Article ID:1007-6492(2000)02-0032-05

Intergrated Service Communication Model with IP

WANG Kai, ZHANG Jia-yi

(College of Electrical & Information Engineering, Zhengzhou University of Technology, Zhengzhou 450002, China)

Abstract: Future network is thought to be a unitive intergrated service network, and IP is a promising unitive network protocol. However, there is a flaw with IP itself that it can not effectively guarantee communication quality. So the improvement on IP is needed. IETF working group intserv and diffserv proposed IP intergrated service communication model IS and DS respectively. This paper analyses two models, and suggests that it is suitable to employ IS modle at the edge of network and DS model in the core of network.

Key words: IP; intergrated service; intserv; diffserv; IS; DS; RSVP; QoS

CLC number: TP 393 **Document code**: A

Introduction

Network, as the infrastructure of society information exchange, is developing rapidly. Its loads are changing from single voice or data traffic to multimedia information including voice, data, video. And current separately deployed single - function networks is evolving to a unitive network. Now people reckon that IP and ATM will be the alternative fundamental protocols of the unititve network. Considering that future communication terminals tend to be intelligent and IP has been being used in computer network for over 20 years, IP is more promising.

There is a research trend currently that "Every thing over IP", namely IP must guarantee transmissions of multimedia information. IP does well when data transfer, but not when voice, video transfer, and the maximum support traditional IP provided is that there is a TOS field in Ipv4 head[1,2] (currently most of network devices neglect it). Therefore the improvement on IP is in the face, and many organizations are striving for it.

Internet is the biggest IP network currently in the world. IETF working group intserv and diffserv has been working for realizing dream of unitive network in Internet.

1 Intergrated Service Model

Intserv proposed Intergrated Service Model (IS Model)[3], hoping the model can adapt Internet to multimedia information transfer by altering it.

1.1 Services Provided by IS Model

Intserv analyzed current applications's needs to the quality of service (QoS) and administration, and then proposed three classes of service in IS model:

(1) service for real - time application[4], in which Delay control is the emphasis;

(2) service for elastic application, namely our familiar best - effort service Internet now provids;

(3) service for resource - sharing (namely controlled - link sharing service)[5]. The service indicates how the resource should be allocated according to users' demands (e.g. according to different departments, protocols, applications). The service is provided for administration.

1.2 Two Requirements of IS Model

IS model suggests two requirements to fulfill three former classes of services. Firstly, Internet must be a common infrastructure supporting real - time and non real - time traffic. Intserv believes that carrying real - time and non real - time traffic on a uni-

Received date:1999-08-17; **revision received date**:1999-12-30

Biography:WANG Kai(1975-), male, born in Taihe County of Jiangxi province, master postgraduate of Zhengzhou University of Technology.

form infrastructure can give us the benefits on management and statistical resource sharing between real－time and non real－time traffic. Farther，there can only be one internet layer protocol，namely IP. Secondly Network resources must be explicitly managed. It follows the idea of call－connection control in traditional circuit－switching.，and then makes uses of resource reservation and admission control. Users must propose resource reservation request to network，if they want the traffics should be guaranteed. Receiving the request，network must consider its feasibility and make admission control to ensure the QoS of existing sessions. If all things are ok，network accepts the request and allocate a proportion of resources for it. However，if the user 's traffic exceeds his proportion，others may be disturbed. So network must condition traffics committed by users. Conditioning include：

(1) Classifying，divides incoming traffic into different micro－flows according to some characteristics of packets (e. g. source/destination address，protocol port number)；

(2) Shaping，regulates burst traffics into smooth data flows conforming to the contracts between the users and network；

(3) Policing，deals with excess traffic according to the contract (e.g. marking or dropping).

1.3 Considering Resource Reservation

Resource reservation，as signaling in telephone network，plays an important role in IS model，and it depends on reservation setup protocol. Intserv figures that a protocol must fulfill the followings：

(1) It must be designed for multicast environment radically. It is because that there are more and more group communications，and using unicast wastes bandwidth，but multicast is a perfect solution；

(2) It must meet different services needed by heterogeneous users. For example，assuming there is a group in which members 'connections are from 28.8kbps dialing link to 10Mbps LAN link，so the reservation setup protocol must cope accordingly；

(3) There must have flexible methods on controlling styles of reservation；

(4) It must meet dynamic changes of group members；

(5) It must be robust，otherwise resources may be lost and network 's ability goes down. RSVP is such a protocol[6].

