Recent progress in the study of the next generation Internet in China

Ying Liu\textsuperscript{1,2}, Jianping Wu\textsuperscript{1,2}, Qian Wu\textsuperscript{1,2} and Ke Xu\textsuperscript{1,2}

\textsuperscript{1}Tsinghua University, Beijing 100084, People’s Republic of China
\textsuperscript{2}Tsinghua National Laboratory for Information Science and Technology (TNList), Beijing 100084, People’s Republic of China

The Internet has become a major part of the global communications infrastructure supporting modern-day socio-economic development, social progress and technological innovation. Invented 30 years ago, today the Internet is facing severe challenges. Many countries have funded research projects on the next-generation Internet, such as GENI, FIND, FIRE and CNGI, in an effort to solve these challenges. In addition, over the past few years, the networking research community has engaged in an ongoing conversation about how to move the Internet forward, and there are now two different approaches towards Internet research. The first approach is based on using the existing Internet architecture to solve the major technical challenges—this is called ‘evolutionary’ research. The other, which is called the ‘clean slate’, involves the design of an entirely new Internet architecture. In the first part of this paper, the basic features of the next generation Internet and its principal contradictions are analysed. Then a survey of recent progress in the study of the next generation Internet in China is discussed. Finally, the focus and direction for the next step in research are presented as based on fundamental research into the international next generation Internet architecture, and the many new innovative demands placed on Internet architecture in recent years.

1. Introduction

As the twentieth century’s key invention, the Internet has led to an unprecedented sharing and communication of information, technology and resources. It has profoundly changed people’s working, living and learning modes and has become an important element of
The infrastructure supporting modern-day socio-economic development, social progress and technological innovation. The way in which a country uses the Internet has become an important criterion in judging its fundamental strength and economic competitiveness.

With the rapid development of technology such as high-speed optical communications, wireless mobile communication and high-performance/low-cost computing and software, as well as the continuous emergence of innovative applications, users’ demands on the performance of the Internet have also increased. The Internet, which was invented 30 years ago and relies on IPv4 as its core technology, is facing increasingly severe technical challenges: an insufficient number of network addresses (thus limiting the large-scale expansion of the network); poor credibility (resulting from a large number of security flaws); a poor ability to control Internet quality of service (thus leaving high-quality network services unguaranteed); network bandwidth and performance levels which cannot always meet the needs of users; and difficulties in achieving a highly efficient mobile Internet, owing to the different systems of traditional wireless mobile communications technology and Internet technology.

In order to solve these technical challenges, the USA and other countries began research on the next generation of the Internet in the mid-1990s. In China, next generation Internet research began in earnest in the year 2000. Many countries have funded research projects on the next generation Internet, such as the Global Environment for Networking Innovations (GENI; http://www.geni.net); Future Internet Design (FIND; http://www.nets-find.net); Future Internet Research and Experimentation (FIRE; http://cordis.europa.eu/fp7/ict/fire/); and the China Next-Generation Internet Demonstration (CNGI; http://www.cngi.cn) (as shown in figure 1). After 10 years’ work, researchers have become increasingly aware of the importance, complexity and difficulties in developing a long-term next generation Internet.

In the first part of this paper, the basic features of the next generation Internet and its principal contradictions are analysed. We then discuss a survey of recent progress in the study of the next generation Internet in China. Finally, the focus and direction for the next step in research,
stemming from fundamental research into the international next generation Internet architecture and the many new innovative demands placed on the Internet architecture in recent years are presented.

2. The analysis of the demands of next generation Internet research

Over the past several years, the networking research community has engaged in an ongoing conversation about how to propel the Internet forward, and there are now two different approaches which have been recognized for research into the Internet. The first approach is based on using the existing Internet architecture to solve the major technical challenges—this is called ‘evolutionary’ research. Evolutionary research aims to address fundamental Internet problems without breaking the architecture that is in place today, while respecting constraints that relate to partial deployment, backwards compatibility and implementation feasibility. A typical example of the evolutionary approach is the Internet Engineering Task Force’s (IETF) IPv6. The second approach, which is referred to as the ‘clean slate’, involves the design of an entirely new Internet architecture. This clean-slate paradigm aims to address fundamental problems and limitations of the Internet without being constrained by the architecture or protocols currently used. The typical examples of the clean-slate approach are the National Science Foundation’s FIND and GENI programmes and the Community Research and Development Information Service’s FIRE programme.

We have studied the demands on the next generation Internet and compared our results with a number of key projects, including IETF, GENI, FIND and FIRE. Our analysis and findings presented by these research projects are broadly consistent (as shown in table 1), suggesting that the demands of users on the next generation Internet are relatively uniform worldwide.

(a) Next generation Internet and its basic features

It is widely recognized that advocating the use of simple and practical technologies such as layered and distributed architecture, connectionless packet switching and scalable routing has been the basis of the rapid development and expansion of the Internet during the past several decades. Such technologies have emerged gradually as a result of long-term, large-scale experiments. Practice has shown that the Internet architecture itself has good multi-dimensional scalability. Accordingly, we think of Transmission Control Protocol/Internet Protocol (TCP/IP) as the DNA of the Internet. We should, to the greatest possible extent, both preserve and move forward with the technological essence of the current Internet architecture. This preservation should also allow for expansion, evolution and innovation. The study of architecture as the keystone of next generation Internet basic research is necessary to maintain its central role in next generation Internet research. The IPv6 protocol and related technologies should be considered the main components in the evolution towards the next generation Internet.

Our analysis leads to the conclusion that the next generation Internet should address the major technical problems of its predecessor—the current Internet—in the areas of scalability, high performance, real time, mobility, security, manageability and economy.

