

The impact of Active Networks on established Network Operators

A Submission for IWAN 99, International Working Conference on Active Networks, 30 June -2 July 1999 in Berlin

Feb 1999

Ian Marshall British Telecom

Stefan Covaci Deutsche Telecom

Thomas Velte Deutsche Telecom

Arto Juhola Helsinki Telephone Corporation/FINNET

Seppo Parkkila Helsinki Telephone Corporation/FINNET

Mike Donohoe Telecom Eirean

Abstract

A collaborative case based study has established that Active Networks will have a very significant impact on Network operators. Active Networking appears to be the only route to adding integrated mobility, security, QoS and management services to existing networks. In the short to medium term operators will be keen to use Application Layer Active Networking since the risks are relatively low. However the benefits of moving to stronger forms as research progresses appear compelling.

1 Introduction

One of the most serious operational problems facing large public network operators, is the difficulty of adding new features and technologies to their large installed network base. Since Active Networking was originally proposed [1] as a means of overcoming this very problem, established network owners such as the European Telcos are understandably extremely interested in using it to solve their problem. One result of this interest is that the BT, DT, TE and AF have collaborated in a 3-month strategic study examining the likely impact of active networks on network operators. The study was carried out under the Eurescom framework and is fully reported in the P844 deliverable. Eurescom is a Heidelberg-based co-operative research organisation of the member TelCos. Eurescom organises research projects in "pre-competitive" telecommunication study areas.

The main aim of the work was to establish the relevance of active networks for network operators and to determine what actions they should take to maximise the benefits. The operators are primarily interested in business impact within a five year planning cycle, but current research does not usually deliver business solutions within 3 years. The project team therefore chose to focus on business solutions that might emerge on a 3-5 year timescale and on the immediate research and development plans needed to ensure delivery.

We explicitly did not consider using active networks to solve problems, such as ipv4 -ipv6 migration, that are expected to be solved in other ways within 18 months nor did we address issues such as service interaction for which no complete solution is expected anytime within the next ten years.

Five usage cases were selected which illustrate the potential range of applicability of active networks and enable the elucidation of impact on the widest possible range of services and operational processes. The selected case studies range from extending the most basic network services (Mobility) through value add services (QoS Routing, Security) to operational support services (Management). The cases were analysed and evaluated separately to see what advantages the introduction of active network technology would bring compared with known "non-active" solutions.

All current active network proposals [1-8] were considered in the study. However, the project team felt that Application layer active networking [5] was the most immediately realisable proposal, and also that it carried the lowest level of risk. In addition it was considered relatively easy to introduce as an overlay, initially in small area of the network. ALAN was therefore prominent in the case studies.

In this paper we summarise the case studies and present the key conclusions and recommendations of the project regarding the impact of Active Networks on operators.

2 Case Study Results

2.1 QoS Routing (BT)

Modern networks must optimise the management of communication among nodes that are interconnected by diverse and alternate paths. These paths may be based on heterogeneous, technologies with divergent properties. This makes smart path choice an essential feature in order to provide quality services (QoS). For our purposes QoS can be characterised in terms of bandwidth, latency, security, strength of guarantee and uni or bi-directionality.

An interesting application that highlights many of the issues is the aircraft services application used in the COIAS project and illustrated in the figure. There are two basic services; communication with the flight deck and e-commerce/www access for the passengers. The aircraft (Node 1) has a 64kbit/s radio based bi-directional link to the control tower and a VSAT based downlink (2Mbit/s). The control tower can use wide-area connectivity to establish a high bandwidth link to the aircraft using ATM and Satellite combined. Path selection is performed at the Control Tower by examining the meta-information of the packets, deciding which is the most appropriate path and adding security if needed. There will also need to be QoS management in the plane to ensure life critical info from the flight deck is delivered into the bandwidth-restricted downlink before any traffic originating from passengers.

