

Network Working Group
Request for Comments: 3466
Category: Informational

M. Day
Cisco
B. Cain
Storigen
G. Tomlinson
Tomlinson Group
P. Rzewski
Media Publisher, Inc.
February 2003

A Model for Content Internetworking (CDI)

Status of this Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (2003). All Rights Reserved.

Abstract

Content (distribution) internetworking (CDI) is the technology for interconnecting content networks, sometimes previously called "content peering" or "CDN peering". A common vocabulary helps the process of discussing such interconnection and interoperation. This document introduces content networks and content internetworking, and defines elements for such a common vocabulary.

Table of Contents

1.	Introduction	2
2.	Content Networks	2
2.1	Problem Description	3
2.2	Caching Proxies.	4
2.3	Server Farms	5
2.4	Content Distribution Networks.	6
2.4.1	Historic Evolution of CDNs	8
2.4.2	Describing CDN Value: Scale and Reach.	8
3.	Content Network Model Terms	9
4.	Content Internetworking	12
5.	Content Internetworking Model Terms	12
6.	Security Considerations	15
7.	Acknowledgements	15
8.	Normative References	16

9. Authors' Addresses	16
10. Full Copyright Statement	17

1. Introduction

Content networks are of increasing importance to the overall architecture of the Web. This document presents a vocabulary for use in developing technology for interconnecting content networks, or "content internetworking".

The accepted name for the technology of interconnecting content networks is "content internetworking". For historical reasons, we abbreviate this term using the acronym CDI (from "content distribution internetworking"). Earlier names relied on analogy with peering and interconnection of IP networks; thus we had "content peering" and "CDN peering". All of these other names are now deprecated, and we have worked to establish consistent usage of "content internetworking" and "CDI" throughout the documents of the IETF CDI group.

The terminology in this document builds from the previous taxonomy of web caching and replication in RFC 3040 [3]. In particular, we have attempted to avoid the use of the common terms "proxies" or "caches" in favor of more specific terms defined by that document, such as "caching proxy".

Section 2 provides background on content networks. Section 3 introduces the terms used for elements of a content network and explains how those terms are used. Section 4 provides additional background on interconnecting content networks, following which Section 5 introduces additional terms and explains how those internetworking terms are used.

2. Content Networks

The past several years have seen the evolution of technologies centered around "content". Protocols, appliances, and entire markets have been created exclusively for the location, download, and usage tracking of content. Some sample technologies in this area have included web caching proxies, content management tools, intelligent "web switches", and advanced log analysis tools.

When used together, these tools form new types of networks, dubbed "content networks". Whereas network infrastructures have traditionally processed information at layers 1 through 3 of the OSI stack, content networks include network infrastructure that exists in layers 4 through 7. Whereas lower-layer network infrastructures centered on the routing, forwarding, and switching of frames and packets, content networks deal with the routing and forwarding of requests and responses for content. The units of transported data in content networks, such as images, movies, or songs, are often very large and may span hundreds or thousands of packets.

Alternately, content networks can be seen as a new virtual overlay to the OSI stack: a "content layer", to enable richer services that rely on underlying elements from all 7 layers of the stack. Whereas traditional applications, such as file transfer (FTP), relied on underlying protocols such as TCP/IP for transport, overlay services in content networks rely on layer 7 protocols such as HTTP or RTSP for transport.

The proliferation of content networks and content networking capabilities gives rise to interest in interconnecting content networks and finding ways for distinct content networks to cooperate for better overall service.

2.1 Problem Description

Content networks typically play some role in solving the "content distribution problem". Abstractly, the goal in solving this problem is to arrange a rendezvous between a content source at an origin server and a content sink at a viewer's user agent. In the trivial case, the rendezvous mechanism is that every user agent sends every request directly to the origin server named in the host part of the URL identifying the content.

As the audience for the content source grows, so do the demands on the origin server. There are a variety of ways in which the trivial system can be modified for better performance. The apparent single logical server may in fact be implemented as a large "farm" of server machines behind a switch. Both caching proxies and reverse caching proxies can be deployed between the client and server, so that requests can be satisfied by some cache instead of by the server.

For the sake of background, several sample content networks are described in the following sections that each attempt to address this problem.

