

The Technical Development Trend for Next Generation Internet

Wei Leping, Xu Jianfeng
China Telecom Corporation, Beijing, China

ABSTRACT

This paper discusses the technical development trend for next generation Internet including network scalability, network availability, network manageability, network security, IPv6 technology and QoS control technology. Finally, the overall design principles of China Telecom's CN2 network are presented.

Key words: next generation Internet (NGI), scalability, availability, manageability, security

Internet has become one of the most important information infrastructures in today's society. It is a bearer network for voice, data and video. Due to application popularization and commercialization as well as broadband development, Internet technology has become a bottleneck for future further development. Industry experts are exploring the development and transition issues for NGI. After studying a series of issues and challenges brought by the rapid development of IP service in recent years, we think the main characteristics of NGI should include scalability, high availability,

manageability, high security, direct addressing and sourcing, and relevant key technologies involve semiconductor, router design technology, IPV6/MPLS technology, routing protocol optimization and network management technology, QOS technology, security technology and policy.

I. NETWORK SCALABILITY

The network scalability refers to the capability of the network to scale to match rapid traffic growth. The network scale can be defined from two aspects: numbers of router node and link in term of routing calculation; and gross switching capacity of network equipment and link bandwidth in term of traffic capacity. The key technologies to realize and maintain network scalability include very large capacity routers, high-speed links, load sharing technology and routing stability technology.

There are many solutions for very large capacity routers. The most viable solution is a unified router architecture, also called router matrix technology or Multi-Chassis technology. In such cases, each router is composed of a switching matrix chassis and multiple interface chassis. All of the

chassis are internally connected by the router and controlled centrally by a management and routing control engine, and logically it is a router. Using low-cost light source VCSEL technology enables the cost of chassis connection to drop significantly from normal interface connection without using common interface cards and special ASIC and super speed storage chip array for implementation of QOS control, route forwarding tables, access control list (ACL) and statistics. Interface card chassis and switching matrix chassis are connected via parallel optical fiber with transmission bandwidth up to dozens of Terabit, overcoming the bandwidth bottleneck in using common ports for chassis interconnection, offering better capacity scalability of routers coupled with centralized switching matrix, and truly realizing super large capacity core routers with several Tbit/s and dozens of Tbit/s. The switching capacity of single chassis router developed with this method reaches 1.28Tbps. Switching matrix provides 250% of acceleration ratio. After employing the multi-chassis technology, the maximum switching capacity can hit 92Tbps in support of 1,152 ports at 40Gbps in theory. But the multi-chassis technology with such large scale requires very high reliability of equipment in practical application. Given the disastrous impact on network and service caused by single machine failure, the practical and economical feasibility still needs to be proven in the future. 40Gbps transmission systems still need several years to realize commercial application, and whether the existing optical fiber cables can support 40Gbps transmission or not requires a careful field investigation.

From a long-term development perspective, the speed of electric switching matrix will always be limited by device and microstrip processing techniques, power consumption and crosstalk, at the same time, the scale will be restrained by chip internal logics and pin numbers. Interface speed enhancement will also be impacted by the complexity of header processing. Moreover, continued growing route tables puts higher demands on

line speed processing and switching. In general, there is no optimal answer being found for the long-standing scalability of routers. In-depth researches are still underway.

Increasing network capacity via multiple equivalent links is the basic means to designing large IP network. Currently, link status algorithm-based IGP routing protocol is able to support the loading sharing of up to 16 equivalent links, basically satisfying the requirements for network scalability. Following iBGP routing protocol being introduced into routing reflectors, routing information is selectively forwarded, screening multiple equivalent route information, making BGP fail to utilize IGP routing information to realize loading sharing for equivalent routes and select the shortest route, causing imbalance in flow distribution and seriously impacting network scalability. In addition, MPLS, MPLS VPN and multicast load sharing technology also has some defects. They need further improvement to meet the requirements for large capacity and scalability.

II. NETWORK AVAILABILITY

According to the definition of network availability- "Percent IP service availability" given by ITU-Y1540, network availability refers to the ratio of data transmission duration to the total duration under a prerequisite of quality guarantee between network nodes. The key technologies impacting network availability include rapid routing convergence technology, fast rerouting (FRR) technology, software and hardware online upgrade technology, protocol graceful restart technology, reliability of equipment itself and availability of basic transport network.

The router availability is more than adding backup cards, but a design principle, which should be incorporated into product architecture from the beginning. The major improvement measures for hardware availability include transition from single switching plane to multiple switching plane in combination with support for N: 1 backup design; implementing separation between control plane

and data forwarding plane; and providing redundancy design and hitless switching. The major improvement measures for software availability include transforming to light kernel software, improving availability of operating systems. Software design is based on function module, which enable every software module to operate different protocols and service at various operation spaces, realize software modular online upgrade, and isolate impact produced by software module failures, therefore enhancing the stability and availability of software systems.

