Motivation Analysis for 4via6 Stateless Solutions
draft-chen-softwire-4v6-motivation-00

Abstract

   The growth of Internet obliged to accelerate IPv6 deployment. IPv6-
   only network likely become prevalent due to low costs and higher
   efficiency for operators to employ. Meanwhile, a significant part of
   network will still stay in IPv4 for long time. Regarding IPv4
   depletion, address sharing mode should be deployed in small segmented
   IPv4 networks. Operators expect 4via6 solution with low capital and
   operational expense by balancing tradeoff between simple core network
   and edge-distributed management. This memo discusses several
   motivations, and it recommends the 4via6 stateless solution should be
   specified for deployment models in the guideline document.

Table of Contents

1. Introduction

   With the fast development of global Internet, the demands for IP
   address are rapidly increasing currently. This year, IANA announced
   that the global free pool of IPv4 depleted on 3 February. IPv6 is
   the only real option on the table. Operators have to accelerate the
   process of deploying IPv6 networks in order to address IP address
   strains.

   IPv6 deployment normally involves a step-wise approach where parts of
   the network should properly updated gradually. As IPv6 deployment
   progresses it may be simpler for operators to employ a single-version
   network, since deploying both IPv4 and IPv6 in parallel would costs
   more than IPv6-only network. Operators have to maintain double
   management interfaces and operational support system for dual-stack
   network. Moreover, some additional efforts should be paid for
   troubleshooting. [RFC6180] also has described mitigation issues
   related to this model. Therefore switching to an IPv6-only network
   will become more prevalent. Meanwhile, a significant part of network
   will still stay in IPv4 for long time. There may not be enough
   public or private IPv4 addresses to support end-to-end network
   communication, without segmenting the network into small parts with
   sharing one IPv4 address space. Operators have to choose a IPv6
   transition technology to bridge these IPv4 islands through IPv6
   network. Furthermore, operators have to facilitate the continued
   growth of the deployment of Internet technology at relatively low
   capital and operational expense without destabilizing deployed
   services.

   Currently, DS-Lite[DS-Lite] could serve for hosts in IPv4 network to
   communicate servers located in IPv4 Internet via IPv6 network by
   encapsulating packages in IPv6 tunnel. In order to route inbound
   traffic to correct tunnel appropriately, AFTR needs to keep states
   for per-session. In such stateful solution, network should maintain
   user-session states relying on the dynamic NAT-state management
   function. Targeting to above objectives, operators may expect more
   lightweigh transition approach in terms of less expensive, easier
   scaling, more administrable and flexible. Lightening state-managment
   burden and reducing complexities in core network, operators desire
   high scalability and simplicity in core network. Accordingly,
   operators could balance the tradeoffs by pushing state-management to
   network edge on which CPE are located to achieve holistic
   optimization. It also could give operators high flexibilities to
   rapidly deploy IPv6 in operational network.

   Operators are always looking for an approach to minimize additional
   resources provision with accommodating more new customers during IPv6
   transition. This memo will specifically state several operational

considerations. It stands for seeking solution of which is not only
stateful but also stateless. The goal here is to help operators make
their decision for which solution should be adopted.

2. Terminology

In this document, the following terms are used.

4via6: it supports residual IPv4 services, over an infrastructure
geared for IPv6-only operations, and doing so in the context of IPv4
address depletion.

Border GW (Gateway): it will serve for interconnects two networks
that use different address families. The Border GW could be AFTR in
DS-Lite [DS-Lite] , CGN or LSN [CGN]. On the carrier side, operators
normally deploy several concentrated GWs for scalability reasons.

3. Operational Motivations Analysis

3.1. NAT Logging Consideration

Today, Border GW devices are required to log events like creation and
deletion of translations and information about the resources it is
managing to provide traceability. The logs are required in many
cases to identify an attacker or a host that was used to launch
malicious attacks and/or for various other purposes of accounting.

Some existing 4via6 solutions adopts statefull models. Users are
identified by a dynamic address and port "NAT Log". During the
operation, logging for every dynamic NAT mapping is created. The
logging information on per-session will consume significant parts of
physical resources, such as memory and processors resource. It can
be observed that NAT logging enabling might degrade Border GW
performances according to experimental results. In addition,
operators have to made complicated data inspection to dig up desired
information from inner tunnel header in case of valid IP package
encapsulated by routable outer header. This would require an
additional Deep Packet Inspection (DPI) equipment to perform such
analysis.