2.2 Caching Proxies

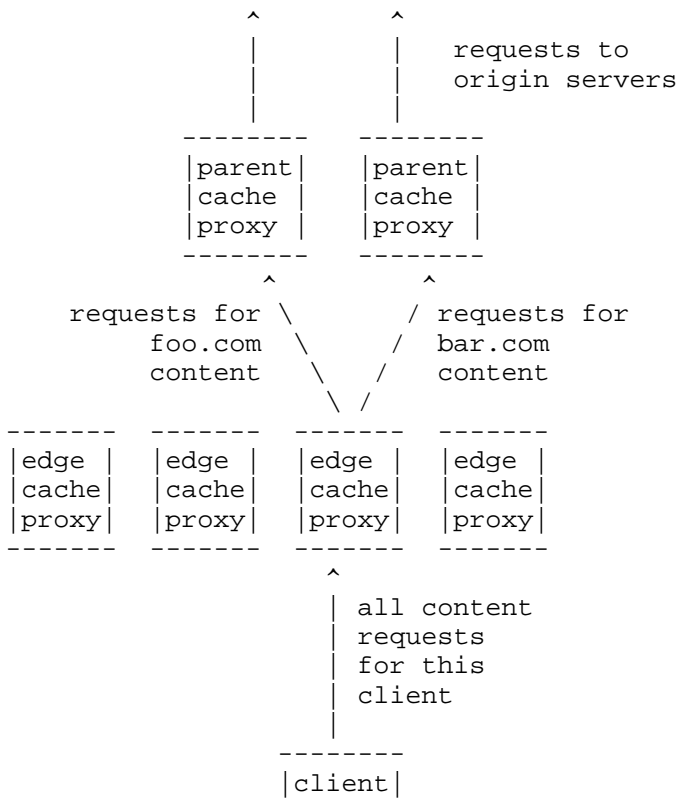
A type of content network that has been in use for several years is a caching proxy deployment. Such a network might typically be employed by an ISP for the benefit of users accessing the Internet, such as through dial or cable modem.

In the interest of improving performance and reducing bandwidth utilization, caching proxies are deployed close to the users. These users are encouraged to send their web requests through the caches rather than directly to origin servers, such as by configuring their browsers to do so. When this configuration is properly done, the user's entire browsing session goes through a specific caching proxy. That caching proxy will therefore contain the "hot set" of all Internet content being viewed by all of the users of that caching proxy.

When a request is being handled at a caching proxy on behalf of a user, other decisions may be made, such as:

- o A provider that deploys caches in many geographically diverse locations may also deploy regional parent caches to further aggregate user requests and responses. This may provide additional performance improvement and bandwidth savings. When parents are included, this is known as hierarchical caching.
- o Using rich parenting protocols, redundant parents may be deployed such that a failure in a primary parent is detected and a backup is used instead.
- o Using similar parenting protocols, requests may be partitioned such that requests for certain content domains are sent to a specific primary parent. This can help to maximize the efficient use of caching proxy resources.

The following diagram depicts a hierarchical cache deployment as described above:



Note that this diagram shows only one possible configuration, but many others are also useful. In particular, the client may be able to communicate directly with multiple caching proxies. RFC 3040 [3] contains additional examples of how multiple caching proxies may be used.

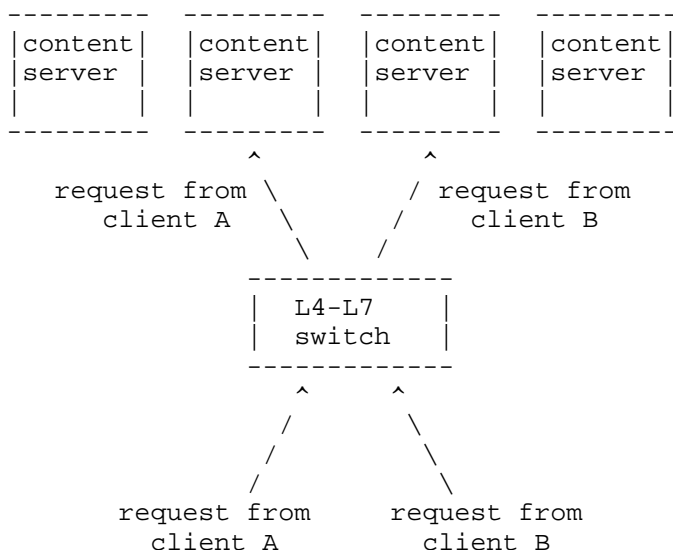
2.3 Server Farms

Another type of content network that has been in widespread use for several years is a server farm. A typical server farm makes use of a so-called "intelligent" or "content" switch (i.e., one that uses information in OSI layers 4-7). The switch examines content requests and dispatches them among a (potentially large) group of servers.