The key factors impacting fast routing convergence and FRR switching time are failure detection and identify technology with Bi-directional Forwarding Detection (BFD) protocol put forward by IETF as the key. Through periodically transmitting failure detection packet, BFD protocol can not only detect and identify interruption failures in transmission links, light interfaces and equipment ports but also detection and identify soft failures such as error and packet loss at transmission layers, data link layers, IP layers and application layers, just offsetting the shortcoming relevant to SDH which can only detect failures at transmission layers. BFD protocol fulfills failure detection for Ethernet links. BFD technology is an indispensable function of new generation router ports to check failures via hardware without impacting equipment performance.

BFD is often used in routers for detecting failures by transmitting UDP packet and triggering routing protocol calculation. Following employment of BFD combined with other technologies, the convergence time of large network routing is less than 500ms and FRR time is less than 50ms. But BFD protocol only defines message formats, which can be encapsulated into multiple messages to transfer such as Ethernet frame and IP/UDP/MPLS etc.. As a result, it offers a key technology for end-to-end failure detection at Ethernet link layers or VLAN, MPLS and LSP. Ethernet is the major aggregation technology for broadband access with tree

like topology. STP with switching time of 10 seconds can not satisfy the service requirements of voice. Through introduction of BFD to equipment such as Ethernet switches, DSLAM and BRAS to support port-based single hop BFD and VLAN sub-port based multi-hop BFD, rapid failover of Ethernet can be realized with milliseconds and BRAS, switch and transmission links can be protected. On the whole, BFD will become a major technology for failure detection.

In order to further improve network availability, IETF brings forward a series of graceful restart protocols relevant to ISIS, OSPF, BGP, LDP and RSVP. Graceful restart ensures normal operation of data forwarding plane and service offering in cases such as router control plane failures, software upgrading and primary standby switch. The graceful protocol restart technology is designed to guarantee service offering as much as possible with no impact on network stability or change in topology. If there are changes in network topology during protocol restart, control engines cannot conduct route calculation and updating in time, which may result in non-synchronization of network routing and produce routing black holes. Therefore, application scenario and design of relevant parameters in practical uses are quite important. Mainstream router manufacturers have offered support for related protocols, but interoperability tests are still underway.

III. NETWORK MANAGEABILITY

Network manageability mainly addresses control on service provision and service admission. Technical bottleneck is the standard model of management protocols and management objects. At present, network management protocols mostly include simple network management protocols (SNMP) and network configuration protocols (NETCONF). SNMP adopts UDP transport with strengths of easy implementation and mature technology, but incapable of meeting manage-

ment requirements in term of security and reliability, management operation efficiency, inter-operation and complex operation. NETCONF protocols use XML as data coding method for configuring data and protocol message contents, carry out transmission with TCP protocol-based SSHv2, implement operation and control with the simple Remote Procedure Call (RPC). XML language can express model-based management objectives with complex and internal logical relationship such as ports, protocols, service and their relationship, greatly increasing operation efficiency and objective standardization, at the same time ensuring reliability, security and interoperability via SSHv2 transmission. However, NETCONF protocols was drafted not long ago, establishment of management objective models is an arduous task and equipment supporting NETCONF protocols needs time. It will take 2-3 years for the entire technology to be fully developed. Generally speaking, NETCONF protocols represent the direction of network management protocol development, especially in equipment configuration and service provision management, while SNMPs will be continually used in term of data collection and failure alarm.

IV. NETWORK SECURITY

Network security is one of the largest challenges faced by today's Internet. The core of network security is to realize sourcing for application layers, network layers and physical layers as well as physical location for attackers. Sourcing is often performed by combining network layer and physical layer sourcing with the final goal of enabling physical layer sourcing. Sourcing is a security protection technology featuring after-event deterring. Currently, PSTN is capable of tracing sources. Similar DDOS attacks are very rare.

Application layer sourcing is carried out by building trust relationship between application systems, or adding network layer information into application layer protocols to transform into net-

work layer sourcing issues.

Network layer sourcing can be performed via the source IP addresses. Physical layer sourcing is based on PVC and VLAN to allocate a special PVC or VLAN to each customer, and realize positioning for user access physical locations. Employment of PPP dial-up or DHCP dynamic address allocation and enterprises Internet connection with NAT brings difficulties to network layer sourcing due to limited number of IPV4 addresses. In order to realize network layer sourcing and finally fulfill physical sourcing, the viable scheme is building a complete address allocation information pool combined with RADIUS recording information including bundled IP addresses and physical port information. Real sourcing cannot be achieved until application of IPV6.

To implement IP layer sourcing, we suggest using automatic addressing for IPV6 and pushing IP address prefix by operator equipment. The prefix is bonded with user physical access ports correspondingly, enabling IP addresses to be determined by operators. A similar process appears in PSTN, where telephone numbers are controlled and allocated by operators and physical tracking can be implemented easily.