In order to make relief for resource consumption and analysis
complexities issues, operators would prefer to directly identify IPv4
users by means of derivatization from IPv6 address. In this case,
Border GW are not required to enable additional NAT logging or DPI
functionalities. It will not only preclude risks of equipment
performance degradation, but also simplify NAT logging analysis.

. **Transition Costs Consideration**

   Minimized resources costs for IPv6 transition is desirable for
   operators. CAPEX(e.g. investments of new added equipments) and OPEX
   (e.g. provisioning costs of network maintenance) should be reduced as
   much as possible so as to make forward progress on IPv6 deployment.

   Investments of Border GW occupy the significant part of 4via6
   transition CAPEX. In stateful case, Border GW serving as a NAT
   concentrator needs to keep mapping states for each session, which
   will require large memory and processors with high performance in
   large scale network. A dedicated NAT board should be integrated into
   Router/FW platform to perform such processing. According to latest
   statistics, the investment of dedicated NAT board supporting
   5,000,000 concurrent sessions is at hundreds of thousands price level
   and 500,000 concurrent sessions is at tens of thousands price level.
   It could be observed that the smaller the number of concurrent
   sessions required, the lower expenditures operators should spend on
   Border GW. Furthermore, high capacity of NAT broad would require
   high-end router platform. The capital expenditures become even more
   expensive. Besides, associated CPEs prices should also be taken
   account into total capital of 4via6 transition. As compared to that,
   operators could rely on the use of fully distributed NAT
   functionality located on CPE. The additional NAT functions doesn't
   increase CPE prices. Each CPE approximately is at hundred
   pricelevel. Yet, the Border GW investment will largely be decreased
   by eliminating dedicated NAT board integration, since Border GW
   doesn't need to keep any session states.

   OPEX is heavily depending on the provisioning for growing users. In
   a stateful "Hubs and Spokes model", a carrier must be able to scale
   the solution to millions of initiators by adding more Border GW
   (e.g., softwire concentrators). Since each Border GW maintains
   specific customers's states, new deployed GW likely requires
   significant manpower to carry out a new round of GW discovery
   planning and addressing layout. With growth of user number, OPEX
   would linearly be increased. It is painful for operators. In order
   to reduce operational costs, operators would require each Border GW
   decouple with specific users states. As a consequence, each Border
   GW can process the packets from all customers. Therefore, it could
   easy to save its maintenance costs.

   Total CAPEX and OPEX of a transition system determine the costs for
   IPv6 transition. According to above analysis, operators should take
   a lightweight approach to minimize stateful costs .

3.3. High Availability(HA) Consideration

   HA is one of key features to facilitate the IPv6 transitions and
   incremental commercial deployment. It will guarantee the reliability
   of deploying IPv6, especially during the IPv6 transition period.
   Border GW is deployed within large-scale networks, where a large
   number of customers are located. These customers within a network
   which is served by a single Border GW device MAY experience service
   degradation due to the presence of the single point of failure.
   Therefore, load-balancing and/or redundancy capabilities of the
   Border GW devices are strongly desired in order to deliver highly
   available services to customers.

   The load-balancing could avoid Border GW falling into the overload
   situations. In stateful cases, the upstream and downstream traffics
   for the same user must go through the same gateway. Load-balancing
   could be achieved only if each Border GW synchronizes up all
   customers states. Several contribution are proposed to satisfy this
   demands[NAT-Redundancy][NAT64-Load-Balancing]. However,
   standardization progress is struggling to proceed. In order to
   facilitate load balance, it is also very important for network
   operators to adopt flexible load sharing mechanism, in which upstream
   and downstream traffics for the same user could go through the
   different gateway. Any requirements of states synchronization
   between Border GW should be avoid in such approach. And GW load-
   balance should be naturally supported when appropriate routing
   information has been set.

   Regarding Border GW redundancy, cold standby and host standby are
   proposed to address stateful GW redundancy issues. The solutions are
   at the cost of a complex election procedure or manual configuration,
   also of a considerable cost and a low reliability. Operators should
   break such tradeoff by eliminating mapping states storing in GW . In
   such case, the backup GW can be replicated automatically if the
   primary GW is out of service. The switching process is smooth since
   any GW doesn't worry about losing mapping information for serving
   users.