Some of the goals of a server farm include:

- o Creating the impression that the group of servers is actually a single origin site.
- o Load-balancing of requests across all servers in the group.
- o Automatic routing of requests away from servers that fail.
- o Routing all requests for a particular user agent's session to the same server, in order to preserve session state.

The following diagram depicts a simple server farm deployment:



A similar style of content network (that is, deployed close to servers) may be constructed with surrogates [3] instead of a switch.

2.4 Content Distribution Networks

Both hierarchical caching and server farms are useful techniques, but have limits. Server farms can improve the scalability of the origin server. However, since the multiple servers and other elements are typically deployed near the origin server, they do little to improve performance problems that are due to network congestion. Caching proxies can improve performance problems due to network congestion (since they are situated near the clients) but they cache objects based on client demand. Caching based on client demand performs poorly if the requests for a given object, while numerous in

aggregate, are spread thinly among many different caching proxies. (In the worst case, an object could be requested n times via n distinct caching proxies, causing n distinct requests to the origin server -- or exactly the same behavior that would occur without any caching proxies in place.)

Thus, a content provider with a popular content source can find that it has to invest in large server farms, load balancing, and high-bandwidth connections to keep up with demand. Even with those investments, the user experience may still be relatively poor due to congestion in the network as a whole.

To address these limitations, another type of content network that has been deployed in increasing numbers in recent years is the CDN (Content Distribution Network or Content Delivery Network). A CDN essentially moves server-farm-like configurations out into network locations more typically occupied by caching proxies. A CDN has multiple replicas of each content item being hosted. A request from a browser for a single content item is directed to a "good" replica, where "good" usually means that the item is served to the client quickly compared to the time it would take fetch it from the origin server, with appropriate integrity and consistency. Static information about geographic locations and network connectivity is usually not sufficient to do a good job of choosing a replica. Instead, a CDN typically incorporates dynamic information about network conditions and load on the replicas, directing requests so as to balance the load.

Compared to using servers and surrogates in a single data center, a CDN is a relatively complex system encompassing multiple points of presence, in locations that may be geographically far apart. Operating a CDN is not easy for a content provider, since a content provider wants to focus its resources on developing high-value content, not on managing network infrastructure. Instead, a more typical arrangement is that a network service provider builds and operates a CDN, offering a content distribution service to a number of content providers.

A CDN enables a service provider to act on behalf of the content provider to deliver copies of origin server content to clients from multiple diverse locations. The increase in number and diversity of location is intended to improve download times and thus improve the user experience. A CDN has some combination of a content-delivery infrastructure, a request-routing infrastructure, a distribution infrastructure, and an accounting infrastructure. The content-delivery infrastructure consists of a set of "surrogate" servers [3] that deliver copies of content to sets of users. The request-routing infrastructure consists of mechanisms that move a client toward a

(usually) highly distributed sites. We refer to increased aggregate infrastructure size as "scale". In addition, a CDN can be constructed with copies of content near to end users, overcoming issues of network size, network congestion, and network failures. We refer to increased diversity of content locations as "reach".

In a typical (non-internetworked) CDN, a single service provider operates the request-routers, the surrogates, and the content distributors. In addition, that service provider establishes (business) relationships with content publishers and acts on behalf of their origin sites to provide a distributed delivery system. The value of that CDN to a content provider is a combination of its scale and its reach.

3. Content Network Model Terms

This section consists of the definitions of a number of terms used to refer to roles, participants, and objects involved in content networks. Although the following uses many terms that are based on those used in RFC 2616 [1] or RFC 3040 [3], there is no necessary connection to HTTP or web caching technology. Content internetworking and this vocabulary are applicable to other protocols and styles of content delivery.

Phrases in upper-case refer to other defined terms.

ACCOUNTING

Measurement and recording of DISTRIBUTION and DELIVERY activities, especially when the information recorded is ultimately used as a basis for the subsequent transfer of money, goods, or obligations.

ACCOUNTING SYSTEM

A collection of CONTENT NETWORK ELEMENTS that supports ACCOUNTING for a single CONTENT NETWORK.

AUTHORITATIVE REQUEST-ROUTING SYSTEM

The REQUEST-ROUTING SYSTEM that is the correct/final authority for a particular item of CONTENT.

