AAA Working Group   Pat R. Calhoun
Internet-Draft   Sun Microsystems, Inc.
Category: Informational   Glen Zorn
<draft-ietf-aaa-diameter-framework-01.txt>   Cisco Systems, Inc.
                                                         Ping Pan
                                                        Bell Labs
                                                     Haseeb Akhtar
                                                   Nortel Networks
                                                       March 2001

                      Diameter Framework Document

Status of this Memo

Abstract

   Current Internet Service Providers (ISPs) scale their networks by
   using the RADIUS protocol, which provides user Authentication,
   Authorization and Accounting (AAA) of Dial-up PPP clients. The recent
   work done in the Roaming Operations (ROAMOPS) Working Group was to
   investigate whether RADIUS could be used in a roaming network, and
   concluded that RADIUS was ill-suited for inter-domain purposes.

   The IETF has formed a new NAS Requirements Working Group, and part of
   their charter is to document the next generation NAS' AAA
   requirements. Recently, the Mobile-IP Working Group also documented
   their own AAA requirements that would help Mobile IP scale for
   Inter-Domain mobility.

   The Diameter protocol is a follow-on to the RADIUS protocol. Diameter
   addresses the known RADIUS deficiencies, and is intended for use with
   the NASREQ, ROAMOPS and Mobile IP application space.

Table of Contents

1.0 Introduction

Historically, the RADIUS protocol has been used to provide AAA
services for dial-up PPP [17] and terminal server access. Over time,
routers and network access servers (NAS) have increased in complexity
and density, making the RADIUS protocol increasingly unsuitable for
use in such networks.

The Roaming Operations Working Group (ROAMOPS) has published a set of
specifications [19, 20, 21] that define how a PPP user can gain
access to the Internet without having to dial into his/her home
service provider's modem pool. This is achieved by allowing service
providers to cross-authenticate their users. Effectively, a user can
dial into any service provider's point of presence (POP) that has a
roaming agreement with his/her home Internet service provider (ISP),
the benefit being that the user does not have to incur a long
distance charge while traveling, which can sometimes be quite
expensive.

Given the number of ISPs today, ROAMOPS realized that requiring each
ISP to set up roaming agreements with all other ISPs did not scale.
Therefore, the working group defined a "broker", which acts as an
intermediate server, whose sole purpose is to set up these roaming
agreements. A collection of ISPs and a broker is called a "roaming
consortium". There are many such brokers in existence today; many
also provide settlement services for member ISPs.

The Mobile-IP Working Group has recently changed its focus to inter
administrative domain mobility, which is a requirement for cellular
carriers wishing to deploy IETF-based mobility protocols. The current
cellular carriers requirements [22, 23] are very similar to the
ROAMOPS model, with the exception that the access protocol is
Mobile-IP [2] instead of PPP.

The Diameter protocol was not designed from the ground up. Instead,
the basic RADIUS model was retained while fixing the flaws in the
RADIUS protocol itself. Diameter does not share a common protocol
data unit (PDU) with RADIUS, but does borrow sufficiently from the
protocol to ease migration.

The basic concept behind Diameter is to provide a base protocol that
can be extended in order to provide AAA services to new access
technologies. Currently, the protocol only concerns itself with
Internet access, both in the traditional PPP sense as well as taking
into account the ROAMOPS model, and Mobile-IP.

Although Diameter could be used to solve a wider set of AAA problems,
we are currently limiting the scope of the protocol in order to

ensure that the effort remains focussed on satisfying the
requirements of network access. Note that a truly generic AAA
protocol used by many applications might provide functionality not
provided by Diameter. Therefore, it is imperative that the designers
of new applications understand their requirements before using
Diameter.

1.1 Requirements language

In this document, the key words "MAY", "MUST", "MUST NOT",
"optional", "recommended", "SHOULD", and "SHOULD NOT", are to be
interpreted as described in [9].

1.2 Terminology

Accounting
   The act of collecting information on resource usage for the
   purpose of trend analysis, auditing, billing, or cost allocation.

Authentication
   The act of verifying the identity of an entity (subject).

Authorization
   The act of determining whether a requesting entity (subject) will
   be allowed access to a resource (object).

AVP
   The Diameter protocol consists of a header followed by one or more
   Attribute-Value-Pair (AVP). The AVP includes a header and is used
   to encapsulation authentication, authorization or accounting
   information.

Broker
   A broker is a business term commonly used in AAA infrastructures.
   A broker is either a proxy or redirect server, and MAY be operated
   by roaming consortiums.

Diameter Client
   A Diameter Client is a device at the edge of the network that
   performs access control. An example of a Diameter client is a
   Network Access Server (NAS) or a Foreign Agent (FA).

Diameter Server
   A Diameter server is a device that is not acting as a NAS or FA.
   Servers can be proxy, redirect, or home servers

Downstream Server
   Diameter Proxy servers identify a downstream server as one that is
   providing routing services towards the home server for a
   particular message.

Home Domain
   A Home Domain is the administrative domain with whom the user
   maintains an account relationship.

Home Server
   A Diameter Home Server is one that authenticates and/or authorizes
   access for users of a particular realm. The same server MAY also
   act as a proxy or redirect server for other realms, in which case
   it is not acting as a Home Server for these realms.

Integrity Check Value (ICV)
   An Integrity Check Value is an unforgeable or secure hash of the
   message with a shared secret.

Interim accounting
   An interim accounting message provides a snapshot of usage during
   a user's session. It is typically implemented in order to provide
   for partial accounting of a user's session in the event of a
   device reboot or other network problem that prevents the reception
   of a session summary message or session record.

Local Domain
   A local domain is the administrative domain providing services to
   a user. An administrative domain MAY act as a local domain for
   certain users, while being a home domain for others.

Network Access Identifier
   The Network Access Identifier, or NAI [3], is used in the Diameter
   protocol to extract a user's identity and realm. The identity is
   used to identify the user during authentication and/or
   authorization, while the realm is used for message routing
   purposes.

Proxy Server
   A proxy server öses the realm portion of the NAI to route Diameter
   messages. Proxy servers are typically used to minimize the number
   of security relationships that are required between Diameter
   servers.

Realm
   The string in the NAI that immediately follows the '@' character.
   NAI realm names are required to be unique, and are piggybacked on
   the administration of the DNS namespace. Diameter makes use of the

      realm, also loosely referred to as domain, to determine whether
      messages can be satisfied locally, or whether they must be
      proxied.

   Real-time Accounting
      Real-time accounting involves the processing of information on
      resource usage within a defined time window. Time constraints are
      typically imposed in order to limit financial risk.

   Redirect Server
      A Diameter redirect server provides realm to address translation,
      by returning information necessary for Diameter peers to
      communicate directly. Redirect servers are different from proxies
      since they do not participate in the routing of messages between
      end Diameter nodes.

   Roaming Relationships
      Roaming relationships include relationships between companies and
      ISPs, relationships among peer ISPs within a roaming association,
      and relationships between an ISP and a roaming consortia.
      Together, the set of relationships forming a path between a local
      ISP's authentication proxy and the home authentication server is
      known as the roaming relationship path.

   Session
      The Diameter protocol is session based. When an authorization
      request is initially transmitted, it includes a session identifier
      that is used for the duration of the session. The Session-
      Identifier AVP contains the identifier and must be globally
      unique.

   Session record
      A session record represents a summary of the resource consumption
      of a user over the entire session. Accounting gateways creating
      the session record may do so by processing interim accounting
      events or accounting events from several devices serving the same
      user.

