CDI
Internet-Draft
Expires: November 30, 2002

M. Green
No Affiliation
B. Cain
Storigen Systems
G. Tomlinson
No Affiliation
M. Speer
Sun Microsystems
P. Rzewski
Inktomi
S. Thomas
TransNexus
June 2002

Content Internetworking Architectural Overview
draft-ietf-cdi-architecture-01.txt

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on November 30, 2002.

Copyright Notice

Abstract

There is wide interest in the technology for interconnecting Content Networks, variously called "Content Peering" or "Content

Internetworking". We present the general architecture and core building blocks used in the internetworking of Content Networks. The scope of this work is limited to external interconnections with Content Networks and does not address internal mechanisms used within Content Networks, which for the purpose of the document are considered to be black boxes. This work establishes an abstract architectural framework to be used in the development of protocols, interfaces, and system models for standardized Content Internetworking.

Table of Contents

1. Introduction

Terms in ALL CAPS, except those qualified with explicit citations are defined in [13].

This memo describes the overall architectural structure and the fundamental building blocks used in the composition of Content Internetworking. Consult [13] for the system model, and vocabulary used in, this application domain. A key requirement of the architecture itself is that it be able to address each of the Content Internetworking scenarios enumerated in [14] . The scope of this work is limited to external interconnections between Content Networks (CN) (i.e. INTER-CN) and does not address internal mechanisms used within Content Networks (i.e. INTRA-CN), which for the purposes of the document are considered to be black boxes. This work is intended to establish an abstract architectural framework to be used in the development of protocols, interfaces and system models for standardized, interoperable peering among Content Networks.

We first present the architecture as an abstract system. Then we develop a more concrete system architecture. For each core architectural element, we first present the structure of the element followed by system requirements. Protocol requirements for individual core elements are presented in accompanying works [17][18][15]. The assumptions and scenarios constraining the architecture is explained in [13] . We intend that the architecture should support a wide variety of business models.

At the core of Content Internetworking are three principal architectural elements that constitute the building blocks of the Content Internetworking system. These elements are the REQUEST-ROUTING INTERNETWORKING SYSTEM, DISTRIBUTION INTERNETWORKING SYSTEM, and ACCOUNTING INTERNETWORKING SYSTEM. Collectively, they control selection of the delivery Content Network, content distribution between peering Content Networks, and usage accounting, including billing settlement among the peering Content Networks.

This work takes into consideration relevant IETF RFCs and IETF works-in-progress. In particular, it is mindful of the end-to-end nature [6][10] of the Internet, and the taxonomy of web replication and caching [11].

1.1 Change Log

1. First pass at integration of the terms from the latest models draft [13].

2. New Editor of document -- Michael Speer (Sun Microsystems).

1.2 Outstanding Issues

1. Need to complete integration of [13], and [17], [18], and [15].

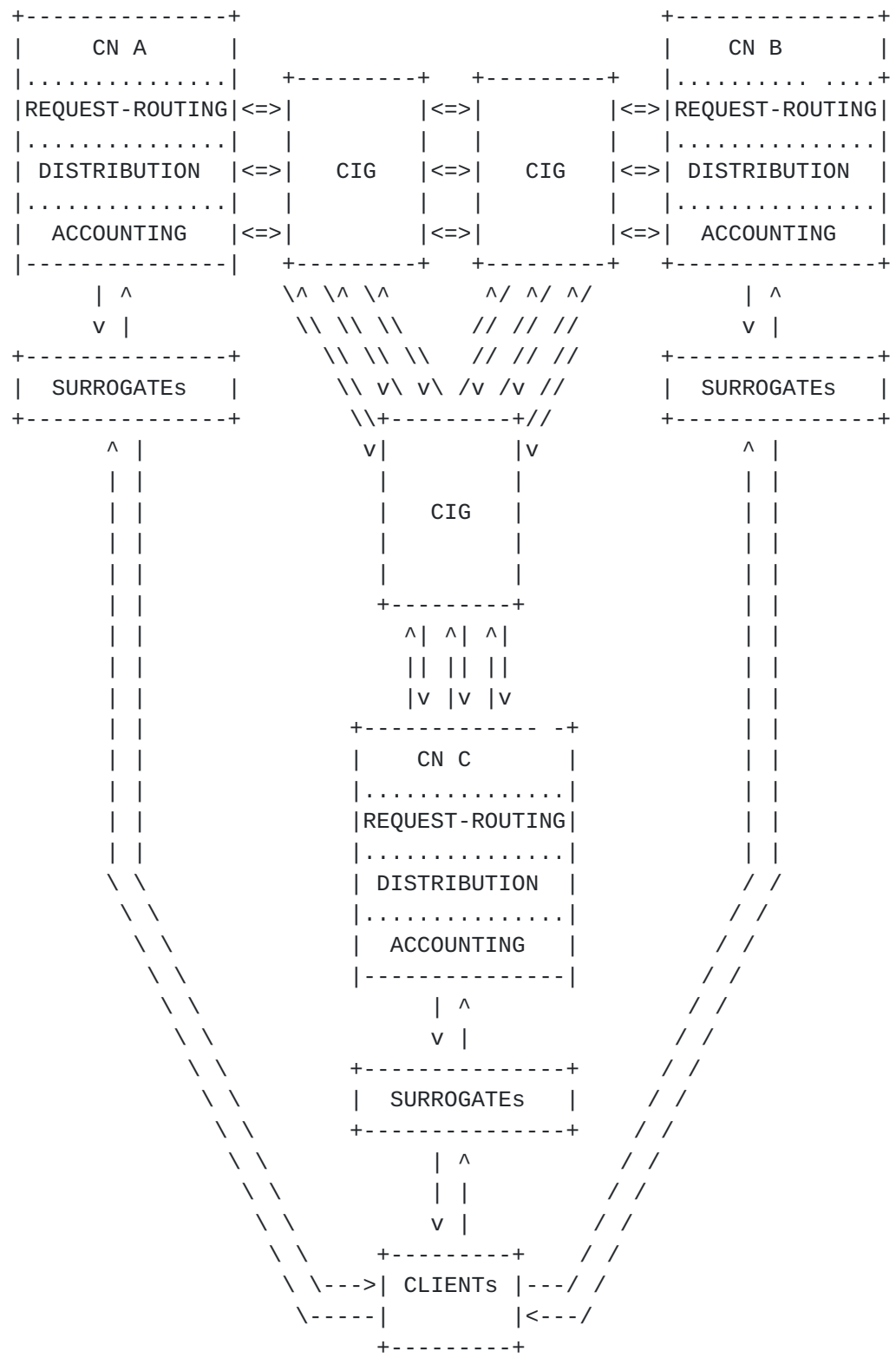
2. Need to address the open and outstanding questions.

2. Content Internetworking System Architecture

2.1 Conceptual View of Peered Content Networks

Before developing the system architecture, a conceptual view of peered CNs is presented to frame the problem space. CNs are comprised principally of four core system elements [13], the REQUEST-ROUTING SYSTEM, the DISTRIBUTION SYSTEM, the ACCOUNTING SYSTEM, and SURROGATES. In order for CNs to peer with one another, it is necessary to interconnect several of the core system elements of individual CNs. The interconnection of CN core system elements occurs through network elements called Content Internetworking Gateways (CIG). Namely, the CN core system elements that need to be interconnected are the REQUEST-ROUTING SYSTEM, the DISTRIBUTION SYSTEM, and the ACCOUNTING SYSTEM.

Figure 1 contains a conceptual peered Content Networks diagram.



CIG = Content Internetworking Gateway

Figure 1 Conceptual View of Peered Content Networks

This conceptual view illustrates the peering of three Content Networks; CN A, CN B, and CN C. The CNs are peered through interconnection at Content Internetworking Gateways. The result is presented as a virtual CN to CLIENTs for the DELIVERY of CONTENT by the aggregated set of SURROGATES.

