

Copyright © 2000 IEEE. Reprinted from (all relevant publication info).

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of MSF's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by sending a blank email message to pubs-permissions@ieee.org.

By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

Multiservice Networking using a Component-based Switch and Router Architecture.

Dave McDysan, MCI WorldCom, USA.
Torbjörn Lundberg, Telia Research, Sweden.
Nils Björkman, Telia Research, Sweden.
Alexander Latour-Henner, ServiceFactory, Sweden.

Abstract

Network operators have long striven to utilize equipment from multiple vendors to drive down costs. Furthermore, in competitive environments, operators have separated out functions of network elements and distributed them in order to differentiate the offered services. This trend began with deregulation in the telephone industry and promises to accelerate in the Internet age.

1 Introduction

The telecommunications world is rapidly changing and the critical success factors for operators are changing as well. In the past, much of the telecommunications market was driven by the market requirements and technology involved with voice services. In the latter half of the 1990s, the explosion of demand for services based on the Internet is a tremendous impetus for change. Furthermore, the demand for non-voice services, such as frame relay and ATM is also significant. Does the move away from a telecommunications infrastructure centred on voice toward one focused on the Internet and data require a revolutionary approach? Or, is an evolutionary approach that adopts the best from the voice architecture and combines it with a twenty-first century data architecture better? This paper asserts that lessons learned from decomposing telephone-switching systems are also applicable to new infrastructures for Internet and data services.

2 Applications of a Generic Switching and Routing Architecture

The Multiservice Switching Forum (MSF) has defined a generic architectural framework for next generation, distributed switches and routers [1]. This architecture decomposes related functions of a network entity into a set of functional planes, as illustrated in **Figure 2-1**. Starting from the bottom of the figure, ports provide bearer protocol interfaces to user devices or other network entities in the adaptation plane. A switching fabric (sometimes shown as an X inside a box in the figures in this paper) interconnects these ports. The ports and switching fabric operate under the supervision of a controller, which communicates using a control protocol with user devices and other network entities. At the top of the figure, an application plane contains processes that implement features and higher-layer services.

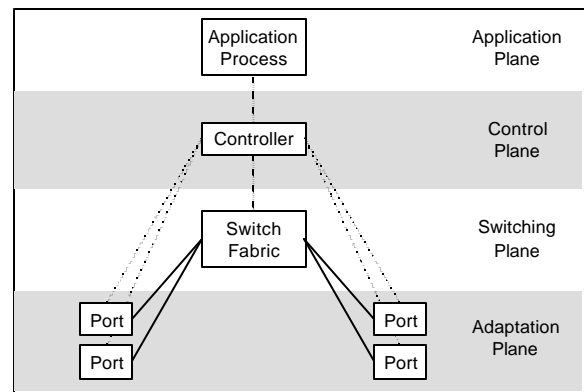


Fig. 2-1. High-Level Multiservice Switching Forum (MSF) Architecture.

The switches and routers of today's network are largely monolithic, that is, an operator must purchase such a machine from a single provider. The operator cannot easily exchange some part for another, and therefore cannot choose the best components to build a network. As bad as this may seem, the situation was actually worse in the past, as shown in **Figure 2-2** [2]. Beginning on the left-hand side, the stored program control switches were truly monolithic with all functional planes implemented in a single machine. However, this situation began to change in the mid 1980's with the development of and eventual standardization of Intelligent Networking (IN), as illustrated in the middle of the figure. Most recently, the industry has begun to decompose this architecture even further, as shown in the right-hand side of the figure. Here, the ports and switching fabric are called

a Media Gateway (MG), and a corresponding Media Gateway Controller (MGC). The IETF and the ITU-T are working to define a control protocol between the MGC and one or more MG's. The IETF version is called Media Gateway Control, MEGACO [5], while the ITU-T version is called H.248 [7]. These standards promise to deliver freedom of choice to network operators for deployment of MGs and MGCs from different suppliers. Furthermore, the logical separation of control from switching should allow more rapid and cost effective introduction of new services and features in comparison with the traditional monolithic approach.