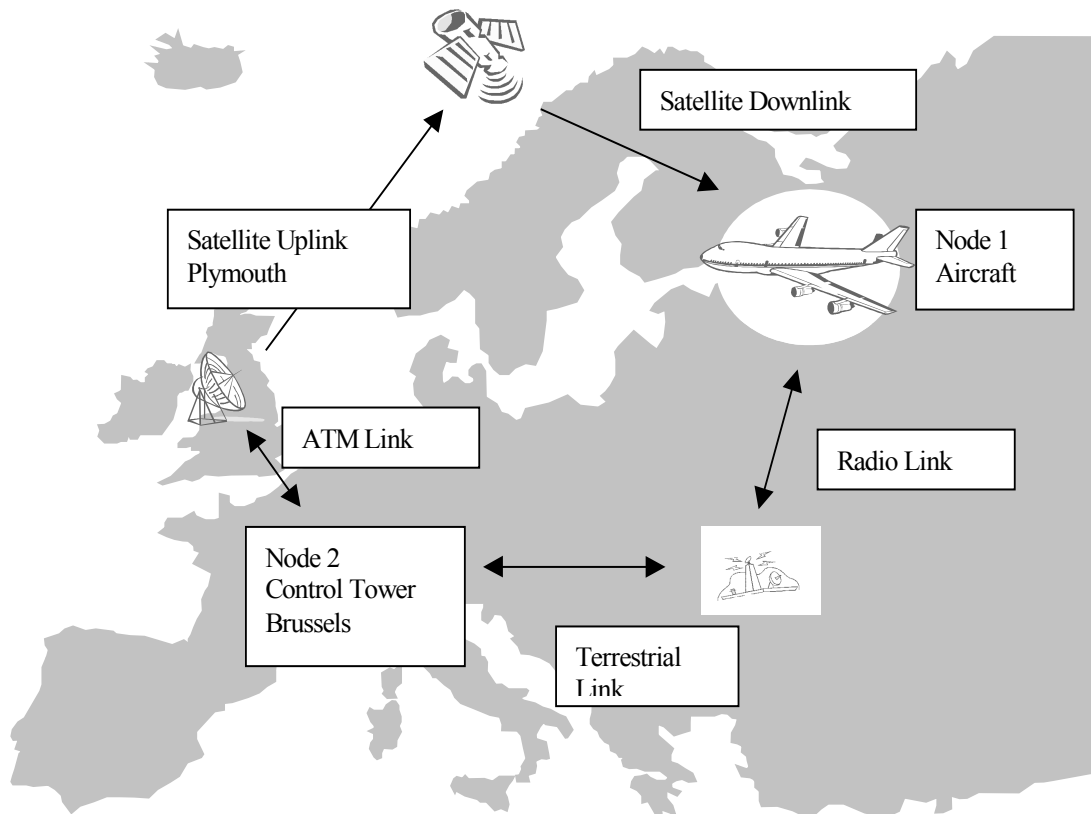


Figure 1. Aircraft application Alternate Path Routing Possibilities

2.1.1 Evaluation

Existing QoS routing schemes, based on load monitoring via the OSPF link state advert mechanism and diffserv packet priority markers in the DS byte, are only applicable within a domain due to the limitations of OSPF. Between domains the appropriate routing protocol is BGP. Even BGP4 (the latest version) does not explicitly provide a scheme to convey load information between domains (although user defined schemes could be used).

The same restriction applies to schemes based on estimating load using RSVP traffic since RSVP is to state intensive to deploy between domains on a large scale.

If an appropriate extension to BGP4 were available we would still be forced to use two network addresses in the application in order to enable return traffic to follow a different route as in the scenario above. Existing applications would need to be recompiled (and extended with a QoS router) for dual homed terminals.

The only viable alternative is some form of active network solution. The capsule approach was dismissed as impractical in the time frame recommended in section 1.3. We have therefore considered two cases; application layer active networking, and transport layer active networking using packet flags. Both approaches solve the basic problem with a low management overhead and good potential for scalability and ease of introduction. The transport layer solution could be thought ideal for this scenario as the added functionality is mostly transport layer. However, at the current state of language development any transport layer solution would be less flexible due to the difficulty of dynamically adding routing instructions to a router kernel safely (without degrading performance and security), and would not necessarily have better performance. The complexity in both cases depends on the complexity of the info gathering system required to assess remote QoS availability.

Table 2.1.1 summarises the evaluation of both approaches.

	Transport layer	ALAN
Utility	High	High
Complexity	Moderate	Moderate
Manageability	High	High
Scalability	Moderate	Moderate
Performance	Potentially high	Low
Flexibility	Moderate – restricted due to risks	High
Risks	High – i.e. not good	Low – i.e. good
Availability	5-10 yr.	18 months
Integration with existing nets	Moderate – must change router software	High – Very easy

Table 2.1.1

We have also highlighted the benefits of considering QoS routing in conjunction with security, mobility and a flexible management system. It seems that the benefits of active network solutions become even more compelling when all aspects of the problem are taken into account. Further analysis of the combinatorial benefits is given in section 2.4

We conclude that active networking is necessary for QoS routing and, that in the short to medium term, application layer active networking provides an acceptable solution.

2.2 Mobility (FINNET)

The case concerns terminal mobility in a broadband environment. It is foreseen that the applications residing in mobile broadband terminals are likely to utilise Internet IP-protocols when communicating with their peers. For mere roaming there are existing solutions that are not considered here. It is the handover/seamless roaming for hosts with wireless access and in motion that we have studied. In the first mobile networks there was no functionality for seamless roaming (between different network technologies). The adoption of this idea in the traditional mobile communication world is pretty recent, and even now the seamless roaming is usually limited to a certain set of network technologies. However, at the Internet-IP layer seamless roaming could be implemented completely regardless of the underlying network technologies. This is possible because Internet IP is not an actual network technology. IP is not designed to transfer data over physical distances via a defined set of media. Instead, IP acts as a "glue" capable of uniting disparate actual network technologies for the creation of an abstract "Network of Networks". This is meaning of the *Internet* layer.

In practice, middleware also uses the IP-layer for network technology transparency. If not, the network technology transparency will be sacrificed, provided by some other (more exotic) protocol or somehow implemented with the middleware (complete with routing etc. protocols). This last option is anything but elegant, simple, efficient and economical. It requires the use of the (very same) middleware by all applications and their components (real-time streams included). From the above follows that the Internet IP-layer is a very promising candidate for the implementation of terminal mobility services instead of the data-link layer (the UMTS/IMT200 etc. approach) or the upper layers (middleware/application). However, the current set of IP-layer protocols does not handle well functionalities required for the handover/seamless roaming.