Note:
Not all Content Networks contain the complete set of core elements. For these Content Networks, peering will be done with only the core elements they do contain.

2.2 Content Internetworking Architectural Elements

The system architecture revolves around the general premise that individual Content Networks are wholly contained within an administrative domain [3] that is composed of either autonomous systems [1] (physical networks) or overlay networks (virtual networks). For the purpose of this memo, an overlay network is defined as a set of connected CN network elements layered onto existing underlying networks, and present the result as a virtual application layer to both CLIENTs and ORIGINs. The system architecture for CN peering accommodates this premise by assuring that the information and controls are available for inter-CN-domain administration . Content Internetworking involves the interconnection of the individual CN administrative domains through gateway protocols and mechanisms loosely modeled after BGP [5].

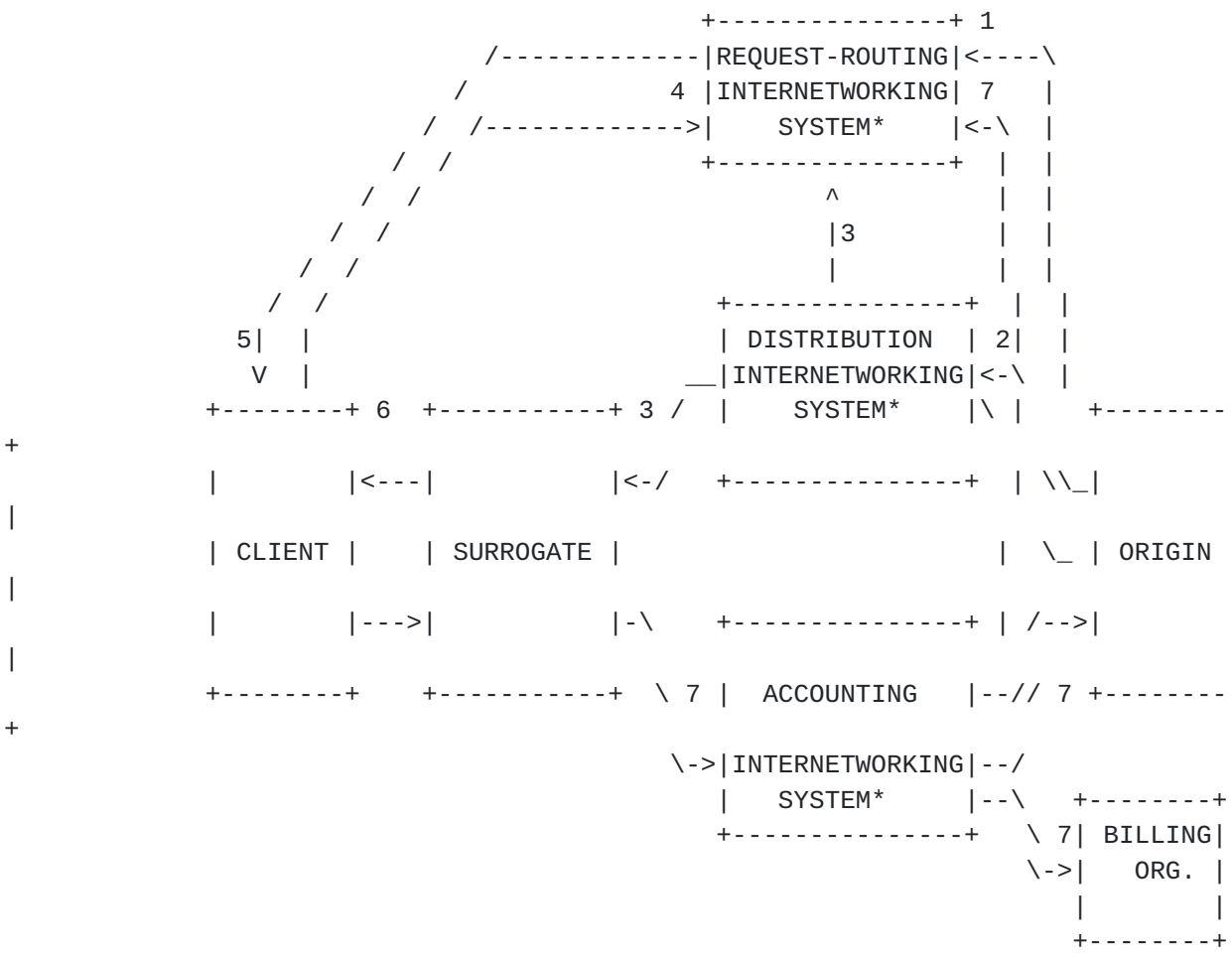
The system architecture depends on the following assumptions:

1. The URI [8] name space is the basis of PUBLISHER object identifiers.

2. PUBLISHERs delegate authority of their object URI name space being distributed by peering CNs to the REQUEST-ROUTING INTERNETWORKING SYSTEM.

3. Peering CNs use a common convention for encoding CN metadata into the URI name space.

Figure 2 contains a system architecture diagram of the core elements involved in Content Internetworking.



Note: * represents core elements of Content Internetworking

Figure 2 System Architecture Elements of a Content Internetworking System

The System Architecture is comprised of 7 major elements, 3 of which constitute the Content Internetworking system itself. The peering elements are REQUEST-ROUTING INTERNETWORKING SYSTEM, DISTRIBUTION INTERNETWORKING SYSTEM, and ACCOUNTING INTERNETWORKING SYSTEM. Correspondingly, the system architecture is a system of systems:

1. The ORIGIN delegates its URI name space for objects to be distributed and delivered by the peering CNs to the REQUEST-ROUTING INTERNETWORKING SYSTEM.

2. The ORIGIN INJECTS CONTENT that is to be distributed and delivered by the peering CNs into the DISTRIBUTION INTERNETWORKING SYSTEM.

Note:

CONTENT which is to be pre-populated (pushed) within the peering CNs is pro-actively injected, while CONTENT which is to be pulled on demand is injected at the time the object is being requested for DELIVERY.

3. The DISTRIBUTION INTERNETWORKING SYSTEM moves content between CN DISTRIBUTION SYSTEMs. Additionally this system interacts with the REQUEST-ROUTING INTERNETWORKING SYSTEM via feedback ADVERTISEMENTs to assist in the peered CN selection process for CLIENT requests.

4. The CLIENT requests CONTENT from what it perceives to be the ORIGIN, however due to URI name space delegation, the request is actually made to the REQUEST-ROUTING INTERNETWORKING SYSTEM.

   Note:
   The request routing function may be implied by an in-path network element such as caching proxy, which is typical for a Access Content Network. In this case, request routing is optimized to a null function, since the CLIENT is a priori mapped to the SURROGATE.

5. The REQUEST-ROUTING INTERNETWORKING SYSTEM routes the request to a suitable SURROGATE in a peering CN. REQUEST-ROUTING INTERNETWORKING SYSTEMs interact with one another via feedback ADVERTISEMENTs in order to keep request-routing tables current.

6. The selected SURROGATE delivers the requested content to the CLIENT. Additionally, the SURROGATE sends accounting information for delivered content to the ACCOUNTING INTERNETWORKING SYSTEM.

7. The ACCOUNTING INTERNETWORKING SYSTEM aggregates and distills the accounting information into statistics and content detail records for use by the ORIGIN and BILLING ORGANIZATION. Statistics are also used as feedback to the REQUEST-ROUTING INTERNETWORKING SYSTEM.

8. The BILLING ORGANIZATION uses the content detail records to settle with each of the parties involved in the content distribution and delivery process.