— Scalability refers to the next generation Internet’s ability to expand and connect with all compatible electronic equipment (whereas the existing Internet mainly links computer systems only). The next generation Internet should connect a greater variety and number of devices, have a larger scale and more extensive applications.
— High performance refers to the next generation Internet’s ability to provide higher transmission speeds; in particular, end-to-end transmission speeds should reach 10 or 100 Mbps in order to better support the next generation Internet applications.
— Security refers to the next generation Internet’s ability to establish a comprehensive security system. This system shall be based on simple sharing characteristics and the provision of authenticity and traceability capabilities, thus providing a more secure and trustworthy network service.
Table 1. Major demands of next generation Internet research plans. FIRE, Future Internet Research and Experimentation; FIND, Future Internet Design; GENI, Global Environment for Networking Innovations; IETF, Internet Engineering Task Force.

<table>
<thead>
<tr>
<th>IETF</th>
<th>GENI</th>
<th>FIND</th>
<th>FIRE</th>
<th>our analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>scalability of large-scale routing</td>
<td>scalability</td>
<td>scalability</td>
<td>scalability</td>
<td>scalability</td>
</tr>
<tr>
<td>network security and trustworthiness</td>
<td>security, high reliability</td>
<td>security and availability; privacy security preservation in an environment of pervasive sensing, computing, content and data mining security</td>
<td>security and availability; privacy security</td>
<td>security and availability; privacy security</td>
</tr>
<tr>
<td>application and transition issues for IPv6</td>
<td>adequate levels of availability; deliver to society the potential of emerging technologies such as wireless communications suitable for tomorrow's technologies suitability for ad hoc/multi-hop/mesh networking, mobility, transparency mobility</td>
<td>suitable for tomorrow's technologies suitability for ad hoc/multi-hop/mesh networking, mobility, transparency</td>
<td>suitable for tomorrow's technologies</td>
<td>mobility</td>
</tr>
<tr>
<td>network management under complex environmental constraints</td>
<td>easy to set up, identify failures and problems, and manage manageable and usable manageability</td>
<td>manageable and usable</td>
<td>manageable and usable</td>
<td>manageability</td>
</tr>
<tr>
<td>control and management of real-time traffic</td>
<td>eliminate barriers to economic investment and enhancement economically viable</td>
<td>capable of supporting applications high performance, real time</td>
<td>capable of supporting applications</td>
<td>high performance, real time</td>
</tr>
</tbody>
</table>
— Real time refers to the next generation Internet’s ability to transform the current ‘best effort’ quality-of-service strategy into a more controllable and reliable quality of service. With this new, more reliable service, next generation Internet applications such as multi-casting, large-scale video and real-time interaction can be supported.

— Mobility refers to the next generation Internet’s ability to use advanced wireless mobile communication technology to achieve an ‘anywhere, anytime’ mobile Internet, and truly become a tool for people’s work, living and learning.

— Manageability refers to the next generation Internet’s ability to provide fine-grained network management elements and means, thus offering reliable and comprehensive management capabilities of the network for users.

— Economy refers to the next generation Internet’s ability to overcome the unreasonable economic model of the current Internet, in which network operators invest heavily in the construction of networks at a loss (while network information content providers offer network-based services at a high profit). Creating a reasonable, fair and harmonious multi-profit model is necessary to ensure the healthy and sustainable development of the Internet.

These objectives are the focus of the next generation Internet research which has been carried out in different countries for over a decade.

(b) The principal contradictions in next generation Internet development

According to the basic features inherent to the next generation Internet, we have concluded that the principal contradictions in its development can be divided into four aspects [1]:

— The contradiction between the complex diversity of network functions and the single-dimensional scalability of network architecture. Although it is believed that the Internet has optimized scalability based on end-to-end arguments and ‘best effort’ service, currently these characteristics focus only on network interconnection. The Internet continues to show increasing deficiencies in its ability to support new applications. For example, the current Internet architecture cannot provide good support for multi-casting and mobility of mass nodes owing to the ‘best effort’ service mode, which only considers the scalability of interconnection rather than other types of scalable service problems. It is difficult to expand the current network architecture further in areas such as address space, addressing and routing and service type.

— The contradiction between the fixed transmission and control goals and the unknown behaviour of the network. The traffic model and the behaviour model of the Internet as based on packet switching have not been studied in sufficient detail. Although some theoretical results concerning self-similarity and long relevance have been gained through the analysis of large-scale network traffic, the scientific principles underlying these results remain unclear. The deficiencies in the traffic and behaviour models lead to a lack of theoretical guidance concerning the control and management of large-scale networks, which remains based on empirical rules and experience. This prevents the network from meeting demands to provide a better quality of service.

— The contradiction between the needs of security and trustworthiness, and the vulnerability of the network. Being a huge artificial system, the Internet has inherent vulnerability. There is an enormous number of hardware systems and numerous applications on the network, and a bug in any hardware or software has the potential to be used in attacks on the network. Methods for analysing the network’s vulnerability in theory while protecting networks in reality remains a problem. The network should not only have its own intrinsic security, but also be able to provide the required security functions for any application. Given the growing size and performance requirements of applications today, it is difficult to ensure the security of the network.
The contradiction between the complexity and variability of network service needs, and the relative stability of the network architecture. The fundamental purpose of network development is to provide services. Because of the complexity of the next generation Internet (having a larger, more complex and heterogeneous structure), as well as the complexity and diversity of the demands on service (inter-operability, speed, availability, scalability, manageability, quality of service, intelligence and individuality), the development of large-scale Internet services lacks a theoretical foundation. Ways in which a service model based on next generation Internet architecture can be constructed, ways in which users can be provided quickly and flexibly with high availability, good performance and inter-operability of services, and ways in which current services can be coordinated to provide end users with reusable services that can be managed are all difficult but important theoretical issues.

3. Progress in next generation Internet research and experience in China

There is a long history of research on the next generation Internet in China in the face of the major technical challenges discussed in this paper. Chinese researchers have strived to solve the major theoretical problems and technical challenges of the Internet, while implementing the technological innovations required for evolution of the next generation Internet.

The Internet was introduced to China in 1994, and from 1994 to 1999 the country learned how to construct and use the Internet while initiating Internet research. From 1999 to 2004, China developed the high-performance IPv4/v6 router, and resulting research focused on its implementation on the Internet. From 2004 to date, Chinese researchers have been focusing on improvement of the current Internet, while studying Internet architecture and its theory. In 2000, China began research into the next generation Internet.