2.2.1 Active Network based solution(s)

MANET based

One approach is to combine the use of MobileIP [9] and MANET (Mobile Ad-Hoc NETWORKS) [10] protocols. The MobileIP would handle the mobility between the access points of the fixed network (no handover) and the MANET would keep the connections alive while the end system is in motion. Since the MANET protocols are likely to remain in a state of constant change awhile, the ability of the active/programmable nodes to accommodate the latest improvements without service disruptions and yet maintain good performance would be an advantage.

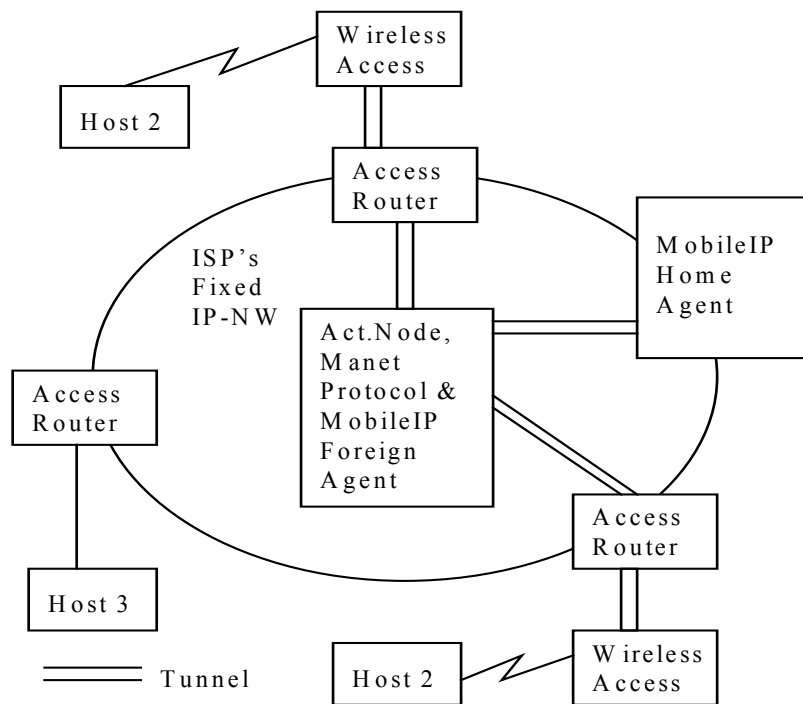


Figure 2.2.1, Handover with MANET

The above figure presents one possible solution. The hosts 1&2 route according to the MANET-protocol, and from their point of view the active node shown in the figure is just another MANET-host (the wireless-access traffic is tunnelled straight to and from it). Although it is capable of routing traffic to/from the fixed part of Internet. To receive traffic through the fixed network the hosts (1&2) may register themselves with the MobileIP Foreign Agent in the Active Node. The fixed-network traffic (e.g. from host 3) will then be routed through the MobileIP Care-of Address owned by the active node.

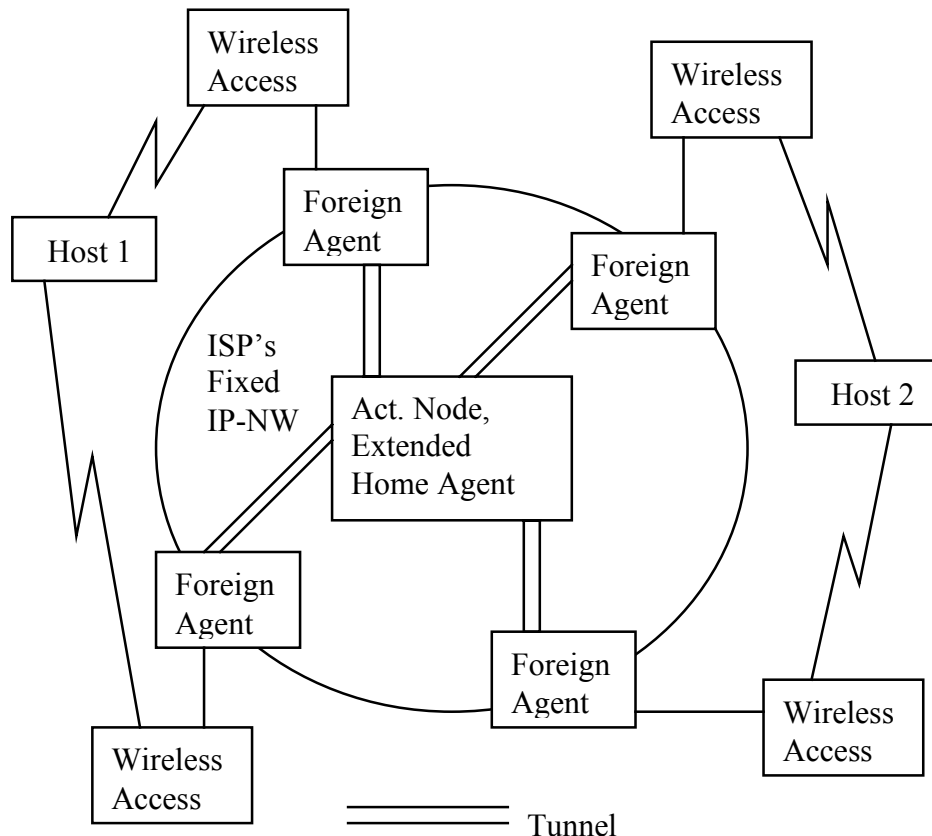
The traffic arriving through the fixed Internet and addressed to the host 1 or 2 arrives first to the MobileIP home agent owning the host 1 & 2 addresses from the point of view of the fixed network.