   Upstream Server
      Diameter Proxy servers identify an upstream server as one that is
      providing routing services towards the Diameter client.

2.0 **Problems to be addressed**

   The RADIUS protocol was designed in the early 1990's as an attempt to
   solve a scaling problem associated with dial-in and telnet servers.
   Over time the networks became more complex (e.g. roaming networks)

   and the Network Access Servers (NAS) increased in complexity and
   density. These changes combined with a massive deployment of the
   protocol uncovered some fundamental issues with the protocol that
   needed to be fixed. The Diameter protocol was designed as a next
   generation RADIUS protocol, designed with roaming and high density
   NASes in mind.

   This section will describe the documented, and undocumented, RADIUS
   problems known today. Further sections will describe how the Diameter
   protocol addresses each one of these problems.

2.1 Strict limitation of attribute data

   One of problems that RADIUS suffers from is its inherent limitation
   on the length of attribute data. This limitation is imposed by the
   fact that the protocol's attribute header only reserves one byte for
   the length field. The RADIUS protocol does specify that larger data
   can be spanned across multiple attributes, however doing so
   introduces a new set of problems. The RADIUS protocol also allows
   multiple attributes of the same type to be included within a message.
   Therefore, it is difficult for a RADIUS server, or client, to
   determine whether multiple identical attributes are in fact multiple
   independent attributes, or a single fragmented attribute.

2.2 Strict limitation on concurrent pending messages

   The RADIUS protocol states that the identifier field, found within
   the header, is used to identify retransmissions. This one byte field
   imposes a strict limitation on the number of requests that can be
   pending at any given time to 255. In the early 1990's, this number
   was sufficient, but the increased density of most NASes today make
   the protocol nearly unusable. Later versions of the protocol
   specification attempts to solve this problem by making use of
   multiple UDP ports, and making use of as many ports as necessary to
   ensure that no more than 255 simultaneous requests are pending.

   The RADIUS protocol also requires that retransmitted request, which
   include changes to the packet, include a new value in the Identifier
   field. Note that most retransmissions do include updated information,
   and therefore typically require a new Identifier field. This further
   reduces the number of sessions that can be supported by the
   Identifier field.

2.3 Inability to control flow to servers

Given the rather bursty nature of the RADIUS protocol, current
servers have no way of properly managing their receive buffers. This
is in part due to the fact that RADIUS operates over UDP, and does
not include any windowing support. This has been known to cause
large bursts of requests to be directed to a server, which can burden
a server's ability to respond in a timely manner. This problem is
most prevalent in cases where a server becomes unavailable and all
requests must be sent to an alternate server, or when an ingress port
on the NAS becomes available (e.g. T3 port on NAS).

2.4 No retransmission procedure

Given that the RADIUS protocol requires that the Identifier field be
changed in retransmissions that have updated information, RADIUS
server developers have had to design clever tricks to identify
retransmissions. One common method is to cache all packets received
in a time window (e.g. 60 seconds). When such servers receive a
packet, it compares the contents of certain attributes, which are
know to be static across retransmissions, with corresponding
attributes in all packets in the cache. When a match is found, a
retransmission has been detected. This burden placed on RADIUS
servers add additional latency, which may cause NAS retransmissions
(see Section 2.5).

2.5 End to end message acknowledgment

The RADIUS protocol requires that a NAS retransmit a request until a
successful or failed response is received, and does not permit a
RADIUS server to retransmit a response. Since RADIUS servers
typically have to perform a database lookup to authenticate the user,
such operations MAY be lengthy, and cause the NAS to assume that the
request was never received, and retransmit (causing further
congestion).

In cases when proxy servers are used, retransmissions are even more
likely since each proxy must identify retransmissions, validate the
request, optionally impose some local policy decision, and forward to
the downstream server.

2.6 Limited server failure detection

The RADIUS protocol, operating over UDP, does not provide a clear
method for a NAS to detect whether the lack of a response for a given
request is the result of congestion, or server failure. In networks
that do not employ proxies, this is not an issue. However, in

networks that do make use of proxies, the lack of a response MAY not
be a local problem, but a problem with a downstream or home server.
The NAS does not have a mechanism to identify that the local server
is still available, and MUST retransmit all pending requests to an
alternate server, including those destined for different downstream
or home servers. This places a burden not only on the offending home
server, but also on the NAS, proxies and all other home servers that
will receive retransmissions.

2.7 Silent discarding of packets

The RADIUS protocol states that messages that do not contain the
expected information, or messages that have errors are silently
discarded. Silently discarding messages causes the NAS to assume that
the local RADIUS server is no longer reachable, and causes it to
retransmit all pending requests to alternate servers (see Section
2.6). Such messages will be retransmitted to alternate servers, and
again silently discarded, and so on. This will occur until the NAS
abandons the request.

2.8 Inefficient Server Fail-Over

Most NAS implementations support a number of RADIUS servers,
consisting of a primary server with a set of alternate servers. When
the NAS detects that the primary can no longer be used, all pending
messages are transmitted to an alternate server. When the alternate
is not available, the next alternate server in the list is used.

Given that the RADIUS server operates over UDP, and has no watchdog
mechanism, the NAS has no way to know in advance whether an alternate
server is reachable. Therefore, if two or more consecutive servers in
the server list are unavailable, denial of service to users can be
very lengthy.

2.9 Inefficient use of RADIUS servers in proxy environments

As previously mentioned, NASes have no method of knowing whether the
lack of a response is due to a failure on the local, downstream
proxies, or the home server. Further, servers do not retransmit
RADIUS requests on behalf of the NAS. Therefore, should a primary
home server become unavailable, the local server does not retransmit
to an alternate server in the home network, but rather waits for the
NAS to timeout and retransmit to the local alternate server,
requiring parallel links between servers (see figure 1).

```
                              Local ISP              Home ISP
                              +--------+            +--------+
                              | Primary|            | Primary|
          +-------+           | Proxy  |----------->|  Home  |
          |       |---------->| Server |            | Server |
          |Network|           +--------+            +--------+
          |Access |
          |Server |           +--------+            +--------+
          |       |---------->|  2 nd  |            |  2 nd  |
          +-------+           | Proxy  |----------->|  Home  |
                              | Server |            | Server |
                              +--------+            +--------+
```
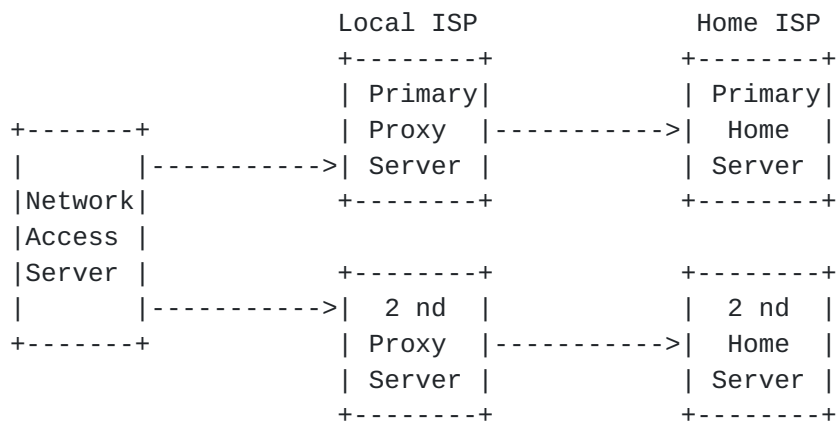
Figure 1: RADIUS Proxy Network

   Take an example where an ISP issues two authentication requests, one
   for abc.net and another for xyz.com. Let's also assume that abc.net's
   primary server is down, while xyz's 2nd server is down. Should such a
   problem occur, all requests for abc.net would cause the NAS to switch
   to the local ISP's 2nd server, while all requests to xyz.net would
   cause the NAS to switch back to the local ISP's primary server.

