

TRAFFIC CONTROL AND RESOURCE MANAGEMENT USING A MULTI-AGENT SYSTEM

Z. Luo, J. Bigham, L.G. Cuthbert and A.L.G. Hayzelden

*Dept. of Electronic Engineering, Queen Mary and Westfield College, University of London
Mile End Road, LONDON E1 4NS, U.K.*

{z.luo, j.bigham, l.g.cuthbert, a.l.g.hayzelden} @elec.qmw.ac.uk

Abstract This paper describes a system that implements agent technology for traffic control and resource management in a telecommunications environment. The adoption of agent technology has allowed more flexible and negotiable resource allocation management procedures, which are relevant to the more open telecommunications service environment. The paper introduces the agent architecture developed but concentrates on the resource agent strategies and service provider negotiation mechanisms for connection admission. The architecture is being validated on a real ATM test bed.

Keywords: Multi-agent System, Resource Management, Connection Admission Control, ATM, Auction.

1. INTRODUCTION

The aim of traffic control and network resource management is to provide the customer with the required Quality of Service (QoS) yet at the same time allow the network operator to run the network in an efficient and economic manner. *Connection Admission Control* (CAC) in an ATM network is the set of actions taken by the network at connection set-up, or during call re-negotiation, in order to establish whether a Virtual Channel or Virtual Path connection can be accepted or not [1]. The purpose of the CAC software is not only to find a free route across the network from source to destination, but also to *find a route with sufficient resources* to ensure that new connections obtain sufficient resources without degrading established connections. This is of fundamental importance in ATM networks and most of the approaches use traffic descriptors (such as average cell rate, peak cell rate, burstiness, source type) as an input to some algorithm that also has

knowledge of established traffic values. This is done on a link-by-link basis across the network.

This paper describes implementation of control strategies on a real ATM test bed as a society of interacting agents and use of the flexibility and brokerage capabilities of agents to provide additional functional and economic benefits. The addition of agent technology to the CAC problem brings a new dimension to traffic control and resource management in ATM networks by adding the *co-ordination between nodes* that is so far lacking in such systems [4],[5]. Benefits from using agents are expected to be as follows:

- allows different policies between user, service provider, network provider,
- faster integration of new services,
- better structure for providing open interfaces for network elements,
- applicable to ATM, IP,
- allows network optimisation,
- connection decision on entry (“one-stop shop”).

The paper is organised as follows; Section 2 will introduce a multi-agent architecture to support a set of agents that manage admission and re-connection. Section 3 details the layering of agent system, including the planning and reactive layers. Section 4 discusses service provider negotiation issues. In section 5, we present the developed agent software system. Finally, section 6 describes ongoing activities and future work.

2. MULTI-AGENT ARCHITECTURE

It is assumed that network resources will be managed by exploiting dynamic bandwidth allocation to Virtual Paths (VPs). A Virtual Path (VP) is a path of specified bandwidth from a source node in the network to the destination node in the network using physical links of the network. Note that only source to destination VPs are considered in the resource management model to be described. (In general VPs can be defined for segments of a path from a source to a destination, but these are not what are being managed here.) Extensions to very large networks makes the issue more complex, but suitable partitioning is a possible means of tackling that problem. No routing is done for individual virtual connections. Rather *all new* connections are allocated to one of the relevant VPs. The bandwidth associated with any VP can change continually and is one of the controllable parameters for the Network Service Provider or negotiation commodity for Service Providers who are not the Network Provider. This is in fact a highly realistic assumption for the management of a complex network.

It is assumed that the set of VPs associated with a source-destination pair are known, fixed in terms of route (though not bandwidth), and are a small

manageable subset of the set of possible VPs for that source-destination pair. Whilst this sounds limiting it is not believed to be so in practice as the set of enumerated VPs could be changed over time. Pre-enumeration of the VPs simplifies the CAC mechanism. Resource Agents (RAs) manage the VP connections. So the routing problem is not ignored but placed at a higher level (arguably where it belongs) leaving the CAC to request the appropriate Resource Agents information on the costs and feasibility of connections. Figure 1 gives a conceptual view of the problem.