This process has been described in its simplest form in order to present the Content Internetworking architecture in the most abstract way possible. In practice, this process is more complex when applied to policies, business models and service level agreements that span multiple peering Content Networks. The orthogonal core peering systems are discussed in greater depth in Section 3, Section 4 and Section 5 respectively.

Note:
Figure 2 simplifies the presentation of the core Content Internetworking elements as single boxes, when in fact they represent a collection of CIGs and interconnected individual CN

core system elements. This has been done to introduce the system architecture at its meta level.

The system architecture does not impose any administrative domain [3] restrictions on the core peering elements (REQUEST-ROUTING INTERNETWORKING SYSTEM, DISTRIBUTION INTERNETWORKING SYSTEM and ACCOUNTING INTERNETWORKING SYSTEM). The only requirement is that they be authorized by the principal parties (ORIGIN and peering CNs) to act in their behalf. Thus, it is possible for each of the core elements to be provided by a different organization.

3. Request-Routing Internetworking System

The REQUEST-ROUTING INTERNETWORKING SYSTEM represents the request-routing function of the Content Internetworking system. It is responsible for routing CLIENT requests to an appropriate peered CN for the delivery of content.

Note:
When the DISTRIBUTION INTERNETWORKING SYSTEM and/or the ACCOUNTING INTERNETWORKING SYSTEM is present, it is highly desirable to utilize content location information within the peered CNs and/or system load information in the selection of appropriate peered CNs in the routing of requests.

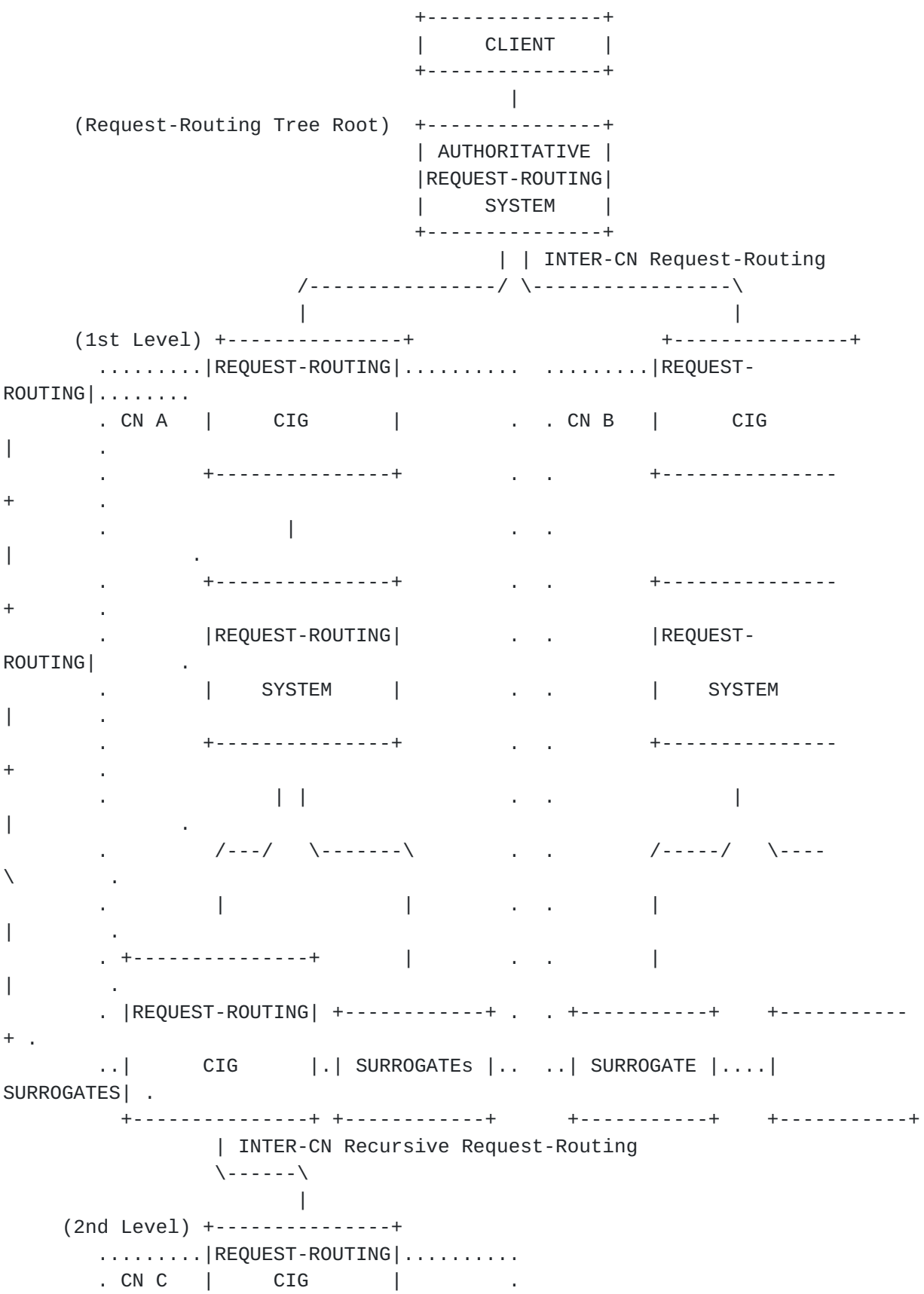
3.1 Request-Routing Overview

REQUEST-ROUTING SYSTEMs route CLIENT requests to a suitable SURROGATE, which is able to service a client request. Many request-routing systems route users to the surrogate that is "closest" to the requesting user, or to the "least loaded" surrogate. However, the only requirement of the request-routing system is that it route users to a surrogate that can serve the requested content.

REQUEST-ROUTING INTERNETWORKING is the 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.

In order for a PUBLISHER's CONTENT to be delivered by multiple peering CNs, it is necessary to federate each Content Network REQUEST-ROUTING SYSTEM under the URI name space of the PUBLISHER object. This federation is accomplished by first delegating authority of the PUBLISHER URI name space to an AUTHORITATIVE REQUEST-ROUTING SYSTEM. The AUTHORITATIVE REQUEST-ROUTING SYSTEM subsequently splices each peering Content Network REQUEST-ROUTING SYSTEM into this URI name space and transitively delegates URI name space authority to them for their participation in request-routing. Figure 3 is a diagram of the entities involved in the REQUEST-ROUTING INTERNETWORKING SYSTEM.

Note:
For the null request routing case (in path caching proxy present), the caching proxy acts as the SURROGATE. In this case, the SURROGATE performs the request routing via its pre-established proxy relationship with the CLIENT and is implicitly the terminating level of request routing. In essence, the SURROGATE is federated into the URI namespace without the need to communicate with the AUTHORITATIVE REQUEST-ROUTING SYSTEM.