(a) Basic theory of next generation Internet architecture

The national Natural Science Foundation of China (NSFC) set up projects to support research into the basic theory of the next generation Internet in 2003. At this time, the national ‘973’ programme (the major state-funded basic research development programme in China) began to support Internet-related research projects. The national ‘973’ project on ‘Research of Future Internet Architecture’ and ‘Research of New-generation Internet architecture and protocol’ undertaken by Tsinghua University, Beijing, People’s Republic of China, and four other research institutes focused on theoretical research into the Internet.

We have defined the multi-dimensional scalability of the next generation Internet by tentatively proposing a multi-dimensional scalable next generation Internet architecture (as shown in figure 2), and defining its five basic elements.

In the above analysis of the four principal contradictions affecting the next generation Internet, resolution of the first contradiction involves a contradiction between the complex diversity of network functions and the single-dimensional scalability of network architecture.

We define the multi-dimensional scalability of next generation Internet architecture as follows.

(b) Scale scalability

Scale scalability refers to the ability of network performance (such as bandwidth utilization and resource utilization of network core equipment) and end-to-end performance to achieve growth corresponding with the growth of network nodes and links. For example, we can evaluate scale scalability by resolving network utility function under the appropriate conditions.

(c) Performance scalability

Performance scalability refers to the ability of the network and end-to-end performance to achieve growth corresponding to the growth of network resources, such as links and nodes. For example,
Figure 2. Multi-dimensional scalable next generation Internet architecture. QoS, quality of service. (Online version in colour.)

we can evaluate performance scalability by solving the maximum problem of network utility function under the appropriate conditions.

(d) Service scalability

Service scalability refers to the ability of network service deployment to achieve growth corresponding with the growth of the overall scale of service. For example, we can construct a network utility function under constraints according to the total amount of services and the proportion of different services, while evaluating service scalability by solving for the maximum.

(e) Security scalability

Security scalability refers to the ability of a network security mechanism’s performance and effectiveness to be enhanced along with the growth of its deployment. For example, we can evaluate the utility function of a security mechanism for different network architectures.

(f) Function scalability

Function scalability refers to the ability of various network functions (such as unicasting, multicasting, tunnels, etc.) to be expanded under a unified architecture framework.

The multi-dimensional scalability of next generation Internet architecture can also be briefly summarized on a higher level: the scalability of network architecture refers to the ability of network characteristics (such as performance, deployment costs, etc.) to achieve improvements corresponding to the change in network-related constraints (such as speed, size, service type, etc.).

In other words, scalability refers to the relation between the overall effectiveness $K$ and network characteristics (such as source rate $x_s$, and network scale $|V| + |E|$). It can be expressed as

$$K = \alpha U + (1 - \alpha)Q,$$
where \( U = \sum_{s \in V_u} k_1s \sum_{w \in W} k_1w U(x_s, w) \) represents the sum of effectiveness of network users with different weights

\[
Q = \sum_{l \in E} Q_1(\mu_l) + \sum_{s \in V_c} Q_2(\eta_s) = \sum_{l \in E} k_2l \sum_{w \in W} k_2w Q_1'(\mu_l, w) + \sum_{s \in V_c} k_3s \sum_{w \in W} k_3w U_c(\eta_s, w)
\]

represents the overall effectiveness. In this formula, \( \alpha \) is a smoothing parameter used to adjust the proportion of network effectiveness to user effectiveness.

On the basis of our definition of scalability, we further propose five basic elements which must be included in the multi-dimensional scalable version of next generation Internet architecture.

— IPv6: this has become the standard of the network layer protocol for the next generation Internet. It helps to achieve the defined scale scalability and security scalability of the next generation Internet.

— Authentic IPv6 addressing: many of the security problems identified in the existing Internet come from the non-authentication of source addresses. The authentic IPv6 addressing in the next generation Internet will help to achieve the desired security scalability and service scalability [2].

— Scalable processing capacity of network nodes: as the demands of users grow, the core exchange node of the next generation Internet should have scalable processing capacity. This will help to achieve the desired performance scalability and scale scalability [3].

— Connectionless quality-of-service control: the quality-of-service control capacity of the Internet is always a hot research topic. Achieving quality-of-service control based on hop-by-hop and connectionless routing is one of the goals of next generation Internet research, and is believed to have the capacity to improve performance scalability and service scalability [4].

— IPv4 over IPv6 network transition strategy: the next generation Internet should cooperate with the existing Internet to provide services to users. However, current transition strategies are suited to only a small-scale IPv6 network. Further research is required to study the strategies allowing a transition from IPv4 to a next generation network using IPv6 as its core protocol. This will help achieve the desired function scalability and service scalability [5].

These five basic elements all support the development of the necessary scale scalability, function scalability, performance scalability, security scalability and service scalability for the next generation Internet.

(g) The evaluation model of Internet scalability

In this section, we first introduce the mathematic description for the scalability of Internet architecture, and then introduce the one- and multi-dimensional scalability evaluation models, respectively.

(i) Mathematic description

For any network architecture, the constraint condition set is defined as \( X = \{x_1, x_2, \ldots, x_n\} \). Because we usually test the system character with a constraint condition in constraint intervals, we set its change bound at \( X \in \Phi = \{x'_i \leq x_i \leq x''_i\} \), and define the quantified evaluation metric set as \( Y = \{y_1, y_2, \ldots, y_m\} \). Then any evaluation metric \( y_i \) must be the evaluation function of \( X \), and we have

\[
y_i = f_i(X) = f_i(x_1, x_2, \ldots, x_n), \quad X \in \Phi.
\]  

Here, \( x_i \) can be a continuous variable or a discrete variable. When \( x_i \) is continuous, we assume that \( f_i \) is a continuous and smooth function, which will be differentiated everywhere.
(ii) Single-constraint, one-dimensional scalability

Single-constraint, one-dimensional scalability refers to the scalability of single-objective single-constraint, whose evaluation function is

\[ y_i = f_{ij}(x_j), \quad x_j \in [x_{j1}, x_{j2}] \].  