2.10 No unsolicited messages

   The RADIUS protocol does not allow a server to send unsolicited
   messages to the NAS. As network services became more complex, this
   limitation has forced manufacturers to deviate from the RADIUS
   protocol, causing interoperability problems. Server initiated
   messages are typically used for accounting purposes and to request
   that a NAS terminate a specific user session.

2.11 Replay Attacks

   Although RADIUS messages contain hop-by-hop authentication, the
   protocol does not include any replay attack prevention. This means
   that a malfunctioning server, or malicious user, can replay an old
   packet without detection. For servers that maintain state
   information, such as those that limit the number of concurrent
   sessions for a given user, a denial of service is very simple by
   replaying old RADIUS messages. For other servers, this problem is
   limited to duplicate accounting messages.

2.12 Hop-by-Hop security

   The RADIUS protocol uses hop-by-hop security, which means that every

hop in a RADIUS proxy network adds authentication data that is used
by the next peer in the chain. RADIUS has no facility for securing
the message between the NAS and the home server, eliminating the
ability for proxy servers to modify critical components in messages.
This has caused opportunities for fraud in RADIUS networks, since
intermediate nodes can easily modify information (e.g. accounting
information), and such events are difficult to traceable.

2.13 No support for vendor-specific commands

Although the RADIUS protocol does support vendor-specific attributes,
it does not allow for vendor-specific commands. This has forced
vendors to abuse the address space, creating interoperability
problems in mixed vendor environments.

2.14 No alignment requirements

Unlike most newer IETF protocols, the RADIUS protocol does not impose
any alignment requirements, which adds an unnecessary burden on most
processors. All fields within the header and attributes must be
treated as byte aligned characters.

2.15 Mandatory Shared Secret

The RADIUS protocol requires that a shared secret exists between two
peers. Therefore, even if IP Security was deployed to secure to
communication, the shared secret would still be required.

3.0 Diameter Architecture

The Diameter architecture consists of a base protocol and a set of
protocol extensions (such as strong security, NASREQ, Mobile-IP and
accounting). Functionality common to all supported services is
implemented in the base protocol, while application-specific
functionality may be provided through the extension mechanism.

The base protocol [18] must be supported for all Diameter
applications, and defines the basic PDU format, a few primitives and
the basic security services offered by the protocol. Unlike RADIUS,
the Diameter protocol operates over SCTP [24], which provides
reliability and an well defined retransmission and timeout mechanism.
Additionally, Diameter defines a fail-over strategy, which is lacking
in the RADIUS protocol. SCTP provides a windowing scheme, which
allows the AAA servers to limit the flow of incoming packets. This

can then be used by the AAA clients to distribute the traffic load
across multiple servers. The transport layer's retransmission and
timeout timers allow clients and servers to detect the reachability
state of peers, allowing for quick transition to back-up servers.

As previously discussed, the ROAMOPS model introduces the proxy, or
broker, which acts as an intermediate server forwarding requests to
user's home ISPs. ROAMOPS also described a set of attacks that one
could mount if such a network was built using the RADIUS protocol
[21]. In order to provide secure broker services, security between
the NAS and the home server is required at the application layer,
preventing such servers from modifying contents of RADIUS messages.

The Diameter Strong Security Extension defines a set of extensions to
the base protocol that provide authentication, confidentiality and
non-repudiation at the Attribute-Value-Pair (AVP) level. With these
extensions, it is possible to secure portions of a Diameter message,
while other parts of the message are not secured. Secured objects are
called protected AVPs; non-secured objects are called unprotected
AVPs. Using Diameter, proxies can add, delete or modify unprotected
AVPs in a message.

The RADIUS protocol provides dial-up PPP AAA services by providing
three commands and many Attributes. Attributes in RADIUS are
analogous to AVPs in Diameter. In order to ease migration from RADIUS
to Diameter, the first 256 AVPs in the Diameter AVP space are
reserved for RADIUS compatibility. This allows both protocols to
share a common dictionary and policy rules for PPP user profiles.

The RADIUS protocol has support for the Extensible Authentication
Protocol (EAP) [10], but RADIUS' lack of support for large attributes
and its inherent unreliability has made the integration of the
protocols very difficult.

The Diameter NASREQ Extension defines a set of
authentication/authorization commands, which can be used for CHAP,
PAP and EAP. Diameter's support for larger AVPs and the SCTP
transport properties have made the use of EAP much more palatable,
allowing for end-to-end user authentication, which reduces many of
authentication replay attacks known to exist with CHAP and PAP.

Unlike PPP, Mobile-IP hosts do not have a long-lived "nailed-up"
connection to a PPP server, but rather get service from routers that
provide service in a particular cell. In the Mobile-IP world, the
router is known as a Foreign Agent, while the moving hosts are known
as Mobile Nodes. The mobile node's home network has a host that
forwards all messages destined to the mobile node through the Foreign
Agent. This router is commonly referred to as the Home Agent.

Mobile-IP [7] allows the mobile nodes to move from one cell (subnet)
to another while retaining the same IP address, minimizing the impact
to applications. Although the Mobile-IP protocol could be deployed in
a small network with any AAA services, a larger network suffers from
many scaling issues such as:

   - Static mobile node home address
   - Static mobile node home agent
   - Requirement to pre-configure mobile node profile on home agents
   - No inter-domain mobility

Both PPP and Mobile-IP require that usage data be collected for uses
such as capacity planning and for accounting purposes. The current
standard protocol for accounting is SNMP [12], but experience
indicates that SNMP often is not the correct protocol for service
accounting. Today many applications and services use RADIUS
accounting [4] as their accounting protocol, however the RADIUS
accounting protocol is not an IETF standard; in addition, it suffers
from similar scaling and security problems. The Diameter accounting
extension [11] is designed to allow accounting information to be sent
across administrative domains (optionally through brokers), and has
been derived from an accounting requirements document [6, 8].

```
                         +-----------+
                         | Mobile-IP |
                         |           |
                         | Extension |
                         +-----------+
          +-----------+        ^      +------------+
          |  NASREQ   |        |      | Accounting |
          |           |        |      |            |
          | Extension |        |      | Extension  |
          +-----------+        |      +------------+
               ^               |            ^
               |               |            |
               v               v            v
          +---------------------------------+--------------------+
          |                                 |                    |
          |     Diameter Base Protocol      |  Strong Security   |
          |                                 |                    |
          +---------------------------------+--------------------+
```
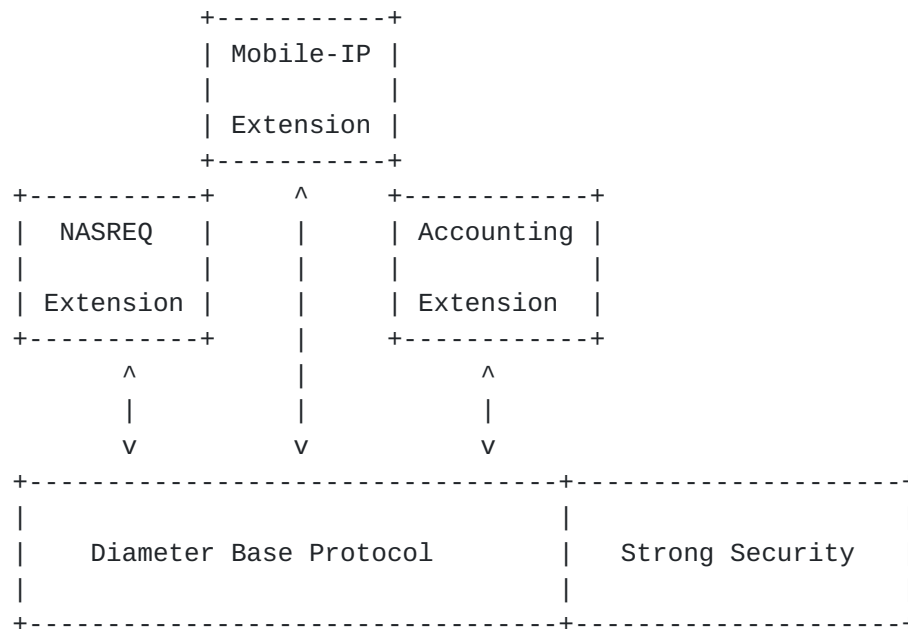   Figure 2: Diameter Protocol Architecture