The home agent redirects the traffic towards the Care-of Address announced of the active node. The active node subsequently routes the traffic towards the hosts 1 & 2.

Extended MobileIP based

A second approach could be to try to "speed-up" the mobileIP Foreign Agent registration processes [11][12] and arrange functionality to prepare for the change of Mobile IP Foreign Agent in advance. The programmable (active) networks could be used to implement the necessary additional functionality.

What is required is an extension to MobileIP Home Agent capable of redirecting incoming messages between user's *several* concurrent Foreign Agent tunnels. By using concurrent foreign agent tunnels the traffic flow will remain uninterrupted even if the host is moving. Since the preparatory



registration can be performed with the mobility agents in advance, the redirection of actual traffic can be immediate.

Figure 2.2.2, Handover with extended MobileIP

There is also work going on to allow a user to dynamically perform the initial registration to MobileIP Home Agent services [13]. In addition, there is a proposal for diminishing the security-related overhead in mobile agent registrations taking place across administrative domains [14]. In short, there will be means for operators to offer Host Agent services for users, if not according to the ideas of [13] and [14], then by some further development.

The operators' extended Home Agents in active nodes could incorporate all the above ideas. The requests for traffic redirection between "live" Foreign Agent tunnels could originate from mobile hosts. The hosts should have the capability to handle at least two concurrent connections and IPv4 addresses (or network parts of address in case of IPv6) at any one time. OSs like UNIX (all variants) or Microsoft NT fulfil this requirement (although with IPv4 the link set-up and configuring time might be long).

As with the MANET case, the active nodes could form an "overlay" network. The figure illustrates this example. Hosts 1 & 2 have both 2 live tunnels between operator's Extended Home Agent service (an active node) and Foreign Agents (access routers) they are capable of reaching.

2.2.2 Evaluation

- Utility: With the presented handover solutions the users will achieve the freedom of movement for their hosts, regardless of the network technology (within the limits of that technology). Provided of course that the IP-level functionalities are present and reachable via the network technologies used. It is a possibility that active nodes could be used in the introductory & testing phases of the new handover software like MANET routing. It seems however that the MANET will reach maturity before the active node technology. Otherwise the presented handover solutions do not themselves benefit much from active networks. They facilitate the use of the "overlay" strategy in implementing many of the discussed combined services, however. And mobility bundled with these services most certainly does require active network technology. Therefore the evaluations below apply specifically to the situation where the mobility-combined services are the actual "active" services.

- Complexity: Neither of the above handover solutions will require complex additional functionality (omitting the intrinsically complex MANET routing protocols).

- Manageability: The management overhead related to the introduction of handover alone (in addition to the roaming management) is not considerable. It is possible to achieve fault tolerance through redundancy, e.g. by having several alternative active nodes available for access routers. The other management aspects are closely tied with the mobility-combined services in question, so we refer here to the appropriate chapters.

- Scalability: In the initial phase, there is no need to deploy excessive amounts of (active) nodes with MANET/extended mobile-IP support. When the capacity needs to be increased, the required performance is the only factor needed to take in account when planning the number of active nodes in both MANET and extended mobile-IP cases. Although global scale handover is a possibility with the extended mobile-IP, better performance is achieved when hosts register to nearest extended mobile-IP agent offering extended functionality.

- Performance: The user experienced performance of handover/seamless roaming could be evaluated against the fastness and the reliability of operation.

It is foreseen that a satisfactory system performance can be achieved at IP-layer. From the network point of view, the redirection (long-distance) traffic can be minimised. This is possible because the hosts can register to nearest handover-supporting nodes available. The performance of the mobility-combined services may fluctuate while a host is in motion (e.g. QoS), but this is the case with all mobile systems.

- Flexibility achieved: Since overlays are possible, changes can be introduced promptly to both service and management functionalities. The complex and recurring behaviour will concern mostly mobility-combined services, and we refer here to the chapters dedicated to other services.

- Fragility & risks involved: There is no need to introduce "active" code into the IP-layer. In case of MANET, the routing is executed by "rigid" software (which can be deployed with nodes with active capability). When considering extended mobile-IP, the redirection of traffic is essentially an end-system function. The need to change functionality and the related risks are dependent on the associated mobility-combined services. The appropriate chapters will provide more specific information.