CDN

Content Delivery Network or Content Distribution Network. A type of CONTENT NETWORK in which the CONTENT NETWORK ELEMENTS are arranged for more effective delivery of CONTENT to CLIENTS. Typically a CDN consists of a REQUEST-ROUTING SYSTEM, SURROGATES, a DISTRIBUTION SYSTEM, and an ACCOUNTING SYSTEM.

CLIENT

A program that sends CONTENT REQUESTS and receives corresponding CONTENT RESPONSES. (Note: this is similar to the definition in RFC 2616 [1] but we do not require establishment of a connection.)

CONTENT

Any form of digital data, CONTENT approximately corresponds to what is referred to as an "entity" in RFC 2616 [1]. One important form of CONTENT with additional constraints on DISTRIBUTION and DELIVERY is CONTINUOUS MEDIA.

CONTENT NETWORK

An arrangement of CONTENT NETWORK ELEMENTS, controlled by a common management in some fashion.

CONTENT NETWORK ELEMENT

A network device that performs at least some of its processing by examining CONTENT-related parts of network messages. In IP-based networks, a CONTENT NETWORK ELEMENT is a device whose processing depends on examining information contained in IP packet bodies; network elements (as defined in RFC 3040) examine only the header of an IP packet. Note that many CONTENT NETWORK ELEMENTS do not examine or even see individual IP packets, instead receiving the body of one or more packets assembled into a message of some higher-level protocol.

CONTENT REQUEST

A message identifying a particular item of CONTENT to be delivered.

CONTENT RESPONSE

A message containing a particular item of CONTENT, identified in a previous CONTENT REQUEST.

CONTENT SIGNAL

A message delivered through a DISTRIBUTION SYSTEM that specifies information about an item of CONTENT. For example, a CONTENT SIGNAL can indicate that the ORIGIN has a new version of some piece of CONTENT.

CONTINUOUS MEDIA

CONTENT where there is a timing relationship between source and sink; that is, the sink must reproduce the timing relationship that existed at the source. The most common examples of CONTINUOUS MEDIA are audio and motion video. CONTINUOUS MEDIA can be real-time (interactive), where there is a "tight" timing

relationship between source and sink, or streaming (playback), where the relationship is less strict. [Note: This definition is essentially identical to the definition of continuous media in [2]]

DELIVERY

The activity of providing a PUBLISHER's CONTENT, via CONTENT RESPONSES, to a CLIENT. Contrast with DISTRIBUTION and REQUEST-ROUTING.

DISTRIBUTION

The activity of moving a PUBLISHER's CONTENT from its ORIGIN to one or more SURROGATES. DISTRIBUTION can happen either in anticipation of a SURROGATE receiving a REQUEST (pre-positioning) or in response to a SURROGATE receiving a REQUEST (fetching on demand). Contrast with DELIVERY and REQUEST-ROUTING.

DISTRIBUTION SYSTEM

A collection of CONTENT NETWORK ELEMENTS that support DISTRIBUTION for a single CONTENT NETWORK. The DISTRIBUTION SYSTEM also propagates CONTENT SIGNALs.

ORIGIN

The point at which CONTENT first enters a DISTRIBUTION SYSTEM. The ORIGIN for any item of CONTENT is the server or set of servers at the "core" of the distribution, holding the "master" or "authoritative" copy of that CONTENT. (Note: We believe this definition is compatible with that for "origin server" in RFC 2616 [1] but includes additional constraints useful for CDI.)

PUBLISHER

The party that ultimately controls the CONTENT and its distribution.

REACHABLE SURROGATES

The collection of SURROGATES that can be contacted via a particular DISTRIBUTION SYSTEM or REQUEST-ROUTING SYSTEM.

REQUEST-ROUTING

The activity of steering or directing a CONTENT REQUEST from a USER AGENT to a suitable SURROGATE.

REQUEST-ROUTING SYSTEM

A collection of CONTENT NETWORK ELEMENTS that support REQUEST-ROUTING for a single CONTENT NETWORK.

SERVER

A program that accepts CONTENT REQUESTS and services them by sending back CONTENT RESPONSES. Any given program may be capable of being both a client and a server; our use of these terms refers only to the role being performed by the program. [Note: this is adapted from a similar definition in RFC 2616 [1].]

SURROGATE

A delivery server, other than the ORIGIN. Receives a CONTENT REQUEST and delivers the corresponding CONTENT RESPONSE. [Note: this is a different definition from that in RFC 3040 [3], which appears overly elaborate for our purposes. A "CDI surrogate" is always an "RFC 3040 surrogate"; we are not sure if the reverse is true.]