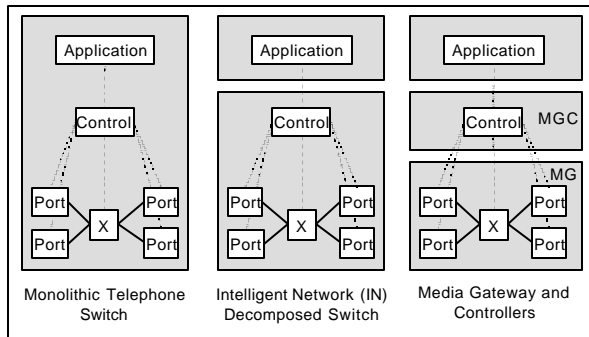


Fig. 2-2. Historical Decomposition of the Telephone Switch

Given that telephone switches and connection-oriented switches can have a decomposed adaptation, switching, control, and application architecture, an important question is whether the same concepts apply to the multiservice connection-oriented switches and connectionless IP routers. This paper describes how such a design could be implemented, but leaves the issue open regarding whether the marketplace will drive operators and vendors to pursue such an approach. **Figure 2-3** illustrates the application of this approach to the monolithic multiservice connection-oriented switches supporting frame relay, ATM, and circuit emulation offered by many vendors shown at the left-hand side of the figure. The principal standards activity involved in this architecture is the General Switch Management Protocol (GSMP), first developed by Ipsilon to allow routers to control ATM switches in response to recognition of long-lived flows of IP data traffic. More recent work by the MSF and the IETF extends GSMP to support not only ATM, but also MultiProtocol Label Switching (MPLS), frame relay, and circuit emulation [6]. It is too early to tell whether the applications will ever be separated from the controller, or whether operators will drive vendors to separate control from switching and port-

level adaptation. Section 3 describes several applications as motivation for this type of network. In part, this uncertainty is driven by the fact that the demand for connectionless IP routing based services is much greater than that for connection-oriented switched services.

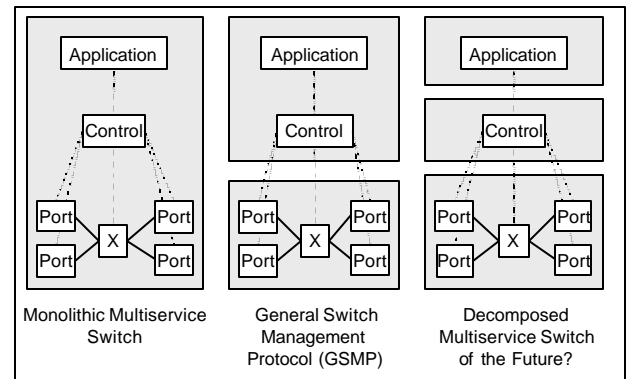


Fig. 2-3. Possible Decomposition of a Multiservice Connection-Oriented Switch

As demonstrated above, the generic MSF architecture applies to connection-oriented voice services as well as connection-oriented data services like frame relay and ATM. Can the same technique describe the architecture of connectionless IP routing based networks? The answer, as documented in [1], is yes. As shown in the left-hand side of **Figure 2-4**, a traditional monolithic router contains all of the same generic elements, but the functions they perform are different. Specifically, the ports implement IP-specific functions involving the packet header, such as longest prefix match, protocol type processing, and Time To Live (TTL) loop prevention. Furthermore, the connectionless controller runs routing protocols instead of signalling protocols that connection-oriented controllers do. Some decomposition of the traditional router has already begun, as illustrated in the middle of the figure using the Common Open Policy Service (COPS) protocol. Will the industry further decompose the traditional router as shown in the right hand side of the figure? One thing is certain; the protocol between a connectionless controller and the switch fabric and ports is not a straightforward extension of the aforementioned connection-oriented control protocols. Section 4 of this paper delves into this subject in greater detail.