- Migration strategies (ease of introduction): It is not necessary to introduce extended functionality into existing routers (just standard mobileIP in case of extended mobileIP). Overlay solutions are possible.

- (Dis)advantages when compared with non-active solutions: The complexity of the considered cases is not enough to warrant the use of programmable behaviour when considered in isolation.

The table below summarises the presented evaluation against the chapter 1.5 criteria. As with the text above, the bias is in considering the merits of the IP-layer handover/mobility in support of mobile-combined services, rather than the plain handover functionality itself.

	MANET	Extended Mobile IP
Utility	High	High
Complexity	Low (MANET protocol excluded)	Low
Manageability	High	High
Scalability	High	High
Performance	Moderate-high	Moderate-high
Flexibility	High	High
Risks	Low – i.e. good	Low – i.e. good
Availability	2-3 yr.	1-2 yr.
Integration with existing nets	High	High – Very easy

Table 2.2.1

2.3 Management (DT)

Network management systems of today rely on relatively rigid ways of handling associated behaviour. Once deployed, there are no smooth ways to alter the management behaviour of the network elements, so in this sense they can be said to be “passive”. This passivity is a major cause of the current interoperability problems.

2.3.1 Active Network based solution(s)

One solution is the NetScript scripting language developed at Columbia University. NetScript is an agent-based middleware for programming functions of intermediate network nodes. In particular, it is targeted to process routing functions and as such can be used to manage network nodes. The idea behind this is that there is already a basic way of programming complex network nodes. It is the use of scripting languages to configure the very complex MIBs of routers, which can incorporate thousands of variables. NetScript is therefore designed to offer interfaces to conventional MIB scripts. NetScript agents can be dynamically dispatched to and executed at a remote system. The NetScript architecture consists of Virtual Network Engines (VNE) and Virtual Links (VL). The authors of NetScript have outlined two scenarios for network management based on NetScript:

- Remote Network Monitor (RMON)
- SNMP Agents

A remote monitor allows an organisation to watch the performance and status of a network and protocols from remote location. RMON supports only a rigid set of operations, and requires that filters be written with low-level bit manipulations. By dispatching NetScript agents throughout the network the nodes can perform high-level filtering. NetScript can interoperate with, or even implement existing standards like SNMP. This is of special importance when a NetScript VNE resides on a network managed by existing non-programmable infrastructure. NetScript agents could

- Provide detailed access to MIB variables,
- Analyse and manipulate MIB status variables,
- Receive SNMP-requests to invoke the appropriate program to get and set a MIB variable.

Another approach to using programmable networks for network management is the „Policy Based Management“ of Imperial College, London. Policies are a means of specifying and influencing management behaviour within a distributed system, without coding the behaviour into the manager agents (i. e. when time > 1/June/1999).

There are authorisation policies and obligation policies. Authorisation policies specify what activities a manager is permitted or forbidden to do to a set of target objects. Obligation policies specify what activities a manager must or must not do to a set of target objects and essentially define the duties of a manager. Automated manager agents interpret policies and so the behaviour of the agents can be modified dynamically by changing policy rather than re-coding. Policies can be combined with directory systems, so that every node can be located with its related policies. The policies thus provide a constrained form of programming of automated agents to change management strategies without shutting down the management system.

2.3.2 Evaluation

Web based management seems to be closely related to „application layer networking“, because in both cases the IP infrastructure is used to provide networking at the application layer. The web based approach in combination with scripting languages enables active behaviour in network elements based on executable code that can move throughout the network. The same is true for policy based management and management with NetScript.

All of these approaches are quite realistic compared to other active network systems by two reasons: They can be restricted to the application layer, and the moveable parts of software can origin from trusted sources with well known, tested and reliable components. Moreover, as the code is mobile, it is possible to build some „Plug and Play“ functionality by moving the management information of a new network node to all relevant parts of the network management system.

2.4 Security (TI)

The biggest obstacle to the widespread use of the Internet for sensitive commercial business is the perceived lack of security. The resulting threat can arise not only from the casual hacker but also from more organised industrial espionage or crime. There’s an increasing number of security breaches in corporate data networks, though most of these may go unreported – assuming they’re detected. These breaches include

- The reading and/or modification of sensitive corporate data
- The interception of transaction data (e.g. credit card transactions)
- The destruction/adding of data or the planting of viruses
- ‘Denial of service’ attacks

The problem is exacerbated by the need for corporate Intranets to allow for (a) remote dial-in and (b) transactions across the public Internet. Exposing Intranets to the public Internet creates many opportunities for security breaches. Furthermore, there is no single system that can be implemented to counter all threats. Nevertheless, the use of the Internet for electronic commerce will continue to grow, creating a need for better security systems that can be implemented speedily at many nodes and that don’t add a sizeable additional processing overhead to transactions.

Another weakness may be the lack of a coherent security policy. Network managers and administrators must have a good overview of all potential users and implement policies that will give

the correct degree of security to all user types. Security measures implemented in isolation may fail to protect key areas and may open up new potential breaches.

Security management is set to become more complex according as IP networks grow and access nodes multiply. Any technique that would help reduce this complexity without compromising security would be welcome.