USER AGENT

The CLIENT which initiates a REQUEST. These are often browsers, editors, spiders (web-traversing robots), or other end user tools. [Note: this definition is identical to the one in RFC 2616 [1].]

4. Content Internetworking

There are limits to how large any one network's scale and reach can be. Increasing either scale or reach is ultimately limited by the cost of equipment, the space available for deploying equipment, and/or the demand for that scale/reach of infrastructure. Sometimes a particular audience is tied to a single service provider or a small set of providers by constraints of technology, economics, or law. Other times, a network provider may be able to manage surrogates and a distribution system, but may have no direct relationship with content providers. Such a provider wants to have a means of affiliating their delivery and distribution infrastructure with other parties who have content to distribute.

Content internetworking allows different content networks to share resources so as to provide larger scale and/or reach to each participant than they could otherwise achieve. By using commonly defined protocols for content internetworking, each content network can treat neighboring content networks as "black boxes", allowing them to hide internal details from each other.

5. Content Internetworking Model Terms

This section consists of the definitions of a number of terms used to refer to roles, participants, and objects involved in internetworking content networks. The purpose of this section is to identify common terms and provide short definitions.

ACCOUNTING INTERNETWORKING

Interconnection of two or more ACCOUNTING SYSTEMS so as to enable the exchange of information between them. The form of ACCOUNTING INTERNETWORKING required may depend on the nature of the NEGOTIATED RELATIONSHIP between the peering parties -- in particular, on the value of the economic exchanges anticipated.

ADVERTISEMENT

Information about resources available to other CONTENT NETWORKS, exchanged via CONTENT INTERNETWORKING GATEWAYS. Types of ADVERTISEMENT include AREA ADVERTISEMENTS, CONTENT ADVERTISEMENTS, and DISTRIBUTION ADVERTISEMENTS.

AREA ADVERTISEMENT

ADVERTISEMENT from a CONTENT NETWORK's REQUEST-ROUTING SYSTEM about aspects of topology, geography and performance of a CONTENT NETWORK. Contrast with CONTENT ADVERTISEMENT, DISTRIBUTION ADVERTISEMENT.

BILLING ORGANIZATION

An entity that operates an ACCOUNTING SYSTEM to support billing within a NEGOTIATED RELATIONSHIP with a PUBLISHER.

CONTENT ADVERTISEMENT

ADVERTISEMENT from a CONTENT NETWORK's REQUEST-ROUTING SYSTEM about the availability of one or more collections of CONTENT on a CONTENT NETWORK. Contrast with AREA ADVERTISEMENT, DISTRIBUTION ADVERTISEMENT

CONTENT DESTINATION

A CONTENT NETWORK or DISTRIBUTION SYSTEM that is accepting CONTENT from another such network or system. Contrast with CONTENT SOURCE.

CONTENT INTERNETWORKING GATEWAY (CIG)

An identifiable element or system through which a CONTENT NETWORK can be interconnected with others. A CIG may be the point of contact for DISTRIBUTION INTERNETWORKING, REQUEST-ROUTING INTERNETWORKING, and/or ACCOUNTING INTERNETWORKING, and thus may incorporate some or all of the corresponding systems for the CONTENT NETWORK.

CONTENT REPLICATION

The movement of CONTENT from a CONTENT SOURCE to a CONTENT DESTINATION. Note that this is specifically the movement of CONTENT from one network to another. There may be similar or different mechanisms that move CONTENT around within a single network's DISTRIBUTION SYSTEM.

CONTENT SOURCE

A CONTENT NETWORK or DISTRIBUTION SYSTEM that is distributing CONTENT to another such network or system. Contrast with CONTENT DESTINATION.

DISTRIBUTION ADVERTISEMENT

An ADVERTISEMENT from a CONTENT NETWORK's DISTRIBUTION SYSTEM to potential CONTENT SOURCES, describing the capabilities of one or more CONTENT DESTINATIONS. Contrast with AREA ADVERTISEMENT, CONTENT ADVERTISEMENT.

DISTRIBUTION INTERNETWORKING

Interconnection of two or more DISTRIBUTION SYSTEMS so as to propagate CONTENT SIGNALS and copies of CONTENT to groups of SURROGATES.