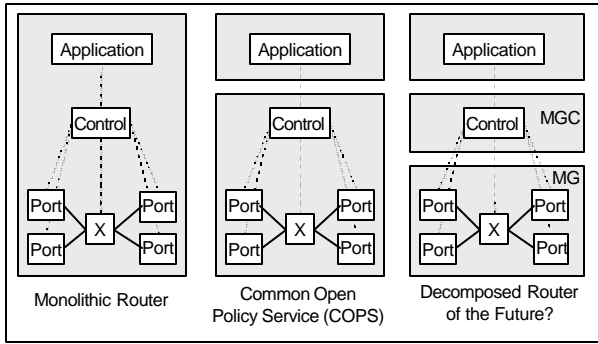


Fig. 2-4. Possible Decomposition of a Multiservice Router

Another further decomposition of the traditional switch or router is also possible, as illustrated in **Figure 2-5**. This type of decomposition has already happened in the personal computer industry, where the user can choose hardware from different vendors. The dynamic marketplace enabled by an open industry standard has driven down costs and given PC users freedom of choice to select best of breed components when assembling a computer system or network. The MSF architecture defines for further work an open interface between the principal components in a router or switch, i.e. switch fabric and the port cards, separately from the control logic. The objective is to empower operators to choose different suppliers for not only software, but hardware as well. This division targets increasing dynamics of the switch and router marketplace in a manner similar to that experienced in the PC industry. For example, new companies may arise that are specialised on some particular niche like producing cheap hardware, or providing attracting software.

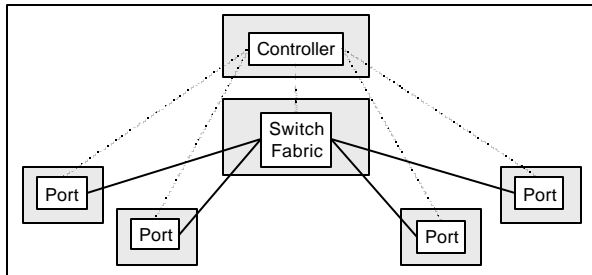


Fig. 2-5. Another Possible Decomposition of Switches and Routers

Furthermore, the MSF architecture describes how the hardware resources of a switch or router may be partitioned into virtual switches, or switchlets. These resources include, but are not limited to, physical ports, label space ranges on a port, broadcast capacity in a switch fabric, scheduling rates, forwarding tables, and buffer capacity. The first release of the MSF architecture defines the use of separate, deterministically defined partitions, each of under the supervision of a dif-

ferent controller. Partitions enable a number of interesting network solutions, for example the creation of virtual private networks running different control protocols over a common set of switching and routing machines.

3. Virtual networks

Construction of a virtual network from a set of network nodes whose resources are deterministically partitioned to a single customer and under control of a separate controller is straightforward. This design makes it easier to customise VPNs according to customer specific requirements as compared with the case where the VPN feature is implemented in a common network. The architecture is general in that a controller can further subdivide a partition amongst many customers. The whole network need not be updated with specific extra functions, they are only provided in the controllers associated with the partitions dedicated to each customer and therefore will not affect other parts of the network. Because of the strict resource-control that exists between different partitions the different virtual networks are strictly independent. Customers who earlier required their own dedicated networks can now trust public network solutions, which is much cheaper than if the customer built his own network with dedicated nodes and private lines.

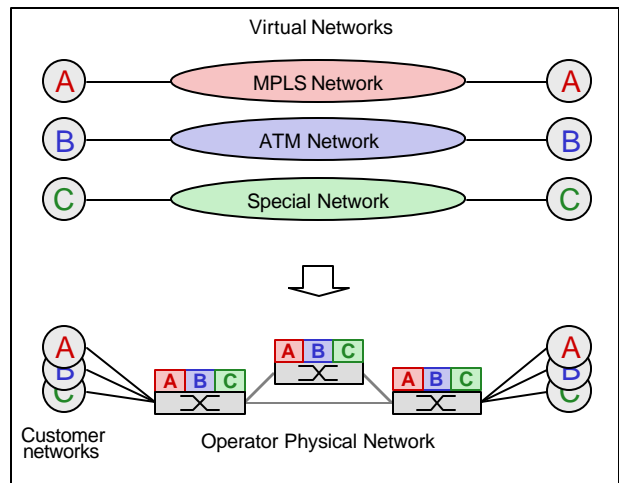


Fig. 3-1: Different virtual networks built upon one common physical infrastructure.

