Wide area fault detection by monitoring aggregated traffic

Kohei OHTA, Glenn MANSFIELD, Nei KATO, Yoshiaki NEMOTO,

Abstract— **The Internet is now an indispensable facility for everyday networked living, communication and commerce. Its stability and reliability must be ensured by proper management. Connectivity is a fundamental aspect in networked communication, but, in the present day Internet architecture, there is no built-in facility to manage reachability. In this paper, we address the reachability management issue. We propose a mechanism for the effective usage of existing ICMP messages in conjunction with new aggregation and mapping techniques based on the Internet hierarchical addressing architecture. This mechanism provides an important piece of information, viz.** *where is the problem*. **We present some applications of the proposed mechanism.**

Keywords—**Internet management, Fault management, Connectivity, ICMP, Network location, IRR**

I. Introduction

The Internet is now an indispensable facility in our daily life. So, a network administrator must maintain connectivity, and handle faults promptly. For the end user, quality of his/her connection is an important concern, so the network administrator is expected to keep users informed about the status of services. Yet, faults at the user level service are related to many underlying components which are likely to be distributed over a wide area. Administrators have control and access over their respective local networks. They cannot, in general, obtain enough information about networks in the large. When a user is unable to access a web page, the problem could be due to some or all of a crashed Web server, a faulty DNS, erratic IP routing, or breaches in physical cabling somewhere in the Internet.

Reachability/connectivity management is one of the fundamental issues of network management. But, it has not been addressed enough yet. In spite of its importance, the network manager is generally unaware of reachability problems until someone actually encounters i.e. the route to the destination is lost.

Loss of reachability is a serious matter. There are many tools and studies to discover routes *traceroute* **and monitor reachability***Ping*. **Almost all Internet hosts have these facilities[?]. Further, to detect and diagnose reachability problems, there are many studies in the field of fault management[RFC1147][Feridun91][DISMAN].**

In general, the steps to manage connectivity problems in a networked system are 1. detect the occurrence of the problem. 2. determine the location of the problem. 3. diagnose the symptoms. Each of these are challenging issues. [Mori98][Akira99] proposed a powerful technique to detect connectivity problems. [Hood97] shows an adaptive statistical approach for determining abnormal situations. [Kätker97] proposes a fault detection and isolation technique in the data link layer. However, locating the fault remains to be difficult problem and the issue has become more complex with the spread of the Internet and ubiquitous access it provides.

Especially, at the service level, various services and network elements depend on each other. For example, the crash of a server hosting a popular web page leads to connectivity failures of many HTTP accesses. In another case, the loss of reachability to a DNS server, results in wide spread loss of application level connectivity as most URLs cannot be resolved. In both cases, there is a single fault point. In the first example it is the HTTP server. In the second example it is the DNS server. The complex relationship of various related elements obscure the location of the actual fault. [Gabi97] proposed a correlation model managing primary services, like DNS. For the WWW, there have been attempts to deal with hyper link connectivity[WebC97][Shark98]. They tried to map connectivity and manage link status at the hyper-link layer. [IPPM] at IETF also defines the metrics of connectivity[RFC2498] as for standard terms and references on connectivity issues.

In the TCP/IP protocol suite, there are provisions for some information to facilitate management of the network. Many MIBs are defined to monitor the various network elements. ICMP[RFC0792] is used for various control messages. In particular, *ICMP Destination Unreachable* **messages convey reachability failure. [Kohei97][Mori98] showed the basic usefulness of ICMP destination unreachable messages to point the location of fault in a network. We propose a technique to synthesize wide area fault information by monitoring ICMP messages. ICMP messages are a rich source of information about network reachability across a wide area network.**

In this paper, we focus on the primary step of network fault management. We propose an efficient mechanism using ICMP messages to locate the point of failure in the network based on the Internet hierarchical addressing architecture. Section II describes the potential of ICMP, and section III discusses the issues involved in ICMP handling. Section IV describes the availability of network configuration information. In section V, we explain the proposed reachability analysis technique using techniques of message aggregation and address information enhancement and, evaluate the algorithm in section VI. In section VII, we show an application of a visual reachability management tool, and conclude this paper in section VIII.

II. Reachability management with ICMP

In the Internet connectivity is not guaranteed, and generally, there is no information about network reachability. So, users can't know of problems until they actually encounter one. In the Internet, there is no special mechanism for communicating reachability information except for control messages like ICMP (Internet Control Message Protocol)[RFC0792][RFC2463].

A standard implementation of ICMP can be very useful for general reachability management. ICMP has various types of messages which are - error notification, routing optimization, security alarms, and so on. And, there is a type to notify the reachability failures, viz. Destination UnReachable (ICMP-DUR) messages. These provide information about source/destination IP address, service port number, ... etc. These are very useful pieces of information that provide hints about the location of the actual fault.