ENLISTED

Describes a CONTENT NETWORK that, as part of a NEGOTIATED RELATIONSHIP, has accepted a DISTRIBUTION task from another CONTENT NETWORK, has agreed to perform REQUEST-ROUTING on behalf of another CONTENT NETWORK, or has agreed to provide ACCOUNTING data to another CONTENT NETWORK. Contrast with ORIGINATING.

INJECTION

A "send-only" form of DISTRIBUTION INTERNETWORKING that takes place from an ORIGIN to a CONTENT DESTINATION.

INTER-

Describes activity that involves more than one CONTENT NETWORK (e.g., INTER-CDN). Contrast with INTRA-.

INTRA-

Describes activity within a single CONTENT NETWORK (e.g., INTRA-CDN). Contrast with INTER-.

NEGOTIATED RELATIONSHIP

A relationship whose terms and conditions are partially or completely established outside the context of CONTENT NETWORK internetworking protocols.

ORIGINATING

Describes a CONTENT NETWORK that, as part of a NEGOTIATED RELATIONSHIP, submits a DISTRIBUTION task to another CONTENT NETWORK, asks another CONTENT NETWORK to perform REQUEST-ROUTING on its behalf, or asks another CONTENT NETWORK to provide ACCOUNTING data. Contrast with ENLISTED.

REMOTE CONTENT NETWORK

A CONTENT NETWORK able to deliver CONTENT for a particular REQUEST that is not the AUTHORITATIVE REQUEST-ROUTING SYSTEM for that REQUEST.

REQUEST-ROUTING INTERNETWORKING

Interconnection of two or more REQUEST-ROUTING SYSTEMS so as to increase the number of REACHABLE SURROGATES for at least one of the interconnected systems.

6. Security Considerations

This document defines terminology and concepts for content internetworking. The terminology itself does not introduce any security-related issues. The implementation of content internetworking concepts does raise some security-related issues, which we identify in broad categories below. Other CDI documents will address their specific security-related issues in more detail.

Secure relationship establishment: CONTENT INTERNETWORKING GATEWAYS must ensure that CONTENT NETWORKS are internetworking only with other CONTENT NETWORKS as intended. It must be possible to prevent unauthorized internetworking or spoofing of another CONTENT NETWORK's identity.

Secure content transfer: CONTENT INTERNETWORKING GATEWAYS must support CONTENT NETWORK mechanisms that ensure both the integrity of CONTENT and the integrity of both DISTRIBUTION and DELIVERY, even when both ORIGINATING and ENLISTED networks are involved. CONTENT INTERNETWORKING GATEWAYS must allow for mechanisms to prevent theft or corruption of CONTENT.

Secure meta-content transfer: CONTENT INTERNETWORKING GATEWAYS must support the movement of accurate, reliable, auditable ACCOUNTING information between CONTENT NETWORKS. CONTENT INTERNETWORKING GATEWAYS must allow for mechanisms to prevent the diversion or corruption of ACCOUNTING data and similar meta-content.

7. Acknowledgements

The authors acknowledge the contributions and comments of Fred Douglass (AT&T), Don Gilletti (CacheFlow), Markus Hoffmann (Lucent), Barron Housel (Cisco), Barbara Liskov (Cisco), John Martin (Network Appliance), Nalin Mistry (Nortel Networks) Raj Nair (Cisco), Hilarie Orman (Volera), Doug Potter (Cisco), and Oliver Spatscheck (AT&T).

8. Normative References

- [1] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P. and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", RFC 2616, June 1999.
- [2] Schulzrinne, H., Rao, A. and R. Lanphier, "Real Time Streaming Protocol", RFC 2326, April 1998.
- [3] Cooper, I., Melve, I. and G. Tomlinson, "Internet Web Replication and Caching Taxonomy", RFC 3040, June 2000.

9. Authors' Addresses

Mark Stuart Day
Cisco Systems
1414 Massachusetts Avenue
Boxborough, MA 01719
US

Phone: +1 978 936 1089
EMail: mday@alum.mit.edu

Brad Cain
Storigen Systems
650 Suffolk Street
Lowell, MA 01854
US

Phone: +1 978 323 4454
EMail: bcain@storigen.com

Gary Tomlinson
Tomlinson Group
14324 227th Ave NE
Woodinville, WA 98072

Phone: +1 425 503 0881
EMail: gary@tomlinsongroup.net

Phil Rzewski
30 Jennifer Place
San Francisco, CA 94107
US

Phone: +1 650 303 3790
EMail: philrz@yahoo.com

10. Full Copyright Statement

Copyright (C) The Internet Society (2003). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.