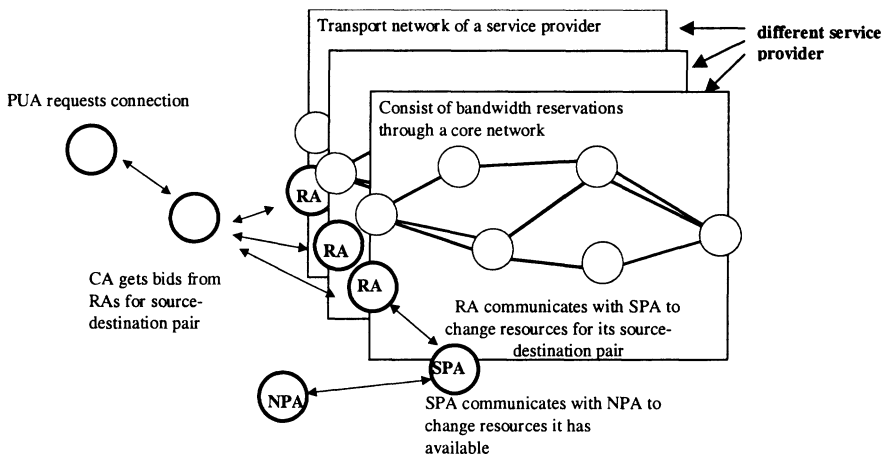


Figure 1 Basic Concepts

The main agents in the system are Connection Agents (CAs), Resource Agents (RAs), Proxy User Agents (PUAs), and Switch Wrapper Agents (SwWAs). Figure 2 shows where these agents would be located in the physical network. Note that control at network edge which is handled by PUAs, CAs and RAs can be done either in edge switches or in separate platforms.

In our model each SPA owns a RA for each source destination pair it services, and this RA manages the resources of the VPs that belong to the source destination pair. So with multiple service providers and multiple source destination pairs, there are many RAs. The user request a connection of a particular CoS along with the associated CoS parameter values via the PUA. The PUA contacts the CA placed at the entry point to the network. The CA queries the RAs for bids and gets the replies from relevant RAs then decides on the preferred service provider and the preferred offer from that service provider and instructs the chosen RA to install the connection. The

chosen RA then interacts with the SwWAs in order to make the necessary connection set up on the source and destination switches. The SwWA constitutes a "virtual" software abstraction of an ATM switch and its resources, and provide a generic, vendor-independent interface for network control and management applications. Interaction between the RA and SwWAs is performed by means of some Agent Language [3], while the SwWAs have access to the switch's proprietary control method (e.g. via SNMP or some form of API).

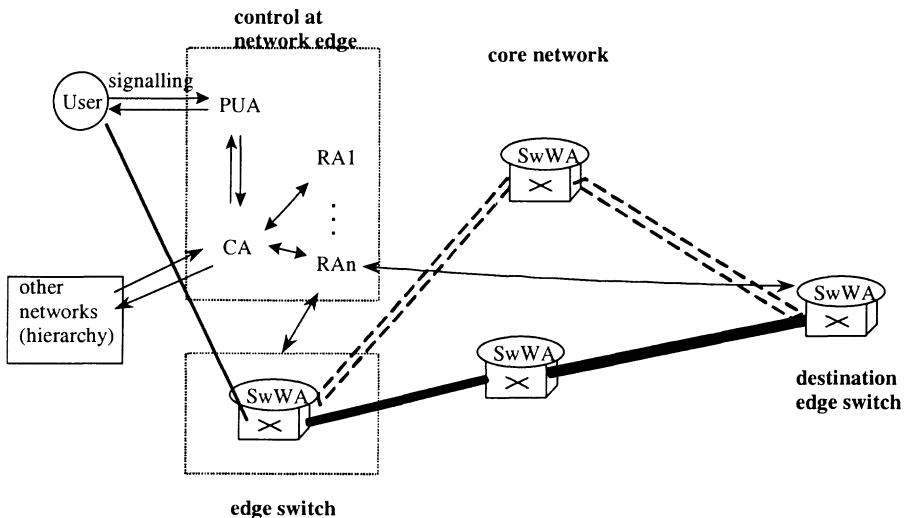


Figure 2 Multi-agent Architecture

One issue that is obviously of interest with real networks (but harder to demonstrate on the testbed) is the problem of scalability. The architecture considered by IMPACT addresses this issue because it allows for networks to be partitioned: a modified form of PUA called a Proxy Connection Agent (PCA) could be used to manage connections between networks instead of connections at user equipment.