2.4.1 Active Network based solution(s)

Firewalls are undergoing constant upgrading and development. This is necessary in order to cater for expanding Intranets and new features being implemented. Some of the above mentioned security systems may be available in proprietary versions, preventing interworking between firewalls from different manufacturers. Newer standardised systems would need to be implemented to overcome this. It would be extremely valuable if some upgrades – such as enabling new protocols- could be carried out on-line. In the active network scenario, applications from approved sources could authenticate themselves to the firewall and then add the appropriate modules. It would also be simpler to implement new security systems and individual user/usage policies simultaneously over all firewalls in an Intranet/Extranet. In the case of authentication, multiple security systems may be operating independently at each protocol layer. The active network concept would permit the development of an integrated mechanism for all resources that can easily be modified and updated.

One of the most important aspects of authentication and encryption is the management and distribution of both public and private keys. A key management architecture is shown in Fig. 2.5.2. Many of the nodes contained within the management and distribution could be automated using active networks.

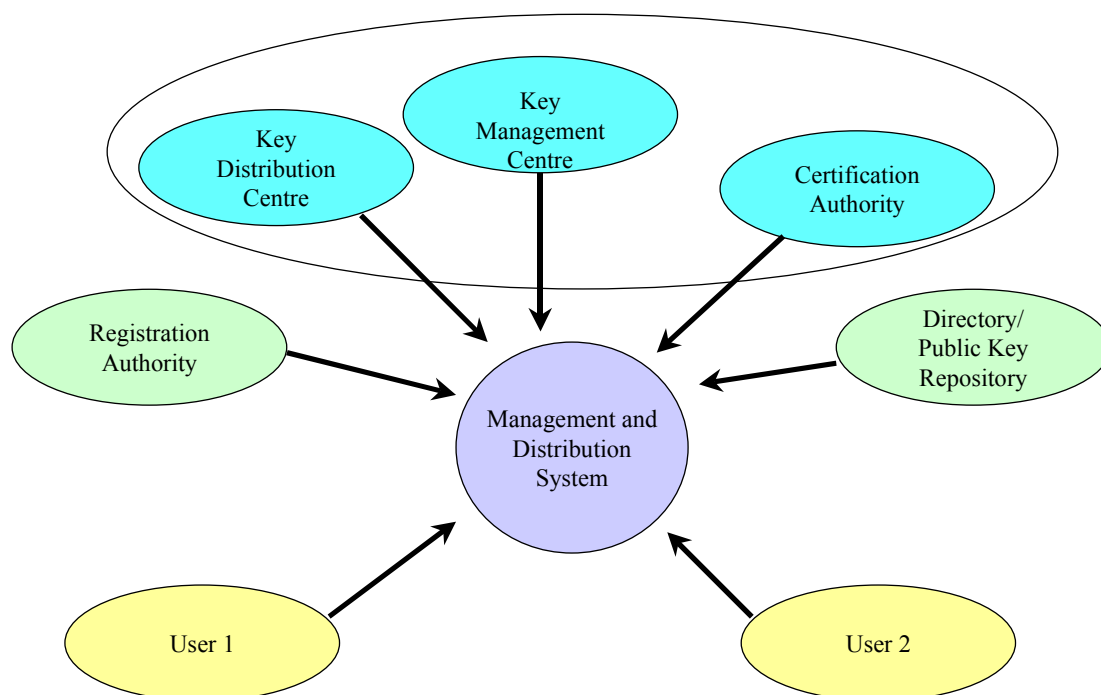


Fig 2.4.2 Key Management Architecture

For example, in the case of a shared trust CA system the process of updating certificate information across several CAs could be automated via active networks. The type of information that could be conveyed includes lapsed certificates, repairing security breaches and user status.

2.4.3 Evaluation

Existing security systems are many and varied. No single system can guarantee complete security and interworking is not always possible. Furthermore, sophisticated security systems introduce

additional processing overhead and hence greater potential delay in communications. This is currently the case for instance with advanced encryption methods. Active networks appear to offer the capability of having a more uniform security policy enforced across an Intranet coupled with live upgrades when necessary. However, there would be a need for some standards to guarantee interworking between equipment from different vendors.

The very act of putting more intelligence into routers may make them more vulnerable to security breaches. A programmable router could be more open to virus attack or unauthorised attempts at control. Careful consideration will have to be given to what active functionalities should be initiated and what should remain implemented as at present.

2.5 Benefits arising from combining the scenarios

As most of the scenarios alluded to additional benefits of using active network techniques when further requirements were taken into account, it was felt worthwhile to consider the application of active networking to a combined service. We have used the concept of a “dynamic managed Extranet” to combine the scenarios. The “dynamic Extranet” envisages a world where some companies are largely virtual – consisting almost entirely of highly mobile workers who are networked into several different companies for the duration of projects (which will typically be run as a collaboration between companies). This scenario adds a strong mobility flavour to conventional Extranet requirements. Hopefully readers can visualise for themselves many of the combined mobility, security, management and QoS features the scenario entails.