K. OHTA and Glenn MANSFIELD are with the Cyber Solutions Inc., Nei KATO and Yoshiaki NEMOTO are with the Tohoku university, Sendai, JAPAN. E-mail: kohei@cysols.com .

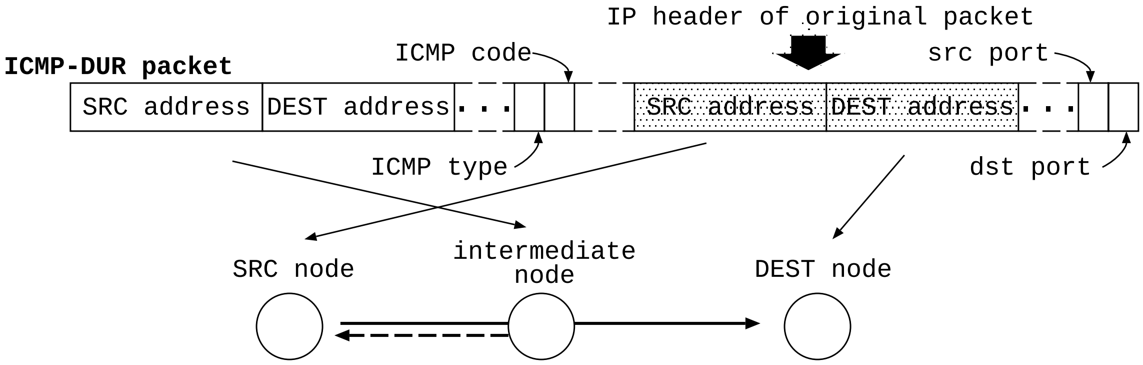

Fig. 1. Information derived from ICMP-DUR message

A. ICMP information

ICMP-DUR packet is generated by an *intermediate node* on the path from source to destination. The *intermediate node* is a router or a host, which can't forward the packet to the next hop or to the destined port(application). The intermediate node embeds the header of the IP datagram, which couldn't be delivered, as the ICMP payload. Fig.1 shows the information and the interpretation of the information contained in a ICMP-DUR message. ICMP messages are generated for every undelivered IP datagram.

A.1 ICMP header

Each ICMP-DUR message, which is defined by ICMP type 3, has its own IP header, whose source address indicates the last point where original IP packet had reached. Four unreachability subtypes, the *ICMP code*, are defined in [**RFC0792**].

Net unreachable
The destination network was unreachable. This type may indicate routing failure or illegal address usage, e.g. leak of private IP addresses into the backbone network.

Host unreachable
The destination network was reachable, but the host in the network was not found by the last router. This type may indicate a failure in stub network on the path.

Protocol unreachable
The destination node was reachable but it did not support the protocol specified in the IP-header. This type may indicate the failure inside a destination node and transport layer.

Port unreachable
The network, host, and the protocol are OK, but the requested port/service is not available. This type may indicate a failure at the application on the destination node.

The different types of unreachable messages indicate different types of failures. *Net unreachable* may be caused by routing failure, *Host/Protocol/Port unreachable* may be caused by node/server down or mis-configuration. Further, it may indicate an attempted illegal access, e.g. a potential intruder is scanning for vulnerabilities.

A.2 IP header as ICMP-DUR payload

ICMP-DUR packets also contain the IP header of original packet(the undelivered IP datagarm) as payload. It has the original source, intended destination, source port and destination port. This information is useful to know the location of users and to estimate the area of impact of the problem.

III. Problems and Potentials of connectivity management with ICMP

ICMP is potentially useful to manage the reachability of the Internet. However, the absence of an integrated and common mechanism to handle ICMP messages renders this rich source of information useless.

Point of monitoring: ICMP messages are essentially independent of each other. They are usually exchanged between end points. Thus the monitoring point is important aspect. A backbone network or the up-link of a network will be a suitable point of monitoring. That will give a good picture of status and performance of the communication as the aggregated traffic of all clients and all requested services from the connected network transit through this point.

Message aggregation: In aggregated traffic, there are many ICMP messages originating at various points of the Internet. Sometimes, there is a burst of ICMP. For example, A fault in a popular web server may generate a large amount of ICMP messages in a short time. To simplify the problem, we aggregate similar ICMP messages over a short duration.

Topological analysis: ICMP message has spatial information, in the form of raw IP addresses. By topological analysis of the variation, we can locate the fault and the region of impact. For example, every packet through a faulty link will generate an ICMP message, which will be from some hosts to various destinations. From the perspective of the global Internet, the summary of the information is more useful. We extracted AS and network number from raw IP address using IRR, WHOIS, and DNS. If the ingress router of an AS is dead, all services in the AS, and all networks beyond the AS will be unreachable. We can diagnose such cases by comparing the ICMP messages with network topology information. This is e useful not only to network administrators, but also to end-users.