Figure 3-1 shows how virtual networks can be created by assigning a partition and a controller to each customer. A, B and C are different customers each attached to a dedicated virtual network, which is realised on one common physical infrastructure. Or alternatively, A, B and C could be different services provided to customers. Figures 3-2 and 3-3 show how the nodes

in this example may be configured. They consist of an ATM switch with a built in PNNI ATM controller for the ATM virtual network. In addition to that, the switch is also attached to two off board controllers, one for the MPLS virtual network, and one 3rd party controller for the special network.

3.1 Support for multiple protocols

One fundamental concept that the partitioned switch enables is the operation of multiple protocols (some of which may be different versions of the same protocol) on the same physical switching machine using a single Switch Control Interface (SCI), for example, GSMP.

Figure 3-2 depicts a possible scenario for a multiservice ATM-based switching machine. Here, the assumption is that the vendor has implemented an on-board controller, for example, implementing the ATM Forum's Private Network-Network Interface (PNNI). Using some standards that are in process, this could control frame relay, voice, circuit emulation, and native ATM over shared ATM trunk ports. Using switch partitioning, several other controllers can share port resources. In the example, there may be two versions of MPLS to provide interworking with different vendor implementations of pre-standard MPLS signalling and routing information interchange. Finally, an entirely separate 3rd party application could use another partition. In this application, sharing resources on trunk side ports is a critical requirement.

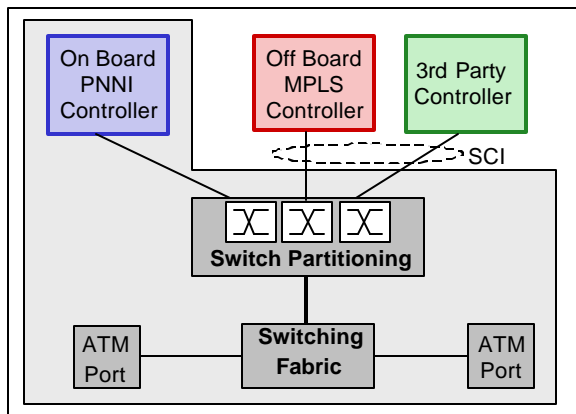


Fig. 3-2: Multiservice ATM-based switch.

The nodes in figure 3-1 need not be ATM switches, and could instead be MPLS based switches. The MSF architecture is general even if for the moment assumes ATM capable switch fabric; it is expected to be extended to cover all different switching technologies. The nodes could in fact alternatively consist of MPLS based switches, attached to a network of Label Switch Routers, but providing similar services as the ATM based node. **Figure 3-3** illustrates this analogous application where the switching machine is an MPLS La-

bel Switch Router (LSR) with an on-board, vendor-provided MPLS controller. Here too, partitioning helps solve interoperability problems with pre-standard MPLS signalling implementations through use of a separate controller. Depending upon the specific capabilities of the LSR, it may also be possible to carry certain types of ATM traffic over MPLS. What is likely missing in LSR implementations is then the requisite ATM control software. Partitioning the switch so that an off-board ATM controller can provide this capability is a logical design choice. Additionally, the uses of 3rd party controllers that allow wholesale resale of MPLS resources, or else empower the deployment of non-standard, innovative application based upon MPLS are also enabled via this architecture.

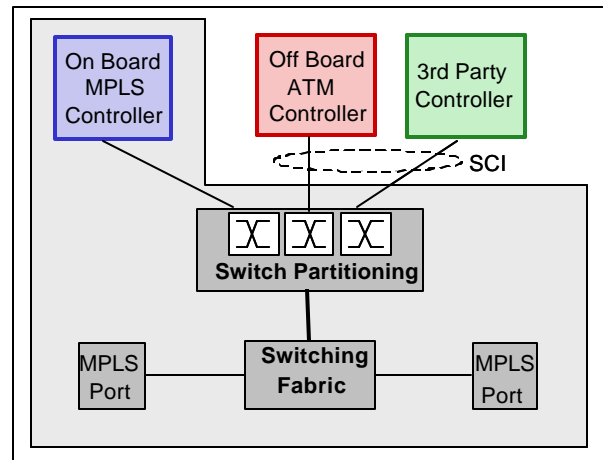


Fig. 3-3: Multiservice MPLS-based switch.