The project team’s view is that this service cannot be offered without some degree of active networking enabling programmes to move between end systems. For example, if a Java environment is assumed for user level tools, new information processing components enabling new working relationships, can be transferred and run during a session. With the addition of Jini join and discover techniques, and some local resource management processes the tools can safely be given access to local disks etc. However network QoS cannot always be maintained for the new tools without some modification of the local stack enabling appropriate QoS decisions. A specific case is that of content streaming which could use proprietary encryption, codecs and packet scheduling protocols, all of which should be implemented in the stack for efficiency reasons. The packet scheduling in particular cannot be implemented as part of the Java VM so the VM must support the use of a dynamic protocol stack, i.e. active networking.

QoS and security can track users needs even more efficiently if intermediates are provided in the network. This is because for mobile users return paths may not be the same as the initial path. If intermediate nodes understand the needs they can perform efficient alternate path routing with appropriate security. Clearly these nodes also need to be able to dynamically acquire and manage the protocols, policies and programmes that are mandated by the clients needs. I.E. for strong guarantees of good QoS strong active networking will be necessary.

The most obvious combination benefit then is enabling services that would otherwise not be possible, however there are other benefits, which can only be obtained if active networks are applied to all services rather than just a particular feature as in the other case studies. A list of some of the most significant which are not highlighted in the individual study evaluations, is provided below

- Combining active QoS routing and active management enables risks to be ameliorated by dynamic distribution of management policies
- Combining active routing and security with mobility enables performance to be maximised through path length minimisation

- Adding active management to reduce risks enables use of low level active code and further performance enhancement
- Combining active management with any active service enables the complexity of the management system to be reduced to only what is currently needed at any node by live services. Active features in an active management system can therefore have specialised management requirements if required without adding to complexity
- Active security and active mobility enables security to be maintained through handover between operators (e.g. when crossing national borders on European mainland)

Of course there are many other benefits, which there is not space to list here. However, we believe that there is enough here to illustrate the idea that active networks should ideally be thought of as a total service package, and not just a neat way of adding one or two smart new network features.

For completeness we have also evaluated the application of active networking to the dynamic Extranet service considered here in the same way as the features considered in the case studies. The table below summarises the comparison of active networks with conventional networks, using the chapter 1.5 criteria, for the dynamic Extranet. As might be expected Active Networks is clearly superior on grounds of Utility, Manageability, Scalability and Flexibility, and no worse on any criteria except risk. Further research should therefore concentrate on risk minimisation

	Active Networks	Existing networks
Utility	High	Low
Complexity	Low	Low
Manageability	High	Low
Scalability	High	High
Performance	Moderate-high	Moderate-high
Flexibility	High	Low
Risks	Moderate	Low – i.e. good
Availability	2-3 yr.	1-2 yr.
Integration with existing nets	High	High – Very easy

Table 2.5.1

3 Implications for the TelCos

The key aspect of active networking is the ability to support the deployment and execution of network programmes as and where required by operators and users. The key benefit will be an improved ability to penetrate lucrative Internet-based markets. Active networks also promise diminished network operating costs.

Some detailed implications of active networking for network operators are:

- Network traffic (revenues) will increase. This is due to inherent mobility (users can be "connected" everywhere), better service availability (no need to stop working/ consuming content because of lack of the right tools), and more services being available.
- Decreased network operating costs. With active networks, the configuration of the management system to reflect the current network can be made automatic. For example, with technologies like SUN Jini [15] a freshly installed network device can announce it's presence and provide management interfaces complete with interface semantics & "device driver" code to the network management system. Thus the management system will be less prone to errors arising from its internal network

model lagging behind the actual network configuration and from incorrect data entry by human operators, and operator intervention will be significantly reduced.

- Reduction of central management overhead. In order to simplify the management system network devices may be given greater autonomy via rule/policy based instructions that enable them to cope with errors, unexpected events, unusual demands, etc. Distribution of policy changes/additions is easily achieved using an active network mechanism. Thus, active networks can relieve central management load and de-centralise network management, with all the associated benefits (fault tolerance, efficiency).

- Simplified service management. Rather than operators being required to manage a huge range of Internet services they could simply offer a managed, virtual machine based, processing platform together with a library of service components. Users would then be free to compose and manage their own services to run on the platform, and operators would only need to manage the platform access service and the library usage service. Clearly this significantly reduces complexity and cost. It also radically improves time to launch for new services since the requirement for novel management development is minimised.

- Novel standardisation requirements. When the "virtual machine" of the managed system is standardised, any interested party can write the 'service' code. If the above idea is applied to an inter-operator environment, the operators in question must agree to "swap" code and rules for handling the code, between their systems. The interactions will be high level, of course. The "virtual machine" and the related management framework need to be standardised.

- Additional flexibility to facilitate fast service introduction & enhancements. The service in question should be of sufficient complexity to warrant programmability (low-complexity services might still be best realised with "rigid" technology).

- 3rd party development of value added services will arise. This does not prevent operator control related to the development if so wished, however (code authorisation).

- The range of services will increase (=application level services). This is a direct consequence of the programmability and 3rd party involvement - when everyone can develop services, rather than just those developers employed by operators, many more services can be developed in response to user needs. At present, without the capability to inject the service code "on demand" where it is needed, even fast paced updates of "rigid" network SW do not fulfil all the user requirements. In the future user needs are likely to become more varied and evolve more rapidly due to technology change so the need will be even greater.