CLIENT
(Request-Routing Tree Root)
AUTHORITATIVE REQUEST-ROUTING SYSTEM
INTER-CN Request-Routing
(1st Level)
REQUEST-ROUTING
CN A
CIG
CN B
CIG
REQUEST-ROUTING
SYSTEM
SYSTEM
REQUEST-ROUTING
SURROGATEs
SURROGATE
SURROGATES
INTER-CN Recursive Request-Routing
(2nd Level)
REQUEST-ROUTING
CN C
CIG

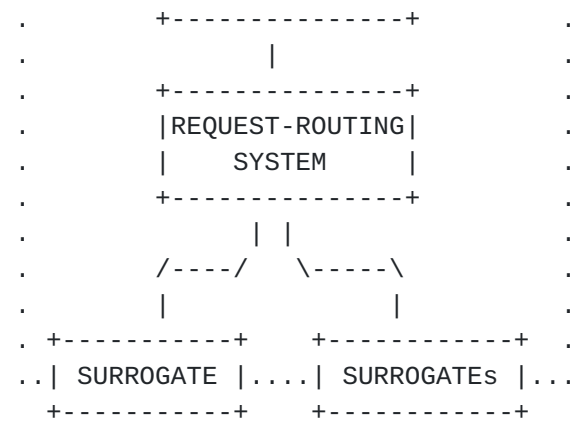

+---------------+
|
+---------------+
|REQUEST-ROUTING|
| SYSTEM |
+---------------+
| |
/----/   \-----\
| |
+-----------+   +------------+
| SURROGATE |....| SURROGATEs |...
+-----------+   +------------+

Figure 3 REQUEST-ROUTING INTERNETWORKING SYSTEM Architecture

The REQUEST-ROUTING INTERNETWORKING SYSTEM is hierarchical in nature. There exists exactly one request-routing tree for each PUBLISHER URI. The AUTHORITATIVE REQUEST-ROUTING SYSTEM is the root of the request-routing tree. There may be only one AUTHORITATIVE REQUEST-ROUTING SYSTEM for a URI request-routing tree. Subordinate to the AUTHORITATIVE REQUEST-ROUTING SYSTEM are the REQUEST-ROUTING SYSTEMs of the first level peering CNs. There may exist recursive subordinate REQUEST-ROUTING SYSTEMs of additional level peering CNs.

Note:
A PUBLISHER object may have more than one URI associated with it and therefore be present in more than one request-routing tree.

3.2 Request-Routing

The actual "routing" of a client request is through REQUEST-ROUTING CIGs. The AUTHORITATIVE REQUEST-ROUTING CIG receives the CLIENT request and forwards the REQUEST to an appropriate DISTRIBUTING CN. This process of INTER-CN request-routing may occur multiple times in a recursive manner between REQUEST-ROUTING CIGs until the REQUEST-ROUTING SYSTEM arrives at an appropriate DISTRIBUTING CN to deliver the content.

Note:
The Client request may be for resolution of a URI component and not the content of the URI itself. This is the case when DNS is being utilized in the request-routing process to resolve the URI server component.

Request-Routing systems explicitly peer but do not have "interior" knowledge of surrogates from other CNs. Each CN operates its internal request-routing system. In this manner, request-routing systems peer very much like IP network layer peering.

3.3 System Requirements

We assume that there is a peering relationship between REQUEST-ROUTING CIGs. This peering relationship at a minimum must exchange a set of CLIENT IP addresses that can be serviced, and a set of information about the DISTRIBUTION SYSTEMs, for which they are performing request-routing.

Request-Routing Requirements

1. Use of a URI name space based request-routing mechanism. The

request-routing mechanism is allowed to use as much of the URI name space as it needs to select the proper SURROGATE. For example, DNS based mechanisms utilize only the host subcomponent, while content aware mechanisms utilize use multiple components.

2. Normalized canonical URI name space structure for peered CN distribution of PUBLISHER objects. The default in the absence of encoded meta data is the standard components as defined by [8]. Encoded meta data must conform to the syntactical grammar defined in [7] .

3. Single AUTHORITATIVE REQUEST-ROUTING SYSTEM for PUBLISHER object URI name space.

4. Assure that the request-routing tree remains a tree -- i.e., has no cycles.

5. Assure that adjacent request-routing systems from different administrative domains (different CNs) use a compatible request-routing mechanism.

6. Assure that adjacent request-routing systems from different administrative domains (different CNs) agree to forward requests for the CONTENT in question.

Editor Note:
System requirements being generated in the request-routing peering protocol design team have not yet been reconciled and integrated into this document.]

3.4 Protocol Requirements

The REQUEST-ROUTING INTERNETWORKING system's protocol has a complete set of requirements that it must implement to solve a variety of internetworking problems to enable REQUEST-ROUTING systems to peer each other.

Some of the internetworking problems that the protocol must address include:

1. Satisfying the need to interconnect a variety of request-routing system types.

2. Satisfying the need to exchange content and associated meta-data.

3. Satisfying the need to exchange attributes and policies assoicated with given pieces of content.

For a complete of set of requirements that the Request-Routing Internetworking protocol needs to satisfy, consult [17].

3.5 Examples

Consult [16] for in-depth information on known request-routing systems.

3.6 Request-Routing Problems to Solve

Editor Note:
This section is being preserved until it has been determined that these issues have been addressed in the request-routing peering protocol requirements draft.]

Specific problems in request-routing needing further investigation include:

1. What is the aggregated granularity of CLIENT IP address being serviced by a peering CN's DISTRIBUTION SYSTEM?

2. How do DNS request-routing systems forward a request? If a given CN is peered with many other CNs, what are the criteria that forwards a request to another CN?

3. How do content-aware request-routing systems forward a request? If a given CN is peered with many other CNs, what are the criteria that forwards a request to another CN?

4. What are the merits of designing a generalized content routing protocol, rather than relying on request-routing mechanisms.

5. What is the normalized canonical URI name space for request-routing? Because request-routing is federated across multiple CNs, it is necessary to have agreed upon standards for the encoding of meta data in URIs. There are many potential elements, which may be encoded. Some of these elements are: authoritative agent domain, publisher domain, content type, content length, etc.

6. How are policies communicated between the REQUEST-ROUTING SYSTEM and the DISTRIBUTION ADVERTISEMENT SYSTEM? A given CN may wish to serve only a given content type or a particular set of users. These types of policies must be communicated between CNs.

7. What are the request-routing protocols in DNS? When a request is routed to a particular REQUEST-ROUTING CIG, a clear set of DNS rules and policies must be followed in order to have a workable

and predictable system.

8. How do we protect the REQUEST-ROUTING SYSTEM against denial of service attacks?

9. How do we select the appropriate peering CN for DELIVERY?

   Note:
   The selection process must to consider the distribution policies involved in Section 4. Investigation into other policy "work in progress" within the IETF is needed to understand the relationship of policies developed within Content Internetworking.

