Network Working Group   L. Amini
Internet-Draft   IBM Research
Expires: June 23, 2003   S. Thomas
                                                        TransNexus, Inc
                                                        O. Spatscheck
                                                             AT&T Labs
                                                         December 2002

Distribution Requirements for Content Internetworking
draft-ietf-cdi-distribution-reqs-01

Status of this Memo

Copyright Notice

Abstract

   This document specifies requirements for interconnecting, or
   peering, the distribution systems of Content Networks (CN).
   Distribution internetworking requires advertising the capabilities
   of a CN offering distribution services, moving content from one CN
   to another, and signaling requirements for consistent storage and
   delivery of content. This document does not address requirements for
   directing user agents to distributed content, nor for aggregating
   access information for distributed content.

[1]. Introduction

   Content Internetworking (CI) assumes an architecture wherein the
   resources of multiple CNs are combined so as to achieve a larger
   scale or reach than any of the component CNs could individually [3].
   A Content Distribution Network (CDN) is an example of a CN.

   At the core of CI are three principal architectural elements. These
   elements are Request Routing, the Distribution and the Accounting.
   The focus of this document, the Internetworking Distribution
   Systems, is responsible for moving content from one Distribution CN
   to another Distribution CN. Note that the original content provider
   is considered a degenerate case of a Distribution CN.

   In any Distribution Internetworking arrangement, the relationships
   between Distribution CNs can always be decomposed into one or more
   pairs of CNs. Each CN pair comprises one CN which has, or has access
   to, content, and another CN which has, or has access to, systems
   capable of providing distribution and/or delivery functions for
   content. The former CN is referred to as the Content Source, while
   the latter is referred to as the Content Destination.

   This document describes the overall architectural structure and
   building blocks of the internetworked Distribution Systems. It also
   defines the protocol requirements for interconnecting two or more
   Distribution CNs via their respective Content Internetworking
   Gateways (CIG). Specifically, it defines the requirements for:

      Distribution Advertising: announcing the distribution
      capabilities of a Content Destination to potential Content
      Sources.

      Content Signaling: communicating content meta-data to enable
      consistent storage and delivery of content to user agents.

      Content Replication: moving content from a Content Source to a
      Content Destination.

   Although this document does not specifically address requirements
   for communicating within a CN, it is plausible that protocols
   developed to meet inter-CN requirements may also be well-suited for
   intra-CN communications.

   Requirements for the remaining CI architectural elements, the
   Request Routing System, which is responsible for directing user
   agents to the distributed content, and the Accounting System, which
   is responsible for aggregating information related to the access of
   distributed content, are detailed in [6], [7].

1.1 Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC-2119 [2].

All other terms in ALL CAPS, except those qualified with explicit citations, are defined in [8].

1.2 Change Log

1. Fixed terminology to comply with updated Models doc.

2. Intro: fixed reach or scale so not synonyms

3. Section 1: clarified definition for Content Signaling

4. Section 2: clarified that fig 1 is a logical view, need not have this strictly hierarchical relationship.

5. Section 2.1-2.2: deleted these sections, feedback indicated they confused more than claried; material is adequately covered in Models doc.

6. Section 3: change Use-initiated to surrogate initiated.

7. Section 3 (para 2): cig is not necessarily a box, could be functionality (protocol conformance) implemented in multiple boxes.

8. Section 4.1: clarified content signalling example.

9. Section 4.3: clarified common base replication vs. content type specific replication

10. Section 4.4: fixed "extensible model..." wording

11. Section 5.1: clarified that content pull for delivery services preparation is optional.

12. Section 5.3 3-4: modified, added request for feedback

13. Section 5.3.5: clarified advertisements as optional

14. Section 5.4: dropped requirement for sending data with content signal.

15. Section 5.4 6: dropped part about source encryption -- if the object is encrypted at source and not decrypted until it

reaches the client, then it is just a (downloadable) blob as
far as the dist system is concerned.