3.4. Avoiding Centralized Traffic Bottleneck

   Statefull takes "Hubs and Spokes model" where there are major CPEs
   connecting a relatively small number of Border GW. When a user
   behind CPE require communications with others behind another CPE,
   data packages have to go through GW to reach proper users. Border GW
   is at risk with traffic bottleneck due to increasing serving users.
   Moreover, it could also bring additional latency during traffic
   delivery. It would become worse when CPE has long distance with
   remote Border GW.

Operators would consider an efficient data path by optimizing routing
to avoid such risks. "Mesh model" would be adopted to optimize data
path when there are communications required between CPEs. The data
packages between different CPEs can be directly delivered bypassing
remote GW. It would not only effectively minimize network latency,
but also relieve traffic burden on centralized GW.

3.5. **Optimizing Encap/Decap and Translation**

Regarding 4via6 transition architecture, a "tunnel" that is created
to cross a segment of IPv6 when communicating from IPv4 to IPv4
[RFC4925]. Both encapsulation/decapsulation and translations methods
for connecting IPv4 networks across IPv6-only networks could be
applied. For encap/decap mode, double IP header is appended to allow
IPv4 package transparently transmit through IPv6 network. As stated
above, operator have to employ DPI to inspect desired information
from inner package in some cases. This would further increase system
complexities and economic costs. In addition to that, additional IP
header would bring overhead on link usages, especially in wireless
network (spectrum is a very valuable and scarce resource). Operators
usually wish to eliminate unnecessary overhead in such environment.
Compared to encap/decap mode, translation based solution using single
IP header allows IPv4 interact with IPv6 by translating IPv4 header
to IPv6 header. However, some elements are missing during
translation since different IP header structures are existing between
IPv4 and IPv6 package.

Operators need to consider optimize data package processing in such
"tunnel" architecture for both encap/decap and translation modes.
The approach should not only minimize overhead introduced by double
IP header, but also avoid infromation missing by IPv4-IPv6
translation.

4. **Conclusions**

Operator should investigate an approach to minimize the costs
spending on IPv6 transition because we believe the ultimate goal of
the transition must be the long-term viability of the Internet and
also the provision of our services. To ensure that, we should take
all operations provisioning and network planning into account. The
considerations yielded conclusion that 4via6 stateless with IP header
compression might reduce IPv6 transition costs and benefit IPv6
network transition. It is recommended that IETF specified the
stateless solution guidance for 4via6 transition.

5. IANA Considerations

   TBD

6. Security Considerations

   TBD

7. References

7.1. Normative References

   [DS-Lite] Durand, A., "Dual-Stack Lite Broadband Deployments
              Following IPv4 Exhaustion",
              draft-ietf-softwire-dual-stack-lite-07.txt (work in
              progress), March 2011.

7.2. Informative References

   [CGN]   Perreault, S., "Common requirements for IP address sharing
              schemes", draft-ietf-behave-lsn-requirements-01.txt (work
              in progress), March 2011.

   [NAT-Redundancy]
              Xu, X., "Redundancy Requirements and Framework for
              Stateful Network Address Translators (NAT)",
              draft-xu-behave-stateful-nat-standby-06.txt (work in
              progress), October 2010.

   [NAT64-Load-Balancing]
              Zhang, D., "Considerations on NAT64 Load-Balancing",
              draft-zhang-behave-nat64-load-balancing-01.txt (work in
              progress), February 2011.

   [RFC4925] Li, X., Dawkins, S., Ward, D., and A. Durand, "Softwire
              Problem Statement", RFC 4925, July 2007.

   [RFC6180] Arkko, J., "Guidelines for Using IPv6 Transition
              Mechanisms during IPv6 Deployment", April 2011.

Authors' Addresses

   Gang Chen
   China Mobile
   53A,Xibianmennei Ave.,
   Xuanwu District,
   Beijing 100053
   China

   Email: chengang@chinamobile.com

   Hui Deng
   China Mobile
   53A,Xibianmennei Ave.
   Beijing 100053
   P.R.China

   Phone: +86-13910750201
   Email: denghui02@gmail.com