4. Distribution Internetworking System

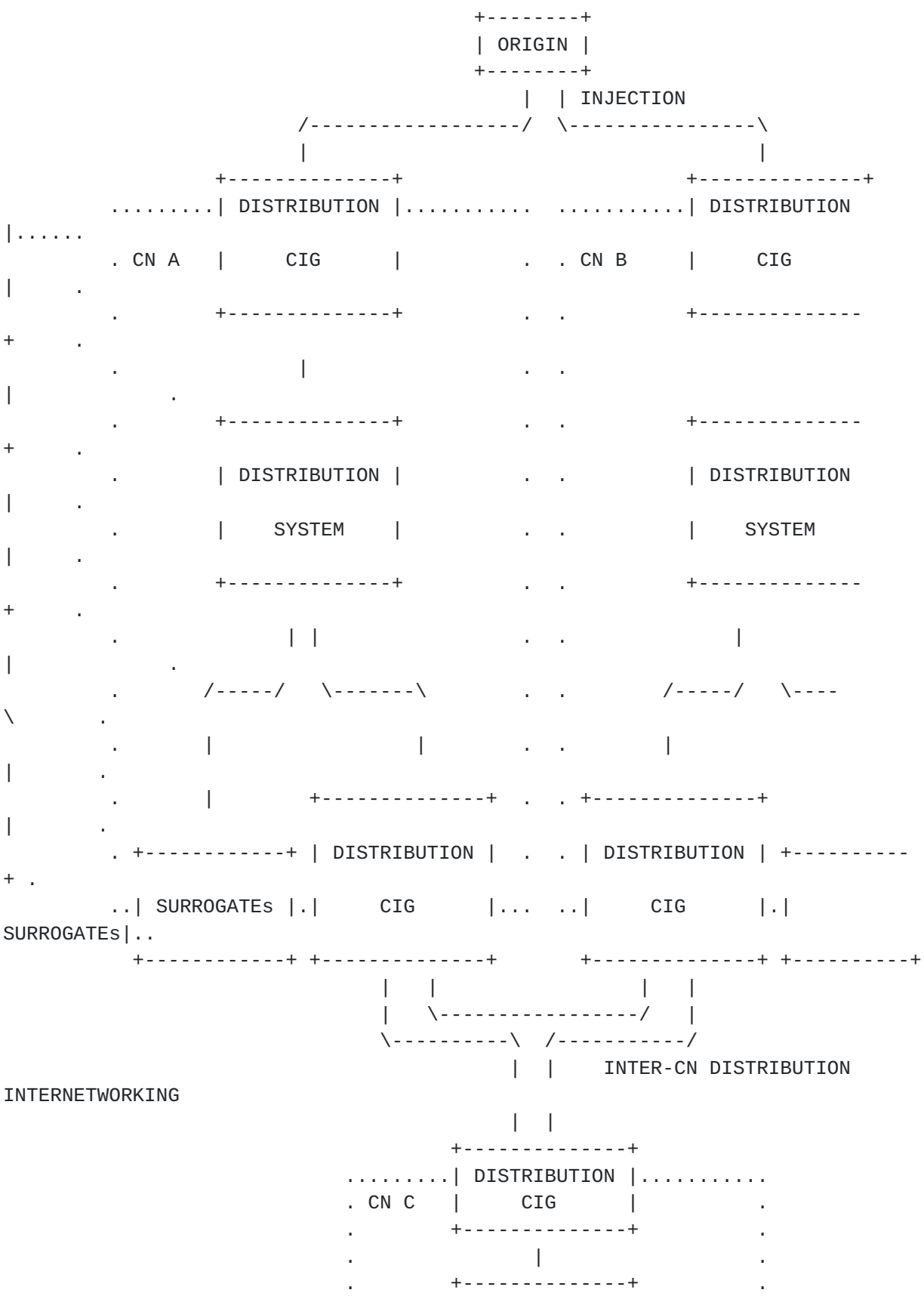
The DISTRIBUTION INTERNETWORKING SYSTEM represents the content distribution function of the CN peering system. It is responsible for moving content from one DISTRIBUTION CIG to another DISTRIBUTION CIG and for supplying content location information to the REQUEST-ROUTING INTERNETWORKING SYSTEM.

4.1 Distribution Overview

One goal of the Content Internetworking system is to move content closer to the CLIENT. Typically this is accomplished by copying content from its ORIGIN to SURROGATEs. The SURROGATEs then have the CONTENT available when it is requested by a CLIENT. Even with a single PUBLISHER and single CN, the copying of CONTENT to a SURROGATE may traverse a number of links, some in the PUBLISHER's network, some in the CN's network, and some between those two networks. For DISTRIBUTION INTERNETWORKING, we consider only the communication "between" two networks, and ignore the mechanisms for copying CONTENT within a network.

In the above example the last server on the content provider's network in the path, and the first server on the CN's network in the path, must contain DISTRIBUTION CIGs which communicate directly with each other. The DISTRIBUTION CIGs could be located in the ORIGIN server and the SURROGATE server. Thus in the simplest form the ORIGIN server is in direct contact with the SURROGATE. However the DISTRIBUTION CIG in the content provider's network could aggregate content from multiple ORIGIN servers and the DISTRIBUTION CIG in the CN's network could represent multiple SURROGATEs. These DISTRIBUTION CIGs could then be co-located in an exchange facility. In fact, given the common practice of independently managed IP peering co-location exchange facilities for layer 3, there exists the distinct opportunity to create similar exchanges for CIGs.

Figure 4 is a diagram of the entities involved in the DISTRIBUTION INTERNETWORKING SYSTEM.


ORIGIN
INJECTION
DISTRIBUTION
CN A
CIG
CN B
CIG
DISTRIBUTION
SYSTEM
DISTRIBUTION
SYSTEM
DISTRIBUTION
SURROGATEs
CIG
CIG
DISTRIBUTION
SURROGATEs
INTER-CN DISTRIBUTION INTERNETWORKING
DISTRIBUTION
CN C
CIG

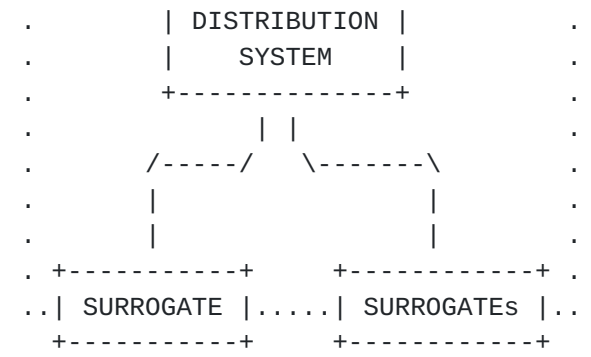


Figure 4 DISTRIBUTION INTERNETWORKING SYSTEM Architecture

While Content Internetworking in general relates to interfacing with CNs, there are two CN distribution peering relationships we expect to be common; INTER-CN distribution peering and INJECTION peering. INTER-CN distribution peering involves distributing CONTENT between individual CNs in a inter-network of peered CNs. INJECTION peering involves the publishing of CONTENT directly into CNs by ORIGINs.

4.2 Distribution Models

Replication ADVERTISEMENTs may take place in a model similar to the way IP routing table updates are done between BGP routers. DISTRIBUTION CIGs could take care of exterior content replication between content providers and CNs, while at the same time performing content replication interior to their networks in an independent manner. If this model is used then the internal structure of the networks is hidden and the only knowledge of other networks is the locations of DISTRIBUTION CIGs.

Replication of content may take place using a push model, or a pull model, or a combination of both. Use initiated replication, where SURROGATEs, upon getting a cache miss, retrieve CONTENT from the DISTRIBUTION SYSTEM, represents the pull model. ORIGIN initiated replication of CONTENT to SURROGATEs represents the push model. DISTRIBUTION CIGs may be located at various points in these models depending on the topologies of the networks involved.

With Content Internetworking it may be desirable to replicate content through a network, which has no internal SURROGATEs. For example add a exchange network between the content provider network and the CN network to the example above. The exchange network could have a DISTRIBUTION CIG co-located with the content provider's DISTRIBUTION CIG, which acts as a proxy for the CN. The exchange network could also have a DISTRIBUTION CIG co-located with the CN's DISTRIBUTION CIG, which acts as a proxy for the content provider. In a consolidated example, the exchange network could have a single DISTRIBUTION CIG that acts as a proxy for both the content provider and the CN.

Replication of CONTINUOUS MEDIA that is not to be cached on SURROGATEs, such as live streaming broadcasts, takes place in a different model from content that is to be persistently stored. Replication in this case, typically takes the form of splitting the live streaming data at various points in the network. In Content Internetworking, DISTRIBUTION CIGs may support CONTINUOUS MEDIA splitting replication, as they likely provide ideal network topologic points for application layer multicasting.

4.3 Distribution Components

The three main components of DISTRIBUTION INTERNETWORKING are replication, signaling and advertising.

The first component of content distribution is replication. Replication involves moving the content from an ORIGIN server to SURROGATE servers. The immediate goal in CN peering is moving the content between DISTRIBUTION CIGs.

The second component of content distribution is content signaling. Content signaling is the propagation of content meta-data. This meta-data may include such information such as the immediate expiration of content or a change in the expiration time of CONTENT. The immediate goal in signaling is exchanging signals between DISTRIBUTION CIGs.