3. RESOURCE MANAGEMENT STRATEGIES

RA, SPA and NPA are responsible for resource management. The user of the telecommunications network tries to make the best kind of connection from A to B at the best price, and the Service Provider marshals its resources

to meet customer demand and provide its contracted QoS with customers. A layering approach including a reactive and a planning layer is taken to implement resource management strategies. The RAs are reactive competences making rapid decisions regarding admission to the network. Speed and appropriate reactivity under a wide range of circumstances can be achieved by having levels of competence [8]. One of the key ideas is that each shell of competence can work with the inner shells working, but without the outer shell functioning – though not necessarily effectively over a wide range of inputs. Here the allocation competence in each RA works within a framework created by the slower planning competence hosted by the Network Service Provider (NSP) or Service Provider (SP).

For clarity we will call a SP that is not a NSP a secondary service provider (SSP). The NSP and each SSP has planning capabilities. We will only consider capacity bandwidth allocation here. The RA's can work without any input from the planning level. The planning level of a NSP or SSP monitors the state of the whole network it owns or rents. The planner changes the view of the RA's world by altering the capacities that they believe they have [5]. The relationship between the NPs and SPs and RAs is shown in Figure 3.

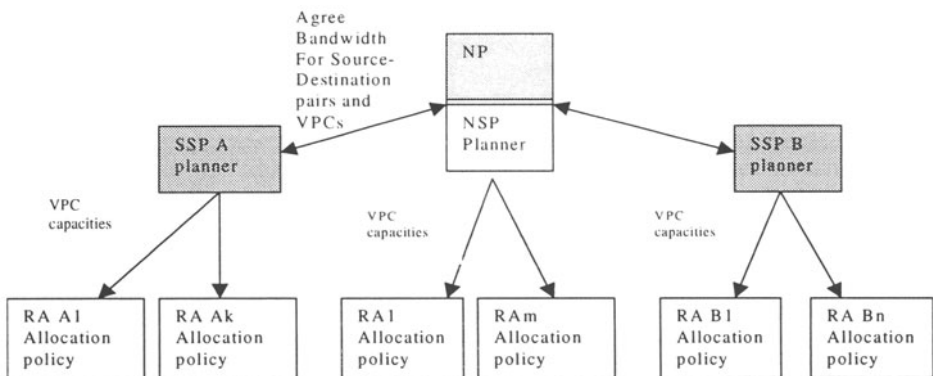


Figure 3 Competence Layers

3.1 The Planning Layer

When network flow indicates that the NP needs to optimise the capacities to allocate to each VPC of the RAs that it owns, it reads the current usage of each VPC it owns and the contracted bandwidth allocation to SSPs and computes an allocation of capacities that maximises the minimum residual in

the links of the network, tie breaking with hop count [6]. This is a conservative strategy and others could be employed. However, the robustness of such a simple criterion is appropriate for this level of planning and its purpose is to ensure overall survivability of the network in a competitive environment. The planning level sets capacities for the reactive allocation policies, i.e., gives each RA *a bandwidth allocation for each VP managed by that RA*. More details of the planning approach can be found in [5]. A similar approach is adopted for SSP planning.

3.2 The Reactive Layer

The reactive layer is embodied in the RAs and has to ensure second by second control of the network. This is distributed through the network and makes decisions based on local knowledge and the most recent input from the planner. It can work without the planner periodically updating capacities, though performance would degrade over time.

A reactive strategy appropriate for a NSP is to allocate to the VPC proportional to the planned capacities and tie break by allocating to the path(s) with maximum residual capacity. A strategy appropriate for a SSP is allocate to the VPC with minimum residual capacity and tie break by allocating to path(s) with minimum hop count. For very simple measures of utilisation, proportional allocation corresponds to minimising network utilisation. So, the first rule is appropriate for a RA that belongs to a NSP as it acts to support the common good of the NSP. The second strategy attempts to minimise fragmentation of bandwidth within the constraints imposed by the higher level planner or as a result of negotiation with the NP. This strategy is also relevant for an NSP at higher levels of utilisation. Such constraints include, for example, the number of physical paths used by the RA. This strategy is appropriate for a SP that is not a NSP. It leaves as large as possible chunks of bandwidth free so that it can accept high bandwidth connections without having to restructure internally or negotiate with other SPs. This is a self-interested strategy and certainly inconsiderate in terms of a common good defined in terms of utilisation or minimum residual capacity of the network.