3.1 Diameter Base Protocol

The Base Protocol defines the Diameter message format, a set of
primitives and how the messages are transmitted in a secure fashion.

The Base Protocol assumes a peer-to-peer communication model, as
opposed to a client-server model. The following goals motivated the
design of the base protocol:

- lightweight and simple to implement protocol
- Large AVP space
- Efficient encoding of attributes, similar to RADIUS
- Support for vendor specific AVPs and Commands
- Support for large number of simultaneous pending requests
- Reliability provided by underlying SCTP
- Well-defined fail-over scheme
- Ability to quickly detect unreachable peers
- No silent message discards
- Support of unsolicited messages to "clients"
- integrity and confidentiality at the AVP level
- Hop-by-Hop security
- One session per authentication/authorization flow
- Provide redirect (referal) services, to allow bypassing of
  broker

The Diameter base protocol is intended to simply provide a secure
transport for the messages defined in the various application-
specific extensions. It is therefore imperative that the base be
lightweight and simple to implement.

In the Diameter protocol, data objects are encapsulated within the
Attribute Value Pair (AVP). An AVP consists of three parts: the
Identifier, Length and Data. A unique AVP Identifier is assigned to
all data objects in order to be able to distinguish the data
contained. The AVP Identifier namespace must be sufficiently large to
ensure that future protocol extensibility is not limited by the size
of the namespace, as in the RADIUS protocol. Furthermore, vendors
wishing to add "proprietary" extensions must be allowed to do so by
using a vendor-specific namespace, managed by IANA.

For many years the question as to whether RADIUS should operate over
UDP or TCP has led to heated discussion. It must be determined
whether the benefits that UDP provides are worth the implementation
complexities. Over time, it has become clear that these benefits are
well worth the cost. The issue with TCP is that an AAA protocol
requires a quick retransmission and fail-over scheme, which TCP
cannot provide. The Diameter protocol must be able to operate over a
transport that has an aggressive retransmission strategy in order to
efficiently switch to an alternate host when the peer in question is
no longer reachable.

Contrary to RADIUS, the Diameter protocol requires that each node in
a proxy chain acknowledge a request, or response, at the "transport"

layer. Since Diameter operates over SCTP, which provides a reliable
transport, each node in a proxy chain is responsible for
retransmission of unacknowledged messages.

The SCTP transport provides retransmission detection, which greatly
simplifies server implementations, and consequently allows a given
server to support a much larger number of transactions per second.
SCTP also provides windowing, which allows the flow of packets to a
specific server to be controlled. Clever implementations can then
decide to send the packets to an alternate server that can handle the
load.

With the exception of a few security related errors, the Diameter
protocol requires that all messages be acknowledged, either with a
successful response or one that contains an error code.

Where the RADIUS protocol is client-server, the Diameter protocol is
peer to peer, allowing unsolicited messages to be sent to NASes.
There are many benefits to peer-to-peer AAA protocols. One example is
the on-demand retrieval of accounting data; another, server-initiated
session termination.

The Base Diameter protocol provides for hop-by-hop security, similar
to the scheme employed by RADIUS today. However, the Diameter
protocol also provides for replay protection through a timestamp
mechanism. This security scheme requires a long lived security
association to be established by peers, or can make use of keying
material negotiated out of band. The Base Protocol also allows the
built-in security measure to be turned off, (i.e., in cases where
IPSec is in use).

The Diameter protocol is a session-oriented protocol, meaning that
for each user being authenticated, there exists a session between the
initiator of the authentication/authorization request and the home
Diameter server. Sessions are identified through a session
identifier, which is globally unique at any given time. All
subsequent Diameter transactions (e.g. accounting) must include the
session identifier to reference the session. A Session termination
message exists in order to end a Diameter session, and all sessions
have a timeout value in order to ensure that they can be cleaned up
properly.

Since today's processors work more efficiently when objects are
aligned on a 32-bit boundary, the Diameter protocol requires 32-bit
alignment of all headers and the data. This has recently become a
common requirement for many new protocols at the IETF.

3.1.1 Proxy Support

   The Diameter protocol was designed from the beginning to support
   roaming networks. This means that every node in the network is
   responsible for it's own retransmissions, and the protocol does allow
   each node to know a priori the reachability state of each peer. This
   allows for a resilient network, and efficient retransmission scheme.
   Figure 3 depicts a network where each Diameter server can communicate
   with all other servers.

   Figure 3 depicts an example of a Diameter network that includes two
   proxy servers in the local network for resilience. Once a message has
   been sent from the NAS to one of its local proxy servers, they are
   responsible for any retransmissions of the message to one of the home
   servers. Since the underlying transport provides quick peer failure
   detection, upon such notification, the local proxies can quickly
   transmit the message to the alternate peer in the home network.

   Figure 3 depicts an example of a proxy network that includes
   alternate servers for resilience. Each node in the proxy chain is
   responsible for its own retransmissions and fail-over detection.
   This provides the following benefits:

     - The number of Diameter nodes in the network is greatly reduced
     - The latency involved in switch-over to an alternate peer is
       greatly reduced
     - Reliability is increased

```
                                  local ISP              Home ISP
                                  +--------+              +--------+
                                  | Primary|              | Primary|
           +-------+              | Proxy  |------------->|  Home  |
           |       |------------->| Server |<-----+----->| Server |
           |Network|              +--------+      |       +--------+
           |Access |                              |
           |Server |              +--------+      |       +--------+
           |       |------------->|  2 nd  |<-----+----->|  2 nd  |
           +-------+              | Proxy  |------------->|  Home  |
                                  | Server |              | Server |
                                  +--------+              +--------+
```
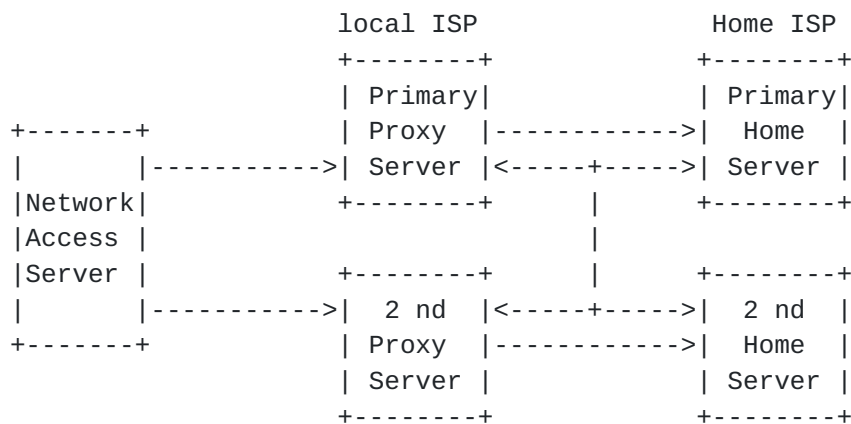
                    Figure 3: Diameter Proxy Network

3.1.2 Redirect Support

   A redirect server is one that provides simple Diameter message
   "routing" functions. Redirect servers are generally deployed in order

to reduce the configuration information that would otherwise be
necessary on all servers owned members of a roaming consortium.