In this function, \( y_i \) is the evaluation metric, \( x_j \) is a single-constraint condition, \( f_{ij} \) is the function of the evaluation metric \( y_i \) with the changing constraint condition \( x_j \), and \([x_{j1}, x_{j2}]\) is the constraint interval.

According to the scalability’s classification on examining objectives, we are able to define three different scalabilities independently:

**Static scalability.** Static scalability refers to the good or bad performance of the architecture’s current evaluation metric. Therefore, it can be denoted by an evaluation function

\[ S_{ij} = y_{ij} = f_{ij}(x_j), \quad x_j \in [x_{j1}, x_{j2}] \].  

**Dynamic scalability.** Dynamic scalability refers to a function’s changing character at a certain moment. We define this as the first derivative of the evaluation function to constraint condition,

\[ D_{ij}(x_{j0}) = \frac{dy_{ij}}{dx_j} \bigg|_{x_j = x_{j0}} = f'_{ij}(x_{j0}). \]

That is, when \( x_j = x_{j0} \), the dynamic scalability of architecture is \( D_{ij}(x_{j0}) \). This is the definition when the constraint condition represents a continuous constraint condition. Researchers may also encounter some discrete constraint conditions, such as the number of hosts. We define its dynamic scalability as follows:

\[ D_{ij}(x_{j,k}) = \frac{\Delta y_{ij}}{\Delta x_j} = \frac{f_{ij}(x_{j,k+1}) - f_{ij}(x_{j,k})}{x_{j,k+1} - x_{j,k}}, \quad x_{j,k} \neq x'_{j}, \]

in which \( x_{j,k} \) denotes that the constraint condition \( x_j \) takes the \( k \)th value.

From figure 1a, we can see that the geometric significance of single-constraint, one-dimensional dynamic scalability is the slope of \( x_j \) on the constraint point \( x_{j0} \) or \( x_{j,k} \).

**Cumulative scalability.** Cumulative scalability reflects the cumulative character of the evaluation function in the constraint interval. We define the single-constraint, one-dimensional cumulative scalability of a continuous condition as

\[ C_{ij} = \int_{x'_{j}}^{x_j} f_{ij}(x_j) \, dx_j, \]

while the one-dimensional cumulative scalability of a discrete constraint condition is defined as

\[ C_{ij} = \sum_{x_{j,k}} f_{ij}(x_{j,k})(x_{j,k+1} - x_{j,k}), \quad x_{j,k} \neq x'_{j}. \]

(iii) Multi-constraint, one-dimensional scalability

Multi-constraint, one-dimensional scalability refers to the single-objective, multi-constraint scalability, whose evaluation function is

\[ y_i = f_i(X) = f_i(x_1, x_2, \ldots, x_n), \quad X \in \Phi, \]

in which \( y_i \) is the evaluation metric, \( X \) is the vector of the multi-constraint condition, \( x_1, x_2, \ldots, x_n \) are each of the constraint conditions, \( f_i \) is the function of the evaluation metric, \( y_i \) changes with constraint condition \( X \) and \( \Phi \) represents the constraint intervals.
Static scalability. Static scalability refers to the good or bad character of a system’s current evaluation metric. Therefore, it can be denoted by an evaluation function

\[ S_i = y_i = f_i(X), \quad X \in \Phi. \quad (3.7) \]

Dynamic scalability. We first consider the situation in which constraint conditions are all representative of continuous constraint. The definition of multi-constraint dynamic scalability is more complicated than that of single-constraint scalability; as such, the relativity of each constraint condition should be considered.

We assume that the constraint conditions are independent; that is, any value of \( x_i \) is completely independent of \( x_j \) itself. Then, its dynamic scalability can be defined as

\[
D_i(X_0) = \left. \frac{\partial^2 y_i}{\partial x_1 \partial x_2 \cdots \partial x_n} \right|_{X=X_0} = \left. \frac{\partial^2 f_i(X)}{\partial x_1 \partial x_2 \cdots \partial x_n} \right|_{X=X_0}. \quad (3.8)
\]

On the other hand, we assume that all \( x_i \) are related to each other, and that both are functions of another variable \( t \). As such, we have

\[ x_j = z_j(t) \]

and

\[ y_i = f_i(X) = f_i(x_1, x_2, \ldots, x_n) = f_i(z_1(t), \ldots, z_n(t)) = f_{it}(t). \]

In this case, dynamic scalability can be simplified into single-constraint, one-dimensional scalability

\[ D_i(X_0) = \left. \frac{df_{it}}{dt} \right|_{t=t_0} = f'_{it}(t_0), \]

of which \( X_0 = (x_{10}, x_{20}, \ldots, x_{n0}) = (z_1(t_0), z_2(t_0), \ldots, z_n(t_0)) \).

Generally, among the constraint conditions, if the first \( t(1) \) ones are related with \( t1 \), the next \( t(2) \) ones are related with \( t2, \ldots, t(t) \) ones are related with \( t_l \), and the \( r-q+1 \) ones have nothing to do with each other. Lastly, then let \( (x_{i1,1}, x_{i2,2}, \ldots, x_{i(t(i))}) \) be the ones related with \( t_l \) and \( x_{i,j} = z_{ij}(t_l) \) \( x_q, x_{q+1}, \ldots, x_r \) have nothing to do with any other \( x \), of which, \( i = 1, 2, \ldots, l, j = 1, 2, \ldots, t(i) \). We have

\[ (r-q+1) + \sum_{i=1}^{l} t(i) = n. \]

Then, \( f_i(X) \) can be written as

\[ f_i(X) = f_{it}(t_1, t_2, \ldots, t_l, x_q, x_{q+1}, \ldots, x_r), \]

meaning that the dynamic scalability can be defined as

\[
D_i(X_0) = \left. \frac{\partial^{(l+r-q+1)} f_{it}}{\partial t_1 \partial t_2 \cdots \partial t_l \cdot \partial x_q \cdots \partial x_r} \right|_{t_l=t_0, x_i=x_{i0}} = f^{(l+r-q+1)}(t_{10}, t_{20}, \ldots, t_{l0}, x_{q0}, \ldots, x_{r0}), \]

when \( t_l = t_{0l}(i = 1, 2, \ldots, l), x_i = x_{i0}(i = q, q+1, \ldots, r) \) and \( X = X_0 \).