Fig.2 shows the statistics of the reachability related problems. The observation was carried out on a regional backbone network, which has a FDDI ring with a throughput of 30 ∼ 40 Mbps. A large number of ICMP-DUR messages are observed almost all the time. They are independent by source-destination-service set, so exhaustive processing is needed for every message, and the result is confusing and meaningless.

- It is necessary to handle the ICMP messages effectively.

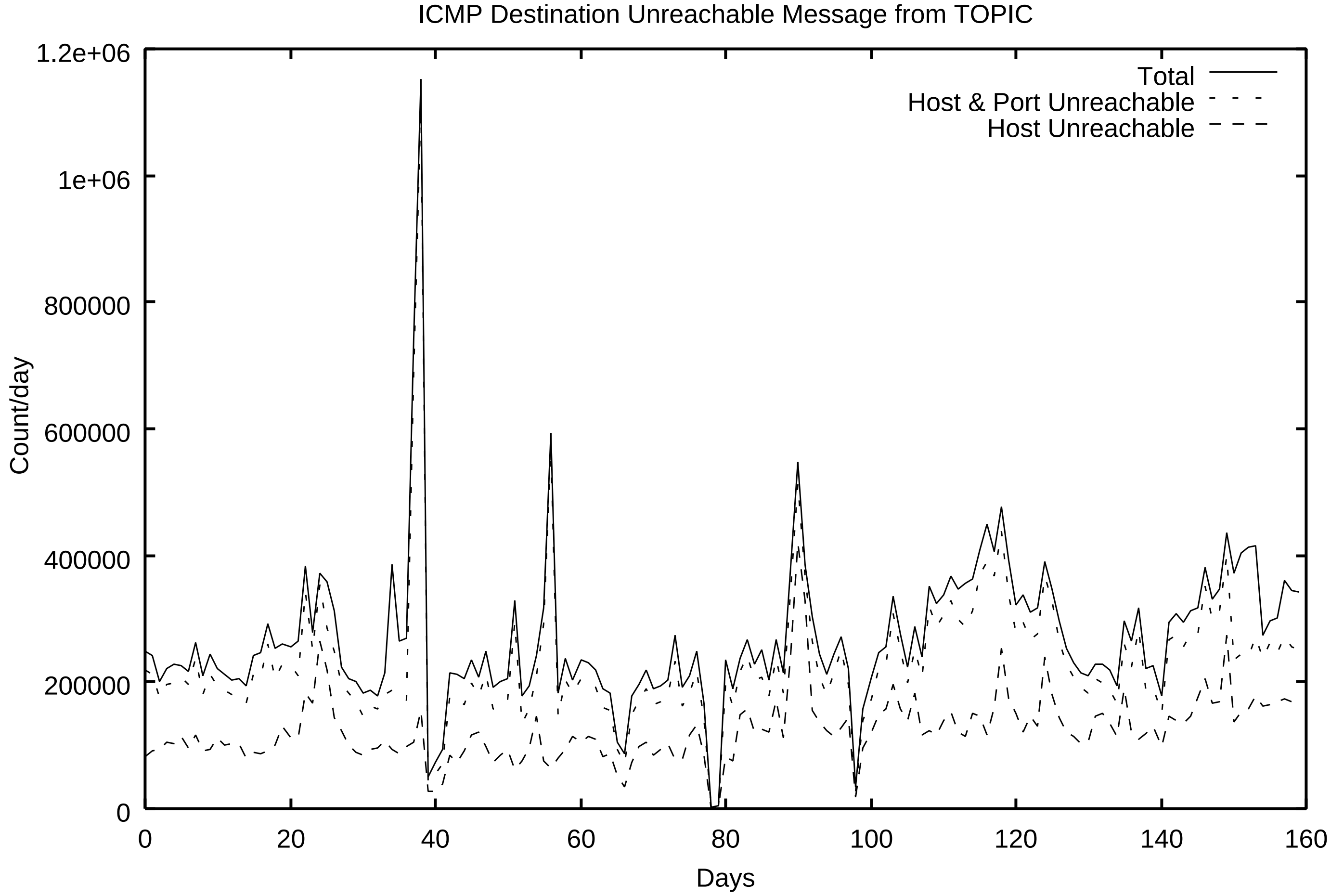

Fig. 2. Number of unreachable queries

Network elements and services are likely to be related to each other. For example, when a primary DNS service is down due to the fault of the server daemon, *port unreachable* messages are generated by the server for all requests irrespective of the service being sought.

In another case, when a ISP has some faults due to a local reason, many accesses to the services hosted by the ISP will be unreachable. At that time, ICMP messages may be generated for all accesses to all the services hosted by the ISP.

As mentioned above, one fault may interact with various elements distributed in the network. Thus, unreachable messages, generated at various point of network, may be caused by a single fault. To diagnose such faults and symptoms, the location information of the related element, which is most important, is hidden in each ICMP message.