Redirect servers allow Diameter entities to communicate directly by
providing NAI realm to home server translation services. When a
request is received by a redirect server, a redirect response is
returned to the initiator of the request with the information
necessary to communicate directly with servers in the home domain.

A broker, owned by a roaming consortium, MAY also provide Certificate
Authority services, by issuing certificates to all Diameter servers
within the consortium (or alternatively sign existing certificates).
This eliminates the need for long lived shared secrets between
Diameter servers, and enables protocols such as IP Security to be
used. In the event that non repudiation is required, public key
cryptography can be used to sign usage information in accounting
messages.

If deemed necessary, a redirect server MAY include the home server's
certificates in the redirect response to the requesting Diameter
server.

```
                      +-----------------+ +---------+
                      |     Redirect    | | CRL DB/ |
                      |      Server     | |  OCSP   |
                      +-----------------+ +---------+
                       ^
                Request |  Response with
                       |  Result Code =
                       |  Redirect
                       v
           +---------+                   +----------+
           |  Local  |                   |   Home   |
           | Diameter|<------------>| Diameter |
           |  Server |                   |  Server  |
           +---------+    Direct     +----------+
                      Communication
```
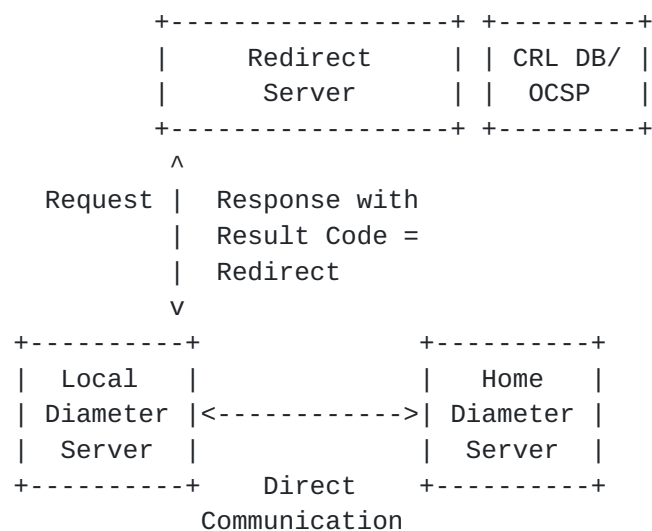Figure 4: Diameter Broker Returning Redirect Indication

It is important to note that redirect servers MAY forbid direct
communication of accounting messages. This may be required in cases
where the server needs such information to provide such services as
auditing and settlement services. Such servers MAY also required that
both parties sign accounting messages in a serial fashion, as
specified in [26].

3.2 **Strong Security Extension**

The Diameter base protocol allows Diameter servers to communicate
securely, using hop-by-hop authentication. Hop-by-hop authentication
means that the requesting server has secure communication with a
proxy or redirect server, and the proxy has secure communicate with
the home server.

The Strong Security extension [26] provides strong authentication of
selective AVPs, which MAY be used for repudiation purposes. This
extension also allows for secure communication through intermediate
Diameter proxies.

The extension achieves this functionality by allowing the
Cryptographic Message Syntax (CMS) S/MIME object to be encapsulated
within a Diameter AVP. The CMS object MAY be used for authentication,
confidentiality and to carry certificates and certificate revocation
lists (CRLs). The extension also provides for multi-party signatures,
which is useful in environments where two or more parties must sign
information, such as an accounting record.

Diameter clients (e.g. NAS, FA) aren't required to implement strong
security. It is possible for the local Proxy server to provide this
functionality, and MAY require that strong security only be used when
messages traverse administrative domain boundaries.

The strong security extension MUST only be used in networks that
include a Public Key Infrastructure (PKI).

3.3 Mobile-IP Extension

The Mobile-IP protocol is used to manage mobility of an IP host
across IP subnets [7]. Recent activity within the Mobile-IP Working
Group has defined the interaction between Mobile-IP and AAA in order
to provide:

    - Better scaling of security associations
    - Mobility across administrative domain boundaries
    - Dynamic home agent assignment

The Mobile IP protocol [7] works well when all mobile nodes belong to
the same administrative domain. Some of the current work within the
Mobile IP Working Group is to allow Mobile IP to scale across
administrative domains. This work requires modifications to the
existing Mobile IP trust model.

Figure 5 depicts the Diameter trust model for Mobile-IP. In this
model each network contains mobile nodes (MN) and a Diameter server.
Each mobility device shares a security association (SA) with the

Diameter server within its own home network. This means that none of
the mobility devices initially share a security association. The
Diameter servers in both administrative domains can either share a
direct security association, or can have a security association with
an intermediate proxy.

```
                        Proxy AAA
                        +--------+
                        |        |
                        |Diameter|
               +------->|        |<------+
               |        +--------+       |
         Local | SA                 SA | Home
           AAA   v                   v   AAA
         +--------+                 +--------+
         |        |      SA4        |        |
         |Diameter|=================|Diameter|
         |        |(when no proxy is used)|        |
         +--------+                 +--------+
              ^                      ^       ^
              |                      |       |
         SA1  |                 SA2  |       | SA3
              |                      |       |
              v                      v       v
         +--------+            +--------+  +--------+
         |        |            |        |  |        |
         |   FA   |            |   MN   |  |   HA   |
         |        |            |        |  |        |
         +--------+            +--------+  +--------+
```
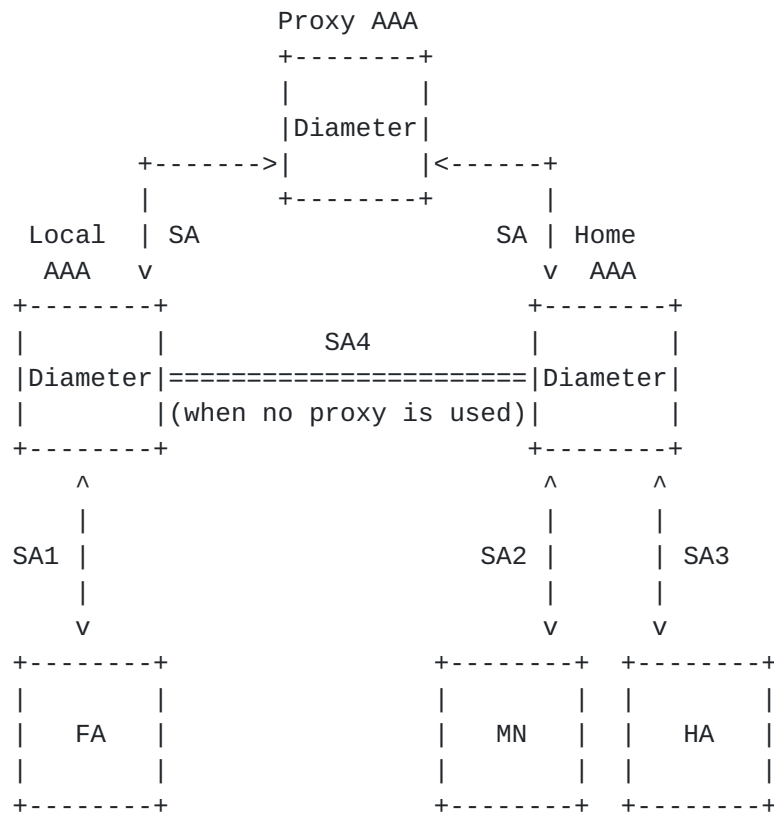
                Figure 5 - Mobile-IP AAA Trust Model