4. SERVICE PROVIDER NEGOTIATION

Deciding how to model the negotiation between service providers and customers is difficult as there can be various degree of lock-in, where good substitutes prior to a commitment become less good substitutes later, and

different degrees of countervailing power between the interested parties when negotiating pricing. The intention here is not to presume the form of such agreements, but use such agreements as the basis of some of the bids in the auctions. We assume that the user has no user-defined procedures for managing the call for proposals but appeal to an open brokerage mechanism, currently a first price sealed bid auction. Allowing the user to upload customer specific auction code would introduce security and procedural issues, which may not be acceptable at the present time. Many would argue that such security problems can be addressed [9], but these arguments are not pursued here. Rather the user (as well as its ID and connection requirements) can provide parameters that allow specialisation of the announcer and the awarder in the auction mechanism (see Figure 4). The FIPA 97 Specification Part 2 [3] gives standards for call for proposals (a form of first-price sealed-bid), iterated call for proposals, English auction and Dutch auction. However, dynamic uploading of code by the SP is less problematic and has many attractive features, not only for the auction, but for procedures for bandwidth allocation by the RA.

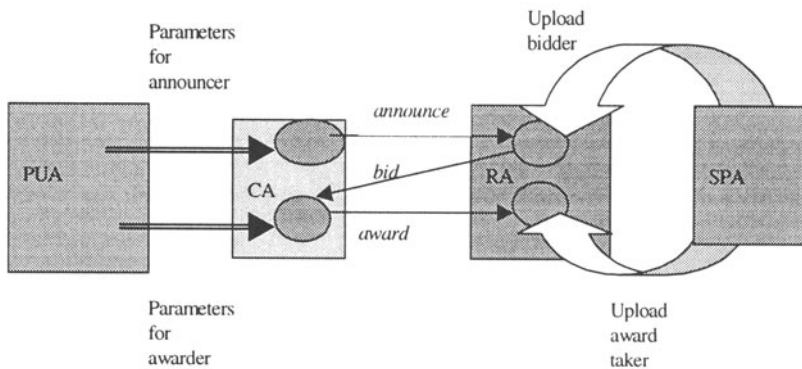


Figure 4 Auction Mechanism

In some cases the user-ID may only be used to classify the customer and bids made by the SPs would then be based only on the classification rather than individual terms. For example, the customer ID allows the most recent negotiated rates for a SP to be used so the prices quoted can be different from customer to customer as well as being history dependent for a particular customer. The awarder can be given a vector of price adjustments for named SPs, so that selection of the best price is done in a user provided context, even though the auction mechanism has not changed.

The more complex the negotiated terms the more complex the nature of the look ahead required by the PUA. Whatever the look-ahead problems, it is

assumed that the calculations map into a parameter set for the awarder, such as a list of (adjustment, service provider ID) pair.

It is worth emphasising that there is no attempt here to emulate 'negotiation' in the full sense of the word. In general, negotiation between Users and Service Providers may involve rather complex packages with 'bundling' and scale reductions. It is assumed that this kind of negotiation is done off line. Connection agent 'negotiation' with service providers will use the *results* of such negotiations. The former are conducted at connection request time and done on an individual connection request basis. The *scope* of the connection request time 'negotiation' could vary from announcements to all service providers (see later) to more constrained announcements to subsets of service providers, based on a user's existing longer term contractual arrangements with SPs, if any. The procedure a User wants the connection agent to apply on its behalf depends greatly on the extent of any such prior arrangements and will be, in many cases, private to the customer.

An important assumption is that bids are taken as binding and so if awarded a connection must be made. Problems could arise if after responding with a bid that exhausts our capacity, more profitable connection request announcements are received. The capacity constraints require them to be rejected. There is then a temptation to renege on the bid. If reneging is allowed then more sophisticated protocols are needed which include penalties for such reneging [7]. It is assumed that if such a temptation arises the SP will arrange to offload the award or another less profitable award to another SP, or will resist and honour the bid. In the latter case this could be because of regulatory penalties or, more subjectively, for good customer relations it would be irrational to do otherwise. Inter SP trade and cross charging is being investigated.

5. AGENT SOFTWARE SYSTEM

To support the multi-agent system described above a purpose-designed agent software system, called the basic agent template (BAT) has been implemented. BAT is written entirely in Java and relies on Java's Remote Method Invocation (RMI) for the agent communication and Java's Reflection mechanism for the implementation of the Agent Communication Language (ACL) protocol, see Figure 5. The implementation of specific agents for special tasks (such as controlling ATM switches or computing new load schemes for a network) is very simple.