When the constraint conditions are all of the discrete constraints, dynamic scalability can be defined as

\[
D_i(X_0) = \frac{f_{it}(t_{1k_1+1}, t_{2k_2+1}, \ldots, t_{jk_j+1}, x_{q_{k_1}+1}, x_{q_{k_2}+1}, \ldots, x_{r_{k_j}+1}) - f_{ij}(t_{1k_1}, t_{2k_2}, \ldots, t_{jk_j}, x_{q_{k_1}}, x_{q_{k_2}}, \ldots, x_{r_{k_j}})}{(t_{1k_1+1} - x_{1k_1})(t_{2k_2+1} - x_{2k_2}) \cdots (x_{r_{k_j}+1} - x_{r_{k_j}})},
\]

in which, when \( t_l = t_{ik}(i = 1, 2, \ldots, l), x_i = x_{ik}(i = q, q+1, \ldots, r) \) and \( X = X_0 \).

When the constraint conditions comprise both the continuous and discrete constraints, the method of thinking is comparable to when the constraint conditions are all continuous. As such, this situation is not re-illustrated here.
Cumulative scalability. According to the fact that the relativities of constraint conditions differentiate them from each other, we consider two situations. First, suppose that all the constraint conditions are independent. Then the cumulative scalability can be defined as

$$C_i = \int_{X \in \Phi} f_i(X) dX = \int_{x_{11}}^{x_{12}} \int_{x_{21}}^{x_{22}} \cdots \int_{x_{n1}}^{x_{n2}} f_i(x_1, x_2, \ldots, x_n) dx_1 dx_2 \cdots dx_n. \quad (3.9)$$

From the definition, we can see that the volume of the cylinder surface surrounded by function $f_i(X)$ is the bound of $X \in \Phi$. When all the constraint conditions relate with $t$, we define its cumulative scalability as

$$S_t = \int f(X) dl,$$

where $l$ is the formula of curve $x_j = z_j(t)$ (whose geometric significance is to carry out the first form curve integral to the curve). Because

$$dl = \sqrt{dx_1^2 + dx_2^2 + \cdots + dx_n^2} = \sqrt{\sum_{j=1}^{n} z_j'(t)^2} dt,$$

we have

$$C_i = \int_t^{t'} f_i(X) \sqrt{\sum_{j=1}^{n} z_j'(t)^2} dt.$$

When the constraint conditions are constituted entirely by discrete conditions that are non-related, the cumulative scalability can be defined as

$$C_i = \sum_{x_1} x_1 \sum_{x_2} x_2 \cdots \sum_{x_n} x_n f_i(X) \Delta x_1 \Delta x_2 \cdots \Delta x_n.$$

When the constraint conditions comprise discrete conditions and all relate with $t$, the cumulative scalability can be defined as

$$C_i = \sum_t f_i(X) \Delta t.$$

For a more general case, when the constraint conditions comprise both discrete and constraint conditions, having both relative and independent relationships, the method of thinking remains the same as in the preceding analysis. As such, it is not illustrated again here.

(iv) Multi-dimensional scalability

On the basis of multi-constraint, one-dimensional scalability, we consider multi-dimensional scalability. This type of scalability applies to numerous evaluation metrics. The methods used for solving multi-objective optimization problems mainly include the linear weighted sum, main objective, layered sequences, ideal point methods, etc. [6], with each having its own applicable situation. With regards to our research questions, linear weighted sum solutions seem to be more workable, introducing the weighted parameters at the time of comprehensive consideration. We define the multi-dimensional evaluation function or static scalability of Internet architecture as

$$S = k_1 \cdot S_1 + k_2 \cdot S_2 + \cdots + k_m \cdot S_m, \quad (3.10)$$

in which $S_i$ represents the function of multi-constraint, one-dimensional static scalability, and $k_1, k_2, \ldots, k_m$ represent the weight of $m$ evaluation metrics, respectively. These factors can be
determined by objective factors, the evaluator’s own preferences, or general policy-making. Different weight setting will lead to different results. According to the above definition, we can set dynamic scalability and cumulative scalability on point $X_0$.

**(h) Next generation Internet experience in China**

**(i) China Next-Generation Internet project overview**

In September 2003, CNGI project was launched to empower the research community and industry in conducting research on the implementation of IPv6 in China. The project is supervised and coordinated by eight ministries, including the China Reform and Development Commission, Ministry of Industry and Information Technology, Ministry of Education, China National Science Foundation Commission, etc.

The overall structure of the CNGI project consists of five components: the CNGI demonstration network infrastructure, CNGI key technology development, CNGI software and equipment development and industrialization, CNGI applications demonstration and promotion, and basic research and standardization.

Almost all major Internet service providers (ISPs) in China have participated in this project. China Telecom, China Unicom, China Netcom (now merged with Unicom), China Mobile and China Railcom (now merged with China Mobile) have built their own IPv6 backbone networks based on IPv6/IPv4 dual-stack technologies. As a research-oriented ISP, the China Education and Research Network (CERNET) chose to build an IPv6-only backbone, (CNGI-CERNET2).

Under the support of the CNGI grant, there have been 59 giga-points of presence (PoPs) developed in Chinese IPv6 backbone networks, extending the IPv6 network to over 22 major cities. More than 270 access networks are connected to this IPv6 backbone. Two IPv6 international exchange centres have been established, the CNGI-6IX and CNGI-SHIX. CNGI-6IX was constructed by CERNET at the Tsinghua University in Beijing, while CNGI-SHIX was constructed by China Telcom in Shanghai. These two exchange centres connect IPv6 backbone networks comprising different Chinese ISPs with each other, and also connect Chinese IPv6 networks with IPv6 ISPs in the USA, European and Asia-Pacific regions.