Figure 6 provides an example of a Mobile-IP network that includes
Diameter. In the integrated Mobile-IP/Diameter Network, it is assumed
that each mobility agent shares a security association between itself
and its local Diameter server. Further, the Home and Local Diameter
servers both share a security association with the broker's Diameter
server. Lastly, it is assumed that each mobile node shares a trust
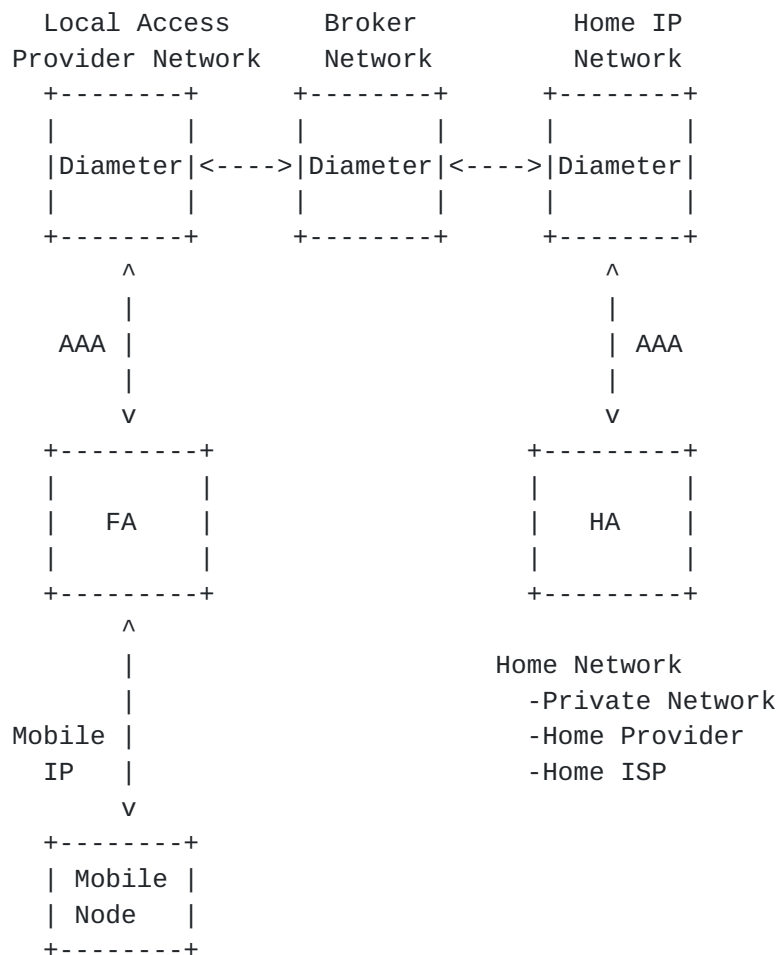relationship with its home Diameter Server.

```
         Local Access        Broker           Home IP
        Provider Network     Network          Network
          +--------+      +--------+       +--------+
          |        |      |        |       |        |
          |Diameter|<---->|Diameter|<---->|Diameter|
          |        |      |        |       |        |
          +--------+      +--------+       +--------+
              ^                                ^
              |                                |
            AAA |                            | AAA
              |                                |
              v                                v
          +---------+                      +---------+
          |         |                      |         |
          |   FA    |                      |   HA    |
          |         |                      |         |
          +---------+                      +---------+
              ^
              |                        Home Network
              |                           -Private Network
          Mobile |                        -Home Provider
            IP   |                        -Home ISP
              v
          +--------+
          | Mobile |
          | Node   |
          +--------+
```

           Figure 6 - General Wireless IP Architecture for Mobile-IP AAA

   In this example, a Mobile Node appears within a local network and
   issues a registration to the Foreign Agent. Since the Foreign Agent
   does not share any security association with the Home Agent, it sends
   a Diameter request to its local Diameter server, which includes the
   authentication information and the Mobile-IP registration request.
   The Mobile Node cannot communicate directly with the home Diameter
   Server for two reasons:

      - It does not have access to the network. The registration
        request is sent by the Mobile Node to request access to the
        network.
      - The Mobile Node may not have an IP address, and may be
        requesting that one be assigned to it by its home provider.

   The Local Diameter Server will determine whether the request can be
   satisfied locally through the use of the Network Access Identifier
   [3] provided by the Mobile Node. The NAI has the form of user@realm
   and the Diameter Server uses the realm portion of the NAI to identify
   the Mobile Node's home Diameter Server. If the Local Diameter Server

does not share any security association with the Mobile Node's home
Diameter Server, it may forward the request to a proxy or redirect
server. If the server has a relationship with the home network, it
can forward the request (or redirect), otherwise a failure indication
is sent back to the Local Diameter Server.

When the home Diameter Server receives the Diameter Request, it
authenticates the user and begins the authorization phase. The
authorization phase includes the generation of:

   - Dynamic session keys to be distributed among all mobility agents
   - Optional dynamic assignment of a home agent
   - Optional dynamic assignment of a home address (note this could
     be done by the home agent).
   - Optional assignment of QOS parameters for the mobile node [22]

Once authorization is complete, the home Diameter Server issues an
unsolicited Diameter request to the Home Agent, which includes the
information in the original Diameter request as well as the
authorization information generated by the home Diameter server. The
Home Agent retrieves the Registration Request from the Diameter
request and processes it, then generates a Registration Reply that is
sent back to the home Diameter server in a Diameter response. The
message is sent to the Local Server, through the proxy if one was
used, and finally to the Foreign Agent.

The Diameter servers maintain session state information based on the
authorization information. If a Mobile Node moves to another Foreign
Agent within the local administrative domain, a request to the local
Diameter server can be done in order to immediately return the keys
that were issued to the previous Foreign Agent. This eliminates an
additional round trip through the internet when micro mobility is
involved, and enables smooth hand-off. In order for the Diameter
server to be able to provide the keying information to the new
Foreign Agent, they must have a pre-existing security association.

Note that smooth hand-off is really a mobility function, and it is
not clear that Diameter should be involved. However, this example is
provided for completeness.

If the Mobile Node enters a service area owned by a new service
provider, the authentication and authorization request will have to
be sent back to the home Diameter server, which will create new
keying information.

3.3.1. Minimized Internet Traversal

Although it would have been possible for the Diameter interactions to
be performed for basic authentication and authorization, and the
Registration flow to be sent directly to the Home Agent from the
Foreign Agent, one of the key Mobile-IP Diameter requirements is to
minimize Internet traversals. Including the Registration Request and
Replies in the Diameter messages allows for a single traversal to
authenticate the user, perform authorization and process the
Registration Request. This streamlined approach is required in order
to minimize the latency involved in getting wireless (cellular)
devices access to the network. New registrations should not increase
the connect time more than what the current cellular networks
provide.