The third component of content distribution is content advertising. Content providers must be able to advertise content that can be distributed by CNs and its associated terms. It is important that the advertising of content must be able to aggregate content information. The immediate goal in advertising is exchanging advertisements between DISTRIBUTION CIGs.

4.4 Distribution System Requirements

Replication systems must have a peering relationship. This peering relationship must exchange sets of aggregated content and its meta-data. Meta-data may change over time independently of the content data and must be exchanged independently as well.

4.4.1 Replication Requirements

The specific requirements in content replication are:

1. A common protocol for the replication of content.

2. A common format for the actual content data in the protocol.

3. A common format for the content meta-data in the protocol.

4. Security mechanisms (see Section 6).

5. Scalable distribution of the content.

4.4.2 Signaling Requirements

The specific requirements in content signaling are:

1. Signals for (at least) "flush" and "expiration time update".

2. Security mechanisms (see Section 6).

3. Scalable distribution of the signals on a large scale.

   Editor Note:
   We have to start being quantitative about what we mean by "large scale". Are we thinking in terms of the number of content items, the number of networks, or the number of signals? For each of those, how big is "large scale"?

4. Content location and serviced CLIENT IP aggregate address exchanges with REQUEST-ROUTING CIGs.

4.4.3 Advertising Requirements

The specific requirements in CONTENT ADVERTISEMENT are:

1. A common protocol for the ADVERTISEMENT of CONTENT.

2. A common format for the actual ADVERTISEMENTs in the protocol.

   Editor Note:
   The following requirements need further discussion. As it stands now, there isn't sufficient information to substantiate them.

3. A well-known state machine.

4. Use of TCP or SCTP (because soft-state protocols will not scale).

5. Well-known error codes to diagnose protocols between different networks.

6. Capability negotiation.

7. Ability to represent policy.

   Editor Note:
   System requirements being generated in the distribution peering protocol design team have not yet been reconciled and integrated into this document.]

4.5 Protocol Requirements

Consult [18] for distribution internetworking protocol requirements.

4.6 Distribution Problems to Solve

[Editor Note:
This section is being preserved until it has been determined that these issues have been addressed in the distribution peering protocol requirements draft.]

Some of the problems in distribution revolve around supporting both a push model and a pull model for replication of content in that they are not symmetric. The push model is used for pre-loading of content and the pull model is used for on-demand fetching and pre-fetching of content. These models are not symmetric in that the amount of available resources in which to place the content on the target server must be known. In the fetching cases the server that pulls the content knows the available resources on the target server, itself. In the pre-loading case the server that pushes the content must find out the available resources from the target server before pushing the data.

4.6.1 General Problems

General problems in distribution peering needing further investigation include:

1. How would a single distribution peering protocol adequately support replication, signaling and advertising?

2. Should a single distribution peering protocol be considered, rather than separate protocols for each component?

3. How do we prevent looping of distribution updates? That is to say, detect and stop propagating replication, signaling and advertisement of events a DISTRIBUTION CIG has already issued. Looping here has the possibility of becoming infinite, if not bounded by the protocol(s). IP route updating and forwarding has faced similar issues and has solved them.

4.6.2 Replication Problems

Specific problems in replication needing further investigation include:

1. How do replication systems forward a request?

2. How do we keep pull based replication serviced within the DISTRIBUTION CIGs in order to prevent it from inadvertently bleeding out into REQUEST-ROUTING SYSTEM and potentially getting into a recursive loop?

3. How are policies communicated between the replication systems?

4. What are the replication protocols?

5. Does replication only take place between CIGs?

4.6.3 Signaling Problems

Specific problems in content signaling needing further investigation include:

1. How do we represent a content signal?

2. What content meta-data needs to be signaled?

3. How do we represent aggregates of meta-data in a concise and compressed manner?

4. What protocol(s) should be used for content signals?

5. What is a scalable architecture for delivering content signals?

6. Do content signals need a virtual distribution system of their own?

4.6.4 Advertising Problems

Specific problems in CONTENT ADVERTISEMENT needing further investigation include:

1. How do we represent aggregates of content to be distributed in a concise and compressed manner?

2. What protocol(s) should be used for the aggregation of this data?

3. What are the issues involved in the creation of CIG exchanges? This is actually a broader question than just for distribution, but needs to be considered for all forms of CIGs {REQUEST-ROUTING, DISTRIBUTION, ACCOUNTING}.

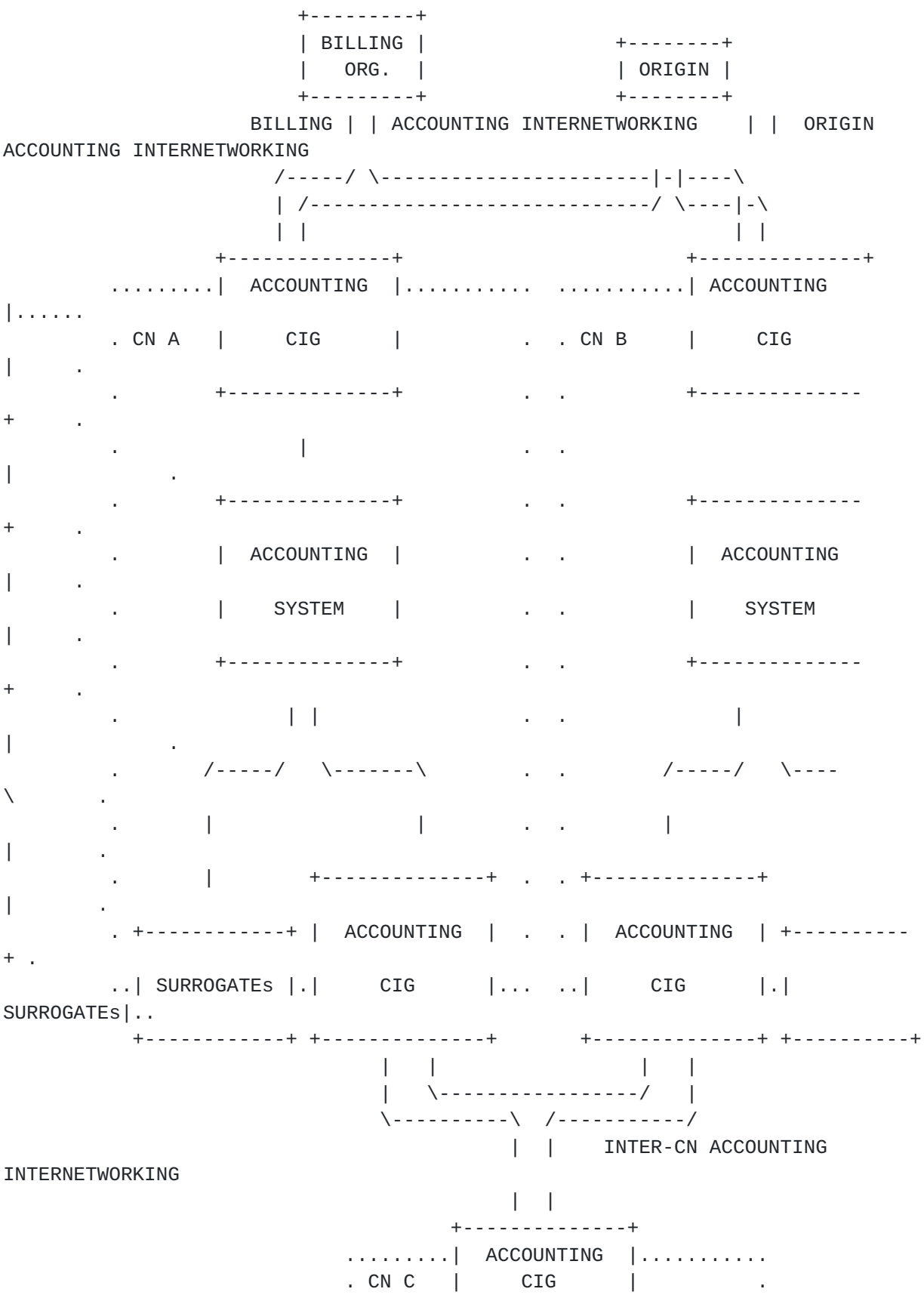
5. Accounting Internetworking System

The ACCOUNTING INTERNETWORKING SYSTEM represents the accounting data collection function of the Content Internetworking system. It is responsible for moving accounting data from one ACCOUNTING CIG to another ACCOUNTING CIG.

