PROPOSAL FOR JOINT WORK

NETWORK OPTIMISATIONS

Stephen Cassidy, Dan Dahle
27 Oct 05

Introduction

We are moving from the ancient days of one network for each application, to a converged world where a huge and growing number of applications share the same IP network. This trend gives us large efficiencies and greater freedom to innovate. This freedom to innovate has resulted in a growing diversity of requirements on the network which it is increasingly unable to meet. This diversity of application requirements, and the wish to optimise for scale, pull in different directions. The one IP network is expected to deliver PSTN-like sessions, a host of tiny signalling messages, vast one-way file downloads, video stream delivery, peer-peer, fast reaction gaming; and to exhibit a whole range of as yet un-thought of behaviours. Whilst these applications were historically a small overhead in overprovided networks, the model of a single undifferentiated network did work – technically. However, as any or all of them may become mainstream – indeed it is our hope – building a network which can simultaneously deliver lossless small urgent messages in the mix with priority sessions, in turn mixed with vast media streams by this method is technically possible, but not commercially feasible today. Virtually any single product (hence pricing) model will favour one traffic type more than others – restricting the viability (unintentionally or not) of many of the innovative services we are trying to support.

For example, pricing by bandwidth will mean that the small valuable signalling-based services will never “punch their weight” in a network – restricting the viability of RFID or health monitoring applications. As another example, pricing by usage will make the commercials of bulk media very difficult.

From the technical perspective similar to the one network for one application, end to end control of services delivery meant end-to-end ownership of the network by a single service provider. Today with the mobility of customers literally requiring global service areas, end-to-end ownership of a network is no longer practical. This model is being replaced by a hybrid of networking technologies owned by independent parties. Internet access is now almost universal given the wide adoption and flexibility of Internet Protocols; however, this has been implemented across networks only for basic, best effort IP services. This is the case even though there are a number of standardized Quality of Service (QoS) optimisation technologies available and implemented in a number of different networks.

The technical challenge then is not to invent an end-to-end QoS technology, but rather to develop a networking model and solution that supports a generalized set of defined business needs by intentionally applying the standardized QoS optimisation technologies across the different network domains of the networking service value chain. What makes this challenge even more interesting is defining the model such that it can (and will) be adopted virally across IP networks.

It is likely that there is a discrete number of network level optimisations, where the network behaviour and commercials are tuned to support an open-ended collection of user applications, which will be more efficient. This will maintain the accepted mantra that the network will remain application unaware; i.e. that changes in the application can happen independently of the underlying network and that there is no "policing" of the network use for different applications; but rather will support fully open commercial terms for different service levels ordered.

The key question is: can a functional network model be defined to provide the right number of optimisations for maximising innovation, the number of successful business models which can be supported and how can these be communicated for consistent and reliable implementation across the network chains that form end-to-end services?

Aims

The work aims to determine the characteristics – commercially and technically – of the minimum number of network optimisations, which can initiate the ability to innovate successfully the commercial provisioning of advanced services in the market.

Methodology

There are three elements which need to be joined together.

1. understanding developing business needs
2. mapping these needs to the functions of the network (and intelligence features)
3. translating these functions into network solutions

To drive this, three traffic types are proposed as case studies. These are based on commonly understood network modes, and also correspond to recognised types of value proposition. This supports the aim of looking at the same time, but independently, the technology and business implications. Network intelligence features are also implied, and other capabilities such as cacheing etc. can be used to augment the basic examples examined. (The figures are examples and up for debate.)

Traffic types

1. "face-to-face" (or "voice-to-voice"): 2-way, low latency for voice/video collaboration/conversations. Latency<100mS, BW=64k or 1Mbit or 2Mbit
2. "file delivery": 1Mbyte, 100Mbyte, 10Gbyte files delivered such that the last bit arrives in 10s, 15mins, 1hr, 12hrs
3. "assured message": a short (10k?) message arrives within 1 sec with guarantee / acknowledgement

…and of course a "best efforts" class for lowest cost and to help early, open innovations.

Modelling
1. Develop the framework to be able to study the simultaneous presence of these traffic types, in varying ratios starting from basic, best effort IP service.
2. Optimise (where the criteria for “optimisation” need to be defined): Examples might include independence of the classes, network efficiency, deployability, simplicity of interfaces (apply Occam’s Razor to the interfaces requiring standardization).

Outputs

Sets of possible network models and deterministic optimisations, characterised by

1. Their support of existing main product types
2. support of example fringe services: an indication of the ability to support unknown future services
3. technical and commercial health assessments
4. testing for different traffic mixes
5. view of the resulting requirements at user/device/access/core/intelligence interfaces (commercially and technically)