In the case of the former example, all ICMP-DUR messages related to the faulty DNS should be handled as a single issue to avoid exhaustive examination. And, in case of the latter example, all destination IP addresses should be treated as belonging to some higher level location indicator like AS number. We should see some indications information like AS-unreachable.

- It is useful to aggregate the littered ICMP messages in the context of the location from both points of management efficiency and load reduction of manager.

IV. Pointing the location of fault in the Internet

In this section, we describe a model of location information, which is follows the hierarchy of the Internet addressing architecture. We define a notation to describe the reachability issue.

A. Availability of Internet location information

At network and transport layer of TCP/IP, the point of a network is indicated by some numbers, e.g. port number, IP address, network address, and AS number. Internet addressing architecture is hierarchical. So, to discover the location of reachability problems, our proposed approach handles location information in a hierarchical fashion.

Fig.3 shows a typical hierarchical structure of Internet. The location of router $1 \sim 4$ are in AS0, and host $1 \sim 2$ are located under router 2. Also, the web service is on host 1.

B. Notation of the network location and the reachability

The Internet location can be represented as combination of one or more numbers mentioned above. For example, the location including application information in the Internet can be represented by an IP address and port number as follows:

$$l = (IP\ address, Port\ number) \qquad (1)$$

The order of elements are the same as in the addressing hierarchy. $l = (192.161.0.1,\ 52)$ means the DNS server at host

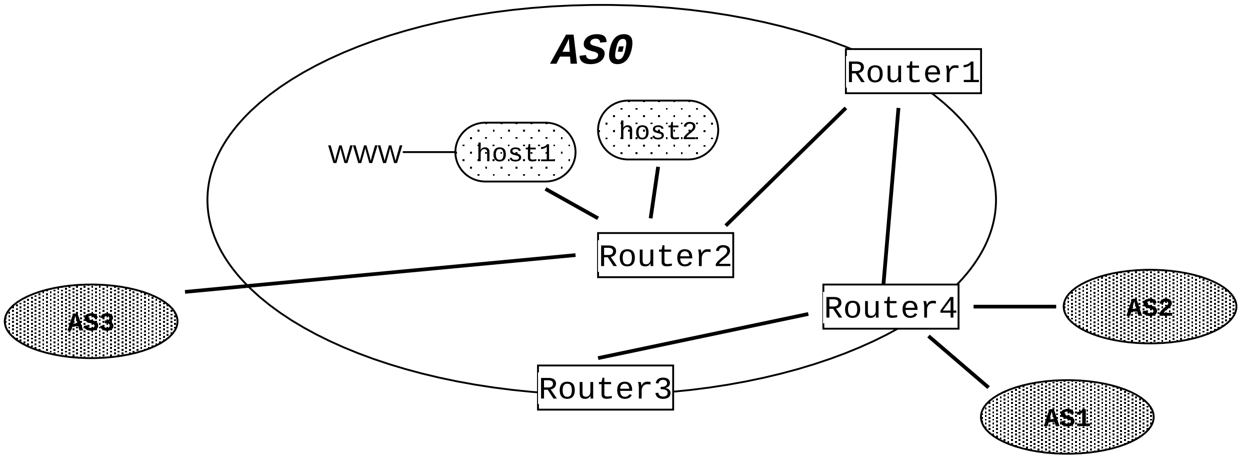

Fig. 3. Internet addressing hierarchy

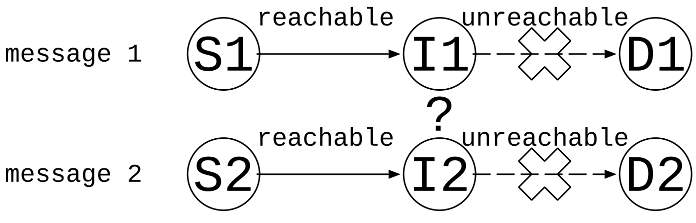

Fig. 4. Message correlation and aggregation

192.161.0.1. We define a boolean value r, which represents the unidirectional reachability from l_1 to l_2 as follow:

$$r(l_1, l_2) = true/false$$

Information derived from a ICMP-DUR message includes three pieces of location information represented by three IP addresses and two port numbers as follows:

$$m = (S,\ D,\ I,\ p_s,\ p_d)$$

where, S, D, and I are IP addresses. As illustrated in fig.1, S and D respectively represent *source* and *destination* node of original IP packet, which was not delivered due to connectivity failure. I is IP address of intermediate node, which is same as the ICMP sender. And p_s and p_d are numbers of the source and destination ports indicating the type of applications.