3.2 Partitions in the Access Network

While the MSF architecture allows for partitioning of network nodes, the network access from customer equipment can also be partitioned. Existing techniques like ADSL and VLAN/Ethernet can be used to partition the access, as illustrated in **Figure 3-4**. A, B and C can be different customers connected via dedicated lines to a concentrating function, from which common ADSL line is connecting to a partitioned access node. Or A, B and C can alternatively stand for different network services that one customer is using. The partition concept further subdivides physical ports into logical components at a network switch. Asymmetric Digital Subscriber Lines (ADSL) serving customers A, B, and C would all be multiplexed by a DSL Access Multiplexer (DSLAM) into a single ATM port on an ATM switch. Either the virtual channel or virtual path level of ATM could be employed to perform this multiplexing. ATM port resources, such as buffers and shapers, would be partitioned amongst these virtual channels (or paths).

Alternatively, the Virtual Local Area Network (VLAN) capability of Ethernet could be used to share a single port on an Ethernet switch serving customers A, B, and C. This could be carried over ADSL transmission facilities to the operator's point of presence, where another Ethernet switch converts the layer 2 encapsulation from Ethernet VLANs to MPLS Label Switched Paths (LSPs). Other protocols can be used to logically partition a physical port amongst multiple customers.

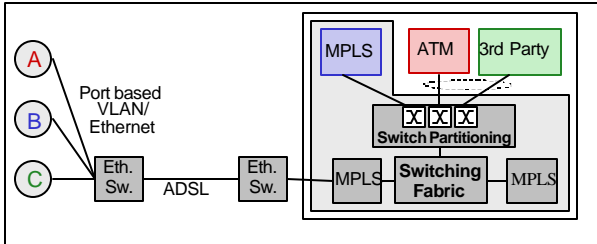


Fig. 3-4: Use of xDSL and Ethernet in local line

3.3 Wholesale of network capacity

Other forms of partitioning are also of interest. For example, an operator may resell switching and transmission capacity to other operators that have not invested in such a hardware infrastructure in a particular geographic region. As shown in Figure 3-6, partitions

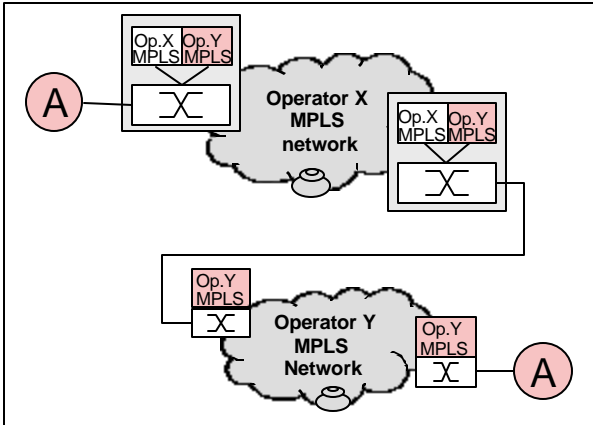


Fig. 3-6: Use of partitions across multiple operator domains to offer a virtual MPLS service

can create VPNs that cover more than one operator domain. As shown in the figure, the VPN control logic must be used in all involved nodes in the different domains. In this specific example, an operator Y's MPLS service is supported natively in operator Y's domain, but as a virtual network in operator X's domain. Even if the networks are based on different

technology (e.g. ATM and frame-based IP) a VPN can be created in this way. Of course, if the bearer protocols are different, then interworking between these technologies must be performed. The MSF architecture enables this division of the market into two sub segments, one that focuses on hardware infrastructure, and another that one focuses on providing network services to customers.

