>36</b>
5.1	Security Considerations for Implementers .....	36
5.2	Index of Security Parameters .....	36
<b>6</b>	<b>Appendix A: Product Behavior .....</b>	<b>37</b>
<b>7</b>	<b>Change Tracking.....</b>	<b>39</b>
<b>8</b>	<b>Index.....</b>	<b>40</b>

# 1 Introduction

The Peer Content Caching and Retrieval: Retrieval Protocol reduces bandwidth consumption on branch-office wide-area-network (WAN) links by having **clients** retrieve content from distributed caches when available instead of the **content servers**, which are often located remotely from branch offices over the WAN links. It is based on a peer-to-peer discovery and distribution model, where the **peers** themselves act as caches from which they serve other requesting peers. The **framework** also supports the mode of using pre-provisioned hosted caches in place of peer-based caching. The main benefit of the framework is to reduce operation costs by reducing WAN link utilization, while providing faster downloads from the local area networks (LANs) in the branch offices.

The Retrieval Framework defines four protocol message exchanges: for querying the protocol version of the **server**, for querying the server for the availability of certain content (two message exchanges), and for retrieving content from a server. The framework incorporates both the Retrieval Protocol and the Discovery Protocol [\[MS-PCCRD\]](#) together to enable a client to discover and retrieve content from multiple peers that have the content instead of the original content server.

Sections 1.5, 1.8, 1.9, 2, and 3 of this specification are normative. All other sections and examples in this specification are informative.

## 1.1 Glossary

This document uses the following terms:

**block:** A chunk of content that composes a **segment**. Each **segment** is divided into one or more **blocks**. Every **block** belongs to a specific **segment**, and within a **segment**, **blocks** are identified by their progressive index. (Block 0 is the first **block** in the **segment**, block 1 is the second, and so on.) See [\[MS-PCCRC\]](#) for more details.

**block hash:** A hash of a content block within a segment. Also known as a block ID.

**block range:** A set of consecutive **blocks** within a **segment** described by a pair of integers, the first being the index of the first **blocks** in the range, and the second the number of consecutive **blocks** in the range.

**client:** For the Peer Content Caching and Retrieval Framework, a client is a client-role peer; that is, a peer that is searching for content, either from the server or from other peers or hosted caches. In the context of the Retrieval Protocol, a client is a peer that requests a block-range from a server\_role\_peer. It acts as a Web Services Dynamic Discovery (WS-Discovery) [\[WS-Discovery\]](#) client.

**client-role peer:** A **peer** that is looking for content, either from the server or from other **peers** or hosted caches.

**content server:** The original source of the content that peers subsequently retrieve from each other.

**distributed mode:** A mode of operation for the **client-role peer** in the **Peer Content Caching and Retrieval Framework**, in which it discovers and obtains content **blocks** from other **peers**, and shares content **blocks** it has with other **peers** in the network.

**encryption key:** One of the input parameters to an encryption algorithm. Generally speaking, an encryption algorithm takes as input a clear-text message and a key, and results in a cipher-text message. The corresponding decryption algorithm takes a cipher-text message, and the key, and results in the original clear-text message.

**higher-layer application:** An application that uses the Peer Content Caching and Retrieval: Retrieval Protocol, either by itself or as part of the Peer Content Caching and Retrieval Framework or other applications.

**HoHoDk:** A hash that represents the content-specific label or public identifier that is used to discover content from other **peers** or from the hosted cache. This identifier is disclosed freely in broadcast messages. Knowledge of this identifier does not prove authorization to access the actual content.

**hosted cache mode:** A mode of operation for the client-role peer in the Peer Content Caching and Retrieval Framework, in which it obtains and shares content (only) with a single server whose location is preconfigured on the client-role peer.

**index:** The **block** number within a **segment**.

**initialization vector:** A data block that some modes of the AES cipher block operation require as an additional initial data input. For more information, see [\[SP800-38A\]](#).

**peer:** An instance of the Retrieval Protocol for the Peer Content Caching and Retrieval Framework running on a host. A peer can be both a client and a server in the Retrieval Protocol operations.

**Peer Content Caching and Retrieval Framework (or Framework):** The framework that creates Peer Content Caching and Retrieval Discovery Protocol instances to discover client-role peers and download the content blocks from either client-role peers (distributed mode) or hosted cache (hosted-cache mode).

**segment:** A subdivision of content. In version 1.0 Content Information, each segment has a size of 32 megabytes, except the last segment which can be smaller if the content size is not a multiple of the standard segment sizes. In version 2.0 Content Information, segments can vary in size.

**segment ID (HoHoDk):** A hash that represents the content-specific label or public identifier that is used to discover content from other peers or from the hosted cache. This identifier is disclosed freely in broadcast messages. Knowledge of this identifier does not prove authorization to access the actual content.

**segment retrieval session:** A session that defines a set of operations on a client-role peer that use the Discovery Protocol (in distributed mode) and the Retrieval Protocol to discover and retrieve ranges of blocks (partial or complete) of a segment.

**server:** For the Peer Content Caching and Retrieval Framework, a server is a server-role peer; that is, a peer that listens for incoming block-range requests from client-role peers and responds to the requests.

**server-role peer:** A **peer** that listens for incoming **block**-range requests from **client-role peers** and responds to the requests.

**simple download:** A GetBlocks request/response that is carried out without an associated GetBlockList request/response.

**target segment:** The segment for which the client-role peer is requesting a particular block range in a segment retrieval session, identified by the segment ID.

**MAY, SHOULD, MUST, SHOULD NOT, MUST NOT:** These terms (in all caps) are used as defined in [\[RFC2119\]](#). All statements of optional behavior use either MAY, SHOULD, or SHOULD NOT.

## 1.2 References

Links to a document in the Microsoft Open Specifications library point to the correct section in the most recently published version of the referenced document. However, because individual documents in the library are not updated at the same time, the section numbers in the documents may not match. You can confirm the correct section numbering by checking the [Errata](#).

### 1.2.1 Normative References

We conduct frequent surveys of the normative references to assure their continued availability. If you have any issue with finding a normative reference, please contact [dochelp@microsoft.com](mailto:dochelp@microsoft.com). We will assist you in finding the relevant information.

[FIPS197] FIPS PUBS, "Advanced Encryption Standard (AES)", FIPS PUB 197, November 2001, <http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf>

[MS-DTYP] Microsoft Corporation, "[Windows Data Types](#)".

[MS-PCCRC] Microsoft Corporation, "[Peer Content Caching and Retrieval: Content Identification](#)".

[MS-PCCRD] Microsoft Corporation, "[Peer Content Caching and Retrieval: Discovery Protocol](#)".

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997, <http://www.rfc-editor.org/rfc/rfc2119.txt>

[RFC2616] Fielding, R., Gettys, J., Mogul, J., et al., "Hypertext Transfer Protocol -- HTTP/1.1", RFC 2616, June 1999, <http://www.rfc-editor.org/rfc/rfc2616.txt>

[SP800-38A] National Institute of Standards and Technology., "Special Publication 800-38A, Recommendation for Block Cipher Modes of Operation: Methods and Techniques", December 2001, <http://csrc.nist.gov/publications/nistpubs/800-38a/sp800-38a.pdf>

### 1.2.2 Informative References

None.

### 1.3 Overview

The Retrieval Protocol defines four request/response exchanges between a client and a server on top of an HTTP [\[RFC2616\]](#) transport to query the supported version range of the server, to query the availability of specific content, and to retrieve specific content. The protocol assumes that the client identifies both the specific content it is looking for and the server it will contact. The discovery of the content information and the server address is outside the scope of the Retrieval Protocol. The request/response exchanges are:

- Content Availability Request: The client initiates a query to the server for the availability of the specified content. The server responds with the ranges (subsets or all) of the requested content it has. There are two types of content availability requests:
  - Segment Availability Request: The client initiates a query to the server for the availability of a set of segments of content. The server responds with the ranges (subsets or all) of the requested segments of content available in the server's local cache.
  - Block Availability Request: The client initiates a query to the server for the availability of a set of ranges of blocks within a single segment of content. The server responds with the ranges (subsets or all) of the requested block of content it has within the specified segment.
- Content Retrieval Request: The client initiates a request to the server for the specified content. The server either replies with the requested content or with content of zero length when the requested content is not available.
- Version Negotiation Request: The client initiates a request to the server to query the supported Retrieval Protocol version range. The server replies with its supported Retrieval Protocol version range.

The exchanges can be utilized in conjunction or independently, as described in the following examples:

- The client can query the server for the availability of the content, identify what content the server has, and then retrieve only the available content from the server; or
- The client can query the server for the availability of the content, identify what content the server has, and decide not to retrieve the content; or
- The client can retrieve the content directly from the server without querying for the availability of the content first.
- For all scenarios described earlier, the client can optionally query the server for its supported version range first before querying for content availability or retrieving **blocks**.

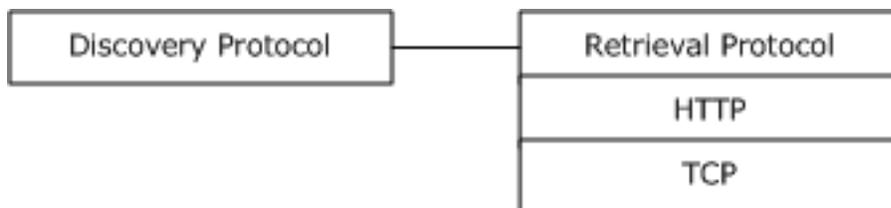
The Retrieval Protocol does not mandate the relationship between these exchanges, as shown in the examples. As a result, in the case where they are used in conjunction, the **higher-layer applications** invoking the Retrieval Protocol retain the availability list from the availability query, and use it to retrieve part or all of the available content in the subsequent retrieval request(s).

Peers within the Peer Content Caching and Retrieval Framework use the Retrieval Protocol in one of two ways, depending on whether they are in **distributed mode**, retrieving content from each other, or **hosted cache mode**, retrieving it only from a single preconfigured server. In the distributed mode case, a peer uses the framework's Discovery Protocol (see [\[MS-PCCRD\]](#)) to locate peers who have the desired content, and then initiates exchanges with the discovered peers to obtain the content. In hosted cache mode, a peer directly initiates exchanges with the hosted cache to obtain the desired content.

#### 1.4 Relationship to Other Protocols

The Retrieval Protocol uses HTTP [\[RFC2616\]](#) as a transport.

The Peer Content Caching and Retrieval Framework uses the Retrieval Protocol [\[MS-PCCRR\]](#) and Discovery Protocol [\[MS-PCCRD\]](#) to discover peers when in distributed mode, and query and download content from other peers. The framework also uses the data structures as described in [\[MS-PCCRC\]](#).



