

ARMD  
Internet-Draft  
Intended status: Informational  
Expires: November 29, 2011

B. Schliesser  
Cisco Systems, Inc.  
L. Dunbar  
Huawei Technologies  
May 28, 2011

ARMD Call for Investigation  
draft-ietf-armd-call-for-investigation-00

## Abstract

This document is a call for investigation into the topic of address resolution in massive datacenters. It describes the intended work of the ARMD working group, providing both context and direction for investigating the issues outlined in the working group's charter.

## Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on November 29, 2011.

## Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Internet-Draft

ARMD Call for Investigation

May 2011

## Table of Contents

<a href="#">1.</a>	Call for Investigation . . . . .	<a href="#">3</a>
<a href="#">1.1.</a>	Context . . . . .	<a href="#">3</a>
<a href="#">1.2.</a>	Questions . . . . .	<a href="#">4</a>
<a href="#">2.</a>	Acknowledgements . . . . .	<a href="#">5</a>
<a href="#">3.</a>	IANA Considerations . . . . .	<a href="#">5</a>
<a href="#">4.</a>	Security Considerations . . . . .	<a href="#">5</a>
	Authors' Addresses . . . . .	<a href="#">5</a>

## 1. Call for Investigation

### 1.1. Context

Modern datacenters are increasingly used to support advanced services, such as multi-tenant hosting, cloud, and Internet-scale websites. Many of these datacenter facilities are being built to a much larger scale than previous generations. As a result, datacenter network infrastructure is being stressed in a number of dimensions and traditional limits to scale are being tested. One such aspect, being investigated by the ARMD working group, is the scaling of address resolution between the network (L3) and link (L2) layers of modern datacenter networks.

In many cases, datacenter operators are responsible for provisioning and running everything from routers, switches, load balancers, firewalls, servers, and storage infrastructure. Further, with the introduction of virtualization technology the capacity of these elements is increasing. Each physical device attached to the network may now represent multiple logical instances, and may expose those instances through unique MAC and/or IP addresses. For instance, with the introduction of Virtual Machine technology the operator can reduce wasted resources and achieve greater flexibility by instantiating multiple hosts on each physical server resource. Likewise virtual storage volumes, virtual routers, etc, may all exist in larger numbers.

This virtualization trend contributes to both the increased scale and management complexity of these datacenter environments. The flexibility of VM placement, including migration between different physical resources, has increased datacenter administrators' ability to instantiate VMs where the resources are, i.e. being able to relocate hosts from over-utilized servers to underutilized servers. There is a growing trend towards using resource-aware algorithms (e.g. evaluating energy, bandwidth, memory, CPU, etc) to determine placement that satisfies the processing and redundancy requirements

of each VM while using the minimal number of physical resources. Fundamentally, such datacenter management tools are responsible for making trade-offs between different dimensions of scale, which can be difficult in very large and dynamic environments, and in all cases requires a significantly stronger understanding of platform capabilities.

In this environment, IP subnets can extend throughout multiple racks and/or rows in a data center, sometimes throughout multiple sites. There are cases, such as HPC and cloud datacenters, where the number of hosts in a single subnet (on a single segment) is growing. In addition to the more recent VM environment, traditional organic

growth of physical hosts can also cause L2 segments to be extended throughout massive datacenters. Availability / redundancy requirements, subnet size requirements (versus port density), and cost issues can all contribute to the growth and/or extension of segments.

The business demands and workloads for data centers have changed greatly over last 10 years, however some fundamental networking limitations remain unexplored. Even though deployed networks generally do work, this often is because they're designed around known limitations (and redesigned around newly discovered limitations as time goes on). There are datacenter networks that work fine until something changes, such as scale, at which time they're "fixed". Often the "fix" also introduces undesired limitations. An example of this is a datacenter that moves from a flat L2 topology to a L3 core with multiple segregated L2 domains due to scale limitations, and subsequently is unable to distribute clustered servers beyond the boundaries of the "pod" in which a VLAN scope can be configured.

## [1.2.](#) Questions

The initial goal of the ARMD working group is to document the limiting factors in address resolution scale and the problems associated with exceeding those limits. Subsequently, the working group will identify operational solutions to these problems (to be promoted as Best Current Practice) or will identify gaps in existing solutions (for exploration in subsequent work). Thus the ARMD working group is asked to consider the following questions:

1. What are the scaling characteristics of modern datacenter networks (e.g. "dimensions" of scale and their normal ranges) that are relevant to address resolution?
2. What are the operational problems related to address resolution in the modern datacenter environment?
3. What is the relationship between scaling characteristics of datacenter networks (question #1) and operational problems related to address resolution (question #2)?
4. What, if any, are alternative solutions to the operational problems of address resolution at massive scale?
5. What, if any, are the "gaps" in existing solutions?

## [2.](#) Acknowledgements

The authors would like to thank Ron Bonica for his significant contributions to this text.

## [3.](#) IANA Considerations

This memo includes no request to IANA.

## [4.](#) Security Considerations

This document does not, by itself, introduce any specific security considerations. However, this document calls for further investigation into subject matter that may require significant consideration of security issues. It is anticipated that documents submitted in response to this call for investigation will include appropriate Security Considerations text.

Benson Schliesser  
Cisco Systems, Inc.

Email: bschlies@cisco.com

Linda Dunbar  
Huawei Technologies

Email: ldunbar@huawei.com