16.   Section 5.5 8: dropped subscription fee and next hop.

17.   Section 5.5 8: changed content id to content set id.

18.   Section 5.5: added information on relationship to WEBI/RUP.

2. Distribution System Overview

In Distribution System Internetworking, even the most complex
communication arrangements can be expressed in terms of simple
interactions between a Content Source and a Content Destination.
Figure 1, for example, shows a relationship between four different
administrative authorities. CN A operates a network of SURROGATES,
and "CN D" (actually the original content provider, or ORIGIN) has
content to be distributed. CN A communicates with CN B, which
communications with CN C, which, in turn, communicates with CN D.

```
   +------------+   +-----------+   +-----------+   +------------+
   |   CN A     |   |   CN B    |   |   CN C    |   |   "CN D"   |
   |(SURROGATES)|<->|  (agent   |<->|  (agent   |<->|  (ORIGIN)  |
   |            |   |  for A)   |   |  for D)   |   |            |
   +------------+   +-----------+   +-----------+   +------------+

     CONTENT DST <-> CONTENT SRC
                   CONTENT DST <-> CONTENT SRC
                                 CONTENT DST <-> CONTENT SRC
```
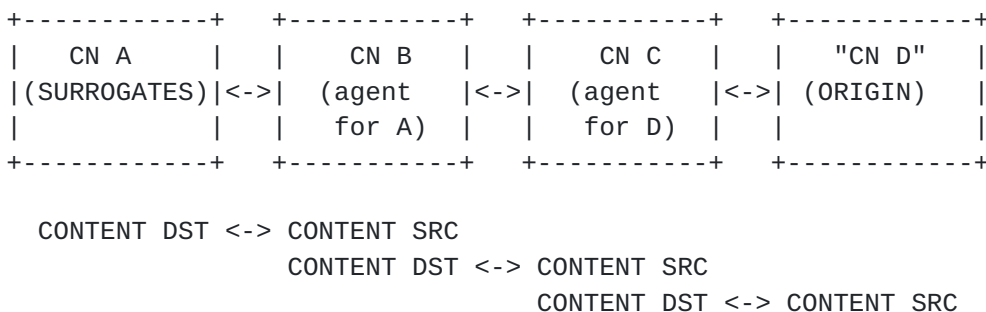
   Figure 1: Distribution System Interworking

In each case, one of the parties in the communications has the role
of Content Destination, while the other party is the Content Source.
Note that a particular CN's role may change, depending on the party
with whom it is communicating. CN B, for example, is a Content
Source when communicating with CN A, but a Content Destination when
communicating with CN C.

Note that a Content Destination which peers directly with the
Content ORIGIN, will interface with the ORIGIN just as it interfaces
with any other Content Source.

Although Figure 1 provides an example of multiple CNs peered in
series, a Distribution CN may serve as the Content Source for
multiple Content Destinations. Likewise, a Distribution CN may
serve as the Content Destination for multiple Content Sources.

Additionally, it is possible for the internetworking relationship
between a single Source-Destination pair to be reciprocal for
different content sets. That is, CN A may request distribution
services from CN B for Content Set A, while CN B requests
distribution services from CN A for Content Set B.

Further, note that Figure 1 is a logical view of the internetworking
relationship between several Content Source-Destination pairs;
actual communications are not restricted to this pair-wise

hierarchy. For example, CDN D may specify a single authoritative
distribution channel from which any distributing CN must retrieve
the CONTENT.

[3](#). **Replication Models**

   Replication of content may take place using a push model, or a pull
   model, or a combination of both.

   o SURROGATE-initiated replication of CONTENT, where a SURROGATE
      retrieves CONTENT based on a cache miss or on a prefetching
      policy specified at the SURROGATE, represents the pull model.
      This is the model currently used by caching proxies.

   o ORIGIN-initiated replication of CONTENT to SURROGATEs represents
      the push model. This model is used to preposition CONTENT in
      anticipation of demand.

   Replication, when between the administrative domains of different
   Distribution CNs must adhere to the CI protocol for content
   replicaiton. Replication within a single Distribution CN is an
   intra-CN communication and therefore, need not adhear to CI
   protocols. Further, the replication model used within a single
   Distribution CN need not be the same as the model used to replicate
   CONTENT between CNs.