**Figure 1: Protocol stack diagram**

#### 1.5 Prerequisites/Preconditions

- A higher-layer application using the protocol has to have the Content Information (see [\[MS-PCCRC\]](#) section 2.3) for the block ranges and **segments** that it is retrieving from the server. The Content Information contains all the relevant information necessary for discovering and verifying the content blocks.
- The client has to identify and use the encryption algorithm and key used by the server to encrypt the content.

#### 1.6 Applicability Statement

The Retrieval Protocol is designed to handle the content availability query and content retrieval parts of the operation. It is also suitable for other types of content or object retrieval tasks because it does not assume any characteristics of the content.

The Peer Content Caching and Retrieval Framework, which uses the Retrieval Protocol, is best suited when there is a need to reduce load on a content server or reduce bandwidth usage on the link between the peers and the content server. This is because the protocol enables downloading data from peers on the high speed link instead of the content server, which might be behind a slow link or might be heavily loaded.

## 1.7 Versioning and Capability Negotiation

This document covers versioning issues in the following areas:

- **Supported Transports:** This protocol is implemented on top of HTTP as discussed in section [2.1](#).
- **Protocol Versions:** The protocol version is 2.0. Supported versions are 1.0 and 2.0. Messages defined in version 1.0 are sent with version specification 1.0, and messages defined in version 2.0 are sent with version specification 2.0.
- **Security and Authentication Methods:** There is no authentication or authorization in the protocol. The blocks served by the **server-role peer**, however, are encrypted as described in section [3.2.5.3](#).
- **Localization:** The protocol does not contain locale-dependent information.

## 1.8 Vendor-Extensible Fields

None.

## 1.9 Standards Assignments

None.

## 2 Messages

The Retrieval Protocol is made up of a limited number of fully defined messages sent on top of the [Peer Download Transport](#). <1>

### 2.1 Transport

#### 2.1.1 Peer Download Transport

The Peer Download Transport is a peer-to-peer transport built on top of HTTP [\[RFC2616\]](#). The client/server HTTP protocol is turned into a peer-to-peer transport by having each peer implement both a client and a server role. In a given transport session between two peers P1 and P2, the initiator peer acts as client, and the other peer acts as server. If P1 is the initiator of the transport session, P1 sends an HTTP request, and P2 replies by sending an HTTP response. Both the Retrieval Protocol request and response message types are included in the body of the HTTP messages. The payload of each such HTTP request or response consists solely of a single Retrieval Protocol message, with the response message prefixed with an additional length field (as defined in section [2.2.2](#)) for reassembly purposes. A transport session between any two peers spans a single request-response sequence, and no context is kept within the transport across different transport sessions between those two peers.

Each peer implements the server role by reserving the URL under the root path of /116B50EB-ECE2-41ac-8429-9F9E963361B7/ and listening for POST requests on it.

The initiating/**client-role peer** P1 at IP address A1 initiates the transport of a given request-type Peer Retrieval Protocol message to peer P2 at IP address A2, by sending an HTTP POST request to the root path of /116B50EB-ECE2-41ac-8429-9F9E963361B7/.

#### 2.1.2 Transport Security

The [Peer Download Transport](#) does not implement any security. There is no peer authentication or authorization, and messages are sent in clear text. At the transport level, peers accept and process all messages coming from any other peer.

### 2.2 Message Syntax

Messages are formed by headers and a message body. Both headers and body are formed by a sequence of fields. Each field is aligned according to the current protocol version's default alignment, currently 4 bytes.

All Retrieval Protocol messages are variable size messages. The valid range of the total message size MUST be from 16 bytes to 98,304 bytes (or 96 KB).

#### 2.2.1 Common Data Types

The protocol supports four field types:

- Integer (DWORD fields as defined in [\[MS-DTYP\]](#) section 2.2.9, transmitted in network byte order).
- [BLOCK\\_RANGE\\_ARRAY](#) ((Integer [2])[count], a count-sized array of [BLOCK\\_RANGE](#) fields).
- [SEGMENT\\_RANGE\\_ARRAY](#) ((Integer [2])[count], a count-sized array of [SEGMENT\\_RANGE](#) fields).
- BYTE array (BYTE[count], a count-sized array of bytes).

### 2.2.1.1 BLOCK\_RANGE

A **BLOCK\_RANGE** is an array of two integers that defines a consecutive array of blocks.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Index																															
Count																															

**Index (4 bytes):** The **index** of the first block in the range.

**Count (4 bytes):** Count of consecutive adjacent blocks in that range, including the block at the **Index** location. The value of this field **MUST** be greater than 0.

**Index** and **Count** are both integer fields in the range of 0x00000000 to 0xFFFFFFFF, but contain a value in the range from 0 to 511 inclusive for the **Index** field, and 1 to (512-Index) inclusive for the **Count** field. For example, a **BLOCK\_RANGE** of [42, 7] represents all the blocks starting from block index 42 to block index 48, including the last one.

### 2.2.1.2 SEGMENT\_RANGE

A **SEGMENT\_RANGE** is an array of two integers that defines a consecutive array of segments.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Index																															
Count																															

**Index (4 bytes):** The index of the first segment in the range.

**Count (4 bytes):** Count of consecutive adjacent segments in that range, including the segment at the **Index** location. The value of this field **MUST** be greater than 0.

**Index** and **Count** are both integer fields in the range of 0x00000000 to 0xFFFFFFFF.

### 2.2.1.3 BLOCK\_RANGE\_ARRAY

Variable-size array containing [BLOCK\\_RANGE](#) entries.

This type is declared as follows:

```
typedef BLOCK_RANGE BLOCK_RANGE_ARRAY[];
```

### 2.2.1.4 SEGMENT\_RANGE\_ARRAY

A variable-size array containing [SEGMENT\\_RANGE](#) entries.

This type is declared as follows:

```
typedef SEGMENT_RANGE SEGMENT_RANGE_ARRAY[];
```

### 2.2.1.5 ENCODED\_SEGMENT\_AGE

An ENCODED\_SEGMENT\_AGE is an array of four bytes that describes the age of a segment of data involved in a Peer Content Caching and Retrieval: Retrieval Protocol message exchange. The age refers to the amount of time that has elapsed since the specified segment was first created in the local BranchCache cache.

0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
SegmentIndex				SegmentAgeLowPart				SegmentAgeMidPart				SegmentAgeHighPart																			

**SegmentIndex (1 byte):** Index of a segment among all of the segments involved in the current message exchange. The index is relative to the first segment addressed in the message containing the specific current ENCODED\_SEGMENT\_AGE structure.

**SegmentAgeLowPart (1 byte):** Low part of the age of the segment.

**SegmentAgeMidPart (1 byte):** Mid part of the age of the segment.

**SegmentAgeHighPart (1 byte):** High part of the age of the segment.

The actual age of the segment is calculated as: **SegmentAgeLowPart** + 256 \* **SegmentAgeMidPart** + 256 \* 256 \* **SegmentAgeHighPart**.

The age of the segment is expressed according to the unit specified in the enclosing object/message. If no enclosing object/message is available, the age MUST be specified (and is assumed to be specified) in hundredths of milliseconds.

### 2.2.2 TRANSPORT\_RESPONSE\_HEADER

The transport adds the following header in front of response-type protocol messages for reassembly purposes:

0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
Size																															

**Size (4 bytes):** Total message size, in bytes, excluding this field. The valid range of the total message size MUST be from 16 bytes to 98,304 bytes (or 96 KB).

### 2.2.3 MESSAGE\_HEADER

All messages associated with this protocol are prefixed by a message header.

Messages can be one of two types: request-type or response-type. Request-type messages include [MSG\\_NEGO\\_REQ \(section 2.2.4.1\)](#), [MSG\\_GETBLKLIST \(section 2.2.4.2\)](#), [MSG\\_GETBLKS \(section 2.2.4.3\)](#), and [MSG\\_GETSEGLIST \(section 2.2.4.4\)](#). Response-type messages include [MSG\\_NEGO\\_RESP \(section 2.2.5.1\)](#), [MSG\\_BLKLIST \(section 2.2.5.2\)](#), [MSG\\_BLK \(section 2.2.5.3\)](#), and [MSG\\_SEGLIST \(section 2.2.5.4\)](#). Request-type messages initiate a communication session between two peers. Response-type messages are sent only on response to a Request-type one (see [Protocol Details \(section 3\)](#) for more details).

A request-type message can be delivered only as an HTTP request. A response-type message can be delivered only as an HTTP response to an incoming HTTP request.

The layout of the message header is as follows.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
ProtVer																															
MsgType																															
MsgSize																															
CryptoAlgoId																															

**ProtVer (4 bytes):** Protocol version number, formed by concatenating the protocol major version number and protocol minor version number, encoded as follows (where MSB is Most Significant Byte and LSB is Least Significant Byte):

1st Byte (Addr: X)	2nd Byte (Addr: X+1)	3rd Byte (Addr: X+2)	4th Byte (Addr: X+3)
Minor version MSB	Minor version LSB	Major version MSB	Major version LSB

The major version number is encoded in the least significant word of the protocol version's **DWORD**.

The minor version number is encoded in the most significant word of the protocol version's **DWORD**.

Both the major and minor version number can express the version range of 0x0000 to 0xFFFF. The value of the **ProtVer** field is determined by the message type specified in the **MsgType** field, rather than the greatest version number supported by the client.

**MsgType (4 bytes):** The type of message in the message body, expressed as a binary integer. MUST be set to one of the following values.

Value	Meaning
MSG_NEGO_REQ 0x00000000	A protocol version negotiation request. The request declares the minimum and maximum version numbers supported by the requesting client-role peer. <2>To use the MSG_NEGO_REQ value, the value of the <b>ProtVer</b> field MUST be set to 1.0.
MSG_NEGO_RESP 0x00000001	A protocol version negotiation response. It is sent in response to any protocol version negotiation request or to any other request with protocol version not supported by the server-role peer.  The response declares the minimum and maximum version numbers supported by the responding server-role peer. To use the MSG_NEGO_RESP value, the value of the <b>ProtVer</b> field MUST be set to 1.0.
MSG_GETBLKLIST 0x00000002	A request for a list of <b>block hashes</b> of blocks in the <b>target segment</b> that are possessed by the destination server-role peer (list expressed as a <b>block range</b> array), and intersecting the list of block hashes specified in the request itself. To use the MSG_GETBLKLIST value, the value of the <b>ProtVer</b> field MUST be set to 1.0.
MSG_GETBLKS 0x00000003	A request for an array of block hashes (specified by a block range array). Since only one block will be returned, a MSG_GETBLKS message SHOULD specify only a single range containing only a single block. To use the MSG_GETBLKS value, the value of the <b>ProtVer</b> field MUST be set to 1.0.
MSG_BLKLIST	A response message containing a list of block hashes of blocks in the target segment that are possessed by the destination server-role peer (list expressed as a block range

Value	Meaning
0x00000004	array), and intersecting the list of block hashes specified in the previous request from the client-role peer. To use the MSG_BLKLIST value, the value of the <b>ProtVer</b> field MUST be set to 1.0.
MSG_BLK 0x00000005	A response message containing the (first) actual block requested by the client-role peer via a block range array in a MSG_GETBLKLIST message. To use the MSG_BLK value, the value of the <b>ProtVer</b> field MUST be set to 1.0.
MSG_GETSEGLIST 0x00000006	A request for a list of segments IDs that are possessed by the destination server-role peer and intersecting the list of segments IDs specified in the request itself. To use the MSG_GETSEGLIST value, the value of the <b>ProtVer</b> field MUST be set to 2.0.
MSG_SEGLIST 0x00000007	A response message containing a list of segments IDs possessed by the destination server-role peer and intersecting the list of segment IDs in the previous request from the client-role peer. To use the MSG_SEGLIST value, the value of the <b>ProtVer</b> field MUST be set to 2.0.

