Open Control of Partitioned Switches

Nils Björkman' Telia Research, Sweden and Editor Multiservice Switching Forum
Alexander Latour-Henner, Sweden and Board Member of Multiservice Switching Forum
Torbjörn Lundberg, Telia Research, Sweden
Dave McDysan, Worldcom, USA and Technical Chair Multiservice Switching Forum

Abstract

In the computer market open interfaces have paved the way for intense competition with dynamic and rapid market evolution as a result. Open switch interfaces could in a similar way change the market conditions for the switch and router industry and in addition result in a more diversified telecommunication market. This Invited Lecture provides an overview of the new system architecture that the MSF, Multiservice Switching Forum, is establishing and describes how the architecture relates to other initiatives for open architectures in the communications industry.

1 Introduction

The evolution of network technology is much faster today than a decade ago. A telephony switch used to be depreciated over a period of 30 years or more, whereas ATM switches and IP routers today typically need to replaced within three years. To be competitive in a rapidly growing market requires rapid upgrades of the performance and functionality of the network. At the same time, a badly chosen network solution may turn out to be a costly cul-de-sac.

One way to manage rapid upgrades of the network at minimum risk is to deploy networking equipment with modular system architecture. Modularity would allow a network operator to mix and match best of breed components rather than relying on a crucial decision for a certain technology. The Multiservice Switching Forum (MSF) [8], the IETF GSMP[7] and MEGACO [10] working groups, Parlay Group [9], the Softswitch Consortium [11] and the IEEE P1520[12] all contribute to standards that support network elements with modular system architecture.

One of the merits of the MSF component based system architecture is that it allows systems to be integrated in many different combinations and to be used in many different ways. The subset of possibilities described in this paper is the understanding of the authors and does not express an official stand-point of the MSF or the companies the authors represent.

2 From monoliths to component based systems

2.1 What's wrong with monoliths?

A network operator of significance offers his services over a variety of separate networks of different types. The nodes building up the network are typically realised in dedicated equipment, even though some of the networks may share the same infrastructure with respect to cables and transmission. In all cases, the logic controlling the traffic flows is indivisible from the switching/ forwarding functionality, as illustrated for some typical network types in Figure 2-1.

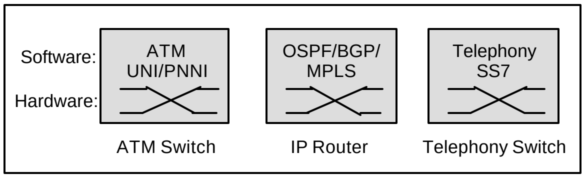


Figure 2-1 Switches of today; control logic and switching or forwarding functionality is indivisible

A consequence of this monolithic approach is that the success of the operator is tightly coupled to the direction taken by the equipment vendor. Once the new network is in place the operator is at the mercy of the selected equipment vendor with respect to:

- Functionality of new S/W releases
- Functionality of new H/W releases
- Pricing policy

If the vendor does not succeed in all of these aspects, the operator is left with no other choice than replacing the whole network – and thus getting stuck again.

2.2 Comparison with PC industry

The decision by IBM to open their PCs to vendors like Microsoft gave the entrepreneur Bill Gates the opportunity to offer his operating system to other PC-vendors than IBM. The principle to separate the control software from the computing hardware using open interfaces was thereby established. Today the monolithic mainframes have been replaced by a large number of PC vendors, a few operating system vendors and a universe of new applications, as depicted in Figure 2-2.

The open PC-architecture allowed more vendors to participate and compete on the market than before with the old monolithic IBM environment. This new multivendor environment resulted in increased competition, rapid development of performance and new functions, reduction in prices and a larger market and an explosion of new applications and software.

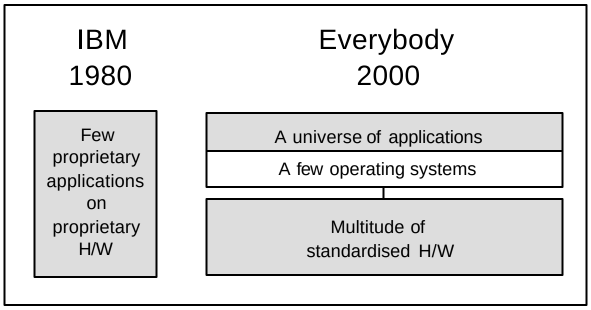


Figure 2-2 The “PC revolution”

3 Initiatives for open interfaces

3.1 Opensig and OpenArch

A wish to acquire network equipment with open interfaces has been expressed by network operators way back in the telephony era. Naturally, developers of networking applications have had a similar desire to be able to explore new ways to use a network. In 1995 the Columbia University [3, 4] invited to the first OPENSIG in order to promote research towards understanding open network control issues. Following a decision in 1997 by IEEE Communications Society sponsor OPENSIG expanded with the yearly IEEE OPENARCH conference.

