

Chapter 4.1

Integrated–Services Architecture for Internet Multimedia Applications

Zhonghua Yang

Nanyang Technological University, Singapore

Yanyan Yang

University of California, USA

Yaolin Gu

Southern Yangtze University, China

Robert Gay

Nanyang Technological University, Singapore

A HISTORICAL PERSPECTIVE

The Internet has gone from near invisibility to near ubiquity and penetrated into every aspect of society in the past few years (Department of Commerce, 1998). The application scenarios have also changed dramatically and now demand a more sophisticated service model from the network. In the early 1990s, there was a large-scale experiment in sending digitized voice and video across the Internet through a packet-switched infrastructure (Braden, Clark, & Shenker, 1994). These highly visible experiments have depended upon three enabling technologies: (a) Many modern

workstations now come equipped with built-in multimedia hardware, (b) IP multicasting, which was not yet generally available in commercial routers, is available, and (c) highly sophisticated digital audio and video applications have been developed. It became clear from these experiments that an important technical element of the Internet is still missing: Real-time applications often do not work well across the Internet. The Internet, as originally conceived, offers only a very simple quality-of-service (QoS), point-to-point, best-effort data delivery. However, for a real-time application, there are two aspects of the problem with using this service model. If the

sender and/or receiver are humans, they simply cannot tolerate arbitrary delays; on the other hand, if the rate at which video and audio arrive is too low, the signal becomes incomprehensible. To support real-time Internet applications, the service model must address those services that relate most directly to the time of delivery of data. Real-time applications like video and audio conferencing typically require stricter guarantees on throughput and delay. The essence of real-time service is the requirement for some service guarantees in terms of timing. In response to these demands of real-time multimedia applications, the Internet Engineering Task Force (IETF) has significantly augmented the Internet protocol stack based on the Internet integrated-services model, which is the focus of this article.

THE INTERNET INTEGRATED-SERVICES MODEL

An Internet *service model* consists of a set of service commitments; that is, in response to a service request, the network commits to deliver some service. The Internet is conventionally designed to offer a very simple service model, best effort, providing no guarantee on the correct and timely delivery of data packets. Each request to send is honored by the network *as best as it can*. This is the worst possible service: Packets are forwarded by routers solely on the basis that there is a known route, irrespective of traffic conditions along that route. This simplicity has probably been one of the main reasons for the success of IP technology. The best-effort service model, combined with an efficient transport-layer protocol (TCP [transmission-control protocol]), is perfectly suited for a large class of applications, which tolerate variable delivery rates and delays. This class of applications is called elastic applications.

However, demanding real-time applications require more sophisticated service models beyond the best effort. There has been a great deal of ef-

fort since 1990 by IETF to add a broad range of services to the Internet service model, resulting in the Internet integrated service model (Braden et al., 1994; Crowcroft, Handley, & Wakeman, 1999). The Internet integrated services model defines five classes of services that should satisfy the requirements of the vast majority of future applications.

1. *Best effort*: As described above, this is the traditional service model of the Internet.
2. *Fair*: This is an enhancement of the traditional model where there are no extra requests from the users, but the routers attempt to partition network resources in some fair manner. This is typically implemented by adopting a random-drop policy when encountering overload, possibly combined with some simple round-robin serving of different sources.
3. *Controlled load*: This is an attempt to provide a degree of service guarantee so that a network appears to the user as if there is little other traffic, and it makes no other guarantees. The admission control is usually imposed so that the performance perceived is as if the network were overengineered for those that are admitted.
4. *Predictive service*: This service gives a delay bound that is as low as possible, and at the same time, is stable enough that the receiver can estimate it.
5. *Guaranteed service*: This is where the delay perceived by a particular source or to a group is bounded within some absolute limit. This service model implies that resource reservation and admission control are key building blocks of the service.

The level of QoS provided by these enhanced QoS classes is programmable on a per-flow basis, and end-to-end QoS commitment for the data flow is built using a unified multimedia protocol stack and through resource reservation.

7 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/integrated-services-architecture-internet-multimedia/27121

Related Content

Video Coding for Mobile Communications

Ferdous Ahmed Sohel, Gour C. Karmakar and Laurence S. Dooley (2008). *Mobile Multimedia Communications: Concepts, Applications, and Challenges* (pp. 109-150).

www.irma-international.org/chapter/video-coding-mobile-communications/26784

Multiplexing Digital Multimedia Data

Samuel Rivas, Miguel Barreiro and Víctor M. Gulías (2009). *Encyclopedia of Multimedia Technology and Networking, Second Edition* (pp. 1023-1032).

www.irma-international.org/chapter/multiplexing-digital-multimedia-data/17512

Dimension Reduction Using Image Transform for Content-Based Feature Extraction

Sourav De, Madhumita Singha, Komal Kumari, Ritika Selot and Akshat Gupta (2018). *Feature Dimension Reduction for Content-Based Image Identification* (pp. 26-40).

www.irma-international.org/chapter/dimension-reduction-using-image-transform-for-content-based-feature-extraction/207226

Dealing with the Uncertainty of Satisfaction Surveys in Organizations That Employ Interactive Multimedia: An Analysis of False Answers Statistical Models through a Digital Music Library Case Study

Stelios Zimeras, Petros Kostagiolas and Charilaos Lavranos (2016). *Experimental Multimedia Systems for Interactivity and Strategic Innovation* (pp. 160-175).

www.irma-international.org/chapter/dealing-with-the-uncertainty-of-satisfaction-surveys-in-organizations-that-employ-interactive-multimedia/135128

Rise of OTT Bengali Originals: Niche Content for the Diverse Audiences

Lokesh Chakma and Mausumi Bhattacharyya (2024). *Exploring the Impact of OTT Media on Global Societies* (pp. 25-41).

www.irma-international.org/chapter/rise-of-ott-bengali-originals/340632