5.1 Accounting Overview

Content Internetworking must provide the ability for the content provider to collect data regarding the delivery of their CONTENT by the peered CNs. ACCOUNTING CIGs exchange the data collected by the interior ACCOUNTING SYSTEMS. This interior data may be collected from the SURROGATEs by ACCOUNTING CIGs using SNMP or FTP, for example. ACCOUNTING CIGs may transfer the data to exterior neighboring ACCOUNTING CIGs on request (push), in an asynchronous manner (push), or a combination of both. Accounting data may also be aggregated before it is transferred.

Figure 5 is a diagram of the entities involved in the ACCOUNTING INTERNETWORKING SYSTEM.


BILLING ORG.
ORIGIN
BILLING ACCOUNTING INTERNETWORKING
ORIGIN ACCOUNTING INTERNETWORKING
ACCOUNTING
CN A CIG
CN B CIG
ACCOUNTING SYSTEM
ACCOUNTING SYSTEM
SURROGATEs
ACCOUNTING CIG
ACCOUNTING CIG
SURROGATEs
INTER-CN ACCOUNTING INTERNETWORKING
ACCOUNTING
CN C CIG

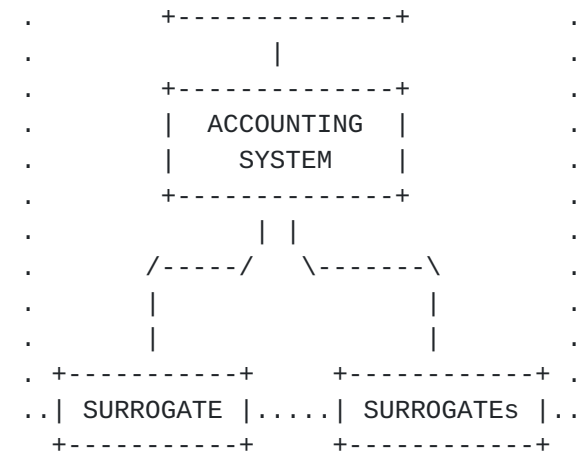

ACCOUNTING
SYSTEM
SURROGATE
SURROGATEs

Figure 5 ACCOUNTING Internetworking Ssystem Architecture

There are three CN accounting peering relationships we expect to be common; INTER-CN accounting peering, BILLING ORGANIZATION accounting peering and ORIGIN accounting peering. INTER-CN accounting peering involves exchanging accounting information between individual CNs in a inter-network of peered CNs. BILLING ORGANIZATION peering involves exchanging to accounting information between CNs and a billing organization. ORIGIN accounting peering involves the exchanging of accounting information between CNs and ORIGINs.

Note:
It is not necessary for an ORIGIN to peer directly with multiple CNs in order to participate in Content Internetworking. ORIGINs participating in a single home CN will be indirectly peered by their home CN with the inter-network of CNs the home CN is a member of. Nor is it necessary to have a BILLING ORGANIZATION peer, since this function may also be provided by the home CN. However, ORIGINs that directly peer for ACCOUNTING may have access to greater accounting detail. Also, through the use of ACCOUNTING peering, 3rd party billing can be provided.

5.2 Accounting System Requirements

[Editor Note:
System requirements being generated in the accounting peering protocol design team have not yet been reconciled and integrated into this document.]

5.3 Protocol Requirements

Consult [15] for accounting internetworking protocol requirements.

6. Security Considerations

Security concerns with respect to Content Internetworking can be generally categorized into trust within the system and protection of the system from threats. The trust model utilized with Content Internetworking is predicated largely on transitive trust between the ORIGIN, REQUEST-ROUTING INTERNETWORKING SYSTEM, DISTRIBUTION INTERNETWORKING SYSTEM, ACCOUNTING INTERNETWORKING SYSTEM and SURROGATES. Network elements within the Content Internetworking system are considered to be "insiders" and therefore trusted.

6.1 Threats to Content Networking

The following sections document key threats to CLIENTs, PUBLISHERs, and CNs. The threats are classified according to the party that they most directly harm, but, of course, a threat to any party is ultimately a threat to all. (For example, having a credit card number stolen may most directly affect a CLIENT; however, the resulting dissatisfaction and publicity will almost certainly cause some harm to the PUBLISHER and CN, even if the harm is only to those organizations' reputations.)

6.1.1 Threats to the CLIENT

6.1.1.1 Defeat of CLIENT's Security Settings

Because the SURROGATE's location may differ from that of the ORIGIN, the use of a SURROGATE may inadvertently or maliciously defeat any location-based security settings employed by the CLIENT. And since the SURROGATE's location is generally transparent to the CLIENT, the CLIENT may be unaware that its protections are no longer in force. For example, a CN may relocate CONTENT from a Internet Explorer user's "Internet Web Content Zone" to that user's "Local Intranet Web Content Zone." If the relocation is visible to the Internet Explorer browser but otherwise invisible to the user, the browser may be employing less stringent security protections than the user is expecting for that CONTENT. (Note that this threat differs, at least in degree, from the substitution of security parameters threat below, as Web Content Zones can control whether or not, for example, the browser executes unsigned active content.)

6.1.1.2 Delivery of Bad Accounting Information

In the case of CONTENT with value, CLIENTs may be inappropriately charged for viewing content that they did not successfully access. Conversely, some PUBLISHERs may reward CLIENTs for viewing certain CONTENT (e.g. programs that "pay" users to surf the Web). Should a CN fail to deliver appropriate accounting information, the CLIENT may

not receive appropriate credit for viewing the required CONTENT.

6.1.1.3 Delivery of Bad Content

A CN that does not deliver the appropriate CONTENT may provide the user misleading information (either maliciously or inadvertently). This threat can be manifested as a failure of either the DISTRIBUTION SYSTEM (inappropriate content delivered to appropriate SURROGATEs) or REQUEST-ROUTING SYSTEM (request routing to inappropriate SURROGATEs, even though they may have appropriate CONTENT), or both. A REQUEST-ROUTING SYSTEM may also fail by forwarding the CLIENT request when no forwarding is appropriate, or by failing to forward the CLIENT request when forwarding is appropriate.

6.1.1.4 Denial of Service

A CN that does not forward the CLIENT appropriately may deny the CLIENT access to CONTENT.

6.1.1.5 Exposure of Private Information

CNs may inadvertently or maliciously expose private information (passwords, buying patterns, page views, credit card numbers) as it transits from SURROGATEs to ORIGINs and/or PUBLISHERs.

6.1.1.6 Substitution of Security Parameters

If a SURROGATE does not duplicate completely the security facilities of the ORIGIN (e.g. encryption algorithms, key lengths, certificate authorities) CONTENT delivered through the SURROGATE may be less secure than the CLIENT expects.

6.1.1.7 Substitution of Security Policies

If a SURROGATE does not employ the same security policies and procedures as the ORIGIN, the CLIENT's private information may be treated with less care than the CLIENT expects. For example, the operator of a SURROGATE may not have as rigorous protection for the CLIENT's password as does the operator of the ORIGIN server. This threat may also manifest itself if the legal jurisdiction of the SURROGATE differs from that of the ORIGIN, should, for example, legal differences between the jurisdictions require or permit different treatment of the CLIENT's private information.