1.4 Security Consideration

IS model also brings up an important administrative problem—identity authentication. Network must know which user can request reservation，and how much he can get. Network also must knows that an incoming packet belongs to the user as it alleges.

1.5 Defects of IS

IS is an elaborate model. It can provide good QoS to applications. But it is toocomplex，and can 't scale well. IS may behave satisfactorily in a small network，however badly in large scale networks. We can simply estimate as the following：Assuming network total node number is N_t，thereinto edge node (connecting with users) number is N_e. Currently assuming there are S_i sessions in edge node S_i ($i \le N_e$)，then total session number is $S = \sum_{i=1}^{Ne} S_i$. Assuming P_{ij} is the length of the path between edge node i and j，namely transitional node number in the path. Assuming the distribution of network sessions is well－proportioned，namely the probability that any session 's destination is edge node i is $\frac{1}{N_e}$，and assuming the cost for maintain a session state in a transitional node is C. Now total state maintaining cost is

$$C_t = C \cdot \sum_{i=1}^{N_e} \sum_{j=1}^{S_i} \sum_{k=1}^{N_e} \frac{1}{N_e} \cdot P_{ik} = \frac{C}{N_e} \cdot \sum_{i=1}^{N_e} S_i \sum_{k=1}^{N_e} P_{ik} (1 \le P_{ik} \le N_t) .$$

Then average cost in every network node is

$$\bar{C} = \frac{C}{Ne \cdot N_t} \cdot \sum_{i=1}^{Ne} S_i \sum_{k=1}^{Ne} P_{ik}$$

Commonly P_{ij} 's order is $\sqrt{N_t}$，we can make a rough computing

$$\bar{C} \approx \frac{C}{\sqrt{N_t}} \cdot \sum_{i=1}^{N_e} S_i = \frac{C}{\sqrt{N_t}} \cdot S ,$$

From the formula，assuming using same network node devices，if supporting sessioncount doubles，network node count must increase three times！In addition，per－flow management can consume lots of processor resources，moreover network core congregates large quantities of data flows, so network core node 's ability is challenged. It is then obvious that IS model can 't scale to the large network.

2 Differentiated Service Model

Diffserv believes that we should adapt Internet to different services through a comparatively simple and easy method. Considering Internet connections are provided by many (not single) ISPs now up to future, diffserv especially takes it into account that affording end – to – end service when spanning several independent administrative domain. Then diffserv proposed DS (Differentiated Service) Model[7].

2.1 Architecture of DS Model

Network consists of several DS domains in DS Mode. A DS domain is a relatively independent administrative domain, composed of a set of network nodes supporting DS. There are common service policies and definitions of PHB in the nodes. PHB, Per – Hop Behavior, is some forwarding actions corresponding to some special classes of traffics in the nodes. Usually a company 's intranet or an ISP network is a DS domain. DS domain must contract a SLA (Service Level Agreement) with the user, negotiating the QoS provided to the user. Among SLA, TCA (Traffic Conditioning Agreement) is one of the most important part, provisioning policies of classifying and conditioning the user 's traffic, and the user 's traffic specification, as peak rate, burst size, etc. Indeed SLA is namely a quality guaranteeing contract between service provider and the user.

Essential idea of the model is: network classifies and conditions incoming traffics and put them into different BAs. A BA, Behavior Aggregate, is a set of packets sharing same dealing in network nodes, and is identified by a DS codepoint. In practice DS codepint is denoted by TOS field in IPv4 or Traffic class byte in IPv6. DS domain core nodes dealing with packets only corresponding to their DS codepoint, not considering per – flow management as conditioning. In this way DS model scale well through pushing flow management, needing lots of memories and processor times, up to network edge, because the more close to edge, the less data flows.

