A Bandwidth-on-Demand system case-study based on GN2 project experiences

M.Campanella , GARR, Via Celoria 16, 20133 Milano, ITALY
R. Krzywania, PSNC, Noskowskiego 12/14, 61-704 Poznań, POLAND
A. Sevasti, GRNET, 56 Mesogion Ave., 11527, Athens, GREECE

Keywords
Bandwidth on Demand, capacity reservation, multi-domain provisioning, end to end service

Introduction

The JRA3 Activity of GN2 project is heading to implement a multi-domain, heterogeneous technologies bandwidth reservation system for end users. During the design and implementation efforts of the last two years, project participants have come across a set of problems and decisions that are common for reservation systems. Having in mind the specific project objectives, supported by the knowledge and experience from similar efforts around the world, the JRA3 team has been working to identify the most suitable, and flexible approaches for inter-domain Bandwidth-on-Demand (BoD) provisioning issues. This paper presents the assumptions and decisions made so far, their background, pros and cons, and their consequences for the further project development.

Technology choice

One of the first decisions to make has been the target network technology for the final system. The choice usually depends on the current infrastructure used by project participants. The spectrum has been quite wide, starting from IP networks in Layer 3 of ISO/OSI model, going through Ethernet or SDH in Layer 2, and lambda management in Layer 1. The configuration can be also done using different signaling protocols of existing standards, like GMPLS RSVP. Although JRA3 is entitled to build a system that manages a heterogenous network environment, the main focus has been placed upon native Ethernet, emulated Ethernet and SDH transport networks. Despite of that fact, the modular architecture allows to extend the JRA3 BoD system, in order to accommodate other technologies, even at lower layers. It is also assumed that some NRENs wish to use their own NMS to manage their domains, and JRA3 system uses them to indirectly configure the network.

Domain privacy

Working in such a diverse environment, raises the issue of of domain privacy. As this is mostly relevant in multi-domain environments, the JRA3 system could not evade it. Two questions lie in the background: what information needs to be shared between domains, and how should the information be shared. As far as the first question is concerned, different approaches are followed by different projects. Firstly the domain topology may be very complex. Secondly the administrators may want to hide some technical or security details from other domains. The common solution is to use a network abstraction and define precisely what information can be advertised to neighbor domains. The JRA3 system assumes that a domain is represented as a graph of ingress/egress ports (as nodes) interconnected with virtual links, thus hiding the details of the internal domain topology. This topology information is constrained to technology agnostic information and time invariant parameters (like links' physical capacity, or reservations' costs). Certainly, this information is insufficient

to perform a reservation, but is enough to check the possibility of a physical connection between two end-points. The parameters required at the inter-domain level during the reservation process are provided by domains on demand. Still, at the intra-domain level, each domain defends its independence and privacy, being responsible for local resource booking and configuration.

Pathfinding

In order to investigate the feasibility of serving a BoD service request in an inter-domain environment, the first step is to define which is the set of domains involved in that particular BoD service instance. Therefore the process of path-finding is necessary. Widely varying approaches can be taken to find a proper chain of domains to successfully undertake a reservation. Starting from EGP like-approaches and looking for the next hoop on the way to the destination end point up to compiling an exhaustive graph based on large number of details about individual domains. After examining the alternatives, JRA3 adopted a model where the end-to-end path is computed on the source domain, based on publicly advertised information. Paths selected in this way are the ones that may physically provide the resources required for a reservation and are called candidate paths. The final parameters for reservation, technical parameters and configuration agreements are determined during the negotiation phase upon one of the candidate paths. The pathfinder module of the JRA3 system provides multiple candidate paths, sorting the list upon a chosen optimization criterion (such as the minimum path delay). It is the Inter-Domain Module (IDM) of the JRA3 system that is responsible to narrow the selection down to a feasible path that meets all domain constraints. This approach requires a protocol for the advertisement of inter- and intra-domain network topology and status. For this purpose the OSPF-TE protocol has been chosen, and implemented using the Quagga routing suite. Other approaches, such as BGP advertisements have also been examined and had their advantages evaluated and the final protocol selection is still open issue for JRA3. For path computation, after a first round of evaluations of available graph searching and routing algorithms like the k-Dijkstra or Eppstein algorithm, Junos routing implementations and more, the Dijkstra algorithm has been selected to operate upon the inter-domain topology graph created by OSPF-TE advertisements,. Experiences from the pilot implementation within JRA3 but also from similar projects are expected to provide useful feedback towards the final decision.