**(ii) CNGI-CERNET2/6IX**

CNGI-CERNET2 (http://www.cernet2.edu.cn), one of the core networks of CNGI, was developed jointly by the network centre of the CERNET with 25 universities. CNGI-6IX (Beijing) is the international exchange point of the CNGI project. Today, CNGI-CERNET2/6IX is playing an important role in the development of China’s next generation Internet.

The CNGI-CERNET2 backbone runs the IPv6 protocol and connects 25 PoPs distributed throughout 20 cities in China, at a speed of 2.5/10 Gbps. Meanwhile, the transmission rate for Beijing–Wuhan–Guangzhou and Wuhan–Nanjing–Shanghai has been measured at approximately 10 Gbps. Each PoP provides the 1/2.5/10 Gbps access capacity needed for the access networks. The domestic/international exchange point is located in Beijing, and interconnects to the six demonstration core networks of CNGI: China Telecom, China Unicom, China Network Communications/CAS (Chinese Academy of Science), China Mobile and China Railway Communications. This connection takes place at speeds of 1/2.5/10 Gbps, and also connects with the US Internet2 network (www.internet2.edu) at a speed of 155 Mbps, the European GEANT2 network (www.geant2.net) at a speed of 622 Mbps, and the Asia-Pacific APAN network at a speed of 1 Gbps.

Since opening in 2004, the CNGI-CERNET2 backbone has connected more than 300 universities with research and development institutes in China and has undergone a number of technical trials and application demonstrations. It has been shown to provide an excellent environment for worldwide, next generation Internet research. CNGI-CERNET2/6IX (figure 3) has become the key fundamental infrastructure for China’s next generation Internet technology.
research, its key applications development and the development of core equipment for next generation Internet industrialization.

**Addressing and routing in CNGI-CERNET2/6IX**

CNGI-CERNET2 was modelled after the innovative concept of establishing large-scale, native IPv6 networks worldwide. CNGI-CERNET2 obtained the IPv6 address block of 2001:0da8::/32. This block was further allocated into access networks according to geographical location. Ten large cities were given a /36 block, while 12 small cities were given a /37 block. As stub nodes, each access network was given a /48 address block.

CNGI-6IX receives an autonomous system (AS) number 23911, while the CNGI-CERNET2 backbone receives an individual AS number of 23910. Each of the 25 regional networks is also allocated an AS number. The CNGI-CERNET2 backbone network runs eBGP4+ to exchange routing information with CNGI-6IX. The core routers of the CNGI-CERNET2 backbone communicate using OSPFv3 (within AS 23910). At the same time, these core routers announce routing advertisements for their customer network using iBGP4+. All access networks are connected to their regional networks using static routing. OSPFv3 is used to exchange routing information within the individual access network.

A large number of IPv6 research activities and experiments have been carried out, with results indicating that the costs to maintain and manage the IPv4/IPv6 dual-stack network are relatively high, while the network itself is not considered to be secure. Moreover, it seems that future work with network development cannot be completely divorced from the influence of IPv4, and that substantial development of an operational IPv6 network is needed if the IPv4/IPv6 dual-stack network is to be retained. At present, establishing a large-scale Internet backbone based on pure IPv6 remains a major global challenge. The CNGI-CERNET2 provides a series of solutions to key technical problems associated with network engineering technology for pure IPv6. These solutions include the topologies and routing design, IP address assignment plan and domain name system registration, network testing and measurement, and integrative network management.

**Security architecture and trustworthy network: SAVA**

In current Internet architecture, data packets are forwarded hop-by-hop to their destination address without any check into their source address. Therefore, it is considered unreliable to use IP source addresses to determine the origins of data packets. Network attackers can spoof IP source addresses to conceal their locations, or even impersonate other network users. To ensure that the source addresses of all packets remain reliable for network operators in diagnosing and locating failures and charge users, and to prevent or trace-back malicious attacks or misbehaving hosts, etc., the CNGI-CERNET2 is based on the Source Address Validation Architecture (SAVA) framework.
It is not to be expected that there will be a single mechanism applied at a single ‘level’ that can solve the source address spoofing problem in the Internet at large. Since the Internet is organized as a hierarchical architecture, it is also natural to consider organizing the SAVA mechanisms in a hierarchical way. Therefore, SAVA has been divided into three levels: (i) first hop with local subnet source address validation, (ii) intra-AS source address validation, and (iii) inter-AS source address validation. This division is demonstrated in figure 4.

Different levels of SAVA are granted different granularity authenticity for source addresses. At each level in the hierarchical architecture, one or more mechanisms are defined to address the problem of source address validation at that level. This particular hierarchy is chosen to maintain a balance between allowing as much freedom as possible for implementers and providers, while keeping the architecture as simple as possible.

(i) **First hop, local subnet source address validation.** Achieving an authenticity of host IP granularity is an important part of the architecture. If there is no special consideration for source validation of this fine granularity, one host can still spoof a source address by sending a packet with the ‘legal’ IP address of another host having the same IP prefix. This level should have the following characteristics:

- all the network devices are governed under the same administrative authority,
- the solution is in compliance with the address allocation and management policy of the local subnet, and
- the solution is in compliance with the end host’s means of accessing the Internet.

In the instance that many end hosts access the Internet through switches, the proposed solution involves creating a dynamic binding between a switch port and valid source IP address, or a binding between the media access control (MAC) address, source IP address and switch port.

(ii) **Intra-AS source address validation.** This is a simple part of the architecture, its goal being achievement of the authenticity of the IP address prefix granularity. All network devices are governed under the same administrative authority. The main idea of the proposed solution
involves building a filtering table that associates each incoming interface of the router with a set of valid source address blocks. Because the AS is under the same administrative authority, this filtering only needs to be deployed in the access router near the subscriber’s network. Ingress filtering is proposed as a solution.