DS functional model is composed of several function elements in nodes, including a set of PHB, packet classifying, traffic conditioning. Packet classifying assorts packets, according to some regulations, as DS codepoint. Traffic conditioning again includes metering, marking, shaping and policing. Thereinto, metering checks whether classified traffic conforms to the TCA. Marking marks packets according to management policies and the feedback from metering (for instance, whether or not exceeding the contract). Shaping regulates the profile of the traffic according to the TCA, commonly using leaky bucket algorithm. Policing copes with excess traffic, as dropping. Complex classifying and traffic conditioning process is done by DS domain edge in DS model. But it must be noticed that unlike IS model, classifying and conditioning need not correspond to every application 's micro – flow and DS domain core don 't maintain per – application – flow states, allowing PHBs compete for resources in proper granularity.

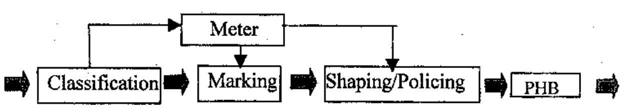

Figure 1 DS node functional model

2.2 Guaranteeing to End – to – end QoS

According to previous discusses, DS model guarantees QoS roughly, compared with IS model. Then can it provide end – to – end QoS? K. Nichols proposed Premium Service[8], trying to solve the problem. And basing their work, diffserv proposed RFC 2598 — An Expedite Forwarding PHB[8] (EF PHB).

Expedited Forwarding 's theory is as following:

In packet switch network, packet 's transmission delay, delay jitter is brought by queue in network nodes, excluding signal propagation delay. If we can reduce queues to the least, we can get the least delay and jitter. As we know, queue service rate must absolutely exceed packets arrival rate if queue occupancy is needed to be the least. So EF PHB provisions TCA between DS domain and users, using peak rate, and network nodes reserve enough bandwidth for it. DS domain edge strictly conditions and polices users 'EF traffics, concretely saying, for token bucket, token arrival rate equals to contracted peak rate, depth is one or two packets and drop excess packets. So the most size of EF PHB queue in every node, guaranteeing no packet losses, is in order of node 's fan - in degree which is commonly small in the core.

With previous description, we can know that every EF traffic rate is ensured, but jitter isn 't easy to control, especially for short packets. Assuming EF queue size is n packets (longest length is MTU), then the most jitter a packet can suffer is $n * MTU/\mu$ (μ is EF queue service rate).

Fortunately common EF PHB queue is small, but if the packet spans lots of node, the total jitter can be great. So EF PHB can only give us relatively rough end - to - end QoS assuring.

2.3 The Negotiation and Configuration Mechanism of SLA[10]

According to the time at which SLA is contracted, there can be two class of mechanism: static and dynamic.

With static mechanism, user delegates consult artificially with ISP delegates, contract the SLA before using. It suits long term、rarely changing and predictable service. With dynamic mechanism, namely through signaling, network devices negotiate with each other. [8] proposed that there is a BB (Bandwidth Broker), which manage the negotiation and configuration of SLA. Using signaling, as RSVP, users dynamically interact and negotiate with BB of local DS domain. In condition of spanning different domain ,which is more complex, local domain 's BB must negotiate with downstream domain 's BB, in turn and finally reach the destination domain.

Once SLA is signed, the domain must configurating corresponding nodes. For instance, the edge devices where user 's traffic comes in must be configured classifying and traffic conditioning according to SLA. With static method, network administrator can do configurations using network management protocol such as SNMP. With dynamic method, configurations are automatically accomplished by signaling.

In practice two methods can be used simultaneously.

2.4 Security Consideration

Like IS model, because of different service level and accordingly different charging, there must be some authentication mechanism.

2.5 Defects of DS model

DS model is relatively simple, but defects are obvious, namely QoS can not be effectively guaranteed. Someone thinks that future bandwidth will be enough, and simple classed service can suffice. It 's false, because QoS does not depend only on network bandwidth, it depend on the method of using bandwidth even more in some sense. And more, bandwidth increasing can 't catch up with the need. In practice, considering performance and cost, congestion in network is inevitable.

In addition, multicast is an important form of future network session. Unfortunately DS model now does not support it well. In essence DS model supplies QoS based on sender, however with heterogeneous user environment and dynamic multicast group, a method basing receiver is more effective, by which bandwidth could be utilized more effective and configuration of network is more easy.