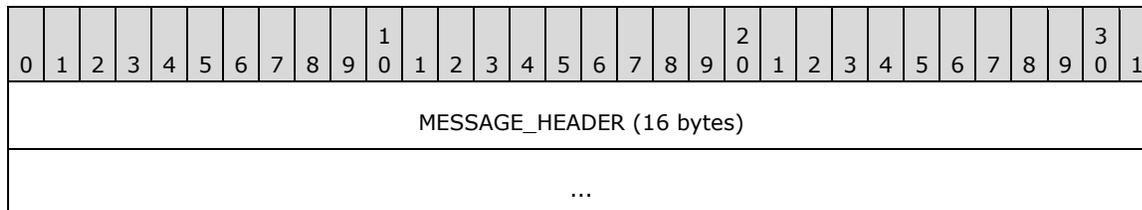
**MsgSize (4 bytes):** Protocol message total size including the MESSAGE\_HEADER, but not including the [TRANSPORT\\_RESPONSE\\_HEADER](#). The valid range of the total message size MUST be from 16 bytes to 98,304 bytes (or 96 KB) for request messages, or from 16 bytes to 393,216 bytes (or 384 KB) for response messages.

**CryptoAlgoId (4 bytes):** The encryption algorithm used by the server-role peer to encrypt data. In the request message to the server-role peer, the client-role peer SHOULD indicate AES\_128 as the preferred encryption algorithm; however, the server-role peer can ignore this value. The **CryptoAlgoId** field MUST be set to one of the following values. Refer to [\[FIPS197\]](#) for the AES standard and [\[SP800-38A\]](#) for the supported block cipher modes listed in the following table.

Value	Meaning
0x00000000	No encryption.
AES_128 0x00000001	AES 128-bit, CBC-mode encryption.
AES_192 0x00000002	AES 192-bit, CBC-mode encryption.
AES_256 0x00000003	AES 256-bit, CBC-mode encryption.

## 2.2.4 Request Message

This protocol specifies four request messages sent by the clients to the servers: [MSG\\_NEGO\\_REQ \(section 2.2.4.1\)](#), [MSG\\_GETBLKLIST \(section 2.2.4.2\)](#), [MSG\\_GETBLKS \(section 2.2.4.3\)](#), and [MSG\\_GETSEGLIST \(section 2.2.4.4\)](#). The complete layout of a request-type Peer Content Caching and Retrieval: Retrieval Protocol message is as follows.



...
MESSAGE_BODY (variable)
...

**MESSAGE\_HEADER (16 bytes):** Message header.

**MESSAGE\_BODY (variable):** Message body, which contains one of the following: Negotiation Request (MSG\_NEGO\_REQ), GetBlockList (MSG\_GETBLKLIST), GetBlocks (MSG\_GETBLKS), or GetSegmentList (MSG\_GETSEGLIST) request message.

### 2.2.4.1 MSG\_NEGO\_REQ

The MSG\_NEGO\_REQ (Negotiation Request) message is a request for the minimum and maximum protocol version supported by the target server-role peer. The message contains the minimum and maximum protocol version supported by the requesting client-role peer.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
MinSupportedProtocolVersion																															
MaxSupportedProtocolVersion																															

**MinSupportedProtocolVersion (4 bytes):** Minimum protocol version supported by the requesting peer. The protocol version is encoded identically to the **ProtVer** field defined in section [2.2.3](#).

**MaxSupportedProtocolVersion (4 bytes):** Maximum protocol version supported by the requesting peer. The protocol version is encoded identically to the **ProtVer** field defined in section [2.2.3](#).

### 2.2.4.2 MSG\_GETBLKLIST

The MSG\_GETBLKLIST (GetBlockList) message contains a request for a download block list. It is used when retrieving a set of blocks defined by one or more [BLOCK\\_RANGE\\_ARRAY](#) items.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
SizeOfSegmentID																															
SegmentID (variable)																															
...																															
ZeroPad (variable)																															
...																															
NeededBlocksRangeCount																															
NeededBlockRanges (variable)																															

...

**SizeOfSegmentID (4 bytes):** Size, in bytes, of the subsequent **SegmentID** field. The syntactic range of this field is from 0x00000000 to 0xFFFFFFFF. The actual value of this field depends on the hashing algorithm used as defined in [\[MS-PCCRC\]](#). Implementations SHOULD support all allowed **SegmentID** lengths, and MUST support content with 32-byte **SegmentIDs**. <3>

**SegmentID (variable):** Public Segment Identifier for the target segment of content (also known as **HoHoDk**). See [\[MS-PCCRC\]](#) for a description of contents, segments, blocks, and identifiers.

**ZeroPad (variable):** Sequence of bytes added (as needed) to restore 4-byte alignment, relative to the beginning of this message. The value of each byte MUST be set to zero. This field is 0 to 3 bytes in length, as required.

**NeededBlocksRangeCount (4 bytes):** Number of items in the subsequent block range array. The syntactic range of this field is from 0x00000000 to 0xFFFFFFFF. The effective range of this field MUST be between 1 and 256 inclusive, because there cannot be more than 256 non-overlapping and non-contiguous ranges in a maximum segment size of 512 blocks.

**NeededBlockRanges (variable):** Block range array listing the block hashes of the blocks within the target segment that the client-role peer is interested in. The server-role peer will reply with a block range array representing the intersection between the list of block hashes in the **NeededBlockRanges** array and the block range array set of blocks within the target segment currently available for sharing in the local cache of the server-role peer. <4>

**2.2.4.3 MSG\_GETBLKS**

The MSG\_GETBLKS (GetBlocks) message contains a request for blocks of content. It is used to retrieve a set of blocks defined by a single [BLOCK RANGE ARRAY](#).

0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
SizeOfSegmentID																																									
SegmentID (variable)																																									
...																																									
ZeroPad (variable)																																									
...																																									
ReqBlockRangeCount																																									
ReqBlockRanges (variable)																																									
...																																									
SizeOfDataForVrfBlock																																									
DataForVrfBlock (variable)																																									

...

**SizeOfSegmentID (4 bytes):** Size in bytes of the subsequent **SegmentID** field. The syntactic range of this field is from 0x00000000 to 0xFFFFFFFF. The actual value of this field depends on the hashing algorithm used as defined in [\[MS-PCCRC\]](#). Implementations SHOULD support all allowed **SegmentID** lengths, and MUST support content with 32-byte **SegmentIDs**. [<5>](#)

**SegmentID (variable):** Public Segment Identifier for the target segment of content (also known as HoHoDk). See [\[MS-PCCRC\]](#) for a description of contents, segment, blocks, and identifiers.

**ZeroPad (variable):** Sequence of bytes added (as needed) to restore 4-byte alignment, relative to the beginning of this message. The value of each byte MUST be set to zero. This field is 0 to 3 bytes in length, as required.

**ReqBlockRangeCount (4 bytes):** Number of items in the subsequent block range array. The syntactic range of this field is from 0x00000000 to 0xFFFFFFFF. The effective range of this field MUST be between 1 and 256 inclusive, because there cannot be more than 256 non-overlapping and non-contiguous ranges in a maximum segment size of 512 blocks.

**ReqBlockRanges (variable):** Block range array representing the blocks requested for the target segment. **ReqBlockRanges** MUST specify a single block range containing only one block.

**SizeOfDataForVrfBlock (4 bytes):** Size in bytes of the subsequent **DataForVrfBlock** field. This field SHOULD be zero.

**DataForVrfBlock (variable):** Not used by the protocol. This field SHOULD be empty.

**2.2.4.4 MSG\_GETSEGLIST**

The MSG\_GETSEGLIST (GetSegmentList) message contains a request for a download segment list. It is used when retrieving a set of segments. This message MUST be formatted as follows: [<6>](#)

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
RequestID (16 bytes)																															
...																															
...																															
CountOfSegmentIDs																															
SizeOfSegmentID																															
SegmentID (variable)																															
...																															
ZeroPad (variable)																															
...																															
SizeOfExtensibleBlob																															

ExtensibleBlob (variable)
...

**RequestID (16 bytes):** Unique identifier among all outstanding GetSegmentList Requests from this peer.

**CountOfSegmentIDs (4 bytes):** Count of the Segment IDs in the current GetSegmentList Request.

The following three fields are repeated N times where N is equal to CountOfSegmentIDs.

**SizeOfSegmentID (4 bytes):** Size, in bytes of the first SegmentID, immediately subsequent to this field. Implementations MUST support all allowed **SegmentID** lengths, and MUST support content with 32-byte **SegmentIDs**.

**SegmentID (variable):** Public Segment Identifier for the first target segment of content (also known as HoHoDk). See [\[MS-PCCRC\]](#) section 2.2 for a description of segment identifiers.

**ZeroPad (variable):** Sequence of bytes added (as needed) to restore 4-byte alignment, relative to the beginning of this message. The value of each byte MUST be set to zero. This field is 0 to 3 bytes in length, as required.

**SizeOfExtensibleBlob (4 bytes):** Size, in bytes, of the **ExtensibleBlob** field. Implementations MAY support extensible blobs in MSG\_GETSEGLIST messages. Implementations that do not support extensible blobs in MSG\_GETSEGLIST messages MUST set **SizeOfExtensibleBlob** to zero and omit the **ExtensibleBlob** field.

**ExtensibleBlob (variable):** An extensible binary large object (BLOB). See [Extensible BLOB \(section 2.2.6\)](#) for the definition of currently defined extensible BLOBs. Implementations MAY support extensible BLOBs in MSG\_GETSEGLIST messages. Implementations that do not support extensible BLOBs in MSG\_GETSEGLIST messages MUST set **SizeOfExtensibleBlob** to zero and omit the **ExtensibleBlob** field. Relative indexes contained in the extensible BLOB are relative to the first segment in the first **SegmentRange** carried by the current MSG\_GETSEGLIST message.

