Chapter 14

Economic Aspects of Quality of Service for Internet Based IT Services

Jochen Wulf
Berlin Institute of Technology (TU Berlin), Germany

Rüdiger Zarnekow
Berlin Institute of Technology (TU Berlin), Germany

ABSTRACT

In academic research there is a long ongoing and vital discussion on the evolution and design of technologies that enable the management and control of quality of service (QoS) for Internet based IT services. Whereas technological aspects have been extensively discussed, from an industry perspective, there is a strong need for a structured research on economic aspects of QoS. A clear understanding of IT service quality and its dependency on the quality of data transmission is a precondition for an effective design of future service distribution technologies. This work provides an overview on the state of the art in economic QoS research. It discusses economic aspects of QoS on four central economic dimensions: the product dimension, the customer interface dimension, the infrastructure dimension and the financial management dimension. Moreover, this article discusses future QoS prospects and challenges.

INTRODUCTION

Internet based IT services have become essential for businesses and private customers. The constant increase in capacities for data transmission, storage, and data processing enable the provisioning of sophisticated software applications, multimedia content and computing resources from remote locations over the Internet (Internet service distribution). A key advantage of this evolution is the opportunity to realize higher operational efficiencies and economies of scale. The downside of this trend is a higher dependency of consumers on the Internet and particularly on data transmission quality. The original design of IP networks and protocols was not aimed at supporting IT services with such high real-time requirements and data rates. Additionally, the integration of wireless networks and the immense growth of wireless data services imposes significant challenges to network operators as one can see for example in the case of complaints of iPhone users about AT&T’s access
quality (Siegler 2009). Internet data transmission bases on the best-effort principle: all data packets are treated equally, regardless of their origin. The quality of transmission is therefore primarily determined by the transmission capacity provided by network operators. Capacities are regularly extended, according to a strategy which is oriented at the peaks of traffic utilization, referred to as over-provisioning. The actual quality of best-effort transport is therefore decided upon by network operators based on economic considerations. Complementarily to best-effort transport, a multitude of technological methods have been developed in order to improve the distribution of IT services over the Internet, few of which have up to date been applied in practice in what we refer to as content distribution business models (Wulf and Zarnekow 2010). The most prominent examples are content delivery networks (CDN) such as AKAMAI, which claims to 20% of total web traffic (AKAMAI 2009). In academic research there is a long ongoing and vital discussion on the evolution and design of technologies that enable the management and control of quality of service (QoS) for Internet based IT services. Whereas technological aspects have been extensively discussed, from an industry perspective, there is a strong need for a structured research on economic aspects of QoS.

In telecommunications research, the term quality of service (QoS) is used heterogeneously to describe concepts of service quality. As discussed by Gozdecki et al. (2003), the term QoS is used to describe the customer’s service quality assessment (IT service quality), technical parameters of service performance (technical service levels) as well as data transmission performance metrics (QoS in the narrower sense). Externally, i.e., in the relationship between service provider and user, service quality stands for the general comparison of inherent service characteristics as expected and perceived by users. Internally, i.e., among parties involved in service production, the service performance comprises all performance related parameters of a service described in technical terms (technical service levels), such as response and transaction times, availability and reliability. The parameters are affected by the inter-working of server, distribution, and client systems. The quality of