3 Conclusion

So there exist different ideas about how IP network provides integrated service in computer network society. The essence of the differences is the style of managing and using resources. One side believes resources must be explicitly managed in exiguous granularity, while the other side thinks it is not needed. The former leans to connection - oriented technology, and the latter agrees connectionless one.

The authors consider both ideas have their reasons and limitations. The former is suitable for services need guaranteed QoS in the network which is obturated in administration, e. g. traditional telecom network. And the latter suits to non - strict - QoS - guaranteeing services in loosely connected open

networks (e. g. Internet). So we must take features of future information network into account when design integrated service model.
Future information network should have following characteristics:

(1) It can provide need – QoS – guaranteeing and non – QoS – guaranteeing service.

(2) Its administrative style is open. Current telecom network is a obturated network mastered by few telecom companies, and society can only passively use the service. In comparison Internet is an open network, having large quantity of users actively participating, and services in it emerge endlessly. Future information should supply a scalable network base, allowing all parts to take part in. In this way information service can just abounds and information network can evolves more quickly.

(3) Network terminals are intelligent.

If strict QoS of every application is needed, network must maintain every application micro – flow in theory. Considering network's scalability, opening and not all sessions need strict quality guaranteeing, it must adopt both virtues of DS and IS simultaneously. We propose the following adaptation model of DS:
The model is based on DS model. End users can request resources for every data flow needing strict QoS towards DS domain's BB, bandwidth broker, through signaling, just like in IS model. For local scale session, the BB does admission control according to conditions of local domain. If the session span several domains, BBs must negotiate each other, also can using signaling. Once passing the control, the source DS domain's corresponding edge nodes must be configured some state parameters, maintaining per – application – flow states for end users (DS core do not maintain these) and mapping flows into corresponding BAs. Like processing in DS model, resource – using contracts between domains do not aim at one end user's request at some time, negotiating in larger granularity.
In the adaptation model, per – application – flow state is managed at DS domain edge. Because forwarding performances of DS domain edge nodes are relatively lower then the core, doing so could improve the performances of applications needing some QoS guaranteeing level. And then the model gets better tradeoff between strictly resource managing and network scalability (See the figure 1).

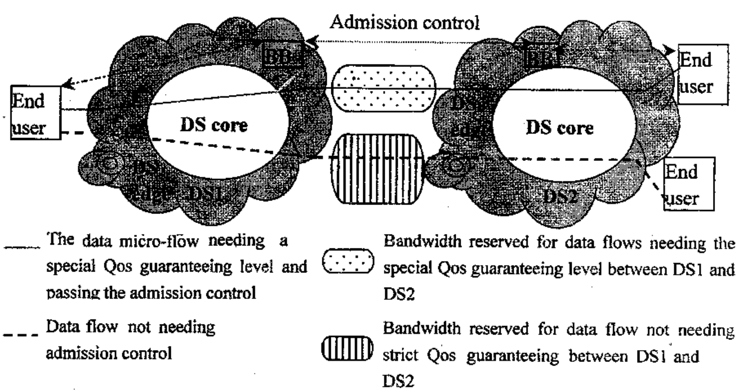

Figure 2 adaptation model of DS

References:

[1] RFC 791, Internet Protocol[S].

[2] RFC 1812, Requirements for IP version 4[S].

[3] RFC 1633, Integrated Services in the Internet Architectuve: An overview[S].

[4] RFC 2212, Specification of Guaranteed Quality of Service[S].

[5] RFC 2211, Specification of the Controlled – load Network [illegible] Service[S].

[6] RFC 2205, Resource Reservation Protocol (RSVP)——version 1Functional Specificafion[S].

[7] RFC 2475, An Architecture for Differentiated Services [S].

[8] NICHOLS K, JACOBSON V, ZHANG L. A Two – bit Diff erentiated Service Architecture for the Internet [R]. New York: IETF, 1999.

[9] RFC 2598, An Expedited Forwarding PHB[S].

[10] BERNETY. A Framework for Differentiated Services [R]. New York: IETF, 1999.