## 2.2.5 Response Message

The Retrieval Protocol defines four response messages sent by the servers in response to client requests: [MSG\\_NEGO\\_RESP \(section 2.2.5.1\)](#), [MSG\\_BLKLIST \(section 2.2.5.2\)](#), [MSG\\_BLK \(section 2.2.5.3\)](#), and [MSG\\_SEGLIST \(section 2.2.4.4\)](#). The complete layout of a response-type Peer Content Caching and Retrieval: Retrieval Protocol message is as follows.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
TRANSPORT_RESPONSE_HEADER																															
MESSAGE_HEADER (16 bytes)																															
...																															
...																															
MESSAGE_BODY (variable)																															
...																															

**TRANSPORT\_RESPONSE\_HEADER (4 bytes):** Transport response header.

**MESSAGE\_HEADER (16 bytes):** Message header.

**MESSAGE\_BODY (variable):** Message body, which can contain one of the following:  
MSG\_NEGO\_RESP, MSG\_BLKLIST, MSG\_BLK, or a MSG\_SEGLIST message.

### 2.2.5.1 MSG\_NEGO\_RESP

The MSG\_NEGO\_RESP (Negotiation Response) message is the response message containing the minimum and maximum protocol version supported by the responding server-role peer. The message is sent in response to a Negotiation Request message or to any other request message with a protocol version not supported by the server-role peer.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
MinSupportedProtocolVersion																															
MaxSupportedProtocolVersion																															

**MinSupportedProtocolVersion (4 bytes):** Minimum protocol version supported by the requesting peer. The protocol version is encoded identically to the **ProtVer** field defined in section [2.2.3](#).

**MaxSupportedProtocolVersion (4 bytes):** Maximum protocol version supported by the requesting peer. The protocol version is encoded identically to the **ProtVer** field defined in section 2.2.3.

### 2.2.5.2 MSG\_BLKLIST

The MSG\_BLKLIST message is the response message containing the download block ranges available on the responding server-role peer. The message is sent by the server-role peer in response to a MSG\_GETBLKLIST message from a requesting client-role peer.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
SizeOfSegmentId																															
SegmentId (variable)																															
...																															
ZeroPad (variable)																															
...																															
BlockRangeCount																															
BlockRanges (variable)																															
...																															
NextBlockIndex																															

**SizeOfSegmentId (4 bytes):** The size, in bytes, of the subsequent **SegmentId** field.

**SegmentId (variable):** The Public Segment Identifier for the target segment of content (also known as HoHoDk). See [\[MS-PCCRC\]](#) for details.

**ZeroPad (variable):** A sequence of N bytes added (only as needed) to restore 4-byte alignment, where  $0 \leq N \leq 3$ . Each byte's value MUST be set to zero.

**BlockRangeCount (4 bytes):** Number of items in the subsequent block range array. The server MUST set the **BlockRangeCount** field to 0 if it does not have any of the requested block ranges.

**BlockRanges (variable):** A block range array describing the blocks currently available for download from the current server-role peer for the target segment, within the boundaries of the list of block ranges of interest (**NeededBlockRanges**) specified by the client-role peer in the previously received GetBlockList request message ([MSG\\_GETBLKLIST \(section 2.2.4.2\)](#)). [<7>](#)

**NextBlockIndex (4 bytes):** The index of the first block after the block sent in the current message, currently available for download from this server-role peer. If no such next block is available, this index MUST be zero.

### 2.2.5.3 MSG\_BLK

The MSG\_BLK message is the response message containing a download block. This message is sent by the server-role peer in response to a MSG\_GETBLKS message from a requesting client-role peer.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
SizeOfSegmentId																															
SegmentId (variable)																															
...																															
ZeroPad (variable)																															
...																															
BlockIndex																															
NextBlockIndex																															
SizeOfBlock																															
Block (variable)																															
...																															
ZeroPad_2 (variable)																															
...																															
SizeOfVrfBlock																															

VrfBlock (variable)
...
ZeroPad_3 (variable)
...
SizeOfIVBlock
IVBlock (variable)
...

**SizeOfSegmentId (4 bytes):** The size, in bytes, of the subsequent **SegmentId** field.

**SegmentId (variable):** The Public Segment Identifier for the target segment of content (also known as HoHoDk). See [\[MS-PCCRC\]](#) for details.

**ZeroPad (variable):** A sequence of N bytes added (only as needed) to restore 4-byte alignment, where  $0 \leq N \leq 3$ . Each byte's value MUST be set to zero.

**BlockIndex (4 bytes):** The index in the target segment of the block sent in the current message.

**NextBlockIndex (4 bytes):** The index of the first block after the block sent in the current message, currently available for download from this server-role peer. If no such next block is available, this index MUST be zero.

**SizeOfBlock (4 bytes):** The size, in bytes, of the subsequent **Block** field. The server MUST set the **SizeOfBlock** field to zero if it does not have the requested block.

**Block (variable):** The actual block of data, encrypted according to the cryptographic algorithm specified in the header of the message itself, not including the **initialization vector**.

**ZeroPad\_2 (variable):** A sequence of N bytes added (only as needed) to restore 4-byte alignment, where  $0 \leq N \leq 3$ . Each byte's value MUST be set to zero.

**SizeOfVrfBlock (4 bytes):** The size, in bytes, of the subsequent **VrfBlock** field, which SHOULD be zero.

**VrfBlock (variable):** Currently not used, and SHOULD be empty.

**ZeroPad\_3 (variable):** A sequence of N bytes added (only as needed) to restore 4-byte alignment, where  $0 \leq N \leq 3$ . Each byte's value MUST be set to zero.

**SizeOfIVBlock (4 bytes):** The size, in bytes, of the subsequent **IVBlock** field.

**IVBlock (variable):** The initialization vector used by the server-role peer when encrypting the block of data (**Block** field) sent with this message.

#### 2.2.5.4 MSG\_SEGLIST

The MSG\_SEGLIST message is the response message containing the segment range array describing the segments currently available for download. This message is sent by the server-role peer in response to a [MSG\\_GETSEGLIST \(section 2.2.4.4\)](#) message from a requesting client-role peer.

The MSG\_SEGLIST MUST be formatted as follows: [<8>](#)

0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
RequestID (16 bytes)																															
...																															
...																															
SegmentRangeCount																															
SegmentRanges (variable)																															
...																															
SizeOfExtensibleBlob																															
ExtensibleBlob (variable)																															
...																															

**RequestID (16 bytes):** Unique identifier matching the **RequestID** received in the associated GetSegmentList Request.

**SegmentRangeCount (4 bytes):** Number of items in the **SegmentRanges** field. The server MUST set the **SegmentRangeCount** field to 0 if it does not have any of the requested segments.

**SegmentRanges (variable):** A [SEGMENT\\_RANGE\\_ARRAY \(section 2.2.1.4\)](#) that describes the segments (full or partial) currently available for download from the current server-role peer. The indexes specified in each range in the response are the relative indexes of the segment in the original array of segment IDs specified in the associated GetSegmentList message.

**SizeOfExtensibleBlob (4 bytes):** The size in bytes of the **ExtensibleBlob** field.

**ExtensibleBlob (variable):** An extensible binary large object (BLOB). For the definitions of extensible BLOBs that are currently defined, see section [2.2.6](#). The relative indexes contained in the extensible BLOB are relative to the first segment in the first [SEGMENT\\_RANGE \(section 2.2.1.2\)](#) carried by the current MSG\_SEGLIST message.

## 2.2.6 Extensible BLOB

Extensible binary large objects (BLOBs) are optional BLOBs that SHOULD be included in MSG\_SEGLIST responses and can be included in MSG\_GETSEGLIST requests.

Extensible BLOBs carry more information about the specific segments associated to the MSG\_GETSEGLIST/MSG\_SEGLIST exchange.

Extensible BLOBs are versioned, and for each version there is a well-known definition. Currently the only defined extensible BLOB version is extensible BLOB version 1. Future protocol versions can define new extensible BLOB versions.

The version of the extensible BLOB is encoded as a network-byte-order unsigned short integer in the first two bytes of the extensible BLOB itself.

Any extensible BLOB smaller than 2 bytes is invalid and MUST be discarded.

Implementations that do not support the specific extensible BLOB version encoded in the first two bytes of the BLOB itself MUST discard the BLOB.

Implementations that do support the specific extensible BLOB version encoded in the first two bytes of the BLOB itself MUST make sure that the size of the BLOB reported through **SizeOfExtensibleBlob** is compatible with the size restrictions for that specific BLOB version; if not, the BLOB MUST be discarded.

### 2.2.6.1 Extensible Blob Version 1

The Extensible Blob Version 1 MUST be formatted as follows.

0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
ExtensibleBlobVersion											SegmentAgeUnits											SegmentAgeCount									
SegmentAges (variable)																															
...																															

**ExtensibleBlobVersion (2 bytes):** Network-byte-order unsigned short integer that contains the version of the extensible blob. It must be equal to 1.

#### SegmentAgeUnits (1 byte):

Unit used to specify the age of the segments in the following ENCODED\_SEGMENT\_AGE structures. Valid values are:

- 1=seconds
- 2=tenths of a second
- 3=hundredths of a second
- 4=milliseconds

When sending a message containing an extensible blob version 1, implementations SHOULD use a value of 3 (hundredths of a second). When receiving a message containing a blob, implementations MUST be able to process any of the above valid values.

**SegmentAgeCount (1 byte):** Count of ENCODED\_SEGMENT\_AGE structures encoded right after this field (acceptable range: 0 - 255).

**SegmentAges (variable):** SegmentAgeCount ENCODED\_SEGMENT\_AGE structures.

#### 2.2.6.1.1 Extensible Blob Version 1 Restrictions and Validation

Before parsing the data in the blob, implementations MUST verify its validity as follows:

1. The blob size MUST be at least four bytes.
2. The blob MUST have a **SegmentAgeUnits** value in the range of 1 to 4.
3. The blob size specified by the enclosing message MUST be equal to or larger than the size needed in order to accommodate the number of ENCODED\_SEGMENT\_AGE structures specified by **SegmentAgeCount**, plus four bytes for the above-described blob layout. That is:

Specified\_size >= 4 + SegmentAgeCount \* sizeof(ENCODED\_SEGMENT\_AGE) = 4 +  
SegmentAgeCount \* 4 = 4 \* (SegmentAgeCount + 1)

The above validations MUST be carried out in the specified order and MUST be aborted at the first failure. If validation fails, the blob MUST be discarded.