   For both ORIGIN- and SURROGATE- initiated replication, the Content
   Source may use replication mechanisms beyond a simple transfer. For
   example, it may be desirable to have the Content Destination join a
   multicast channel on which a set of content is pushed to all
   SURROGATES.

   Another example is for CONTINUOUS MEDIA. In the case of live
   broadcasts, the data need not be cached on the SURROGATES. Instead,
   replication takes the form of "splitting" the live stream at various
   points in the network. Splitting is also referred to as application
   layer multicast.

   Replication of CONTINUOUS MEDIA streams which are not live, and
   therefore may be stored on SURROGATES, also benefits from mechanisms
   beyond in-line replication. For example, CONTINUOUS MEDIA is often
   delivered to CLIENTS over an unreliable channel. However, a CN
   distributing this content to many CLIENTS should work with a full
   replica. Existing proprietary replication protocols enable
   distribution of CONTINUOUS MEDIA objects in which a full or partial
   replica can be propagated, the data may be encrypted and/or
   authenticated, and the SURROGATE can support CONTINUOUS
   MEDIA-related services such as random access and stream
   insertion/splicing.

   It may be desirable to replicate content to a Distribution CN which
   has no internal SURROGATES. For example, a Distribution CN may have
   servers at key exchange points within the network which only serve

content to other distribution systems, and peer with other CNs which
provide SURROGATES which deliver content to CLIENTS.

[4](). **Distribution Internetworking Requirements**

   This section details general requirements for exchange of
   inter-domain distribution information.

[4.1]() **General Requirements**

   The goal of the Distribution Internetworking is to interconnect the
   Distribution Systems of multiple CNs. The intent of this
   interconnection is to effectively position content for fast,
   reliable access by CLIENTs. Generally this is accomplished by
   replicating content on SURROGATEs. While the communications path
   from ORIGIN to CLIENTs may traverse a number of links, some within a
   Distribution CN and some between Distribution CNs, Distribution
   Internetworking is concerned only with those communications between
   Distribution CNs.

   The three main components of Distribution Internetworking are
   advertising, replication, and signaling.

   Advertising: Distribution CNs MAY advertise their capabilities to
      potential Content Source CNs.

   Replication: Distribution CNs MUST be able to move content from a
      Content Source to a Content Destination.

   Content signaling: Distribution CNs MUST be able to propagate
      content meta-data. This meta-data includes information such as
      the immediate invalidation of content or a change in the source
      or distribution method of content.

   Note that these requirements do not necessarily translate directly
   into three distinct Distribution Internetworking protocols.

[4.2]() **Advertising Requirements**

   The following list specifies requirements to enable advertising of
   distribution capabilities.

   1. A common protocol for the Advertisement of distribution
       capabilities.

   2. A common format for the actual distribution capabilities
       Advertisements in the protocol.

   3. Security mechanisms.

[4.3](#) Replication Requirements

The following list specifies requirements to enable content replication.

1. A common base protocol for the replication of content.

2. This common base protocol should specify:

    *   a common format for the actual content data in the protocol.

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

3. Alternate content-specific protocols for the replication of content should be enabled. The replication protocol for a particular content set is specified via content signaling.

4. Scalable distribution of the content.

5. Security mechanisms.

[4.4](#) Content Signaling Requirements

The following list specifies requirements to enable content signaling.

1. A common protocol for signaling content meta-data.

2. An extenisble format for communicating content metadata. Minimally, support is required for "add," "withdraw," and "expiration time update."

3. Scalable distribution of signals on a scale to enable Internet-wide peering.

4. Security mechanisms.

[5]. Distribution Internetworking Protocol Requirements

This section defines protocol requirements for each of the
advertising, replication and content signaling components of
Distribution Internetworking. Note that these requirements do not
necessarily translate directly into either one converged or three
distinct Distribution Internetworking protocols.