4. Toward Distributed Routing

A number of additional difficult problems face the designers of modern routed connectionless networks. The state variable explosion that occurs in routed networks is currently contained by careful address administration that results in significant aggregation. Adding QoS and guaranteed bandwidth requirements by per-flow or connection-level routing and signalling creates a need for per-flow state that markedly increases the state space of routers and switches. What happens to existing machines deployed in the network when the capacity of the on-board processor is no longer adequate? Commensurate with this increased load, the cost of software development and testing and the resulting stability (or lack thereof) become significant operational issues. Furthermore, support for multiple protocol versions from multiple vendors makes interworking and upgrades a daunting challenge. Additionally, since much of the change occurs in software, systems that combine processing hardware and switching face premature technological obsolescence. The proliferation of multiple overlay networks, and in some cases, multiple networks for each service creates operational inefficiencies.

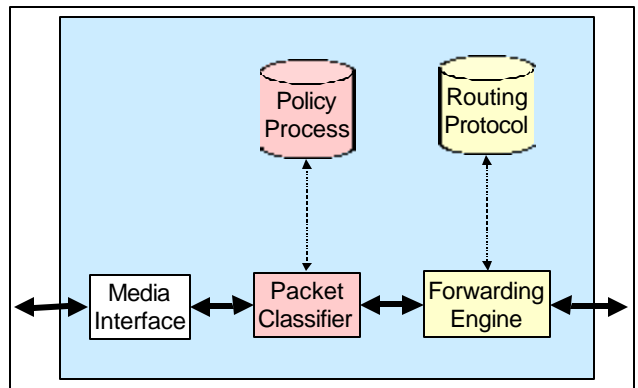


Fig. 4-1 Traditional Monolithic Router

The traditional IP router is a monolithic machine, much like the older telephone switches, as illustrated in **Figure 4-1**. As shown in that figure, a monolithic router consists of a number of functions. The principal ones are media interfaces, such as Ethernet, a packet classifier and a forwarding engine that does address prefix lookups and implements queuing on an interface to trunk side connections. For every data path function, there is a corresponding control function. In the case of packet classification, there is a policy control and in the case of the forwarding engine there is a corresponding routing protocol implementation.

As considered in the MSF architecture work [1], it may be desirable to implement a distributed router along the lines of the high-level functional architecture illustrated in **Figure 4-2**. In this architecture, the implementation of the policy control and routing protocol server functions control a multiplicity of media encapsulation, packet classification, and forwarding engine functions. The policy control interaction with the classifier is like the Network Service Instance Control (NSIC) and Logical Port control of the current MSF architecture. The interaction between the routing protocol and the forwarding engine is like the interaction between Bearer Control and the virtual switch.

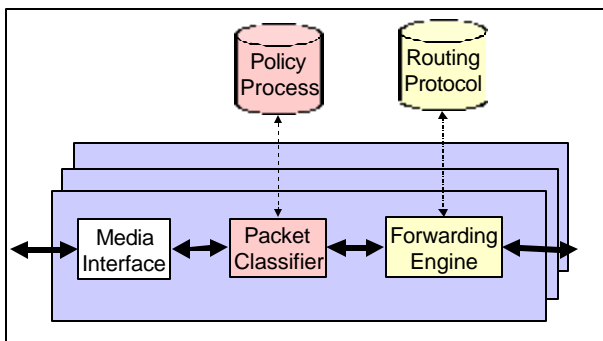


Fig. 4-2 Possible Architecture of a Distributed Router

Figures 4-3 and **4-4** illustrate one potential benefit of such an architectural model. Each figure shows two planes: the switching and adaptation plane, and the routing control plane. In the traditional router model of **Figure 4-3**, the instances of routing control map one-to-one to the instances of switching. This presents scalability problems for routing protocol implementations in very large networks. As discussed earlier, the addition of QoS and traffic management exacerbates these scalability problems.

Figure 4-4 depicts the situation when distributed routing control of label switches is implemented. Although the number of instances of switching remains the same, the number of routing control in-

stances is markedly reduced. This along with other motivations should be considered as potential requirements for future work items for the MSF switch control and media control working groups.

The MSF architecture allows placement of control logic hardware at separate locations. One possible architecture for a smaller network would be to place control logic hardware in two locations (for purposes of reliability), but have the forwarding hardware distributed across many network nodes, as determined by traffic engineering and economic considerations.