In other words, a message m means that IP packet generated at port p_s of node S, which was intended to send it to port p_d of node D, was returned from intermediate node I because of the something wrong at forward area of node I. They are source node, destination node, and intermediate node, represented as $l_s = (S, p_s)$, $l_d = (D, p_d)$, and $l_i = (I, unknown)$ respectively. So, ICMP messages, informing $r(S, D) = false$, are carrying two more reachability information, which can be represented as follows:

$$r(l_s, l_i) = true,\ and\ , r(l_i, l_d) = false \qquad (2)$$

V. Message Aggregation and Address Enhancement

In this section, we propose a fault location mechanism with efficient ICMP handling schemes for the Wide-Area Internet reachability management, which is based on efficient message aggregation and address enhancement. We essentially use the Internet address hierarchy to obtain aggregation.

Message aggregation will aggregate many independent ICMP messages along with the addressing hierarchy. That aids in narrowing down the location of the point of failure and reduces the load of the NMS.

Address enhancement is essentially a value-addition process to the raw detected messages, using IRR information and enabling correlation of ICMP-DURs with higher layer context. That will provide further sophisticated information to the manager for Wide-Area Internet management, and reduce the load for examination significantly.

A. Message aggregation

The reachability to various points of the network may widely vary, and it is difficult to correlate these directly. In a ICMP-DUR message, two reachability information are included (equation (2)). One is *true* (reachable), the other is false (unreachable). This *true* reachability coupled with *false* reachability is useful to narrow down the point of failure and estimate the region of the impact.

Fig.4 shows a correlation and aggregation mechanism. $S1, S2, I1, I2, D1$ and $D2$ are location information as equation (1). "message 1" and "message 2" are reachability information derived from ICMP-DUR messages, which indicate $r_1(S1, D1) = false$ and $r_2(S2, D2) = false$ respectively. r_1 and r_2 are also including two more reachability information along with equation (2) as follow.

$$\begin{cases} r_1\ include\ r_{11}(S1, I1) = true,\ and\ r_{12}(I1, D1) = false \\ r_2\ include\ r_{21}(S2, I2) = true,\ and\ r_{22}(I2, D2) = false \end{cases}$$

From these optional reachability information, two more location information could be obtained, $I1$ and $I2$. In case

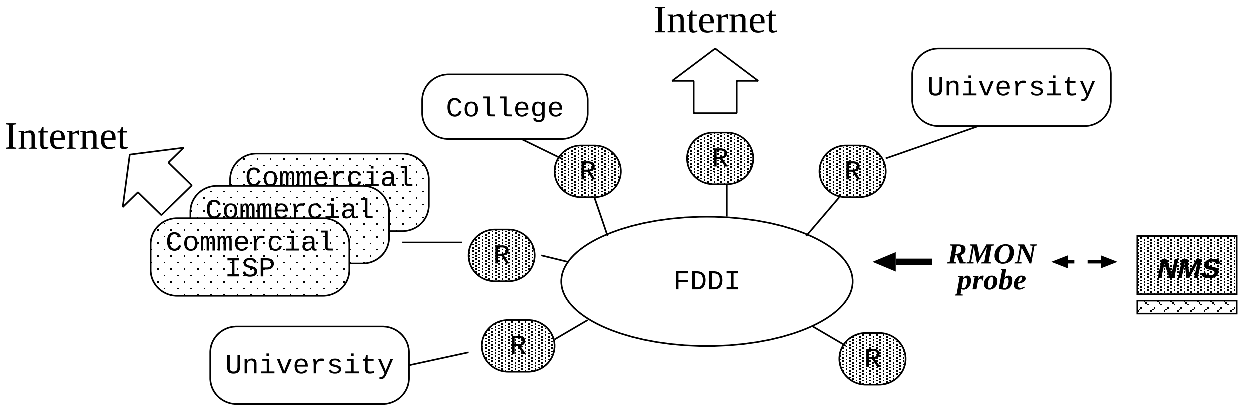

Fig. 5. Network environment for analysis and experiment

TABLE I
Number of total ICMP-DUR messages

day	0	1	2	3	4	5	6
msgs	6592352	36497789	20786164	51156746	7542619	22278982	63704058

$I1$ and $I2$ indicate the same point of network, the two reachability information, r_1 and r_2, could be correlated and aggregated. Then, $S1$ and $S2$ forms a group of clients influenced by this fault, while $D1$ and $D2$ forms a group of services which don't work. Faulty reachability r_{12}, r_{22} and their related location information can potentially indicate the region of degraded services.

For example with fig.3, in case of the down of router 2, all reachability to the services hosted by host 1 and host 2 from anywhere are $false$, but reachability to router 2 may be $true$. So, collection of such reachability information make obvious that the router 2 is a location of bottleneck, and if more ICMP-DUR are monitored, it could be known that reachability to any services in AS3 via router 2 are also influenced.

