Modeling the Internet Routing Scalability: From Qualitative Description to Quantitative Evaluation

[1,2]Wei Zhang, [2,3,4] Jun Bi, [2,3,4] Jianping Wu
[1] *National Academy of Defense Information, Wuhan, China*
[2] *Department of Computer Science and Technology, Tsinghua University, Beijing, China*
[3]*The Network Research Center, Tsinghua University, Beijing, China*
[4] *Tsinghua National Laboratory for Information Science and Technology, Beijing, China*
zw@netarchlab.tsinghua.edu.cn, junbi@tsinghua.edu.cn, jianping@cernet.edu.cn

Abstract

The Internet has been growing rapidly during the past few decades. Its routing system may encounter scalability problem as the size of Internet keeps growing. The Internet Research Task Force (IRTF) highlighted this issue lately and inspired many proposals that attempt to address this problem. However, those proposals do not reach a consensus on a fruitful solution. Instead of presenting another new proposal, this paper deals with specification, analysis and evaluation of the routing scaling of the Internet. The qualitative analysis in this paper reveals gains and losses of different techniques that aim to improve the scalability of the Internet routing system. In particular, the analysis focuses on techniques of route aggregation, forwarding information base (FIB) compression and route cache mechanism, which correspond to tradeoffs between routing scaling and engineering considerations on the functioning, overhead and performance of the routing system respectively. Based on understanding of these tradeoffs, we apply the Analytic Hierarchy Process (AHP) to develop a quantitative evaluation framework that can evaluate alternative routing schemes. This modeling tool can systematically capture various evaluation criteria, where each routing scheme leads to quantified gains and losses under given evaluation criteria.

Key Words: *Routing, Modeling, Analysis, Internet*

1. Introduction

The Internet has become the most influential and largest information infrastructure that connects various computer networks all over the world. Millions of Internet users' networks are connected through the Internet Protocol (IP). The size of the IP routing system keeps growing inexorably. By the year of 2011, the IPv4 address pool is almost draining. According to the measurement on the Border Gateway Protocol (BGP) [1], more than half of the assigned IPv4 addresses have been advertised in the inter-domain routing system. In addition to the growth of the Internet, some ISPs and end users advertise more and more de-aggregated fragmental IP prefixes in BGP, for their legitimate needs of having portable IP addresses, multihoming and Traffic Engineering (TE). BGP routers will have to drastically improve their ability to scale to large numbers of prefixes. Observation shows that the size of the BGP route table increases at a high speed [2]. It is well believed that the momentum will not change in the near future.

The increase of routing information subsequently causes the inflation of the forwarding table (namely, FIB in a router). The FIB size has exceeded 380 000 entries at most BGP routers. On forwarding a packet, a router has to process longest prefix match (LPM) lookup over all of the FIB entries. On the other hand, the core Internet routers are facing challenges brought by the ever-increasing transit bandwidth. The Autonomous System border routers (ASBRs) may have to forward packets on each port at a line rate of 40 Gbps, which means a LPM lookup has to be accomplished within a few nanoseconds.

This requirement can only be met by applying high-performance hardware, such as ternary content addressable memory (TCAM). However, this kind of special-purpose memory is expensive and energy hungry. As the Internet scales up inexorably, more and more routing information has to be stored and maintained in the routing system, and the cost of routing will increase accordingly. ISPs and manufacturers have a strong motive to adopt techniques that can improve the scalability of the routing system.

doi:10.4156/jcit.vol7.issue8.38

The routing scalability problem of the Internet, specifically, the scalability of the inter-domain routing, has been highlighted by IRTF RRG in RFC 4984 [3]. This concern has inspired some new proposals in recent years. As a guideline for designing new routing schemes, RFC 6227 [4] summarizes the design goals for scalable Internet routing and prioritizes the basic requirements of the target architecture. However, both industry and academic communities have not reached consensus on how to meet those requirements. Based on different understandings, people propose various techniques to address this problem. In section 2, we briefly survey those proposals. On modeling the Internet routing scalability, we focus on the following questions: what is the essence of the routing scalability problem, what are critical elements (associated with typical techniques) to address this problem and how to evaluate new routing schemes in terms of their routing scalabilities?

In this paper, we start from modeling the routing scale in the control plane from a perspective of the information theory. Our model of the routing information entropy reveals the basic elements that have different impacts on the routing scale. Then we analyze the FIB compression techniques which is mostly relevant to the data plane. At last we discuss the route caching mechanism with respect to the interacting between the control plane and the data plane. These qualitative analyses help us to depict the scalability of the Internet in terms of considerations on routing functioning, overhead and performance respectively. Different routing schemes may optionally adopt relevant techniques to improve the routing scalability of the Internet. On evaluating the gains and losses of various routing schemes, we develop a generalized evaluation framework through applying the Analytic Hierarchy Process (AHP). This framework leverages systematic, quantitative evaluation on alternative routing schemes under various evaluation criteria.