[5.1] Overview of Distribution Internetworking Flow

In a Distribution Internetworking arrangement, the following
sequence of events is expected:

1. A Content Source may receive distribution capabilities
    advertisements from one or more Content Destinations. A Content
    Source may or may not require receipt of distribution
    capabilities advertisements prior to requesting services. For
    example, a Content Source may request services based on a
    contractual agreement negotiated off-line.

2. The Content Source will make a decision to request content
    distribution services from a Content Destination.

3. The Content Source will send a content signal requesting
    distribution services.

4. The Content Destination will accept or reject the request; no
    partial acceptance or negotiation is defined.

    *   If the request is rejected, the error code SHOULD provide
        enough information for the Content Source to determine if it
        should send a request with modified service requirements.

    *   If the request is accepted, the Content Destination will
        prepare for distribution services. Generally, this
        preparation will entail:

        +   optionally retrieving a copy of the object(s),

        +   joining the content update channel (if any), and

        +   preparing to provide access information to the Accounting
            System

    *   Each of the above steps are according to the Content Source's
        specification, and to the Content Destination's policies and
        configuration.

5. Once the Content Destination is prepared, it will notify the

Request Routing System of the content's availability.

6. The Content Destination will terminate service on first
    occurrence of either:

    * the time frame specified in the Content Source's request for
       distribution services expires or

    * a content signal requesting withdrawal of the content is
       received.

5.2 General Distribution Internetworking Protocol Requirements

Protocols must be scalable, i.e., support Distribution
Internetworking on an Internet-wide scale.

Protocols must prevent looping of advertisements, replication and
content signaling.

Protocols must support the ability to optionally conduct
authenticated and/or encrypted exchanges.

Protocols must support the ability to optionally exchange
credentials.

5.3 Advertising Protocol Requirements

1. Distribution Internetworking protocols MUST enable a Content
    Destination to advertise the capabilities of its distribution
    service in a common format. This common format for distribution
    service capabilities will be referred to as a Service Profile
    for the remainder of this draft.

2. The advertisement protocol must be extensible with the
    restriction that implementation-specific capabilities may be
    safely ignored by Content Source.

3. Distribution Internetworking protocols MUST provide
    low-overhead, mechanism to notify in-line elements (e.g.,
    proxies) of CI support.

4. [Editor's Note: prev item can be as simple has having the
    Content Source include a reference to it's CIG so that inline
    systems could contact the CIG and establish an internetworking
    arrangement. But the feedback has been lukewarm to bad --
    drop?]

5. Distribution Advertising by a Content Destination must be
    optional. That is, a Content Source may not require real-time

advertisement of distribution capabilities in order to
establish a Distribution Internetworking arrangement.
Distribution capabilities may be communicated via
Advertisements or some other agreed upon mechanism such as an
off-line contract negotiated between the parties.

6. Advertised capabilities are those available to the peer,
    potentially based on some off-line contractual agreement, and
    may not necessarily reflect the total capacity of the Content
    Destination.

7. The protocol MUST enable a Content Destination to advertise
    multiple service profiles. Each service profile MUST be
    specifiable by a profile identifier. The profile identifier
    MAY encode Content Source or Content Destination specific
    information, but it has local significance only (i.e., it is
    strictly between the Content Source and Content Destination).

8. The protocol MUST enable a Content Destination to advertise
    multiple services profiles to the same or different potential
    Content Sources.

9. A Content Destination with regional capabilities SHOULD
    advertise capabilities on a per region basis. A Content
    Destination which advertises regional capabilities MUST
    minimally be able to identify regions by network
    addresses/prefixes.

10. By default, advertisements are advisory. A Content Destination
     SHOULD be able to specify whether the capabilities are advisory
     or binding.

11. The protocol MUST provide the ability to specify distribution
     capabilities in terms of one or more of the following
     attributes:

     Profile ID: Identifier for the service profile being
        advertised. The profile identifier is to be used by the
        Content Source when requesting service. This attribute is
        required for all advertisements. The value need not be
        unique across Distribution CNs, and may be used in
        advertisements to multiple Content Sources.