B. Address Enhancement

Location defined in equation (1) is enough to locate at IP layer. However, additional information obtained *IRR (Internet Routing Registry)*[IRR][RFC2622] is useful to correlate and aggregate the messages in Wide-Area Internet environment. Using IRR, the upper-level address hierarchy i.e. AS level, becomes available. AS is a network unit, which has same administrative policy, typically an ISP. Thus, AS is a reasonable unit for fault management in the global Internet. Enhanced location message can be described as follows:

$$L = (AS\ number,\ IP\ address,\ Port\ number) \qquad (3)$$

Here, we can aggregate ICMP messages into AS level aggregates. Using AS information, further correlation and aggregation can be carried out using location information L defined in equation (3), instead of l (equation (1)). For example with Fig.3, it is possible to say that the region affected by an accident of router 4 includes AS1 and AS2 simply, instead of complex IP address wise description.

VI. Evaluation on Operational Network Environment

All evaluation for this proposed system is carried on operational network, which is TOPIC (Tohoku Open Internet Community).

A. Environment for analysis and experiment

The TOPIC network consists of one FDDI loop and about 25 connected academic organizations, e.g. universities, museums, colleges, regional IXs, and so on. A probe to capture ICMP packets is attached to the back bone FDDI loop (fig.5). All evaluation was carried out on the ICMP-DUR messages collected by the probe over a week. Total number of ICMP-DUR monitored is $208,558,710$. Table I shows a transition of total number of ICMP messages a day.

B. Result of Message Aggregation and Address Enhancement

Fig.6 shows the number of groups of aggregated ICMP messages of on a daily basis. The upper line represents the number of unique source addresses, the middle line represents the number of unique destination addresses, and the lower line represents the number of unique intermediate addresses. From a comparison with the initial number of ICMP-DUR messages listed in Table I, the number of messages, which needed to be handled, was reduced significantly by the aggregation process. The reduction rate is around $99\% \sim 99.9\%$. Also, Fig.6, points at the particular site(location) which is likely to have fault.

Fig.7 shows a result of AS level aggregation. The upper, middle and lower lines represents the number of unique source, destination, and intermediate ASs respectively. The figures are much smaller than those in Fig.6. Several reachability problems are caused by a small set of ASs. That means there are weak points in some ASes. The manager can use such information for routing policy and filtering decision in global Internet.

VII. Visualization and forecasting of Internet connectivity

The experiments in the previous section show that relationship among the unreachabilities indicate a potential point of failure. So we tried to visualize the weak points to obtain a better insight from the reachability problems.