6.1.2 Threats to the PUBLISHER

6.1.2.1 Delivery of Bad Accounting Information

If a CN does not deliver accurate accounting information, the PUBLISHER may be unable to charge CLIENTs for accessing CONTENT or it may reward CLIENTs inappropriately. Inaccurate accounting information may also cause a PUBLISHER to pay for services (e.g. content distribution) that were not actually rendered.) Invalid accounting information may also effect PUBLISHERs indirectly by, for example, undercounting the number of site visitors (and, thus, reducing the PUBLISHER's advertising revenue).

6.1.2.2 Denial of Service

A CN that does not distribute CONTENT appropriately may deny CLIENTs access to CONTENT.

6.1.2.3 Substitution of Security Parameters

If a SURROGATE does not duplicate completely the security services of the ORIGIN (e.g. encryption algorithms, key lengths, certificate authorities, client authentication) CONTENT stored on the SURROGATE may be less secure than the PUBLISHER prefers.

6.1.2.4 Substitution of Security Policies

If a SURROGATE does not employ the same security policies and procedures as the ORIGIN, the CONTENT may be treated with less care than the PUBLISHER expects. This threat may also manifest itself if the legal jurisdiction of the SURROGATE differs from that of the ORIGIN, should, for example, legal differences between the jurisdictions require or permit different treatment of the CONTENT.

6.1.3 Threats to a CN

6.1.3.1 Bad Accounting Information

If a CN is unable to collect or receive accurate accounting information, it may be unable to collect compensation for its services from PUBLISHERs.

6.1.3.2 Denial of Service

Misuse of a CN may make that CN's facilities unavailable, or available only at reduced functionality, to legitimate customers or the CN provider itself. Denial of service attacks can be targeted at a CN's ACCOUNTING SYSTEM, DISTRIBUTION SYSTEM, or REQUEST-ROUTING SYSTEM.

6.1.3.3 Transitive Threats

To the extent that a CN acts as either a CLIENT or a PUBLISHER (such as, for example, in transitive implementations) such a CN may be exposed to any or all of the threats described above for both roles.

7. Acknowledgements

The authors would like to acknowledge the contributions and comments of Mark Day (Cisco), Fred Douglis (AT&T), Patrik Falstrom (Cisco), Don Gilletti (CacheFlow), Barron Housel (Cisco) John Martin (Network Appliance), Raj Nair (Cisco), Hilarie Orman (Novell), Doug Potter (Cisco), John Scharber (CacheFlow), Michael Speer (Sun), and Oliver Spatscheck (AT&T).

References

[1] Hawkinson, J. and T. Bates, "Guidelines for creation, selection, and registration of an Autonomous System (AS)", BCP 6, March 1996, <http://www.rfc-editor.org/rfc/bcp/bcp6.txt>.

[2] Postel, J., "Internet Protocol, DARPA Internet Program Protocol Specification", RFC 791, September 1981, <http://www.rfc-editor.org/rfc/rfc791.txt>.

[3] Hares, S. and D. Katz, "Administrative Domains and Routing Domains A Model for Routing in the Internet", RFC 1136, December 1989, <http://www.rfc-editor.org/rfc/rfc1136.txt>.

[4] Postel, J., "Domain Name Structure and Delegation", RFC 1591, March 1994, <http://www.rfc-editor.org/rfc/rfc1591.txt>.

[5] Rekhter, Y. and T. Li, "A Border Gateway Protocol 4 (BGP-4)", RFC 1771, March 1995, <http://www.rfc-editor.org/rfc/rfc1771.txt>.

[6] Carpenter, B., "Architectural Principles of the Internet", RFC 1958, June 1996, <http://www.rfc-editor.org/rfc/rfc1958.txt>.

[7] Schulzrinne, H., Rao, A. and R. Lanphier, "Real Time Streaming Protocol (RTSP)", RFC 2326, April 1998, <http://www.rfc-editor.org/rfc/rfc2326.txt>.

[8] Berners-Lee, T., Fielding, R. and L. Masinter, "Uniform Resource Identifiers (URI): Generic Syntax", RFC 2396, August 1998, <http://www.rfc-editor.org/rfc/rfc2396.txt>.

[9] 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, <http://www.rfc-editor.org/rfc/rfc2616.txt>.

[10] Carpenter, B., "Internet Transparency", RFC 2775, February 2000, <http://www.rfc-editor.org/rfc/rfc2775.txt>.

[11] Cooper, I., Melve, I. and G. Tomlinson, "Internet Web Replication and Caching Taxonomy", RFC 3040, January 2001, <http://www.rfc-editor.org/rfc/rfc3040.txt>.

[12] Volbrecht, J., Calhoun, P., Farrell, S., Gommans, L., Gross, G., de Bruijn, B., de Latt, C., Holdrege, M. and D. Spence, "AAA Authorization Framework", RFC 2904, August 2000, <http://www.rfc-editor.org/rfc/rfc2904.txt>.

[13] Day, M., Cain, B., Tomlinson, G. and P. Rzewski, "A Model for Content Internetworking (CDI)", draft-ietf-cdi-model-02.txt (work in progress), May 2002, <http://www.ietf.org/internet-drafts/draft-ietf-cdi-model-02.txt>.

[14] Day, M., Gilletti, D. and P. Rzewski, "Content Internetworking Scenarios", draft-ietf-cdi-scenarios-01.txt (work in progress), April 2002, <http://www.ietf.org/internet-drafts/draft-ietf-cdi-scenarios-01.txt>.

[15] Gilletti, D., Nair, R., Scharber, J., Guha, J. and D. Frascone, "Content Internetworking Authentication, Authorizaiton, and Accounting Requirements", draft-ietf-cdi-aaa-reqs-01.txt (work in progress), June 2002, <http://www.ietf.org/internet-drafts/draft-ietf-cdi-aaa-reqs-01.txt>.

[16] Douglis, F., Chaudhri, I. and P. Rzewski, "Known Mechanisms For Content Internetworking", draft-douglis-cdi-known-mech-00.txt (work in progress), November 2001, <http://www.ietf.org/internet-drafts/draft-douglis-cdi-known-mech-00.txt>.

[17] Cain, B., Spatscheck, O., May, M. and A. Barbir, "Request-Routing Requirements for Content Internetworking", draft-ietf-cdi-request-routing-reqs-00.txt (work in progress), February 2002, <http://www.ietf.org/internet-drafts/draft-ietf-cdi-request-routing-reqs-00.txt>.

[18] Amini, L., Thomas, S. and O. Spatscheck, "Requirements for Content Distribution Internetworking", draft-ietf-cdi-distribution-reqs-00.txt (work in progress), February 2002, <http://www.ietf.org/internet-drafts/draft-ietf-cdi-distribution-reqs-00.txt>.

Authors' Addresses

Mark Green
No Affiliation

EMail: reserved@pacbell.net

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

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

Gary Tomlinson
No Affiliation

EMail: gary@tomlinsongroup.net

Michael F. Speer
Sun Microsystems, Inc.
4150 Network Circle
UMPK17-103
Santa Clara, CA 95054
US

Phone: +1 650 786 6368
EMail: michael.speer@sun.com

Phil Rzewski
Inktomi
4100 East Third Avenue
MS FC1-4
Foster City, CA 94404
US

Phone: +1 650 653 2487
EMail: philr@intkomi.com

Stephen Thomas
TransNexus, Inc.
430 Tenth Street NW
Suite N204
Atlanta, GA 30318
US

Phone: +1 404 872 4887
EMail: stephen.thomas@transnexus.com

Full Copyright Statement

Acknowledgement

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