3.3.2. **Key Distribution**

In order to allow the scaling of wireless data access across
administrative domains, it is necessary to minimize the security
associations required. This means that each Foreign Agent does not
share a security association with each Home Agent on the Internet.
The Mobility Agents share a security association with their local
Diameter server, which in turn shares a security association with
other Diameter servers. Again, the use of proxies (as defined by
ROAMOPS) allows such services to scale by allowing the number of
relationships established by the providers to be reduced.

After a Mobile Node is authenticated, the authorization phase
includes the generation of Sessions Keys. Specifically, three keys
are generated:

   - K1 Key to be shared between the Mobile Node and the Home Agent
   - K2 Key to be shared between the Mobile Node and the Foreign
     Agent
   - K3 Key to be shared between the Foreign Agent and the Home Agent

Each key is encrypted in two separate methods. K1 is encrypted using
SA3 (for the Home Agent), and using SA2 (for the Mobile Node). K2 is
encrypted using SA4 (for the Foreign Agent) and using SA2 (for the
Mobile Node). Lastly, K3 is encrypted using SA4 (for the Foreign
Agent), and using SA3 (for the Home Agent). When the Foreign Diameter
Server receives the keys, they are decrypted and re-encrypted using
SA1. All of the Security Associations (SAx) are shown in figure 5.
The keys destined for the foreign and home agent are propagated to
the mobility nodes via the Diameter protocol, while the keys destined
for the Mobile Node are sent via the Mobile-IP protocol.

Figure 7 depicts the new security associations used for Mobile-IP
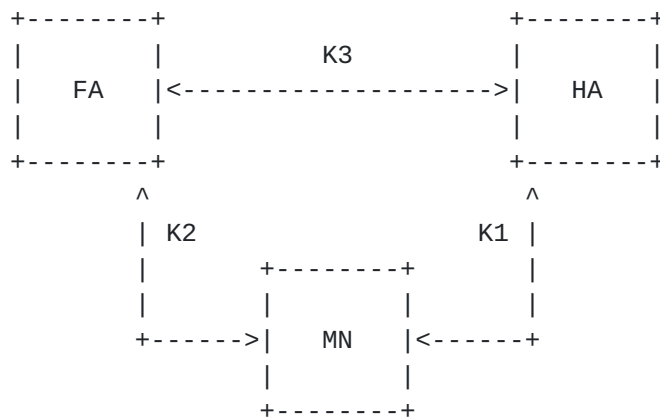message integrity using the keys derived by the Diameter server.

```
        +--------+                    +--------+
        |        |        K3          |        |
        |   FA   |<------------------->|   HA   |
        |        |                    |        |
        +--------+                    +--------+
            ^                             ^
            | K2                     K1  |
            |          +--------+         |
            |          |        |         |
         +------>|   MN   |<------+
                       |        |
                       +--------+
```

          Figure 7 - Security Association after Key Distribution

   Once the session keys have been established and propagated, the
   mobility devices can exchange registration information directly
   without the need of the Diameter infrastructure. However the session
   keys have a lifetime, after which the Diameter infrastructure must be
   used in order to acquire new session keys.

3.4 NASREQ Extension

   The NASREQ extension provides authentication and authorization for
   dial-in PPP users, terminal server access and tunneling applications,
   such as L2TP. The extension makes use of the attributes defined in
   the RADIUS protocol to carry the data objects. This was intended to
   ease migration of existing RADIUS servers to Diameter since they
   could share a single dictionary and user profile. Furthermore, this
   would reduce the amount of processing required for an inter-working
   system that acts as a RADIUS/Diameter bridge.

   Diameter has native EAP support that solves known problems in the
   RADIUS protocol. Furthermore, Diameter takes end-to-end
   authentication one step further by providing for end-to-end
   authentication via PPP's CHAP. This allows for a more secure
   authentication infrastructure without having to replace or modify the
   installed base of clients.

   If end-to-end CHAP is used in bridged Diameter/RADIUS environments,
   the bridge host is responsible for generating the challenge to the
   user.

   The remaining authentication and authorization logic found in RADIUS
   implementations can then be re-used. The basic changes are the
   message formats and the transmission mechanism as defined in the
   Diameter base protocol. This section does not detail RADIUS
   authentication and authorization. The interested reader should refer

to [1].

3.5 Accounting Extension

The Accounting extension provides usage collection to both the
Mobile-IP and the NASREQ extensions. The accounting requirements
specifications [6, 8] define that an accounting protocol must provide
the following functionality:

  - Negotiable transfer mechanism.
  - Provide general purpose AVPs.
  - Flexible to allows new extensions to use the accounting
    extension.
  - Scalable to allows millions to users and thousands of sites.
  - Secure accounting data transfer.

Like the RADIUS protocol, Diameter includes accounting usage
information in AVPs. The Accounting extension defines a set of
accounting AVPs that are used for all services, while each extension
defines their own service specific accounting AVPs.

The Diameter Accounting Extension allows accounting information to be
sent in real-time. Real-time accounting transfers are useful in
environments where timely arrival of the information is required,
such as when debit cards are used.

The Diameter protocol is session oriented, and each session typically
has a finite lifetime. Prior to the timeout of a session, a user
typically needs to be re-authentication and/or re-authorized in order
to extend the life of the session. In the Mobile-IP world, this
equates to the mobility registration lifetime, while in PPP this
means that the PPP authentication must be re-opened. When a re-
authentication and/or re-authorization occurs, a new token is
generated, which is used in the corresponding accounting message.

The Diameter Accounting extension combined with the Strong Security
[26] extension (see section 3.2), provides strong authentication of
accounting data, which MAY be used for repudiation purposes. The
strong security extension also allows multiple parties to sign the
accounting information, which is beneficial in environments that
include a referral broker. The foreign and home servers can both
sequentially sign the accounting record, and submit the result to the
broker. The broker can then use the signatures to ensure that both
parties agreed to the contents of the accounting record.

3.6 Resource Management

Many network access services requiring AAA support have a requirement
for servers that maintain session state information. An example of
such a requirement is in the dial-up PPP world. With the introduction
of flat-rate internet access, there has been a surge in fraud where a
user provides his username/password pair to other people. The end
result is that a single username (account) can have simultaneous
concurrent sessions.

Internet Service Providers have had to implement proprietary
extensions to RADIUS, in order to attempt to identify when such fraud
occurs. Unfortunately, since RADIUS does not provide the necessary
functionality required to maintain state information, these solutions
have been largely unreliable.

The Diameter Base Protocol [18], the Accounting extension [11], the
Mobile IP [13] and NASREQ [23] extensions provide some of the
functionality that is required for servers to maintain state
information, such as:

   - Reliable Transport
   - Indication of the termination of a session
   - A Reboot message
   - Interim Accounting
   - Accounting On/Off message
   - Ability to re-authorize an existing session

Although the above features do allow nodes to maintain state
information, it MAY be necessary for Diameter nodes to request a
snapshot of active sessions from a peer. This may be used when state
information is lost, which could occur after a device failure, or
this may be done periodically in order to ensure that the state is
current.

The Diameter Resource Management extension [5] provides the messages
that are required for a node to request a snapshot of active sessions
from a peer. State information is exchange via the Resource-Token
AVP, which is used to encapsulate a set of AVPs that describe the
session and resources used. There is one Resource-Token AVP for each
active session.