The rest of the paper proceeds as follows: section 2 surveys related works; section 3 qualitatively analyzes the Internet routing scalability with respect to the considerations on routing functioning, overhead and performance; section 4 applies the AHP to establish a quantitative framework on evaluating alternative routing schemes. In the end, we draw a conclusion in section 5.

2. RELATED WORKS

In the field of distributed computing, to address the routing scalability problem, abundance of scalable distributed routing algorithms has been well studied. Peleg [5] surveys compact routing schemes that support routing on large sized networks with limited route entries. A price for reducing the routing table size is the increase in the average forwarding path length. Theoretic analysis has been conducted carefully on those schemes. Bounds have been derived to evaluate the maximum increase in the forwarding path length for a given routing table size. These analyses mainly focus on the tradeoff between the lower bound of maximum path stretch and the routing table size.

However, compact routing algorithms can not be directly applied in the Internet for the following reasons. Firstly, most compact routing algorithms are based on static topology. Both labeled and name-independent routing schemes cannot handle dynamic topology very well. Compact routing algorithms take advantages of the features of the network topology. That’s why compact routing algorithms do well with special topology, e.g. regular topology or trees. To implement compact routing on dynamic topology can be very costly. Secondly, inter-domain routing is compliant with routing policies. Compact routing algorithms can barely support flexible routing policies in a distributed system. Thirdly, compact routing may incur more or less stretch of the forwarding path. In the case of the Internet topology, although the average stretch is not much, the worst case is not acceptable. Lastly, compact routing schemes are mostly clean slate proposals which may not support incremental deployment. Those reasons prevent compact routing techniques from being adopted in the Internet routing system. A practical Internet routing scheme should support dynamic topology updates as well as complex inter-domain routing policies. And it would be acceptable with low stretch bound and admirable if it is compatible with the current Internet routing implementation. In general, the compact routing schemes are mostly less viable in the Internet routing scenario.

There have been a dozen of more viable proposals attempt to curb the inflation of the routing/forwarding table (RIB/FIB) in a more pragmatic way. For example, LISP [6] and other core/edge split routing schemes [7, 8] suggest separating the transit core of the Internet from the edge networks. The addresses used in the core can be aggregated. The edge networks have to map the local source/destination addresses to the access point and egress point in the core before establishing transit tunnels between the source and destination. Some proposals [9, 10] suggest applying Provider

Aggregateable (PA) IP addresses, which can facilitate route aggregation in a routing hierarchy. But these proposals require topology dependant addressing and face frustrating renaming problem consequently. Virtual Aggregation (VA) [11] proposes a mechanism to virtually aggregate prefixes, which may reduce the FIB size if being applied Internet-wide. Some [12, 13] propose FIB aggregation algorithms to reduce the FIB size. Others [14, 15] propose route caching mechanism that caches a subset of the RIB into the FIB. There are many other clean slate proposals [16, 17, 18, 19] that also address the routing scalability problem with *clean slate* routing architectures. However, aside from striving on fruitful engineering efforts, we also need a generalized analytical model to facilitate our evaluation on new proposals.

In this paper, we are going to analyze the gains and losses of relevant techniques that attempt to improve the routing scalability of the Internet. Our analyses mainly deal with route aggregation, FIB compression and route cache mechanisms. Based on our theoretical analyses, it would be convenient to establish a quantitative framework for evaluating new routing schemes.

3. QUALITATIVE ANALYSES

To improve the routing scalability of the Internet, there are three typical techniques: route aggregation, FIB compression and route caching mechanism.

3.1. Route aggregation

Route aggregation is relevant to the routing information in the control plane. In this subsection, we analyze the essence of route aggregation from a perspective of information theory.

On analyzing the routing information of the Internet, we simply model the distributed routing scheme as follows. Given a connected graph *G(V, E)*, where *V* is the set of vertices and *E* represents the set of edges, each vertex *v* in *V* may exchange routing information of numerous routing objects in *G* (e.g. agglomeration of destinations that share the same route) with each other. Each *v* has to store routing information for every reachable routing object. To measure the routing information that should have been stored in an individual route table, we apply Shannon's entropy defined in the information theory.