- Possibilities to offer customised services will be improved. To support a high level of service flexibility it is crucial to use techniques that avoid undesirable interactions between the programmes. There are two main possibilities; use of a sandbox that eliminates any possibility of interaction, and extensive testing and monitoring of the programmes. The second possibility allows greater flexibility but at the cost of increased management overhead.

3.1 Research

Some outstanding research challenges the study identified in active networks are:

- Active network topology - deciding the location and number of active nodes, location change handling, configuration of the topology
- Protocols between active nodes and programs (routing etc.), measuring & diagnostics.

- The programming paradigm (languages, specification techniques, distribution model, co-operation models, frameworks).
- Methods to alleviate the feature interaction problem.
- Development of standardised proof for active network software. The proof of the properties of the active network software is a problem related to the satisfying the code safety and code/node security.
- Code development for active node Operation Systems - resource controls, scheduling, active kernels.
- Flexible hardware (firmware) in support of active nodes.
- Changes to operator workflows & processes (service provisioning, billing etc.).
- Management of Active networks, active FCAPS (Fault, Configuration, Accounting, Performance and Security management), Agents.
- Tools test & debug.
- I/O for Mobile Agents, Network aware applications and humans.
- Integration with existing systems

3.2 Marketing and Services

Active network dependent services may prove to be an invaluable marketing spearhead for TelCos, since they enable rapid deployment of 'value add' that is apparently part of the network, but which can, in fact, be sourced from 3rd parties. This will enable market response to be quickly tested, and market presence to be rapidly established with minimal development costs. There will also be real benefits to users compared to "plain vanilla" ISP services, which will tend to markedly increase network traffic. In addition there will be many opportunities for novel forms of commercial activity by both operators and their customers.

Some key market possibilities are highlighted below:

- Provision of active network services
- Service & code brokerage, directories for active programs & their sources (advertisement/searching/advising)
- Service "Operating System" leasing (processing time to be charged)
- Active network program distribution and storage (information logistics)
- Facilities management of third party services, e.g. advertising, access control, usage logging, billing, revenue collection

3.3 Infrastructure/Roll-out

Our initial view of the probable evolution is as follows:

- First generation: Active nodes at the edge of networks, code running in user-space and at application layer, 2 years from now.
- Second generation: Sparse overlay of fixed interdomain active nodes, some processing at kernel level, new routing technology at active nodes, 5 years from now.
- Third generation: Dense overlay, dynamic re-configurable/self-configurable topology, active firmware, new technology in all nodes (not replacing old technology but extending it), 10-15 years from now.

4 Summary

The results from the case studies indicated that the application of active/programmable network ideas is useful for some individual network features, and necessary in creating flexible wide area services such as the dynamic Extranet. There are also a large number of operational benefits to network

owners who deploy active networks. Although there is a need for much further research, the near-term technological possibilities and the revealed service provision prospects are extremely promising. We are therefore confident that network operators will be enthusiastic participants in the research and will also be willing customers of the results as the research matures. Early adoption is likely to follow the application layer active networking approach as this is both more immediately realisable and less risky than other alternatives

5 References

- [1] D.Tennenhouse, D.Wetherall, "Towards an active network architecture" Computer communication Review, 26, 2 (1996), pp5-18
- [2] Alexander, Shaw, Nettles and Smith "Active Bridging" Computer Communication Review, 27, 4 (1997), pp101-111
- [3] A.Lazar "Programming Telecommunication Networks" IEEE Network Oct 1997 pp 2-12
- [4] D.S.Alexander et al "A secure active network environment architecture" IEEE Network 1998
- [5] M.Fry and A.Ghosh "Application layer active networking" HIPPARCH '98 Workshop
- [6] E.Amir, S.McCanne, R.Katz "An active service framework and its application to real time multimedia transcoding" Proc SIGCOMM '98 pp178-189
- [7] G.Parulkar et.al "Active Network Node Project" Washington University St Louis
- [8] P. Cao, J. Zhang and K. Beach "Active Cache: Caching Dynamic Contents (Objects) on the Web".
- [9] IETF RFC 2002, Network Working Group, Editor C. Perkins, IP Mobility Support
- [10] IETF RFC 2501, S. Corson, and J. Macker, Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations
- [11] IETF Draft, Luis A. Sanchez, Gregory D. Troxel, Rapid Authentication for Mobile IP <draft-ietf-mobileip-ra-00.txt>
- [12] IETF Draft, Route Optimization in Mobile IP, Chales Perkins, David B. Johnson, draft-ietf-mobileip-optim-07.txt
- [13] P. Calhoun, C. Perkins, "Mobile IP Dynamic Home Agent Allocation", draft-ietf-mobileip-ha-alloc-00.txt, November 1998.
- [14] P. Calhoun, C. Perkins, "Mobile IP Foreign Agent Challenge/Response Extension" draft-ietf-mobileip-challenge-00.txt, November 1998
- [15] SUN Jini: <http://www.sun.com/jini/>