## 3 Protocol Details

The Retrieval Protocol consists of four types of exchanges: the Negotiation request/response, the BlockList request/response, the Block request/response, and the SegmentList request/response. (See section 2 for message formats and field definitions.)

- **Protocol Version Negotiation:** A client-role peer (referred to here simply as "client") initiates a protocol version negotiation with a server-role peer (referred to here simply as "server") by sending a Negotiation Request message ([MSG\\_NEGO\\_REQ \(section 2.2.4.1\)](#)), declaring the minimum and maximum protocol versions it supports. The server responds with a Negotiation Response message ([MSG\\_NEGO\\_RESP \(section 2.2.5.1\)](#)), declaring the minimum and maximum protocol versions it supports. Implementation of the client side of the protocol version negotiation is optional. The server side of the protocol version negotiation **MUST** be implemented.
- **BlockList request/response:** A client initiates a GetBlockList request ([MSG\\_GETBLKLIST \(section 2.2.4.2\)](#)) to a server in order to query the list of content blocks available on the server for a given segment ID, and a list of block ranges within the segment, by sending a MSG\_GETBLKLIST request. The server responds with a BlockList response ([MSG\\_BLKLIST \(section 2.2.5.2\)](#)) containing the list of block ranges for the specific segment ID that are within the ranges of the request. If the server does not support the client's protocol version, it treats the request as a Negotiation Request and responds accordingly (see section [3.2.5.1](#)).
- **Blocks request/response:** A client initiates a GetBlocks request ([MSG\\_GETBLKS \(section 2.2.4.3\)](#)) to a server to retrieve a specific block of a given segment, which is identified by the segment ID and the index of the block in the segment. It does this by sending a MSG\_GETBLKS request. The server responds with the requested content blocks in a Block response ([MSG\\_BLK \(section 2.2.5.3\)](#)). If the server does not support the client's protocol version, it treats the request as a Negotiation Request, and responds accordingly (see section [3.2.5.1](#)).
- **SegmentList request/response:** SegmentList exchanges are restricted to Hosted Cache client/server interaction only. A SegmentList request **MUST** be sent only by a Hosted Cache Client and it **MUST** always be directed to a Hosted Cache Server. A client initiates a SegmentList request (MSG\_GETSEGLIST (section [2.2.4.4](#))) to a server in order to query the availability on the server of an array of segments, each of them identified by its segment ID. The server responds with a SegmentList response (MSG\_SEGLIST (section [2.2.5.4](#))) containing the list of segment ranges for the specific segment IDs that are within the requested array. The indexes specified in the ranges are the indexes of segments in an array of Segment IDs in the originating MSG\_GETSEGLIST request. If the server does not support the client's protocol version, it treats the request as a Negotiation Request and responds accordingly (see section [3.2.5.1](#)).

The Peer Content Caching and Retrieval Framework (also referred to as simply "the framework") then uses the Retrieval Protocol to retrieve and assemble complete segments of a content from a combination of sources, including either a set of server peers or a hosted cache, plus the original content server if the former does not have the complete content.

### 3.1 Client Details

#### 3.1.1 Abstract Data Model

The Retrieval Protocol client maintains the following data:

- **Outstanding Request List:** A list of request messages sent for which responses have not yet been received, along with the addresses of the peers to which they were sent.

### 3.1.2 Timers

- **Request Timer:** A per-request-message timer set by the client whenever it sends a [MSG\\_NEGO\\_REQ \(section 2.2.4.1\)](#), [MSG\\_GETBLKLIST \(section 2.2.4.2\)](#), [MSG\\_GETBLKS \(section 2.2.4.3\)](#), or [MSG\\_GETSEGLIST \(section 2.2.4.4\)](#) request message. When the timer expires before the exchange is completed, the client **MUST** cancel the current exchange. The default timeout value **MUST** be set to 2 seconds.<9>

### 3.1.3 Initialization

The Retrieval Protocol requires no explicit initialization for clients.

### 3.1.4 Higher-Layer Triggered Events

#### 3.1.4.1 MSG\_NEGO\_REQ Request

An implementation of the Retrieval Protocol **MAY** support the sending of a Protocol Version Negotiation Request message ([MSG\\_NEGO\\_REQ \(section 2.2.4.1\)](#)), when triggered by a higher-layer application. The following description only applies to the implementations that support this feature.

When initiating a Retrieval Protocol query for the supported protocol versions, the higher-layer applications **MUST** specify a server address. The Retrieval Protocol implementation **MUST**:

1. Construct an MSG\_NEGO\_REQ message (as specified in section 2.2.4.1).
2. Fill in the minimum and maximum protocol versions it supports.
3. Send the message to the server.
4. Store the message in the **Outstanding Request List**.
5. Start the **Request Timer**.

#### 3.1.4.2 MSG\_GETBLKLIST Initiation

To initiate a Retrieval Protocol query for the list of block ranges on a server, the higher-layer applications **MUST** specify a server address, a segment ID, and a set of block ranges within the segment identified by the segment ID. The client instance of the Retrieval Protocol instantiation **MUST** construct and send a GetBlockList message ([MSG\\_GETBLKLIST \(section 2.2.4.2\)](#)) to the server, store it in the **Outstanding Request List (3.1.1)**, and start the **Request Timer (3.1.2)**. The **SegmentID** and **NeededBlocksRanges** fields of the GetBlockList message correspond to the segment ID and the set of block ranges supplied by the higher-layer applications.

#### 3.1.4.3 MSG\_GETBLKS Initiation

To initiate a Retrieval Protocol request for specific block ranges, the higher-layer applications **MUST** specify a server address, a segment ID, and a set of block ranges with the segment identified by the segment ID. The client instance of the Retrieval Protocol **MUST** construct and send a GetBlocks message ([MSG\\_GETBLKS \(section 2.2.4.3\)](#)) to the server, store it in the **Outstanding Request List (3.1.1)**, and start the **Request Timer (3.1.2)**. The **SegmentID** and **ReqBlockRanges** fields correspond to the segment ID and the block ranges of the request.

The Retrieval Protocol **MUST** only request and retrieve one block per exchange of MSG\_GETBLKS request and [MSG\\_BLK \(section 2.2.5.3\)](#) response messages. If the higher-layer applications need to retrieve more than one block, multiple GetBlocks messages **MUST** be sent with one block per request.

A peer SHOULD perform a **simple download** if it involves a limited number of consecutive blocks in a single block range. <10> This implies that the blocks are consecutive in the segment.

#### 3.1.4.4 MSG\_GETSEGLIST Initiation

To initiate a Retrieval Protocol request for an array of segment IDs, the higher-layer application MUST generate a sequential or random non-repeating Request ID, and it MUST specify a server address and an array of segment IDs. The client instance of the Retrieval Protocol MUST construct and send a GetSegmentList message (MSG\_GETSEGLIST (section 2.2.4.4)) to the server, store it in the Outstanding Request List (3.1.1), and start the Request Timer (3.1.2). The **RequestID** and **SegmentID** fields correspond to the ID of the current request and to the segment IDs contained in the request.

### 3.1.5 Message Processing Events and Sequencing Rules

#### 3.1.5.1 MSG\_NEGO\_RESP Received

On receiving a [MSG\\_NEGO\\_RESP \(section 2.2.5.1\)](#) response message from a server, the client MUST first determine if this is a response to a previously sent request by checking the **Outstanding Request List** for the address of the server. If it is a response to either a [MSG\\_GETBLKLIST \(section 2.2.4.2\)](#) or [MSG\\_GETBLKS \(section 2.2.4.3\)](#) message, the client MUST compare the ranges of protocol versions and select a protocol version based on the rules described later in this section. It must then resend the original MSG\_GETBLKLIST or MSG\_GETBLKS message, using the selected version.

If it is a response to an [MSG\\_NEGO\\_REQ \(section 2.2.4.1\)](#), the client MUST return the highest protocol version supported by both the client and the server to the higher-layer applications.

If the client and server are incompatible, then the client MUST silently discard the MSG\_NEGO\_RESP message and MUST abort any current exchange with the server, including exchanges for previously sent MSG\_GETBLKLIST, MSG\_GETBLKS, or MSG\_NEGO\_REQ messages.

The rules for determining compatibility and selecting a version are listed below:

1. The client's and the server's major version ranges are calculated from the MSG\_NEGO\_REQ and MSG\_NEGO\_RESP messages, respectively. In both cases, they are defined as the inclusive range between the major version from the **MinSupportedProtocolVersion** field and the major version from the **MaxSupportedProtocolVersion** field. The highest common major version is the highest value that is included in both ranges. If these ranges do not contain any common values, then no highest common major version exists.
2. The minor versions within the same major version do not affect protocol compatibility. For instance, a client sending a version 3.2 request message and a server replying with version 3.0 message are fully compatible. The client and the server MUST each select their own highest minor version supported within the highest common major version. For example, if the client supports protocol version range [3.2, 5.0] and the server supports protocol version range [2.0, 4.3], then the highest common major version is 4, and the client will be sending messages with version 4.8 (assuming the highest minor version number for major version 4 is 4.8), whereas the server will be replying with messages with version 4.3. Another example: a client with a supported version range of [1.0, 2.1] and a server with a supported range of [2.5, 2.9] will result in a highest common major version of 2, with the client using version 2.1 and the server using version 2.9.
3. If no highest common major version exists, then the client and the server are incompatible.

If there is no existing request message previously sent to the server stored in the **Outstanding Request List**, the client MUST silently discard the received message.

### 3.1.5.2 MSG\_BLKLIST Response Received

On receiving a [MSG\\_BLKLIST \(section 2.2.5.2\)](#) response message from a server, the client MUST verify that it is well-formed and corresponds to a GetBlockList request message ([MSG\\_GETBLKLIST \(section 2.2.4.2\)](#)) in its **Outstanding Request List**. The client then performs the following checks:

- The client SHOULD verify if the segment ID matches any request in the **Outstanding Request List**. If the client performs the segment ID check, it MUST silently discard the MSG\_BLKLIST message and abort the exchange if the segment ID does not match the segment ID of any request. <11>
- The client MUST check if the block ranges overlap with the ranges specified in any request with a matching segment ID in the **Outstanding Request List**. The client MUST silently discard the MSG\_BLKLIST message and abort the exchange if the check fails.

If this verification is successful, then the peer MUST:

- Delete the corresponding request message from the **Outstanding Request List**, and cancel its **Request Timer**.
- Return the segment ID and block range from the MSG\_BLKLIST message, as well as the server address, to the higher-layer applications.

Otherwise, the response message MUST be silently discarded.