According to Shannon's theory [20], the amount of information can be quantified with the expectation of uncertainties. In our above Internet routing model, the probability of forwarding a packet to a certain outlet interface is uniformly $p=1/k$, where k is the number of neighbors. The packets heading for the same routing object will share the same routing information. In this context, for each routing object the vertex has to store an entry to eliminate the uncertainty of forwarding. According to Shannon, the information entropy contained in each route entry will be:

$$\sum_{1}^{k} -\frac{1}{k}\log(1/k) = \log(k) \quad bits$$

Assuming that the total addressing space is *M*, and the aggregation ratio is *a*, which represents the routing granularity, the number of routing objects will be *M/a*. The sum of routing information entropy for each vertex can be formulated as

$$H = \frac{M}{a}\log(k) \qquad (1)$$

In this formulation, we can derive the following essential elements (namely *M*, *a*, and *k*) that are relevant to the routing scale.

Firstly the addressing space *M* determines the fundamental routing scale of the network. For instance, in IPv4, the addressing space may contain 2^{32} individual IP addresses. In the case of IPv6, *M* extends to 2^{128} maximally. The size of the network, in terms of the addressing space, is the base of the routing space. But sometimes the routing space can be extended to 2 or more dimensions that we will explain in the next paragraph.

In practice, both IPv4 and IPv6 routing are based on the destination address. Imaging a new routing scheme that refers to both the source and destination addresses when forwarding a packet, there will be two orthogonal dimensions in the routing space. Consequently, the routing scale may increase exponentially.

Theoretically, when the routing scheme deals with high dimension routing objects, the measurement of the routing space will be

$$M = \prod^{i} m_i ,$$

where m_i denotes the size of the addressing space on a given dimension.

We can modify formulation (1) as following:

$$H = \frac{\prod^{i} m_i}{a} \log(k) \qquad where \quad i \geq 1 \qquad (2)$$

In fact, Access Control List (ACL) table may define more than one dimension for the purpose of access control. Another example is OpenFlow [21], which defines a dozen of fields for the sake of fine-grained flow control. Nevertheless, in practice, fundamental Internet routing is defined only on one dimension (i.e. the destination), and its routing scale is determined by the addressing space of routable objects.

Secondly the aggregation factor *a* is another decisive element to the scalability of a routing scheme. In the routing space *M*, a group of routing objects that share the same routing information may be agglomerated together. The aggregation factor can be roughly estimated in each routing scheme. It is worth noting that the minimum aggregation unit determines the granularity of routing policies. For instance, in the case of IPv4, BGP conventionally refrains to advertise prefixes longer than /24. Therefore, the IPv4 BGP can barely impose routing policy on prefixes longer than /24. There is a tradeoff between the routing scale and the routing granularity. In other words, the routing granularity provides a knob to tune the routing scalability.

Thirdly, parameter *k* has something to do with the topology, where *k* is the connection degree of a vertex. According to the Internet topology measurement, the distribution of *k* obeys a power law. A few vertices may have denser connections than others. This uneven distribution of degree suggests that when the Internet encounters routing scalability problem, different ISPs may have different pressure. The ISPs at the transit core may have a strong incentive to be separated from the edge networks because normally they have a higher degree of connection. However, *k* imposes a logarithmic impact on the routing entropy, which is relatively less influential.

The routing information entropy model depicts the essential elements that are relevant to the Internet routing scale. This model reveals the following facts:

- The scale of the routing objects is the most decisive element which is normally determined by the size of the network and mostly a rigid demand of a routing system. It is worth noting that when the routing functions extend to multiple dimensions, the routing space may increase exponentially.
- The granularity of the routing objects plays critical roles on tuning the scalability of the routing system. As the network scales up, normally the granularity of routing should become coarser accordingly. This is a common trade off in practice.
- The degree of a vertex is relatively less influential than other elements in dealing with the routing scalability. But the densely connected vertices may encounter routing scalability problem severer than the others.

The routing information entropy is closely related to the size of the routing table which is in the control plane of a routing system. To improve the routing scalability by escalating the routing granularity is to address the problem from the stem. However, route aggregation may partly sacrifice the function of the routing system. In the data plane, we have alternative techniques to curb the size of the forwarding table (FIB).

3.2. FIB compression and route update overhead

In contrast to RIB, a FIB is optimized for fast forwarding hardware. Earlier implementations cached only a subset of the RIB entries into the FIB. However, modern routers used for accessing the entire Internet moved to having FIBs in one-to-one correspondence with the RIB (in terms of the best routes selected from routing protocols in this context). In this case, the FIB size equals to the RIB size. Conceptually a RIB is in the control plane while FIB has something to do with the data plane. RIBs are optimized for efficient updating by routing protocols and other control plane methods, and contain the full set of routes learned by the router. A router may distribute the RIB to a FIB from time to time. Typically the FIB size is determined by the RIB size (the routing scale).