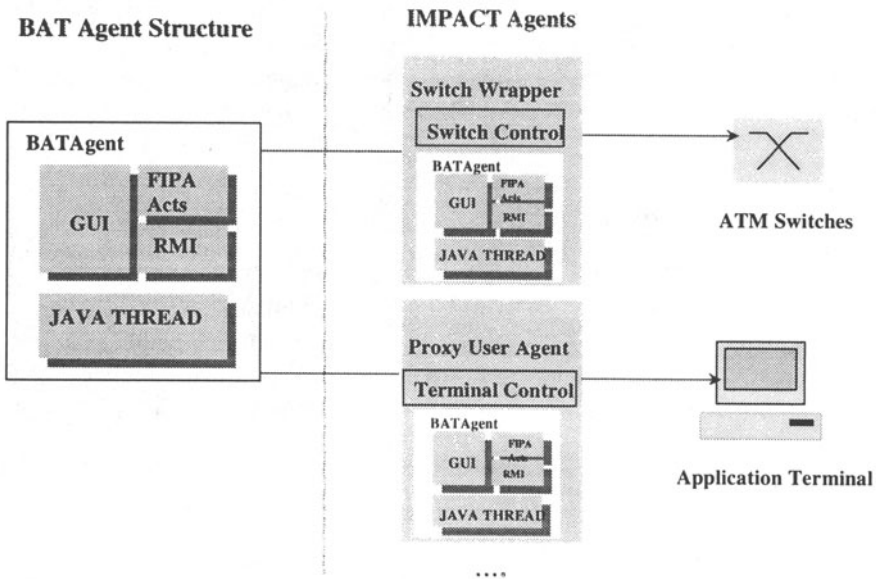


Figure 5 BAT Agent Structure

6. CONCLUSIONS

We have presented a traffic control and resource management framework using a multi-agent system. This new approach is expected to provide simple call set up procedures and better utilisation of the network through exploitation of the agents' intelligence. These expectations stem from the following features of the project's approach:

1. The CAC decision will be taken at the edge switch, thus alleviating the burden of executing the CAC procedure in every individual switching node in the network.
2. The agents will be engineered to take proper advantage of the provided resources using a more global network perspective.
3. Features of agents (such as autonomy, social ability, responsiveness and pro-activeness) will be exploited to achieve flexibility, which will be based in intelligence distributed in the network.

Although an ATM network is used as the basic infrastructure, many of the concepts are generic and the system can handle IP traffic equally well. In fact the overall concept is of a fairly dumb ATM core network with management-plane switching (cross connects) and intelligent edge switches

controlling resources. Such a network represents a structure used by operators around the world.

The prototype of the agent system has been designed and implemented. The trial of a set of scenarios on an ATM test bed is in progress.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge support from the European Commission under the ACTS Project AC324 "*Implementation of Agents for CAC on an ATM Test bed*". The authors also acknowledge valuable help and contributions from its partners Swisscom (the implementation of BAT system), Tele Danmark, National Technical University of Athens, Flextel, Teltec, and ASPA.

References

- [1] J. L. Adams, "Asynchronous transfer mode - an overview", BT Technology Journal, Vol. 13, No. 3, pp. 9-14, 1995.
- [2] M. N. Huhns, M. P. Singh, (Editors), "Readings in Agents", Morgan Kaufman, ISBN 1-55860-495-2, 1998.
- [3] Foundation for Intelligent Physical Agents, <http://drogo.cselt.stet.it/fipa>.
- [4] G. P. Kumar, P. Venkataram, "Artificial Intelligence approaches to network management: recent advances and a survey", Computer Communications, Vol. 20, pp.1313-1322, 1997.
- [5] A. Hayzelden, J. Bigham. "Heterogeneous Multi-Agent Architecture For ATM Virtual Path Network Resource Configuration", Proceedings of IATA 98, Lecture Notes in Artificial Intelligence, pp. 45-59, Paris, France. 1998.
- [6] J. Bigham, A. Hayzelden, "Using Survivability to Manage Uncertainty in a Telecommunications Network Management Application", Internal Report, Dept. of Electronic Engineering, Queen Mary and Westfield College, University of London, 1998.
- [7] T. Sandholm, V. Lesser, "Issues in Automated Negotiation and Electronic Commerce: Extending the Contract Net Framework", Proceedings of the International Conference on Multi-agent Systems, pp.328-335, 1995.
- [8] R. Brooks, "A Robust Layered Control System for A Mobile Robot" IEEE Journal of Robotics and Automation, pp.14-23, Vol. RA-2, No.1, March 1986.
- [9] E. R. Harold, "Java Network Programming", Published by O'Reilly, ISBN 1-56592-227-1, 1997.