### 3.1.5.3 MSG\_BLK Response Received

On receiving a [MSG\\_BLK \(section 2.2.5.3\)](#) response message from a discovered peer, the client MUST verify that it is well-formed and corresponds to a GetBlocks request message ([MSG\\_GETBLKS \(section 2.2.4.3\)](#)) in its **Outstanding Request List** (the segment ID and block index would match that of an outstanding GetBlocks request). The client MUST silently discard the message if this verification is unsuccessful. Otherwise, it MUST:

- Delete the corresponding request message from the **Outstanding Request List**, and cancel its **Request Timer**.
- If an encryption algorithm is specified in the MSG\_BLK message (the **CryptoAlgoId** field does not equal 0x00000000 as specified in section [2.2.3](#)), decrypt the block using the pre-provisioned key.
- Pass the segment ID, block index, and (decrypted) block up to the higher-layer applications.

Otherwise, the response message MUST be silently discarded and the exchange aborted.

### 3.1.5.4 MSG\_SEGLIST Response Received

On receiving a MSG\_SEGLIST (section [2.2.5.4](#)) response message from a server, the client MUST verify that it is a well-formed GetSegmentList request message (per section [2.2.4.4](#)) in its **Outstanding Request List**. The client then performs the following checks:

- The client MUST verify that the **RequestID** matches any request in the **Outstanding Request List**. If a request with a matching **RequestID** is not found in the **Outstanding Request List**, the client MUST silently discard the MSG\_SEGLIST message and abort the exchange.
- The client MUST make sure that the all index entries from the ranges reported in the MSG\_SEGLIST received are contained within the array of segment IDs specified in the request matching the **RequestID** in the **Outstanding Request List**. The client MUST silently discard the MSG\_SEGLIST message and abort the exchange if the check fails.

If this verification is successful, then the peer **MUST** do the following:

- Delete the corresponding request message from the **Outstanding Request List**, and cancel its Request Timer.
- Return the list of segment IDs identified by the ranges from the MSG\_SEGLIST message, as well as the server address, to the higher-layer applications.

Otherwise, the response message **MUST** be silently discarded.

### 3.1.5.5 Other Messages Received

All malformed messages received by the client and messages of unknown type sent to the Retrieval Protocol URLs specified in section [2.1.1](#) **MUST** be silently discarded.

### 3.1.6 Timer Events

#### 3.1.6.1 Request Timer Expiration

When the **Request Timer** expires before the exchange (GetBlockList ([MSG\\_GETBLKLIST \(section 2.2.4.2\)](#)), GetBlocks ([MSG\\_GETBLKS \(section 2.2.4.3\)](#)), Negotiation Request ([MSG\\_NEGO\\_REQ \(section 2.2.4.1\)](#)), or GetSegmentList ([MSG\\_GETSEGLIST \(section 2.2.4.4\)](#))) is completed, the client **MUST** abort the current exchange.

#### 3.1.7 Other Local Events

None.

## 3.2 Server Details

### 3.2.1 Abstract Data Model

- **Content Cache:** This is the local content cache on the server. It consists of a list of **segment IDs** and associated block ranges, along with their Content Information (see [\[MS-PCCRC\]](#) section 2.3) and corresponding content blocks that the client or server has previously obtained either from other peers or from the content server. The server replies to client queries with the information and content blocks stored in its content cache; the client retrieves the content from the server using the Core Retrieval Protocol.
- **Active Client Count:** This counter keeps the number of active clients the server is currently serving. The counter is incremented by 1 when the server receives a request (GetBlockList ([MSG\\_GETBLKLIST \(section 2.2.4.2\)](#)) or GetBlocks ([MSG\\_GETBLKS \(section 2.2.4.3\)](#))), and is decremented by 1 when the server sends back a response or discards the request. This counter is used to limit the number of concurrent clients for a server to a maximum value. The default maximum threshold **SHOULD** be set to 64<12><13>, and it **MUST** be configurable based on the processing capability of the server. If this counter reaches the threshold, the server will send back an empty response (empty block range in BlockList ([MSG\\_BLKLIST \(section 2.2.5.2\)](#)) or empty block in Block ([MSG\\_BLK \(section 2.2.5.3\)](#))) to the client.

### 3.2.2 Timers

- **Upload Timer:** A per-instantiation timer set by a server when the protocol is instantiated. The server **MUST** abort the protocol instance when the timer expires before the request/response exchange is completed. The default timeout value **MUST** be set to 15 seconds.<14>

### 3.2.3 Initialization

The server is initialized by starting to listen for incoming HTTP requests on the URL specified in section [2.1.1](#). The server MUST set the **Active Client Count** to zero.

### 3.2.4 Higher-Layer Triggered Events

There are no explicit higher-layer triggered events for the server, other than waiting for the client messages as enabled by the initialization.

### 3.2.5 Message Processing Events and Sequencing Rules

#### 3.2.5.1 MSG\_NEGO\_REQ Received

On receiving a valid [MSG\\_NEGO\\_REQ \(section 2.2.4.1\)](#) message from a client, the server MUST construct a [MSG\\_NEGO\\_RESP \(section 2.2.5.1\)](#) message with the maximum and minimum protocol versions that it supports, set the **Upload Timer**, and send the response message back to the client.

#### 3.2.5.2 MSG\_GETBLKLIST Request Received

On receiving a valid [MSG\\_GETBLKLIST \(section 2.2.4.2\)](#) request message from a client, the server MUST perform the following actions in the order specified:

1. The server MUST first check if the protocol version is supported, based on the version range comparison rules specified in section [3.1.5.1](#). If the major protocol version is outside the range of the server implementation, the server MUST construct an [MSG\\_NEGO\\_RESP \(section 2.2.5.1\)](#) message, fill it in with the maximum and minimum protocol versions it supports, and send the MSG\_NEGO\_RESP message back to the client.
2. If the major version is supported by the server, the server MUST select a compatible protocol version based on the same rules specified in section 3.1.5.1 for the following reply message.
3. The server MUST check if its **Active Client Count** is greater than or equal to the maximum number allowed. If the server is already serving more than or equal to the maximum number of clients, the server MUST reply to the client using a [MSG\\_BLKLIST \(section 2.2.5.2\)](#) message with an empty block range.
4. Otherwise, the server MUST increment the **Active Client Count** by 1, set the **Upload Timer**, and compute the intersection of the block ranges (for the segment specified) in the MSG\_GETBLKLIST request with the block ranges for the same segment in the server's **Content Cache**. The server MUST then send the client a MSG\_BLKLIST response message containing the segment ID listed in the MSG\_GETBLKLIST request message, and the computed intersection block ranges (possibly empty).
5. Once the MSG\_BLKLIST response message is sent, the server MUST decrement the **Active Client Count** by 1.

#### 3.2.5.3 MSG\_GETBLKS Request Received

On receiving a valid [MSG\\_GETBLKS \(section 2.2.4.3\)](#) request message from a client, the server MUST perform the following actions in the order specified:

1. The server MUST first check if the protocol version is supported, based on the version range comparison rules specified in section [3.1.5.1](#). If the major version is outside the range of the server implementation, the server MUST construct a [MSG\\_NEGO\\_RESP \(section 2.2.5.1\)](#) message, fill in the maximum and minimum protocol versions it supports, and send the MSG\_NEGO\_RESP message back to the client.

2. If the major version is supported by the server, the server MUST select a compatible protocol version based on the same rules specified in section 3.1.5.1 for the following reply message.
3. The server MUST check if its **Active Client Count** ([Abstract Data Model](#), section 3.2.1) is greater than or equal to the maximum number allowed. If the server is already serving more than or equal to the maximum number of clients, the server MUST reply to the client using a [MSG\\_BLK \(section 2.2.5.3\)](#) message with an empty block.
4. Otherwise, the server MUST increment the **Active Client Count** by 1, set the **Upload Timer** ([Timers](#), section 2.2.5.3), and construct and send the client a MSG\_BLK response message containing a block that is selected based on the following rules:
  - If the block ranges in the MSG\_GETBLKS request message contain only one block, the server MUST select the requested block.
  - If the block ranges contain more than one block, the server SHOULD select the first (smallest-index) block from the block ranges and the segment that is specified in the request message.

The server then MUST check whether the selected block exists in the server's **Content Cache** ([Abstract Data Model](#) section 2.2.5.1). If it does, then the server MUST include this block in the MSG\_BLK response message it sends. Otherwise, the response MUST contain an empty MSG\_BLK response message. The **SegmentID** field in the response message MUST be set to the segment ID of the request, and the **BlockIndex** field MUST be set to the index of the block sent in this message. The server MUST also calculate the value of the **NextBlockIndex** field (section 2.2.5.3).

The server MUST apply the encryption algorithm chosen by the upper-layer application to the block in MSG\_BLK response message. The list of permissible encryption algorithms is given by the **CryptoAlgoID** value table in section [2.2.3](#).

5. Once the MSG\_BLK message is sent, the server MUST decrement the **Active Client Count** by 1. If the resulting value is negative, the server MUST set the counter to zero.

#### 3.2.5.4 MSG\_GETSEGLIST Request Received

On receiving a valid MSG\_GETSEGLIST (section [3.1.5.4](#)) request message from a client, the server MUST perform the following actions in the order specified:

1. The server MUST first check if the protocol version is supported, based on the version range comparison rules specified in section [3.1.5.1](#). If the major protocol version is outside the range of the server implementation, the server MUST construct an MSG\_NEGO\_RESP (section [2.2.5.1](#)) message, including the maximum and minimum protocol versions it supports, and send the MSG\_NEGO\_RESP message back to the client.
2. If the major version is supported by the server, the server MUST select a compatible protocol version based on the same rules specified in section 3.1.5.1 for the following reply message.
3. The server MUST check if its Active Client Count is greater than or equal to the maximum number allowed. If the server is already serving the maximum number of clients or more, the server MUST reply to the client using a MSG\_SEGLIST (section [2.2.5.4](#)) message with an empty segment range.
4. Otherwise, the server MUST increment the Active Client Count by 1, set the Upload Timer, and retrieve the availability in the server's Content Cache for each of the segments specified in the GetSegmentList request. Based on the availability of these segments, the server MUST build an array of ranges identifying the segments partially or fully available in the local cache, by identifying each segment with the index of its segment ID in the array of segment IDs contained in the originating GetSegmentList request. The server MUST then send the client a MSG\_SEGLIST response message containing the array of segment ranges just calculated (possibly empty).

5. Once the MSG\_SEGLIST response message is sent, the server MUST decrement the Active Client Count by 1.

### **3.2.5.5 Other Messages Received**

All malformed messages received by the server and messages of unknown types sent to the Retrieval Protocol URLs specified in section [2.1.1](#) MUST be silently discarded.

### **3.2.6 Timer Events**

#### **3.2.6.1 Upload Timer Expiration**

When the **Upload Timer** expires, the server-role peer MUST abort the protocol instance.