FIBs may be implemented with fast hardware lookup mechanisms, such as ternary content addressable memory (TCAM). TCAM, however, is quite expensive and energy eager. To keep the FIB size scalable with the rapid growth of the global routing table led to practical FIB-reduction solutions. In the Internet, the simplest FIB-reduction solution is FIB compression, whereby individual routers locally reduce the FIB size without any changes to external operation.

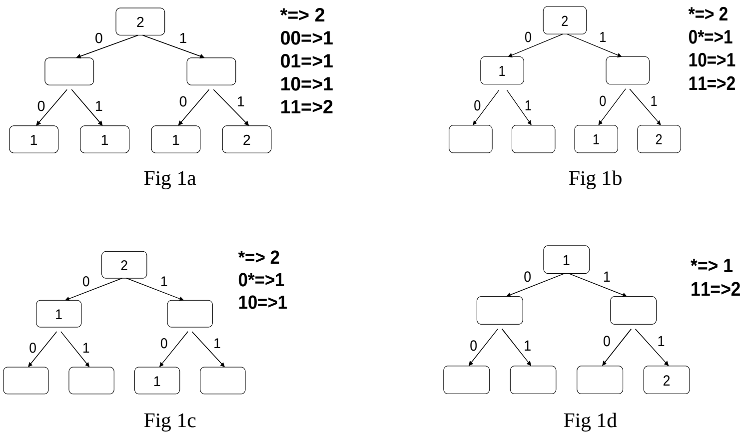

Figure 1. FIB compression.

Figure 1 illustrates the basic idea of FIB compression algorithms. The IP addresses can be organized in a binary tree structure. The less specific IP prefix (located at the root) is the parent node of the more specific prefixes. Normally in a FIB, there are some prefixes that share the same forwarding next hop (e.g. outlet port). For example, figure 1a illustrates a FIB with 5 entries, where the root node in the tree represents the null prefix, with a route to port 2. The leaves of the tree represent prefix 00, 01, 10 and 11 respectively and have been marked with the number of their forwarding ports. Note that there exists some redundant forwarding information in this FIB with the longest prefix match (LPM) lookup mechanism. Firstly, sibling prefixes 00 and 01 share the same forwarding port. These two routes can be aggregated into their parent prefix 0*, as shown in figure 1b. Secondly, prefix 11 has the same forwarding port with the root. Therefore, prefix 11 can be removed from the FIB as shown in figure 1c. At last, although prefix 0* and 10 cannot be aggregated since they are not sibling prefixes, we can adjust the next-hops of the root and prefix 11 accordingly without affecting the forwarding paths taken by data traffic. Figure 1d shows the final result of FIB compression: the FIB size shrinks to 2 entries.

According to researches conducted by [12], the gains of FIB compression depend on the structure of the RIB and the local forwarding environment (with respect to the types of forwarding behavior). Experiments conducted on BGP routers show that there would be approximately 30~70% reduction of the FIB size after FIB compression.

Nevertheless, FIB compression cannot acquire such gains for free. The FIB compression algorithm may remove some prefixes or add new prefixes that do not exist in the RIB. When the router needs to update the FIB according to changes in the RIB, the compressed FIB has to be recalculated and may incur much more overhead than that of updating un-compressed FIBs.

The FIB compression algorithm sabotages the relationship of the prefixes in the RIB. For instance, in figure 1a, prefix 00 and 01 are two independent routing objects although they share the same forwarding port. In figure 1b, these two routing objects disappear after route aggregation. Once prefix 00 is withdrawn, the compressed FIB may introduce non-routable entries into the data plane, since any packet heading for prefix 00 should be dropped instead of being forwarded to next-hop 1. It is worth noting that when FIB compression is not used, a change in the primary route of a prefix is translated to precisely one RIB to FIB download. But in the case of FIB compression, the number of FIB updates corresponding to a RIB change can be as high as a few hundred.

SMALTA [22] develops algorithms for updating a compressed FIB to best fit the BGP dynamics. Their experiments show that, on average, the number of RIB to FIB downloads in FIB compression is comparable to the number of RIB to FIB downloads when no compression is used. But the worst case may have more than 300 RIB to FIB downloads in the FIB compression scenario. SMALTA update algorithm may weaken the efficiency of FIB compression by introducing new prefixes at each FIB update process. Moreover, the compressed FIB needs to be recalculated periodically to sweep off *unnecessary* prefixes.

FIB compression allows another tradeoff between routing scalability and FIB update overhead. If the router has adequate computational resources and the RIB update is infrequent, FIB compression can be adopted. However, the overall gains of FIB compression depend on the structure of the RIB and the forwarding environment. In a BGP router, hopefully the FIB size will shrink about 50% after FIB compression.