The distributed router architecture shown in **Figure 4-4** is reminiscent of the ATM Forum's Multi-Protocol Over ATM (MPOA) protocol specification. The reason that MPOA has not been successful is that it assumed that the cost of variable length IP header prefix lookups was much more expensive than that of establishing virtual connections on demand. Advances in routing hardware design and implementation rendered this assumption invalid. However, the separation of routing control from forwarding need not assume the switched connection paradigm that MPOA did. Instead, the population of the distributed forwarding tables could be done by traditional routing protocols, external policy controls, or pre-configured information. For example, as described in [1], the forwarding engines can intercept routing messages and forward these to the route processors, which compute the forwarding tables and download the information to the forwarding engines.

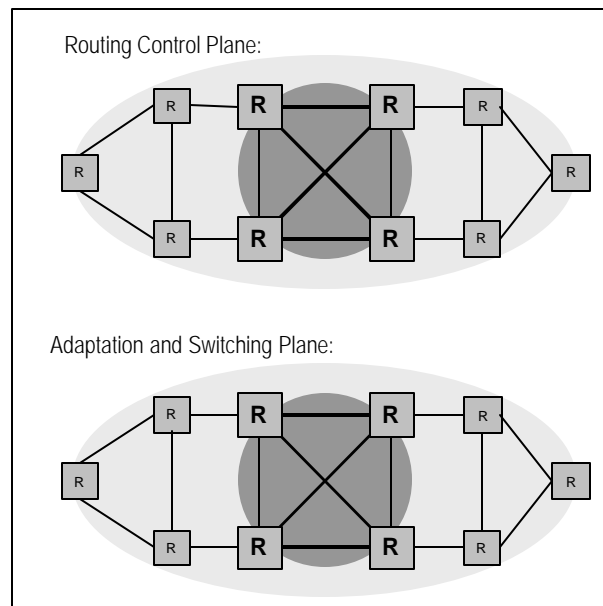


Fig. 4-3 Network View of Traditional Routers

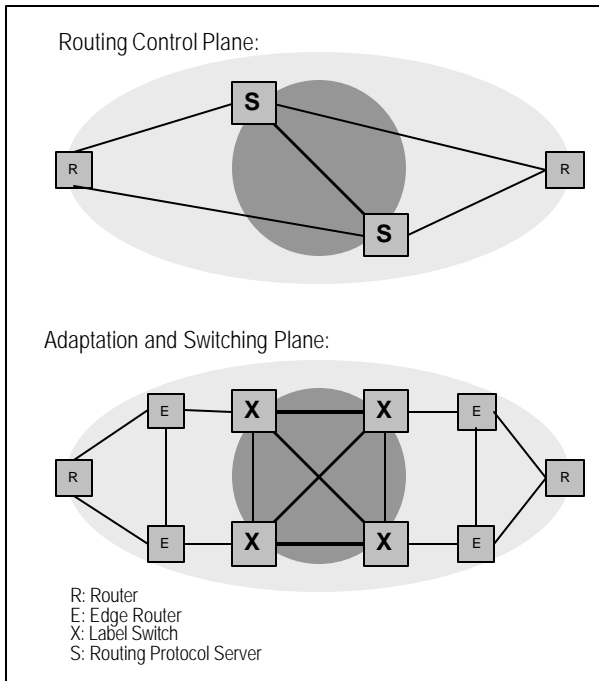


Fig. 4-4 Network View of Distributed Routers

In addition to basic IP level forwarding, the separation of control from forwarding enables the development of value-added IP services. These include service differentiated based upon fields within the packet, enforcement of sophisticated policies, and implementation of value-added routing algorithms specifically tailored to user applications. In the example shown in **Figure 4-5**, a customer is connected to two different backbone networks, one for best effort traffic, and one for high quality traffic. The controller logic may be different for each customer, or may be shared across many customers as controlled by configuration parameters. For example, the best effort network may provide access to the public Internet, while the QoS-aware IP network is used for internal VPN traffic. The customer may be attached by use of xDSL or Ethernet to a common access node. In the case where a separate controller is allocated to each customer, that customer effectively has his or her own virtual router that determines how to handle his or her traffic. The virtual router divides traffic from the customer into two categories based upon fields in the received IP packets, including network, transport, and application layer information. For example, the virtual router could direct web-browsing to non-intranet sites to the best effort IP network, while directing traffic to intranet web sites over the QoS-aware IP network. This means that users accessing intranet sites should receive better performance than those accessing public web sites.