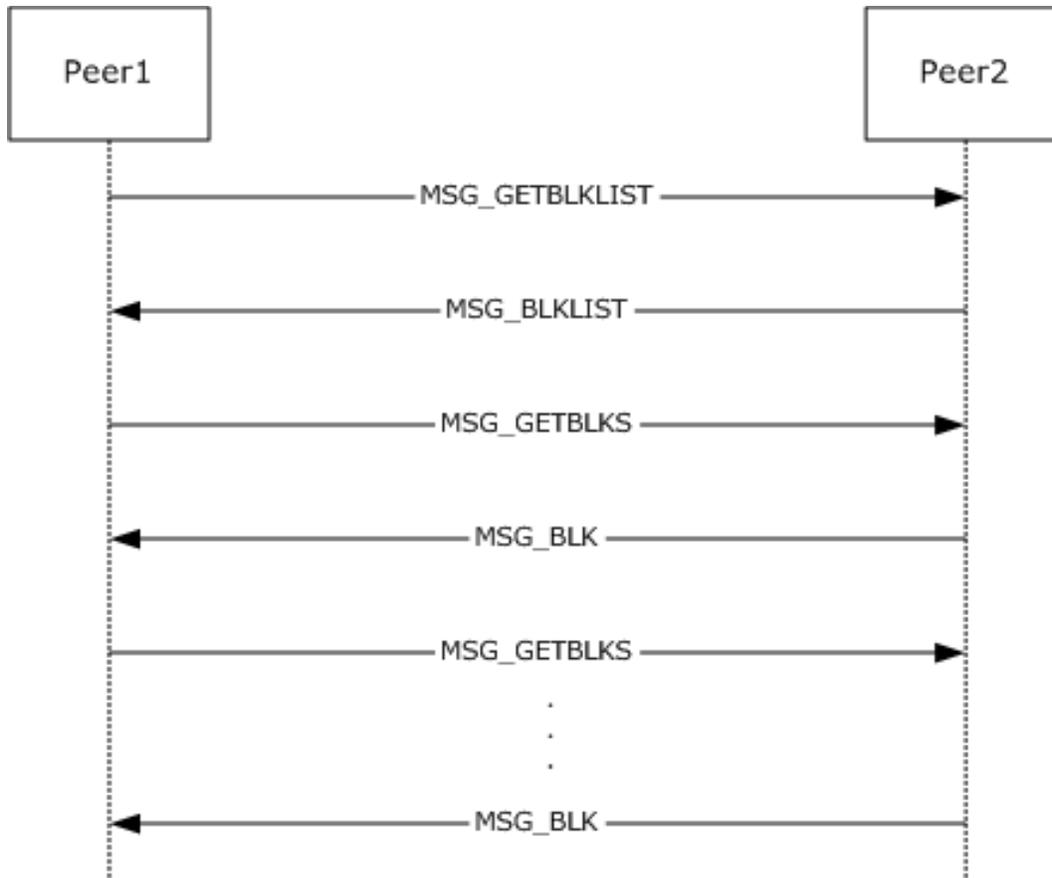
#### **3.2.7 Other Local Events**

None.

## 4 Protocol Examples

### 4.1 Download with GetBlockList and GetBlocks Exchanges

Scenario: Peer P1 is trying to download blocks BN0 :- BN1 and BN2 :- BN3 of segment S1 from peer P2.



**Figure 2: Download using GetBlockList and GetBlocks request/response pairs**

Sequence of events:

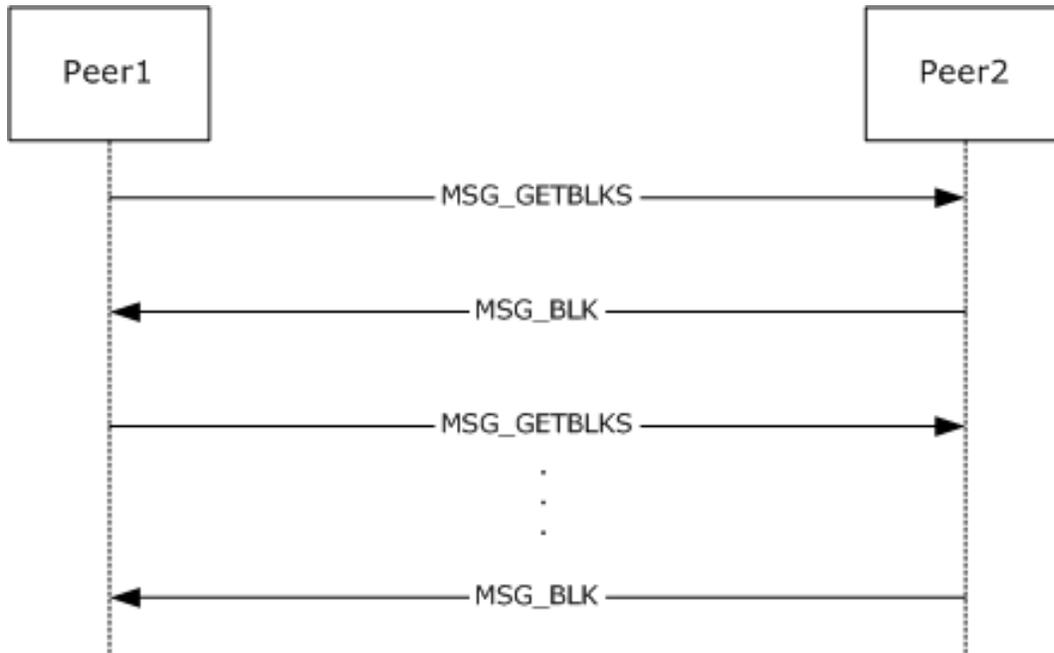
1. Peer P1 sends a GetBlockList message ([MSG\\_GETBLKLIST \(section 2.2.4.2\)](#)) to P2, specifying two block ranges of interest, one for BN0 :- BN1 and one for BN2 :- BN3.
2. Upon receiving the GetBlockList message, P2 gathers the list of blocks it currently has for the target segment S1, it intersects that with the list of needed blocks specified by P1 in the GetBlockList message, and it sends back to P1 a BlockList message ([MSG\\_BLKLIST \(section 2.2.5.2\)](#)) containing the set resulting from the previous intersection.
3. Upon receiving the BlockList message from P2, P1 starts downloading blocks by sending a GetBlocks message ([MSG\\_GETBLKS \(section 2.2.4.3\)](#)) for one block at a time.
4. Upon receiving the GetBlocks message for a given block, P2 replies with a Block message ([MSG\\_BLK \(section 2.2.5.3\)](#)) containing the actual block of data encrypted using the crypto algorithm selected locally on the server-role peer P2. The **encryption key** was previously

provisioned on both P1 and P2. Refer to section [5.1](#) for the detailed requirements on encryption and initialization vector generation.

The encrypted block and initialization vector are added to the Block message and sent back to the client-role peer. The requesting peer will be able to decrypt the data only if it knows the hash of data of the segment.

## 4.2 Simple Download with GetBlocks Download Sub-Sessions only

Scenario: Peer P1 is trying to download two consecutive blocks, BN and BN + 1, of segment S1 from peer P2.



**Figure 3: Simple Download using GetBlocks request/response**

Sequence of events:

1. Since the download involves only two blocks, P1 decides to skip the GetBlockList message ([MSG\\_GETBLKLIST \(section 2.2.4.2\)](#)). It sends a GetBlocks message ([MSG\\_GETBLKS \(section 2.2.4.3\)](#)) for BN, and later for BN+1.
2. Upon receiving the GetBlocks message, P2 replies with a Block message ([MSG\\_BLK \(section 2.2.5.3\)](#)) containing the encrypted block of data and the initialization vector used during the encryption.
3. Once received the Block message, P1 decrypts and stores it, then it proceeds asking for the second block by sending a new GetBlocks message.

## 5 Security

### 5.1 Security Considerations for Implementers

A higher-layer application provides the server-role peer with the encryption algorithm, key size and the encryption key. The choice of the encryption algorithm and key size is one of the **CryptoAlgoId** field values specified in section [2.2.3](#).

The server-role peer generates an initialization vector suitable for the chosen encryption algorithm and uses the encryption key to encrypt the block using the chosen encryption algorithm. The server-role peer then records the chosen algorithm and the initialization vector in the message, as described in section [2.2.5.3](#).

Server-role peers and client-role peers never exchange/share/send each other the encryption key.

The client-role peer has a-priori knowledge of the encryption key. Using the encryption algorithm and initialization vector it received from the server-role peer, it decrypts the block.

There is no other explicit authentication or authorization built into the protocol, except for the Utility Index strategies above described that can lead to deny service to peers currently considered untrustworthy.

### 5.2 Index of Security Parameters

None.

## 6 Appendix A: Product Behavior

The information in this specification is applicable to the following Microsoft products or supplemental software. References to product versions include released service packs.

- Windows Vista operating system
- Windows Server 2008 operating system
- Windows 7 operating system
- Windows Server 2008 R2 operating system
- Windows 8 operating system
- Windows Server 2012 operating system
- Windows 8.1 operating system
- Windows Server 2012 R2 operating system
- Windows 10 operating system
- Windows Server 2016 operating system

Exceptions, if any, are noted below. If a service pack or Quick Fix Engineering (QFE) number appears with the product version, behavior changed in that service pack or QFE. The new behavior also applies to subsequent service packs of the product unless otherwise specified. If a product edition appears with the product version, behavior is different in that product edition.

Unless otherwise specified, any statement of optional behavior in this specification that is prescribed using the terms SHOULD or SHOULD NOT implies product behavior in accordance with the SHOULD or SHOULD NOT prescription. Unless otherwise specified, the term MAY implies that the product does not follow the prescription.

[<1> Section 2:](#) For Windows Vista and Windows Server 2008, support for the client-side elements of this protocol is available only via the optional installation of the Background Intelligent Transfer Service via Windows Management Framework. Support for the server-side elements of this protocol is not available for Windows Vista or Windows Server 2008.

[<2> Section 2.2.3:](#) This message is never sent by Windows, but it is handled by the code if received, by responding with an [MSG\\_NEGO\\_RESP \(section 2.2.5.1\)](#) message.

[<3> Section 2.2.4.2:](#) By default, Windows implementations use SHA-256 as the hashing algorithm to generate the **SegmentID**, which corresponds to a **SegmentID** length of 32 bytes. Windows Server 2008 R2 operating system is capable of generating **SegmentIDs** using SHA-384 and SHA-512 in addition to SHA-256, but the Windows implementation of the Retrieval Protocol only supports **SegmentIDs** generated using SHA-256.