3.3. Route caching mechanism

Another simple method to reduce the FIB size is to load a small portion of the RIB into the FIB, that is, the route caching mechanism. Earlier implementations of the FIB cached only a subset of the routes most frequently used in actual forwarding, and this worked reasonably well for enterprises where there is a meaningful *most frequently used* subset. However, route caching mechanism may experience severe performance degradation in refreshing a small cache, and various implementations moved to having FIBs in one-to-one correspondence with the RIB.

Some proposals [14, 15] revisit the route caching mechanism and propose alternative cache mechanism to improve performance on cache misses. The routes not cached in the FIB can be stored in a slow memory or even stored in other routers (a distributed cache mechanism as described in the proposal of VA [11]). In the case of a cache miss, instead of refreshing the cache instantly, the router may send the packet to a slow but larger memory for further LPM lookup or tunnel it to another router who has knowledge of forwarding this packet. A cache miss may introduce extra forwarding delay, but the general performance degradation will be acceptable if cache misses happen infrequently.

The good news is that a low cache miss rate is possible. Previous studies of ISP traffic patterns from as early as 1999 have observed that a small fraction of Internet prefixes carries a large majority of ISP traffic [23, 24, 25 and 26]. Our study [15] and other researches [11, 14] find that the distribution of the Internet traffic across IP prefixes is becoming increasingly skewed, making route caching more appealing. Our observation shows that approximately 2~5% of the most popular prefixes carry 80% of the traffic while 5~10% of the most popular prefixes carry 90% of the traffic at the Internet core routers. A proper route cache strategy may selectively cache the most popular routes that contribute major traffic loads. In this case, only a small portion of the traffic will be influenced by the route caching mechanism.

The cache miss rate has something to do with two fundamental elements: the cache size and the cache updating interval. We have studied the dynamics of the popularity of BGP prefixes and analyzed the marginal benefit on increasing the cache size in our previous work [15]. Empirically, the cache size and its updating interval should be set according to the memory and computational resources at hand.

The route caching mechanism may take advantages of the skewed distribution of traffic across prefixes. A smart route cache strategy may dynamically load popular prefixes into the FIB. Typically, if the FIB caches 10% most popular BGP routes, the majority of the traffic (more than 90%) will not be influenced, and the rest traffic will take some acceptable forwarding delay. In this context, route caching mechanism may considerably reduce the FIB size at some cost of forwarding performance.

The route caching mechanism provides a tradeoff between the routing scalability and forwarding performance. Anyway, this technique is the last resort to address the routing scalability problem in the Internet.

To summarize the above qualitative analyses, we have an analogy from an image processing application (shrinking the size of an image file) to the techniques we have discussed in this section. Imaging if we need to reduce the size of an image file, we can optionally take the picture at lower resolution and apply image compression encoding, or even directly trim off some parts of the picture. Similarly, in order to curb the RIB/FIB size, a routing scheme may combine techniques of escalating routing granularity (reducing the routing information entropy via route aggregation), FIB compression and route caching mechanism. How to properly process an image file depends on the usage of the picture. How to design a routing scheme depends on the demands of the routing system. In order to systematically evaluate different routing schemes under various criteria, we present an evaluation framework in section 4.

4. AHP MODEL IN THE EVALUATION FRAMEWORK

Our above analysis in section 3 reveals the essence and relevant techniques about the routing scales of the Internet. On evaluating the scalability of different routing schemes, we apply the analytic hierarchy process (AHP) to establish an evaluation framework.

Before presenting an overview of the AHP in this section, we emphasize the primary use of the AHP in the framework of evaluating routing scalability as a modeling tool, to capture evaluation criteria (through weighted influential factors) and design freedom (in alternative routing schemes). Comparing various routing schemes under different evaluation criteria, rather than developing an optimum routing scheme, is the main goal of this study. As evidenced by the large number of proposals that address the routing scalability problem of the Internet over the past few years, we expect that there will be no shortage of good ideas based on piecemeal approaches. However, what seems to be lacking is a level ground for fair comparison among the variety of Internet routing schemes. Our evaluating framework provides a systematic judgment on gains and losses of candidate routing schemes.

4.1. AHP introduction

The AHP is a structured technique for organizing and analyzing complex decisions. Based on mathematics and psychology, it was developed by Thomas L. Saaty in the 1970s and has been extensively studied and refined since then. It is used around the world in a wide variety of decision and evaluation situations.

Rather than prescribing a "correct" decision, the AHP helps decision makers find one that best suits their goal and their understanding of the problem. It provides a comprehensive and rational framework for structuring a decision problem, for representing and quantifying its elements, for relating those elements to overall goals, and for evaluating alternative solutions.