3.7 **Diameter Command Naming Conventions**

The following conventions are proposed for the naming of Diameter
messages. Diameter commands typically start with an object name, and
end with one of the following verbs:

[3.7.1](#) Request/Answer

   Request is used when the command is asking the peer to do something
   for it, for example, authorize a user, or terminate a session. The
   Answer MUST contain either a positive or negative result code,
   telling the requester whether or not the request successfully
   occurred. Other information can also be returned in the Answer.

   For example, AA-Request asks the peer device to authorize and/or
   authenticate a user in order to set up a session. The request may
   fail, thus the answer may be positive or negative.

[3.7.2](#) Query/Response

   Query is used when the command is asking for information that it
   expects the peer to have. An example would be querying for current
   configuration information, or querying for information on resources
   or sessions in use. The Response usually contains a positive result
   code and the information, or a negative result code with the reason
   for not answering the query.

   For example, Resource-Query requests the peer device to return
   specific information about one or more resources. The answer is
   returned in a Resource-Response.

[3.7.3](#) Indication

   Indication is used when the command is giving information on
   something that is about to or has already occurred. The peer
   receiving the message does not respond to the message, but a
   transport level acknowledgement must be done in order to ensure that
   the message was reliably delivered.

[4.0](#) Why not LDAP?

   One common question is whether LDAP would provide the functionality
   required.

   A Server MAY wish to access policies using LDAP, but the use of LDAP
   between the client and the server is not possible. The use of LDAP in
   this case would require that all routers have read/write access to
   the directory. Most customers would not accept this requirements and
   it is not efficient.

   In the case of roaming, customers would have to open up their

directory so outside routers have writable access. The security
implications set aside, having 1000's of routers constantly
read/write to the directory would cause some additional problems to
the Directory Service.

Finally, LDAP does not provide server initiated messages which is a
requirement for an AAA protocol.

5.0 References

[1] Rigney, et alia, "RADIUS", RFC-2138, Livingston, April 1997

[2] Veizades, Guttman, Perkins, Kaplan, "Service Location Protocol",
     RFC-2165, June 1997.

[3] Aboba, Beadles, "The Network Access Identifier", RFC 2486, Janu-
     ary 1999.

[4] Rigney, "RADIUS Accounting", RFC-2139, April 1997.

[5] P. Calhoun, "Diameter Resource Management", draft-calhoun-
     diameter-res-mgmt-07.txt, IETF Work in Progress, February 2001.

[6] B. Aboba, J. Arkko, D. Harrington. "Introduction to Accounting
     Management", RFC 2975, October 2000.

[7] C. Perkins, Editor. IP Mobility Support. RFC 2002, October
     1996.

[8] S. Kent, R. Atkinson, "Security Architecture for the Internet
     Protocol", RFC 1825, November 1998.

[9] Bradner, "Key words for use in RFCs to Indicate Requirements
     Levels", BCP 14, RFC 2119, March 1997.

[10] L. Blunk, J. Vollbrecht, "Extensible Authentication Protocol
     (EAP)", RFC 2284, March 1998.

[11] J. Arkko, P. Calhoun, P. Patel, G. Zorn, "Diameter Accounting
     Extension", draft-ietf-aaa-diameter-accounting-01.txt, IETF work
     in progress, March 2001.

[12] J. Case, D. Harrington, R. Presuhn, B. Wijnen, "Message Process-
     ing and Dispatching for the Simple Network Management Proto-
     col:", RFC 2572, April 1999.

[13] P. Calhoun, C. Perkins, "Diameter Mobile IP Extensions", draft-
     ietf-aaa-diameter-mobileip-01.txt, IETF work in progress, March
     2001.

[14] M. Baum, H. Perritt, "Electronic Contracting, Publishing and EDI
     Law", Prentice-Hall, ISBN 0-471-53135-9.

[15] P. Calhoun, C. Perkins "Mobile IP Foreign Agent
     Challenge/Response Extension", RFC 3012, November 2000.

[16] D. Harkins, D. Carrell, "The Internet Key Exchange (IKE)" RFC
     1409, November 1998.

[17] W. Simpson, "The Point-to-Point Protocol (PPP)", RFC 1661, STD
     51, July 1994.

[18] P. Calhoun, A. Rubens, H. Akhtar, E. Guttman, "Diameter Base
     Protocol", draft-ietf-aaa-diameter-01.txt, IETF work in pro-
     gress, March 2001.

[19] B. Aboba, G. Zorn, "Criteria for Evaluating Roaming Protocols",
     RFC 2477, January 1999.

[20] B. Aboba, J. Lu, J. Alsop, J. Ding, W. Wang, "Review of Roaming
     Implementations", RFC 2194, September 1997.

[21] B. Aboba, J. Vollbrecht, "Proxy Chaining and Policy Implementa-
     tion in Roaming", RFC 2607, June 1999.

[22] T. Hiller and al, "CDMA2000 Wireless Data Requirements for AAA",
     draft-hiller-cdma2000-aaa-02.txt, IETF work in progress, Sep-
     tember 2000.

[23] P. Calhoun, W. Bulley, A. Rubens, J. Haag, "Diameter NASREQ
     Extension", draft-ietf-aaa-diameter-nasreq-01.txt, IETF work in
     progress, March 2001.

[24] R. Stewart et al., "Simple Control Transmission Protocol", RFC
     2960, October 2000.

[25] Myers, Ankney, Malpani, Galperin, Adams, "X.509 Internet Public
     Key Infrastructure Online Certificate Status Protocol (OCSP)",
     RFC 2560, June 1999.

[26] P. Calhoun, W. Bulley, S. Farrell, "Diameter Strong Security
     Extension", draft-calhoun-diameter-strong-crypto-06.txt, IETF
     work in progress, February 2001.

6.0 **Acknowledgements**

   The Authors would like to thanks Bernard Aboba and Jari Arkko for
   their Accounting Requirements contribution. Thanks also goes to Erik
   Guttman for some very useful comments in helping make this draft more
   readable. The Mobile-IP Extension section was text originally writ-
   ten by Pat Calhoun for another Internet-Draft, which was subsequently
   cleaned up by Dave Spence. The authors would like to thank Nenad
   Trifunovic, Tony Johansson and Pankaj Patel for their participation
   in the Document Reading Party. A final thanks to Stephen Farrell for
   his security review.

7.0 Author's Addresses

Questions about this memo can be directed to:

     Pat R. Calhoun
     Sun Laboratories, Network and Security
     Sun Microsystems, Inc.
     15 Network Circle
     Menlo Park, California, 94025
     USA

      Phone: +1 650-786-7733
        Fax: +1 650-786-6445
     E-mail: pcalhoun@eng.sun.com

     Glen Zorn
     Cisco Systems, Inc.
     500 108th Avenue N.E., Suite 500
     Bellevue, WA 98004
     USA

      Phone: +1 425 438 8218
     E-Mail: gwz@cisco.com

     Ping Pan
     Bell Laboratories
     Lucent Technologies
     101 Crawfords Corner Road
     Holmdel, NJ 07733
     USA

      Phone: +1 732-332-6744
     E-mail: pingpan@dnrc.bell-labs.com

     Haseeb Akhtar
     Wireless Technology Labs
     Nortel Networks
     2221 Lakeside Blvd.
     Richardson, TX 75082-4399
     USA

      Phone: +1 972-684-8850
     E-Mail: haseeb@nortelnetworks.com

[8.0](#) Full Copyright Statement

[9.0](#) Expiration Date

   This memo is filed as <[draft-ietf-aaa-diameter-framework-01.txt](#)> and
   expires in August 2001.