     FootPrint: The areas served by the CN. Minimally, a Content
        Destination should support expressing footprint according to
        IP network addresses/prefixes.

     Content Type: The type of content (e.g. static Web pages,
        streaming media, etc.) that the CN is able to distribute.

Capacity: The storage capacity that the CN can provide.

Bandwidth: Maximum outbound bandwidth available from the CN.

Object Bandwidth: Maximum outbound bandwidth supported for a
   single object.

Distribution Method: The distribution methods that the CN
   supports; one or more of push, pull, and alm. "alm" refers
   to application layer multicast, or splitting, of CONTINUOUS
   MEDIA; if specified, supported protocols must also be
   specified.

      [Editor's Note: Specifying support for splitting
      requires refinement.]

Request Routing Type: Type(s) of request routing supported for
   CI Request Routing Systems.

Accounting System Format: Supported protocol(s) and format(s)
   for sending accounting and access feedback to a specified CI
   Accounting System.

Time Frame: Time frame for which this advertisement is valid.

Client Protocols: Indicates the client protocols supported.
   Currently only HTTP and RTSP are valid. However, this field
   must be further qualified than just the transport or
   signalling protocol. The protocol must clearly define a
   level of support implicated by a given Client Protocol
   value.

Distribution Fee: Indicates the fee charged for distribution
   services. The value may be expressed in Mbps
   (megabits/second) or in MB (megabytes of storage).

Advertisement Type: Indicates whether the advertisement is
   advisory or binding. By default, all advertisements are
   advisory.

Private Extensions: Additional metrics that the communicating
   parties may agree to use, but are not part of the IETF
   standard. Extensions must be defined such that if not
   understood by the Content Source, they can be safely
   ignored.

[5.3.1](#) Advertising Examples

To be provided.

[5.4](#) Replication Protocol Requirements

1. A common (base) replication protocol MUST be defined which is
    supported by all CIGs, for any content type which can be used to
    transport meta-data and a full replica of content data.

2. Replication MUST support the ability for a Content Source to
    specify a replication channel from which content may be
    retrieved.

    1. [Editors Note: I am using channel as a generic term which
        would provide a contact point and protocol, and any
        additional info required to establish a connection. E.g.
        wcips://invalidation.com/content_set for signaling; will
        provide clarification later]

3. Replication MUST enable specifying optionally supported,
    alternative replication protocols which may be better suited
    than the common base protocol for specific content types or
    configuration scenarios.

4. A Content Source SHOULD be able to specify an authoritative
    source for content as well as distribution points.

5. The protocol MUST enable replication that is secured (encrypted)
    across the communications channel.

[5.4.1](#) Replication Examples

To be provided.

[5.5](#) Content Signaling Protocol Requirements

1. A Content Source MUST be able to request distribution services
    for one or more content objects.

2. A Content Destination MUST explicitly accept or reject a request
    for distribution services.

3. A Content Source MUST be able to withdraw (cancel) a request for
    content services for one or multiple content objects.

4. Rejected requests for distribution services MUST include error
    codes. Partial rejections or negotiations are not supported. A
    Content Source may follow a rejection with a request for

distribution services under alternate service requirements.

5. A Content Source MUST be able to signal consistency meta-data.
    Minimally, Content Sources SHOULD support weak consistency
    mechanisms. Content Sources MAY support mechanisms for strong
    consistency.

6. Content signaling SHOULD include mechanisms to aggregate content
    information.

7. Content Signaling SHOULD be decoupled from the content ORIGIN.
    I.e., a Content Source should be able to specify a content
    signaling channel.

8. The following attributes are defined for content signals:

    Content Set ID: A unique identifier for this specific content
       set, so that future references (e.g. to modify the content or
       to withdraw it from distribution) may be resolved. This value
       can also be used to avoid loops. The Content Set ID MUST be
       global and unique, i.e., a given content set MUST have the
       same ID across all Distribution Systems, and this ID MUST be
       unique across *all* Distribution Systems.

    URI: The uniform resource identifier for the content. It
       identifies how CLIENTS will request delivery services from
       the Distribution CN. This attr