Users of the AHP first decompose their decision/evaluation problem into a hierarchy of more easily comprehended sub-problems, each of which can be analyzed independently. The elements of the hierarchy can relate to any aspect of the decision problem—tangible or intangible, carefully measured or roughly estimated — anything at all that applies to the decision at hand.

Once the hierarchy is built, the decision makers systematically evaluate its various elements by comparing them to one another, two at a time, with respect to their impact on an element above them in the hierarchy. In making the comparisons, the decision makers can use concrete data about the elements, but they typically use their judgments about the elements' relative meaning and importance. It is the essence of the AHP that human judgments, and not just the underlying information, can be used in performing the evaluations.

The AHP converts these evaluations to numerical values that can be processed and compared over the entire range of the problem. A numerical weight or priority is derived for each element of the hierarchy, allowing diverse and often incommensurable elements to be compared to one another in a rational and consistent way. In the final step of the process, numerical priorities are calculated for each of the decision alternatives. These numbers represent the alternatives' relative ability to achieve the decision goal.

The procedure for using the AHP to evaluate the scalability of alternative routing schemes can be summarized as:

- Model the problem as a hierarchy containing the decision goal (i.e. ranking the scalability of different routing schemes), the alternatives routing schemes, and the criteria for evaluating the alternatives.
- Establish priorities among the elements of the hierarchy by making a series of judgments based on pair-wise comparisons of the elements. The judgment matrixes will be established after deriving a series of numerical scales of measurement for the pair-wise comparisons.
- Synthesize these judgments to yield a set of overall priorities for the hierarchy.
- Check the consistency of the judgments.
- Come to a final evaluation based on the results of this process

In the following subsection, based on our qualitative analysis, we present a case study of applying AHP to establish an evaluation framework for alternative routing schemes.

4.2. The evaluation framework with AHP

This case study describes the use of the AHP to evaluate the routing scalability of alternative routing schemes under given evaluating criteria. The decision goal is to single out the best routing scheme that can improve the routing scalability of the Internet. Suppose we have candidate routing schemes R_a, R_b and R_c that optionally apply route aggregation, FIB compression and route caching mechanism.

R_a: has route aggregation factor a_1 and applies route caching mechanism (cache 10% RIB).

R_b: has route aggregation factor a_2 and suggests using FIB compression.

R_c: has route aggregation factor a_3 and applies route caching (cache 5% RIB).

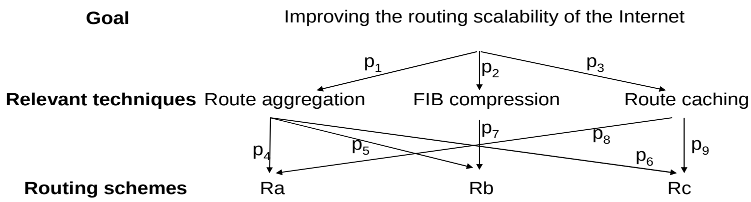

Figure 2. AHP analytic hierarchy

By using AHP, we establish the analytic hierarchy as shown in figure 2. At the top of the hierarchy, the decision goal is to improve the routing scalability of the Internet. Then at the second level we decompose the decision goal into three typical techniques that may reduce the RIB/FIB size. The significance of each technique should be determined by pair-wise comparison. At the end of the hierarchy, the alternative routing schemes may optionally adopt some techniques to improve the routing scalability (linked with corresponding techniques).

After establishing the analytic hierarchy, we need to determine the weight of each relevant technique with respect to the decision goal and priorities for the alternative routing schemes with respect to each of the techniques adopted. The priorities will then be combined throughout the hierarchy to give an overall evaluation for each alternative routing scheme and the ranking of their priorities will indicate their relative strengths in reaching the Goal.

Table 1 Fundamental Scales for Pair-wise Comparisons

Intensity of Importance	Definition	Explanation
1	Equal importance	Two elements contribute equally to the objective.
3	Moderate importance	Experience and judgment moderately favor one element over another.
5	Strong importance	Experience and judgment strongly favor one element over another.
7	Very strong importance	Experience and judgment very strongly favor one element over another; its dominance is demonstrated in practice.
9	Extreme importance	The evidence favoring one element over another is of the highest possible order of affirmation.
Intensities of 2,4,6 and 8 can be used to express intermediate values.		

The priorities will be derived from a series of measurements: pair-wise comparisons involving all the nodes at the same level. Table 1 lists the fundamental scales of the pair-wise comparisons. Normally the value ranged from 1 to 9 can be used to represent the intensity of importance. The results of these comparisons will be entered into a matrix which is processed mathematically to derive the priorities for all the nodes on the same level.