(iii) Inter-AS source address validation. This is the most complex part of the architecture, its goal being the achievement of authenticity of AS-level granularity. This level should have the following characteristics:

— it should cooperate among different ASs having different administrative authorities and interests and
— it should be lightweight to support high throughput while not impacting forwarding efficiency.

Thus, the problem of source address validation has been divided into three parts. Different parts are given different granularities of authenticity for source address. In the edge network, fine granularity source address validation is deployed. In the core network, however, only AS granularity source address validation is deployed. These different granularities will prevent the source address validation mechanism from becoming a bottleneck of the network.

Tsinghua University promoted the foundation of a new IETF working group called Source Address Validation Improvement (SAVI). One RFC5210 on the SAVA testbed has been published [2].

In this model, the SAVI device monitors control packets sent by a host for a legitimate IP address, and then binds the IP address to a host (as specified by a particular link layer property of the host’s network attachment or binding anchor), filtering out the packets found to be inconsistent with the binding entry. Obviously, the implementation of SAVI would vary according to the IP address assignment method and binding anchor. SAVI can be deployed at any location to achieve different levels of granularity for validation and is designed to be purely network based. In other words, SAVI does not require the cooperation of hosts. In CNGI-CERNET2, we choose to deploy SAVI on all access switches between hosts in IPv6 subnets and their corresponding default routers, as this is the closest location to the hosts. This deployment is regarded to be the most effective deployment with the ability to provide the finest-grained source address validation. In other words, packets have to undergo IP source address validation even when exchanged locally on the link.

Our current SAVI implementation accommodates two legitimate IP address assignment methods: stateless address auto-configuration and Dynamic Host Configuration Protocol. The binding anchor is determined as the host’s MAC address switches with the port of the Ethernet, where the IPv6 host attaches. Currently, we are still working to improve the source address validation solution in the scenario that access switches cannot be upgraded to easily enable the SAVI function. The basic idea of the solution is as follows. We first analyse the network topology to identify necessary check points. The devices on these check points are referred to as key devices. Next, we collect information on the address prefix configuration of these devices. On the basis of this information, we can derive and configure filter rules automatically for network operators. We refer to this solution as Intra-AS SAVA [2].

According to our research, the ability to determine check points is the most important part of this framework. The selection of check points must satisfy the following requirements: (i) a packet with a source address of SAVI-enabled subnets is trustworthy, meaning that computers located within SAVI-disabled subnets cannot spoof SAVI-enabled subnet addresses; (ii) a packet with a source address for SAVI-enabled subnets can be reliably traced back to its corresponding host through SAVI-enabled access switches; and (iii) a packet containing a source address for SAVI-disabled subnets can be reliably traced back to its corresponding subnet. Figure 5 shows an example of SAVA deployment.
IPv4/IPv6 transition solution

IPv6 is not backwards-compatible with IPv4, meaning that intercommunication between IPv4 and IPv6 users is not naturally supported. Meanwhile, network users still rely heavily on IPv4 today because of the wide deployment of related contents, services and infrastructure. IPv6 cannot be expected to replace IPv4 overnight; instead, the two systems should be expected to coexist for a long period of time. Thus, it is critical to maintain both IPv4 and IPv6 network availability while achieving IPv4–IPv6 intercommunication during the transition period.

Current transition techniques supporting IPv4–IPv6 coexistence and inter-operation can be divided into two categories: tunnelling and translation. Tunnelling is used to connect IPvX islands across an IPvY network by means of encapsulation and decapsulation. (IPvX, IPvY) represents {IPv4, IPv6}.

---

**Figure 5.** An example of Source Address Validation Architecture deployment. (Online version in colour.)

**Figure 6.** Softwire structure. (Online version in colour.)
In contrast, translation achieves direct interconnection between IPv4 and IPv6 by converting addresses and semantics between IPv4 and IPv6 packets. Tsinghua University has promoted Softwire tunnelling techniques and IVI translation techniques. Additionally, these two techniques have been used in a series of IETF Requests For Comments (RFCs), such as RFC4925 [5], RFC5565 [7], RFC5747 [8], RFC6052 [9], RFC6144 [10], RFC6145 [11] and RFC6219 [12]. They have also been deployed on the CNGI-CERNET2.

Tsinghua University promoted the foundation of a new IETF working group called Softwire (as shown in figure 6), which is dedicated to handling the IPv4/v6 transition problem. The University also formulated a summary Softwire Problem Statement (RFC4925 [5], 5565 [7], 5747 [8]), which divided the problem into two categories: ‘hubs and spokes’ and ‘mesh’. More specifically, for the ‘mesh’ problem, we have specified the standardization of discovery, control and encapsulation methods for connecting IPv4 networks across IPv6-only networks, and for IPv6 networks across IPv4-only networks. This standardization will encourage multiple inter-operable vendor implementations, providing a solution to the compatibility, manageability, expansibility, dependability and auto-configuration of the transition from IPv4 to IPv6.

We propose an IPv4 networking-interconnection mechanism in Softwire frame called 4over6. Generally, our 4over6 solution is mainly composed of two parts: the control plane and the data plane. The control plane is based on border-gateway protocol and designed to advertise 4over6 tunnels and IPv4 network prefixes. It is also in charge of maintaining routing and encapsulation information. The stateless softwire tunnels are built between provider edge routers that exchange the IPv4 network layer reachability information by MP-BGP. The data plane uses standard IP encapsulation and decapsulation performed at IPv4–IPv6 dual-stack routers. We achieve IPv4 network interconnection by using routing transport on the control plane and packet transport on the data plane. As it avoids explicit tunnels and manual configuration, the 4over6 mechanism is lightweight, adaptive to dynamic routes and transparent for network end systems. It is designed for the purpose of interconnecting large-scale networks owing to its high scalability and ease of deployment.

Tsinghua University also proposed IVI (IETF RFC6052 [9], 6144 [10], 6147 [13], 6219 [12]), a systematic translation scheme. The IVI scheme consists of three types of address translation: stateless 1:1 translation, stateless 1:N translation and stateful translation. In stateful translation, an address is dynamically translated to a valid address from the available pool. In RFC6052 [9] specifically, a fundamental framework of stateless address translation is defined. The IPv4-translatable IPv6 address composes a variable-length network-specific prefix (NSP) from the ISP, with the original IPv4 address and suffix available to provide more information about the translation (illustrated in figure 7).