Failure handling

Once all pathfinding and negotiations succeed, the resources are booked and a reservation is ready to realize. At this stage, a problem to face is that of unexpected situations like network failures. Usually BoD systems face such events in two ways – ignore the problem and drop a reservation or try to come up with an alternative connection. In the JRA3 system the user may request the required level of resiliency during the submission of his reservation request. If no resiliency is requested, the system simply ignores a failure and informs the reservation owner about it. Otherwise, the system reacts to restore the connection as requested. An alternative path may be sought for or the traffic may be redirected to an already defined backup connection. These operations may take place either at the inter-domain level (including inter-domain communication, parameters negotiations, etc.) or at the intra-domain level, where a single domain silently takes care about critical failures. Despite of the solution chosen for a particular failure, the JRA3 transaction system is monitoring the overall resource booking process. A failure of the JRA3 system operating within a single domain, at any stage of a reservation process, can be faced and solved. This management of failures preserves the dynamic nature of network configurations.

Advanced reservations

As opposed to other BoD solutions, JRA3 reservations can be scheduled for any time in the future, instead of instant provisioning. This implies the complexity of provisioning in a limited-knowledge environment. Most BoD systems do not tackle this issue, as reservations are realized instantly, or resources are firmly pre-booked and can't be used by anyone until a reservation expires. The JRA3 system allows allocation of resources along the time axis by using a calendar module and solving overlapping issues. Scheduling is performed on both inter- and intra-domain levels, and guarantees availability of resources at the reservation time, unless unexpected events occur. Such events may be a network reconfiguration or failure, which are not predictable, but in fact are not very frequent.

Summary

The number of open issues and decisions made throughout the JRA3 system design and implementation is constantly growing. Going over the JRA3 conclusions and choices, while taking into consideration equivalent BoD systems, results in gaining valuable experience in building distributed infrastructures, not only for BoD service provisioning. Many of the issues presented here, related to global transactions and communications, resources provisioning and advance reservations, are common in multi-domain environments or scenarios involving interactions of independently operated systems.

Authors' Biographies

Mauro Campanella is working for the Italian research Network (GARR) as senior engineer.
He started working on computers and networks in 1984 as system and network manager and is participating since then to European research and developments in networking.
He is one of the creator of the Premium IP QoS an he is at present working on the definition and testing of advanced services for the GÉANT European backbone particularly in the bandwidth on demand area.
He holds a laurea in physics, has spent one year at CISCO systems between 1999 and 2000 and he lives north of Milano in Italy.

Radosław Krzywania received the M.Sc. degree in Computer Science – Software Engineering from the Poznań University of Technology in 2003. He is working in Poznań Supercomputing and Networking Center as a network applications engineer. His research involves network measurements, transport protocol enhancements for fast networks, and bandwidth on demand services. He is also interested in software design, development and software production management.

Afrodite Sevasti received her Diploma and MSc degree in Computer Engineering and Informatics, Patras University (Greece). She also holds an MSc degree in Information Networking from the Information Networking Institute, Carnegie Mellon University. She has worked as R&D Engineer for several years at the RA Computer Technology Institute (Greece) and currently holds the position of Network Services Development Manager at the Greek Research and Technology Network (GRNET). She has published 20 papers in refereed conferences and journals and participated in several R&D projects. Her main responsibility in GÉANT2 is the coordination of the Bandwidth on Demand provisioning pan-European initiative.

Full Paper

The full paper will be provided for the Terena 2007 conference.