We start from comparing the relevant techniques with respect to their significance to reaching the Goal. This part of the process requires best understanding of those techniques. A decision maker should have full awareness of the gains and losses of those techniques under different conditions. As we have analyzed in section 3, route aggregation provides a tradeoff between routing scaling and the granularity of policy. FIB compression may introduce extra computational overhead and route caching mechanism may influence some traffic. A decision maker has to consider these side effects before weighting the significance of each technique. For instance, when the granularity of routing policy is not regarded as a rigid demand in the inter-domain routing, aggressive aggregation is favored. Consequently route aggregation may be valued with a high priority. If the computational resource at modern routers is stringent, and the FIB compressing overhead is considerably high in most cases, the FIB compression will be less appealing and should be valued as less significant on improving the routing scalability. The judgment on route caching mechanism depends on the extent of tolerance of the forwarding delay. Different people have different understandings. Therefore, it is reasonable to invite a group of experts to figure out the comparison matrix. In case of a group decision, there would be much discussion and debate among the decision makers. This process captures the weight of each relevant technique in the second level, which formulates our criteria of evaluation. The follow comparison matrix gives an example of the evaluation criteria.

Table 2 pair-wise comparison of three techniques

Techniques	R.A.	F.C.	R.C.	Priority
R.A.	b_{11}	b_{12}	b_{13}	p_1
F.C	b_{21}	b_{22}	b_{23}	P_2
R.C	b_{31}	b_{32}	b_{33}	P_3
Accepted condition: Inconsistency <0.1				

The priority can be calculated as follows:

$$p_i = \frac{w_i}{\sum_{j=1}^{n} w_j} \quad \textit{where} \quad w_i = \sqrt[n]{m_{ij}}; \quad m_{ij} = \prod_{j=1}^{n} b_{ij} \quad (i = 1,2,...,n; j = 1,2,...,n)$$

Next we need to compare alternative routing schemes with each other. For each comparison, we will judge which scheme of the pair is weaker with respect to the technique under consideration. Then we will assign a relative weight to the other candidate. Consequently, similar to the previous process, we have comparison matrixes, and we can calculate priorities (i.e. p_4~p_9 in figure 2) for each scheme afterwards.

Now that we know all the priorities (i.e. p_1~p_9) in the hierarchy, we can synthesize the final priorities of the alternatives with respect to the goal. This is a straightforward matter of multiplying and adding, carried out over the whole of the hierarchy. The final evaluation of the candidate routing schemes has been quantified as follows:

R_a: p_4*p_1 +p_8*p_3

R_b: p_5*p_1 +p_7*p_2

R_c: p_6*p_1 +p_9*p_3

And ranking R_a, R_b *and* R_c is a piece of cake.

There are many proposals to improve the routing scalability of the Internet, each of which corresponds to a human engineering effort. These efforts have different characteristic impacts on routing functioning, overhead and performance. Some are better than others depending on the evaluation criteria set by network operators. An evaluation framework is needed, where each particular routing scheme leads to quantified gains and losses.

Evaluating the routing scalability with AHP provides a unified framework. It decomposes the goal of improving the routing scalability into typical techniques. This decomposition makes complicated case by case comparison easy to be handled. It is also particularly convenient to change the criteria for different scenarios.

5. CONCLUSIONS

This paper presents theoretic analyses to reveal the essence of the routing scale of the Internet. The analyses cover both the control plane and the data plane of the routing system, as well as the interacting between them. In the control plane, from a perspective of the information theory, the routing scale is associated with the addressing space, routing granularity and degree of connection of each vertex in the network. In the model of routing information entropy, the routing granularity, which is related to route aggregation, is singled out as a convenient tuning knob on controlling the routing scale in the control plane. In the data plane, the FIB compression technique is associated with the computational overhead on route updates. As for the interaction between the control plane and the data plane, the route caching mechanism is analyzed, together with its corresponding side effect of performance degradation. These analyses depict the gains and losses of those techniques that attempt to improve the routing scalability of the Internet.

Evaluating the routing scalability of the Internet requires weighing the gains and losses of alternative routing schemes. In order to make the complicated evaluation easier, we use the AHP method which provides a unified framework that facilitates quantitative evaluation. The AHP captures the flexibility of the evaluation criteria and the freedom of design in alternative routing schemes by constructing analytic hierarchies. Qualitative analysis may help us to establish the AHP hierarchy rationally. The AHP quantifies the value of evaluation in a systematic way, and finally synthesizes the priority of the candidate routing schemes with respect to the top decision goal.

This paper presents a synthetic methodology that facilitates systematically analyzing the routing scalability of the Internet. The modeling work in this paper reveals the impact of routing scale on the functioning, overhead and performance of a routing system. The AHP-based evaluation framework provides a common tool that allows researchers to evaluate the routing scalability of alternative routing schemes.