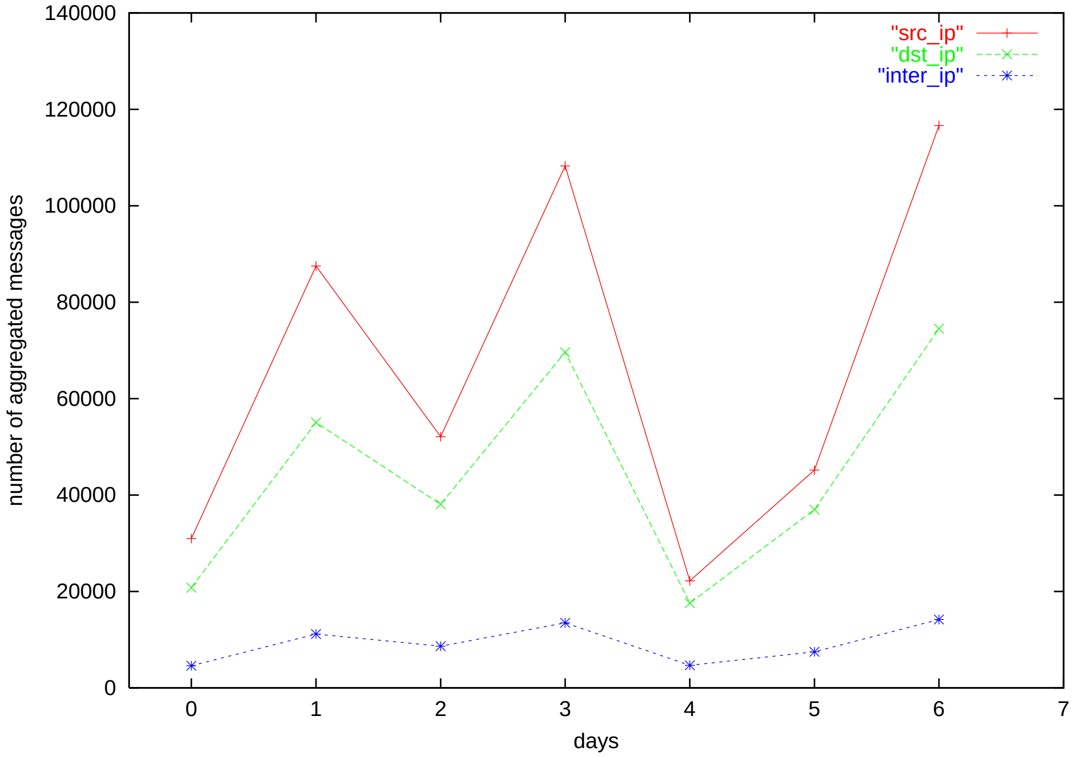

Fig. 6. Message aggregation by IP address

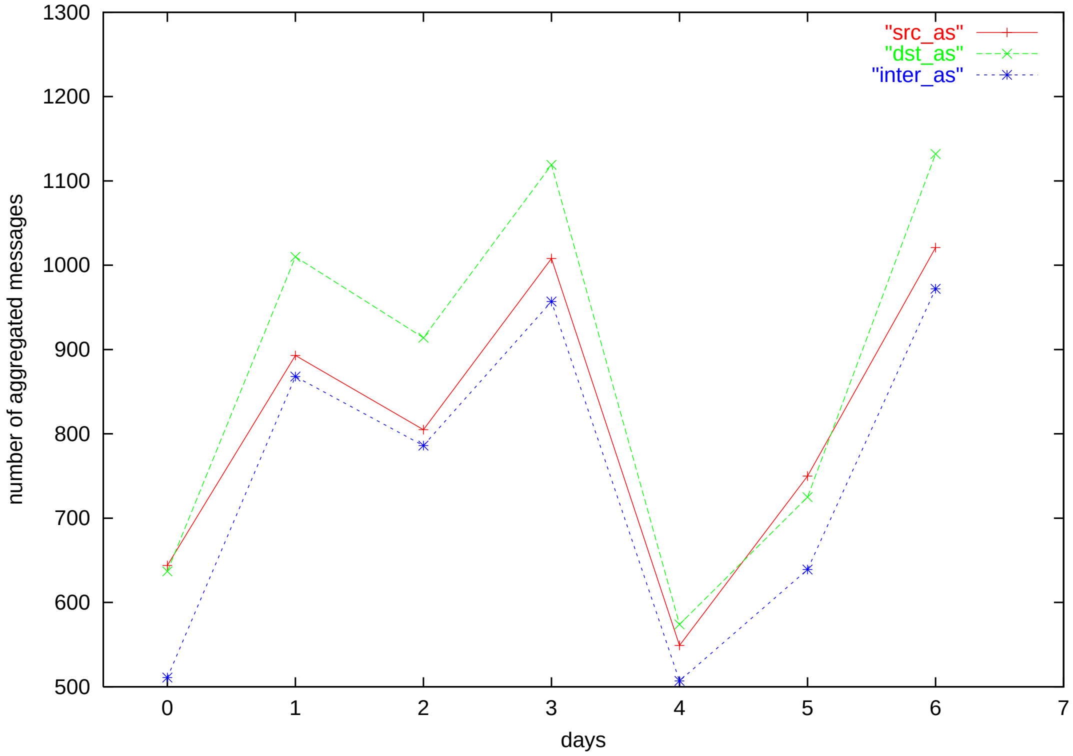

Fig. 7. Message aggregation by enhanced address (AS number)

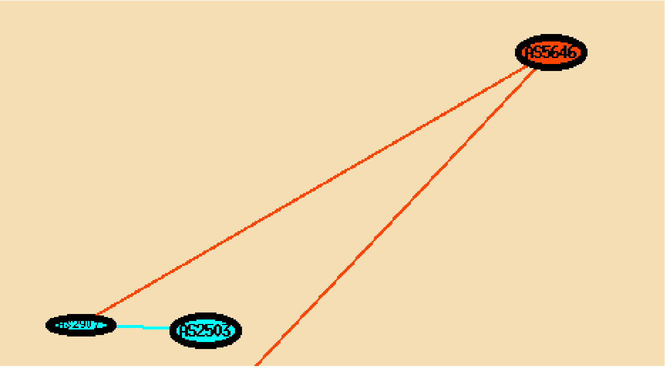

Fig. 8. Reachability visualization

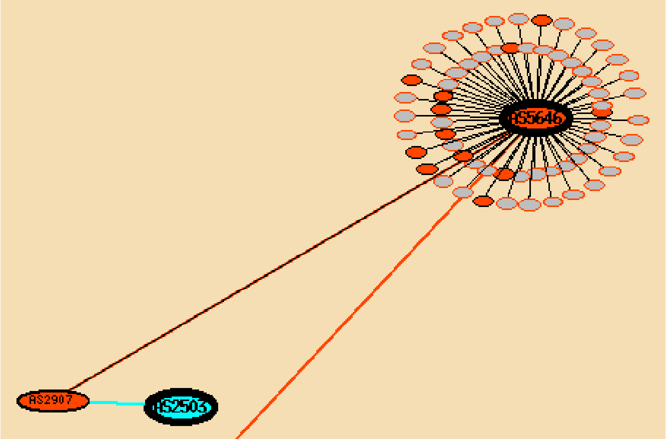

Fig. 9. Predicted unreachable region

A. Visualization with configuration information

IRR is a good information source for network configuration visualization. It gives AS level connectivity, so enhanced location information can be mapped to it. Fig.8 shows a sample result of the visualization, which shows a result of AS level aggregation of multiple ICMP-DUR messages. It shows $r(AS2503, AS2907) = true$ (blue line), $r(AS2907, AS5646) = false$ (red line), then we can know that there is bottleneck of reachability around AS5646.