The underlying idea of this address translation scheme involves embedding an IPv4 address behind an NSP, allowing the routing information of the IPv4 prefix to be merged into an NSP. Thus, IPv4 routing information is effectively isolated from IPv6 routing information.

The stateless translation of IVI is an update of Stateless IP/ICMP Translation (SIIT) algorithm [14]. The packet translation process includes: translation between IPv4, IPv6 and transport layer headers; translation between ICMPv4 and ICMPv6 headers; translation between ICMP error messages; and checksum recalculation. The translation between IPv4 and IPv6 headers complies with SIIT, although source and destination addresses are translated according to the new mechanism described earlier.

IVI could run in the following modes: stateless 1:1, stateless 1:N and stateful. The first mode is used primarily in scenarios where the IPv4 address pool is sufficient for hosts in the site. In this case, one IPv6 address is assigned to a host that used to have an IPv4 address. The stateless 1:N mode is used when the quantity of IPv4 addresses is limited, thereby leaving one IPv4 address to be translated into several IPv6 addresses. The address is translated algorithmically instead of dynamically, resulting in low computational and memory costs when compared with stateful translation. Stateless 1:1 IVI and stateless 1:N IVI could be combined to form more complex flavours of IVI, such as dual IVI. For connections at the backbone, stateless translation
Figure 7. An illustration about the RFC6052 address format.

Figure 8. The coexistence and communication of IPv6 users with IPv4 users. (Online version in colour.)

should be deployed over the stateful approach based on its high throughput. Stateless IVI enables
the connectivity between IPv4 space and the component of IPv6 address space which is IPv4-
translatable. To enable the IPv4 hosts to connect with the entire IPv6 Internet, stateful translation
must be used. The stateful translation device should be deployed adjacent to the edge of the
Internet owing to its higher computational and memory cost.

Figure 8 illustrates Softwire and IVI implementation in CNGI-CERNET2. The 4over6 tunnel,
which includes the 4over6 initiator and 4over6 concentrator, is used to transit a portion of the IPv4
traffic flows over the IPv6 backbone. Since our IPv6 backbone can provide a superior quality of
service when compared with the IPv4, the 4over6 tunnel service is attractive for campus network users. The campus IVI translator is used for communication between IPv6 and IPv4 users on the same campus, while the backbone IVI translator allows IPv6 users to visit IPv4 resources in other campus networks or ISP networks. Note that the deployment of the 4over6 tunnel and IVI translator requires a special arrangement of IP address allocations and routing configurations.

(iii) Major applications in China Next-Generation Internet

Owing to the characteristics expected of the next generation Internet (scalability, high performance, real time, mobility, security, manageability and economy), many new applications that cannot be implemented in the existing Internet can expect to be supported. The CNGI project has supported the development of many new applications. These include

— large-scale high-performance grid applications over IPv6. The grid over IPv6 covers 20 universities in 13 cities, with an aggregation capability exceeding 150,000 times $s^{-1}$ and a storage capability exceeding 150 TB.

— 6PlanetLab-CNGI. 6PlanetLab–CNGI is a distributed, extendable and open-ended computing environment based on CNGI-CERNET2. This application connects with PlanetLab, the worldwide distributed computing platform.

— ngMaze: a P2P file-sharing system under a mixed environment. ngMaze, a P2P file sharing system which supports the IPv4 and IPv6 network environment, has 1 million registered users with more than 200 million indexes.

— WiFi/WiMax and SIP-based IPv6 mobile communication system. The IPv6-distributed multi-media communication system is based on WIFI/WIMAX, while SIP supports large-scale point-to-point video communications.

— Digital home. The home gateway based on IPv6 supports the unified control of distributed household appliances.

— IPv6-based IPTV. 3TNet, a high-performance broadband information network, is funded under the national ‘863’ High Tech Development Plan. This network is connected to CNGI-CERNET2, and provides high-definition TV programme access services with an image resolution of 1080i and 20 Mbps per stream.

— Large-scale video conferencing system. A large-scale high-resolution video conferencing system based on CNGI-CERNET2 is deployed at 25 universities in 20 cities, with an image resolution of $720 \times 480$ pixels and 30 Mbps per video stream.

— No-border art performance, remote medical services. No-border art performance and remote medical services have been implemented over CNGI-CERNET based on a high-resolution video conferencing system, which includes the Sino–USA violin teaching programme, drama performance and other cultural exchange activities, and remote observations of operations by Asia-Pacific medical fellows.

— IPv6 transportation monitoring. An intelligent traffic management system based on IPv6 supports real-time traffic data acquisition and video monitoring.

4. Summary and prospects

In the last 10 years, people have become increasingly aware of the importance, complexity and difficulties as well involved in the development of the next generation Internet. Developed countries have successively incorporated next generation Internet research as a key component of information technology research. In recent years, next generation Internet research has attracted significant attention from China’s government, and it is listed in the ‘Outline of the National Plan for Medium to Long-term Scientific and Technological Development’.

Faced with the new demands for innovative applications placed on Internet architecture in recent years, we propose that the next generation Internet must solve the following five key scientific issues: (i) the scalability and evolution problems of existing Internet architecture, (ii) the trustworthiness and convergence problems involved in large-scale routing, (iii) the
network transmission problems involved in sending massive amounts of data, (iv) real-time transmission problems of connectionless networks, and (v) complex autonomic network management problems preventing users from making cross-border visits.

We suggest that a ‘persevere with evolution, encourage innovation’ approach is necessary. In order to facilitate the evolution of IPv4–IPv6, we need to solve the major technical challenges inherent to an Internet based on the IPv6 platform, while actively participating in international standard setting. We will strive to make our research results into IETF international standards, to join the leading edge of global research, and to gradually form a theoretical system underpinning the next generation Internet architecture and protocols.

References