Acknowledgement

This work was supported by the Key Project of Chinese National Programs for Fundamental Research and Development (973 program) No. 2009CB320501 and Specialized Research Fund for the Doctoral Program of Higher Education (SRFDP) No 200800030034.

References

[1] CAIDA, http://www.caida.org.
[2] G. Huston, BGP Routing Table Analysis Reports, http://bgp.potaroo.net/.
[3] D. Meyer, L. Zhang, K. Fall, Report from the IAB workshop on routing and addressing, RFC 4984, 2007.
[4] T. Li, Design Goals for Scalable Internet Routing, RFC 6227,2011.
[5] D. Peleg, Distributed computing – A locality-sensitive approach, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.
[6] D. Farinacci, V. Fuller, D. Meyer, D. Lewis, Locator/ID separation protocol (LISP), Internet-Draft, 2009. http://tools.ietf.org/html/ draft-ietf-lisp-08.txt.
[7] eFIT project. Enabling Future Internet innovations through Transit wire. http://irl.cs.ucla.edu/efit/index.html
[8] R. Atkinson, ILNP concept of operations, Internet-Draft, 2010. http://ilnp.cs.st-andrews.ac.uk/docs/id/draft-rja-ilnp-intro-03.txt.
[9] E. Nordmark, M. Bagnulo, Shim6: Level 3 multihoming shim protocol for IPv6. RFC 5533, 2009. http://tools.ietf.org/html/rfc5533
[10] X. Yang, D Clark, A.W. Berger, NIRA: A new inter-domain routing architecture, IEEE/ACM Trans. on Networking, 2007, 15(4) pp.775 788.
[11] H. Ballani, P. Francis, T. Cao, J. Wang, Making routers last longer with ViAggre, in: Proc. of the 6th USENIX Symp. on Networked Systems Design and Implementation. Berkeley: USENIX Association, 2009, pp. 453 466.
[12] B. Zhang, L. Wang, X. Zhao, Y. Liu, L. Zhang, FIB aggregation. Internet-Draft, 2009. http://tools.ietf.org/html/draft-zhang- fibaggregation-02.
[13] R. Draves, J. Doe, Optimal Routing Table Constructor, INFOCOM 1999, July 1999.
[14] C. Kim, M. Caesar, A. Gerber, J. Rexford, Revisiting Route Caching: The World Should Be Flat, in: Proc of the PAM 2009.

[15] W. Zhang, J. Bi, J. Wu, B. Zhang, Caching Popular BGP Prefixes with Grey Modeling Prediction, in: Proc. of the ICCCN 2011.
[16] K. Calvert, J. Griffioen, L. Poutievski, Separating routing and forwarding: A clean-slate network layer design. BROADNETS 2007
[17] M. Caesar, T. Condie, J. Kannan, K. Lakshminarayanan, I. Stoica, S. Shenker, ROFL: Routing on flat labels, in: Proc. of the SIGCOMM 2006. New York: ACM, 2006, pp. 363-374.
[18] W. Zhang, X. Yin, J. Wu, W. Zhang, "Real aggregation for reducing routing information base size", Journal of Convergence Information Technology, vol.5, no.6, pp. 47-53, 2010.
[19] W. Zhang, J. Bi, J. Wu. "Improving routing scalability with aggressive route aggregation", Journal of Convergence Information Technology, vol.6, no.2, pp.151-160, 2011.
[20] C. Shannon, A Mathematical Theory of Communication, The Bell System Technical Journal, Vol. 27, 1948, pp. 379–423.
[21] Openflow Specification. http://www.openflow.org/documents/openflow-spec-v1.1.0.pdf.
[22] Z. Uzmi, S. Jawad, A. Tariq, P. Francis, Prefix Aggregation with SMALTA, Technical Report. http://www.mpi-sws.org/~zartash/TR-MPI-SMALTA.pdf, July 2010.
[23] W. Fang, L. Peterson, Inter-As traffic patterns and their implications, in: Proc. of Global Internet (1999).
[24] A. Feldmann, A. Greenberg, C. Lund, N. Reingold, J. Rexford, F. True, Deriving traffic demands for operational IP networks: methodology and experience. IEEE/ACM Transactions on Networking 9, 3 (2001).
[25] J. Rexford, J. Wang, Z. Xiao, Y. Zhang, BGP routing stability of popular destinations, in: Proc. of Internet Measurement Workshop (2002).
[26] N. Taft, S. Bhattacharyya, J. Jetcheva, C. Diot, Understanding traffic dynamics at a backbone PoP, in: Proc. of Scalability and Traffic Control and IP Networks SPIE ITCOM (2001).