[<4> Section 2.2.4.2:](#) Windows implementations normalize the ranges in the array of block ranges in the [MSG\\_GETBLKLIST](#) and [MSG\\_BLKLIST](#) messages, using the following rules:

- Ranges in the array never overlap with each other.
- Overlapped or adjacent ranges in the array are always combined into a single range.
- Ranges in the array are always sorted by **Index**.

Windows implementations always send block ranges normalized with these rules, but can accept non-normalized ranges in received messages.

<5> [Section 2.2.4.3](#): By default, Windows implementations use SHA-256 as the hashing algorithm to generate the **SegmentID**, which corresponds to a **SegmentID** length of 32 bytes. Windows Server 2008 R2 is capable of generating **SegmentIDs** using SHA-384 and SHA-512 in addition to SHA-256, but the Windows implementation of the Retrieval Protocol only supports **SegmentIDs** generated using SHA-256.

<6> [Section 2.2.4.4](#): Windows Vista, Windows Server 2008, Windows 7 and Windows Server 2008 R2 do not support the MSG\_GETSEGLIST message.

<7> [Section 2.2.5.2](#): Windows implementations normalize the ranges in the array of block ranges in the MSG\_GETBLKLIST and MSG\_BLKLIST messages, using the following rules:

- Ranges in the array never overlap with each other.
- Overlapped or adjacent ranges in the array are always combined into a single range.
- Ranges in the array are always sorted by **Index**.

Windows implementations always send block ranges normalized with these rules, but can accept non-normalized ranges in received messages.

<8> [Section 2.2.5.4](#): Windows Vista, Windows Server 2008, Windows 7 and Windows Server 2008 R2 do not support the MSG\_SEGLIST message.

<9> [Section 3.1.2](#): Windows uses a 2 second timeout for each request message. The timeout is configurable between 1 millisecond and 1 minute.

<10> [Section 3.1.4.3](#): Windows performs a simple download when it involves less than 4 consecutive blocks in a single block range. When Internet Explorer is used for content retrieval, it reads into 64K buffers. Therefore, in general each read generates a **segment retrieval session** for a single block; in some cases the read could span two blocks if it is not block aligned. This results in a simple download, and no MSG\_GETBLKLIST is generated.

<11> [Section 3.1.5.2](#): Windows implementations do not perform the segment ID verification for any MSG\_BLKLIST message received. Windows implementations rely on the binding handle of the transport from which the MSG\_BLKLIST is received in order to identify which request (and, implicitly, the corresponding segment ID) the MSG\_BLKLIST is for.

<12> [Section 3.2.1](#): By default the server-role peer Windows implementation serves up to 64 simultaneous Upload Sessions per serving-role peer; this limit is configurable between 1 and 16,384.

<13> [Section 3.2.1](#): By default the server-role peer Windows implementation serves up to 1,024 simultaneous Upload Sessions per hosted cache server; this limit is configurable between 1 and 4,294,967,295.

<14> [Section 3.2.2](#): Windows uses a 15 second timeout for each incoming request. The timeout value is configurable between 100 milliseconds and 1 hour.

## 7 Change Tracking

No table of changes is available. The document is either new or has had no changes since its last release.

## 8 Index

### A

Abstract data model  
[client](#) 26  
[client - Retrieval Protocol](#) 26  
[server](#) 30  
[server - Retrieval Protocol](#) 30  
[Applicability](#) 9

### B

[BLOCK\\_RANGE packet](#) 12  
[BLOCK\\_RANGE\\_ARRAY](#) 12

### C

[Capability negotiation](#) 10  
[Change tracking](#) 39  
Client  
[abstract data model](#) 26  
[initialization](#) 27  
[other local events](#) 30  
[timers](#) 27  
Client - Retrieval Protocol  
[abstract data model](#) 26  
higher-layer triggered events  
[MSG\\_GETBLKLIST Initiation](#) 27  
[MSG\\_GETBLKS Initiation](#) 27  
[MSG\\_NEGO\\_REQ request](#) 27  
[initialization](#) 27  
[local events](#) 30  
message processing  
[MSG\\_BLK\\_response received](#) 29  
[MSG\\_BLKLIST\\_response received](#) 29  
[MSG\\_NEGO\\_RESP received](#) 28  
[other messages received](#) 30  
sequencing rules  
[MSG\\_BLK\\_response received](#) 29  
[MSG\\_BLKLIST\\_response received](#) 29  
[MSG\\_NEGO\\_RESP received](#) 28  
[other messages received](#) 30  
[timer events - Request Timer expiration](#) 30  
[timers](#) 27  
[Common data types](#) 11  
[Common Data Types message](#) 11

### D

Data model - abstract  
[client](#) 26  
[client - Retrieval Protocol](#) 26  
[server](#) 30  
[server - Retrieval Protocol](#) 30  
[Data types](#) 11  
[Download with GetBlockList and GetBlocks exchanges example](#) 34

### E

[ENCODED\\_SEGMENT\\_AGE packet](#) 13  
Examples

[download with GetBlockList and GetBlocks exchanges](#) 34  
[simple download with GetBlocks download sub-sessions only](#) 35  
[Extensible BLOB message](#) 23  
[Extensible Blob Version 1 packet](#) 24

### F

[Fields - vendor-extensible](#) 10

### G

[Glossary](#) 6

### H

Higher-layer triggered events  
client - Retrieval Protocol  
[MSG\\_GETBLKLIST Initiation](#) 27  
[MSG\\_GETBLKS Initiation](#) 27  
[MSG\\_NEGO\\_REQ request](#) 27  
[server](#) 31  
[server - Retrieval Protocol](#) 31

### I

[Implementer - security considerations](#) 36  
[Index of security parameters](#) 36  
[Informative references](#) 8  
Initialization  
[client](#) 27  
[Client - Retrieval Protocol](#) 27  
[server](#) 31  
[server - Retrieval Protocol](#) 31  
[Introduction](#) 6

### L

Local events  
[client - Retrieval Protocol](#) 30  
[server - Retrieval Protocol](#) 31

### M

Message processing  
client - Retrieval Protocol  
[MSG\\_BLK\\_response received](#) 29  
[MSG\\_BLKLIST\\_response received](#) 29  
[MSG\\_NEGO\\_RESP received](#) 28  
[other messages received](#) 30  
server - Retrieval Protocol  
[MSG\\_GETBLKLIST request received](#) 31  
[MSG\\_GETBLKS request received](#) 31  
[MSG\\_NEGO\\_REQ received](#) 31  
[other messages received](#) 33  
[MESSAGE\\_HEADER message](#) 13  
[MESSAGE\\_HEADER packet](#) 13  
Messages  
[Common Data Types](#) 11  
[data types](#) 11  
[Extensible BLOB](#) 23

- [MESSAGE HEADER](#) 13
  - [Request Message](#) 15
  - [Response Message](#) 19
  - [syntax](#) 11
  - transport
    - [peer download](#) 11
    - [security](#) 11
- [TRANSPORT\\_RESPONSE\\_HEADER](#) 13
- [MSG\\_BLK packet](#) 21
- [MSG\\_BLKLIST packet](#) 20
- [MSG\\_GETBLKLIST packet](#) 16
- [MSG\\_GETBLKS packet](#) 17
- [MSG\\_GETSEGLIST packet](#) 18
- [MSG\\_NEGO\\_REQ packet](#) 16
- [MSG\\_NEGO\\_RESP packet](#) 20
- [MSG\\_SEGLIST packet](#) 22

**N**

- [Normative references](#) 8

**O**

- Other local events
  - [client](#) 30
  - [server](#) 33
- [Overview \(synopsis\)](#) 8

**P**

- [Parameters - security index](#) 36
- [Peer download transport](#) 11
- [Preconditions](#) 9
- [Prerequisites](#) 9
- [Product behavior](#) 37
- Protocol Details
  - [overview](#) 26

**R**

- [References](#) 7
  - [informative](#) 8
  - [normative](#) 8
- [Relationship to other protocols](#) 9
- [Request Message message](#) 15
- [Request Message packet](#) 15
- [Response Message message](#) 19
- [Response Message packet](#) 19
- Retrieval Protocol
  - message processing
    - [server](#) 31
  - sequencing rules
    - [server](#) 31

**S**

- Security
  - [implementer considerations](#) 36
  - [parameter index](#) 36
- [SEGMENT\\_RANGE packet](#) 12
- [SEGMENT\\_RANGE\\_ARRAY](#) 12
- Sequencing rules
  - client - Retrieval Protocol
    - [MSG\\_BLK response received](#) 29
  - MSG\_BLKLIST response received 29
  - MSG\_NEGO\_RESP received 28
  - other messages received 30
  - server - Retrieval Protocol
    - [MSG\\_GETBLKLIST request received](#) 31
    - [MSG\\_GETBLKS request received](#) 31
    - [MSG\\_NEGO\\_REQ received](#) 31
    - other messages received 33
- Server
  - [abstract data model](#) 30
  - [higher-layer triggered events](#) 31
  - [initialization](#) 31
  - other local events 33
  - timers 30
- Server - Retrieval Protocol
  - [abstract data model](#) 30
  - [higher-layer triggered events](#) 31
  - [initialization](#) 31
  - [local events](#) 33
  - message processing
    - [MSG\\_GETBLKLIST request received](#) 31
    - [MSG\\_GETBLKS request received](#) 31
    - [MSG\\_NEGO\\_REQ received](#) 31
    - other messages received 33
  - sequencing rules
    - [MSG\\_GETBLKLIST request received](#) 31
    - [MSG\\_GETBLKS request received](#) 31
    - [MSG\\_NEGO\\_REQ received](#) 31
    - other messages received 33
    - timer events - Upload Timer expiration 33
    - timers 30
- [Simple download with GetBlocks download sub-sessions only example](#) 35
- [Standards assignments](#) 10
- [Syntax](#) 11

**T**

- Timer events
  - [client - Retrieval Protocol - Request Timer expiration](#) 30
  - [server - Retrieval Protocol - Upload Timer expiration](#) 33
- Timers
  - [client](#) 27
  - [client - Retrieval Protocol](#) 27
  - [server](#) 30
  - [server - Retrieval Protocol](#) 30
- [Tracking changes](#) 39
- Transport
  - [peer download](#) 11
  - [security](#) 11
- [TRANSPORT\\_RESPONSE\\_HEADER message](#) 13
- [TRANSPORT\\_RESPONSE\\_HEADER packet](#) 13
- Triggered events - higher-layer
  - client - Retrieval Protocol
    - [MSG\\_GETBLKLIST Initiation](#) 27
    - [MSG\\_GETBLKS Initiation](#) 27
    - [MSG\\_NEGO\\_REQ request](#) 27
  - [server](#) 31
  - [server - Retrieval Protocol](#) 31

**V**

- [Vendor-extensible fields](#) 10