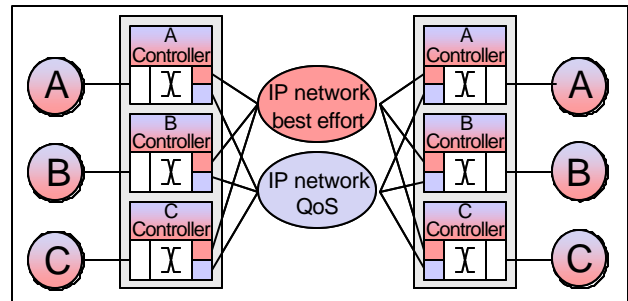


Fig. 4-5: Access node to multiple core networks

5. Conclusions, Open Issues, and Future Directions

The new proposed architecture will increase the available options when choosing network configuration, as shown above. But it will also give operators increased freedom in following ways:

- Profiling of operator services.

New manufacturers specialised on switch/router components (software and hardware) are expected to emerge. Therefore operators can choose from a wider spectrum of network components, hardware and software, than before. There will be greater possibilities for operators to acquire specific unique control software with special facilities, both standardised and non-standardised.

- Flexible network solutions

The new architecture will enable flexible network solutions that allows for independent scaling of hardware and software, flexible provisioning of new networks, or reallocation of resources to other networks. New networks can be deployed much cheaper than before if they can be implemented as a virtual network in partitions in existing infrastructure. Network resources can be reallocated in a flexible way from one virtual network to another.

- Provisioning and operation

Many different forms of provisioning and operating networks are made possible by the MSF architecture, as mentioned below:

Network hotel

VPN's can be provided according to the Web-hotel model. That is, a bigger operator is selling partitions of its infrastructure. And as in ordinary hotels, opera-

tions and maintenance are provided as a part of the service.

Customer controlled VPN

The customer may himself organise control and management of his own VPN. This may be an attractive alternative for bigger customers that already have their own professional IT staff, and wants to include and maybe integrate management of their VPN within their own supervision environment.

3rd party management of VPN

The customer or the provider may not handle the management of the VPN, but some other 3rd party may handle it. This may be attractive arrangement in some special cases.

The influence of the MSF architecture on the telecommunication market will be to further speed up already existing processes. The equipment and software market will broaden its offerings and competition will increase. This will contribute, combined with other factors, to a greater range of options for the operators. Competition between operators will increase, and some operators will profile their products and occupy specific niches, while others may broaden their scope. Lower thresholds for providing telecommunication services will lead to even smaller niches, as contrasted with the past trend of production at a very large scale that only allowed for utilization of broad niches.. All this will further speed up the trends summarized above.

The MSF architecture may inspire to developments of some more innovative networking solutions. The increased flexibility may invite creation of “instant networks”, where networks are established (and released) at the same time-scale as it now takes to establish a connection. Or there is quite possible that virtual networks may turn out to be the kind of infrastructure needed for Active Networks, or that virtual networks will provide a suitable infrastructure for applications using distributed processing [4].

6. References

- [1] Multiservice Switching Forum: System Architecture Implementation Agreement v.1.0, 2000.
- [2] McDysan, D., VPN Applications Guide, Wiley, 2000.
- [3] J. Boyle, R. Cohen, D. Durham, S. Herzog, R. Rajan, A. Sastry, “The COPS (Common Open Policy Service) Protocol,” IETF, draft-ietf-rap-cops-07.txt, August 1999
- [4] Chen, T.M.:Evolution to the Programmable Internet.(To appear in IEEE Networks)
- [5] Cuervo et.al: MEGACO Protocol, Draft-ietf-megaco-protocol-07, 2000.
- [6] Doria,A et.al: General Switch Management Protocol V3, Draft-ietf-gsmp-04, 2000.
- [7] ITU H.248, Gateway control protocol. 2000.