V. IPV6

Employment of IPV6 eliminates address limitation in IPv4 and better supports mobile IP, bringing revolutionary benefits to service delivery and network operation and management. Firstly, IPv6 expands the address space to 128 bits from 32 bits of Ipv4, completely eliminating network barriers and communication barriers produced by Internet addresses, enabling end-to-end addressing and call at network layer, helping operator to extend its networks to enterprise networks and home networks. Secondly, IPV6 avoids dynamic address allocation and NAT application, enabling the network layer sourcing, delivering basic solutions for network security and cleaning out obstacles of NAT to service delivery. Thirdly, IPv6 protocols have been inter-

nally equipped with mobile IPv6 protocols, which enables mobile terminals to freely move between various access media without changing their IP addresses, creating conditions for seamless application of 3G, WLAN and WiMAX etc. Fourthly, IPv6 protocols simplify the management and maintenance of network nodes through a series of automatic discovery and automatic configuration, realize plug and play (PnP), help support mobile nodes and application of many small electrical appliances and communication equipment. Fifthly, a lot of new hot applications can be developed with Ipv6, such as P2P service, online chatting, online game etc. Generally, Ipv6 will become the convergence protocols of service bearer layer for the evolution towards NGN.

VI. QOS SERVICE CONTROL TECHNOLOGY

QOS, a critical enabling technology for NGI, is the key technology for providing network quality guarantee, network quality control, network resource allocation and network security.

Network quality guarantee is implemented through employment of light loading for critical services and DiffServ in combination with traffic engineering (TE) as simple as possible. The Internet traffic model conforms to Poisson distribution according to real-time testing results of Internet traffic. Guangzhou Research Institute of China Telecom conducted tests on new generation mainstream core routers with this model, and made the following conclusions: firstly, the quality of high-level service can be guaranteed even under congestion (150% of loading); secondly, the quality of service of a certain level can be ensured as long as the traffic dose not exceed the distributed broadband resources. Most of IP networks in developed countries were built many years ago without employment of QoS mechanism. Currently, they employ light loading to ensure quality, resulting in resource waste. The development and adoption of QOS technology will effectively enhance network utilization and lower costs. China Telecom CN2

network allocates 8 categories of DiffServ service, equal to dividing CN2 into 8 lanes. providing sufficient bandwidth to key service, ensuring bandwidth “over provision” for key service, filling the remaining bandwidth with Best effort service, enabling higher utilization of bandwidth.

Network quality control is an important component of network control. It is also the key factor in implementation of network layer differentiated service. NGI should realize packet loss ratio and packet loss methods control, packet disorder and packet delay control with manual provision according to different datagram messages, application types and service types. In doing so, real controllable differentiated services will be fulfilled, at the same time illegal application and illegal operation will be blocked. Access routers should provide the function of network quality control.

QOS service embezzlement refers to users enjoying higher-level services after modification of QOS class marks by themselves, and even launching security attack with high-priority traffic. As a result, QOS theft prevention is an issue to be addressed after allocation of QOS. Service classification and class identification should be conducted according to physical ports, which is the safest and most reliable. For example, the highest level service of CN2 network should be labeled with QOS service marks based on physical ports; in addition, service classification should be checked and re-labeled on service access control point equipment. DSL forum TR – 059 framework specifies that CPE should be responsible for the QOS classification of upstream services, labeling and control, while BRAS should be in charge of the QOS functions of downstream services and check QoS of CPE upstream service.

VII. OVERALL DESIGN PRINCIPLES OF CHINA TELECOM'S CN2 NETWORK

China Telecom's CN2 project is a service bearer platform built by China Telecom for next generation network and service. Its major design idea is to construct

a service platform with scalability, high availability, some QoS and security making full use of existing up-to-date technologies. The network is divided into 2 network function layers (including high speed forwarding layers and service delivery layers) and 4 network structure layers (including core layers, aggregation layers, edge layers and service access layers).

The principal characteristics of CN2 are high quality, multi-functions, high security and large capacity. The entire network employs fast route convergence with core layers using MPLS FRR; the whole network provides differentiated services, Multicast, MPLS VPN, and NTP. It features protocol graceful restarts and BFD function and adopts simplified service offering strategy.

In general, the construction of CN2 will lay a solid foundation for building a unified, higher quality service platform consisting of a component of NGN. It is reasonable to believe that CN2 construction will promote the overall development of China's NGN and the transition towards converged networks.

BIOGRAPHIES

Mr. Wei Leping, graduated from the Depart-

ment of E.E. of Tsing Hua University in 1970 in China and received M.S. of E.E. degree from China Academy of Posts and Telecom. Science, Beijing, in 1981. He has published over 100 papers and 6 books. Currently he is a member of Information Tech. Field Committee of 863 National High Tech. Program, fellow member of China Communication Institution and Chief Technology Officer of China Telecom Corporation.



Xu Jianfeng, graduated from NanJing Institute of Technology in 1990 in China and received M.S of Electronic and Communication degree from Southern China University of Technology. Since 1997, he has worked in Guangzhou R&D Center of China Telecom Corporation Limited and engaged in network technology research and planning and design.