B. Forecasting the connectivity problems

By using the connectivity information from the IRR, a manager can figure out the potentially affected sites. Fig.9 shows a sample of visualization of potentially unreachable sites. This information is useful to forecast the problem of the accesses to such sites. In this case, given ICMP notified $r(AS2907, AS5646) = false$, sites connected to the related location are likely to be unreachable (red circles). That enables proactive reachability management.

VIII. Conclusion

In this paper, we proposed a mechanism for reachability management of the Internet, which uses existing ICMP messages in the network traffic. It gleans for and provides a basic piece of information about Wide-Area Internet reachability. For effective handling of the enormous volume of ICMP messages, and for providing more value added information to the manager, we proposed message aggregation and address enhancement technique. We evaluated the proposal with packet data collected from an operational network. We showed a sample application which offers visualization of reachability problems and enables proactive reachability management.

References

[RFC1812] F. Baker, "Requirements for IP Version 4 Routers", RFC 1812, June 1995.

[RFC1147] R. Stine(ed.), FYI on a Network Management Tool Catalog, "Tools for Monitoring and Debugging TCP/IP Internets and Interconnected Devices, RFC 1147, 1990.

[Feridun91] M. Feridun, "Diagnosis of Connectivity Problems in the Internet," In Proc. of IFIP ISINM '91, pp.691-701, 1991.

[DISMAN] IETF, "Distributed Management Working Group", work in progress, http://www.ietf.org/html.charters/disman-charter.html

[Hood97] C.S.Hood, C.Ji, "Automated Proactive Anomaly Detection", proceedings of IM'97, pp.688-699, May 1997

[Kätker97] S.Kätker, M.Paterok, "Fault Isolation and Event Correlation for Integrated Fault Management", proceedings of IM'97, pp.583-596, May 1997.

[Gabi97] G. Dreo Rodosek, "Determining the Availability of Distributed Applications", Proceedings of the Fifth IFIP/IEEE International Symposium on Integrated Network Management(1997-05).

[IPPM] V. Paxson, G. Almes, J.Mahdavi, M. Mathis.,"ramework for IP Performance Metrics.", May 1998.

[RFC2498] J. Mahdavi, V. Paxson.,"IPPM Metrics for Measuring Connectivity", January 1999.

[WebMap94] Peter Doemel, "WebMap - A Graphical Hypertext Navigation Tool", 2nd International WWW conference, Chicago, IL, USA, 1994.

[WebC97] Yoelle S. Maarek, Michael Jacovi, Menachem Shtalhaim, Sigalit Ur, Dror Zernik, Israel Z. Ben Shaul, "WebCutter: A System for Dynamic and Tailorable Site Mapping", 6th International WWW conference, Santa Clara, CA, USA, 1997.

[Shark98] Michael Hersovici, Michael Jacovi, Yoelle S. Maarek, Dan Pelleg, Menachem Shtalhaim, Sigalit Ur, "The shark-search algorithm - An application: tailored Web site mapping", 7th International WWW conference, Brisbane, Australia, 1998.

[RMON WG] S. Waldbusser. "Remote Network Monitoring Management Information Base.", RFC 1757, February 1995.

[RFC0792] J.Postel, "Internet Control Message Protocol", RFC0792, 1981.

[RFC2463] A. Conta, S. Deering.,"Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification.", December 1998.

[Kohei97] Kohei Ohta, Takumi Mori, Nei Kato, Hideaki Sone, Glenn Mansfield and Yoshiaki Nemoto, "Divide and Conquer Technique for Network Fault Management", Proceedings of the Fifth IFIP/IEEE International Symposium on Integrated Network Management(1997-05).

[Mori98] Takumi MORI, Kohei OHTA, Nei KATO, Hideaki SONE, Glenn MANSFIELD, Yoshiaki NEMOTO "The Dynamic Symptom Isolation Algorithm for Network Fault Management and Its Evaluation", IEICE Trans.Commun., Vol.E81-B, No.12, pp.2471-2480, (1998-12).

[Akira99] Akira KANAMARRU, Kohei OHTA, Nei KATO, Glenn MANSFIELD, Yoshiaki NEMOTO, "Simple Aggregation Technique for Fault Detection Based on Packet Monitoring", Proceedings of Symposium on Performance Evaluation of Computer and Telecommunication Systems, July, 11-15, 1999.

[IRR] Tony Bates et al, "RIPE-181, Representation of IP Routing Policies in a Routing Registry", http://www.ra.net/RADB.tools.doc/ripe-181.html

[RFC2622] C. Alaettinoglu et al, "Routing Policy Specification Language (RPSL)", RFC2622, June 1999.