3.2 IEEE 1520

As a consequence of the success of the OPENSIG and OPENARCH conferences IEEE decided to establish a standardisation group for Programmable Networks 1520 [12]. The ultimate goal of this is to make the network as programmable as the PC through a set of standardised interfaces offering different levels of abstractions. Based on proof of concept implementations demonstrated early 1999, the following interfaces are defined:

In the model four interfaces exist and provide the separation between end-user applications and value-added services, network service provider and the underlying network resources:

- User level programming interfaces provide access to the value-added services level. These are collectively called the V interface.
- Programming interfaces between the value-added services level and network generic services level. These are collectively called the upper interface, or U interface.
- Programming interfaces between the network generic services level and the virtual network devices level. These are collectively called the lower interface or L interface.
- Open protocols to access the state of physical elements. These are collectively called the Connection Control and Management interface, or CCM interface.

3.3 The International Softswitch Consortium

The International Softswitch Consortium [11] was created in May of 1999. The consortium recognizes that many applications emulate circuit switching in software, hence the name "softswitch". The consortium focuses on interoperability and certification of voice and other real-time services. Five standard interfaces are adopted for promotion in the Consortium’s charter, including H.323, Session Initiation Protocol (SIP), Real-time Transport Protocol (RTP), Real-time Streaming Protocol (RTSP) and Media Gateway Control Protocol (MGCP). Additional standard interfaces are expected to be adopted by the Consortium.

3.4 Parlay

The Parlay Group [9] has been formed by eleven companies to specify and promote the Parlay open Application Programming Interface (API). The focus of the Parlay group is to enable enterprises outside of the network operators' domain to access network information and control a range of network capabilities.

The phase 1 specification, released in December 1998, primarily addressed call control, messaging and security. The phase 2 programme is focused on the expansion of theAPI in the areas of wireless and IP.

3.5 Multiservice Switching Forum

In November 1998, slightly before the IETF GSMP standardisation initiative, the Multiservice Switching Forum [8] was formed on the initiative of Cisco, Bellcore and MCI Worldcom. MSF has gathered strong industrial support from:

- Nearly all traditional telecom vendors
- Nearly all established datacom vendors
- Several leading network operators in Europe, Japan and the USA.
- Several HW and SW component vendors that see there chance to add novelty to systems from the established vendors

"The MSF mission is to accelerate the deployment of open communications systems that realize economic benefits, which result from the flexible support of a full range of network services using multiple infrastructure technologies. The focus is on development of architectures and industry agreements that enable interoperability and innovation in a rapidly evolving environment [8]."

3.6 IETF: MEGACO and GSMPv3

Two IETF working groups have a direct relationship to the core of the MSF Architecture:

- MEdia GAteway COntrol (MEGACO) protocol being developed by the IETF in cooperation with ITU-T. The working group will develop an informational RFC detailing the architecture and requirements for controlling Media Gateways from external control elements such as a Media Gateway Controller. A media gateway is a network element that provides conversion between the information carried on telephone circuits and data packets carried over the Internet or over other IP networks [10].
- Generic Switch Management Protocol (GSMPv3). GSMP provides an interface that allows a routing logic component to control a label switch [7].

4 The MSF Architecture

4.1 Leveraging from the PC industry

A mapping of the principles established in the computer industry to the communications industry implies:

- a division of the monolithic switch into a controlling software part and an executing hardware part
- an open standardised interface between these parts
- a partitioning of the executing hardware is to be able to execute several controlling software instances in parallel.

Resource	Unit
Label/Address Space	Number of labels
Forwarding Table Space	Number of entries
Bandwidth	Bits/s, cells/s, packets/s
Control Connection Bandwidth	Bits/s, cells/s, packets/s
Buffer Space	Bits, cells, packets
Processing Capacity for e.g. traffic forwarding, signalling, queue schedulers	CPU cycles, MIPS, FLOPS

Table 1 Common resources in switches / routers

In order for this principle to work in a communication network, all network resources, listed in table 1, have to be partitioned to enable simultaneous control of subsets of network resources by multiple network control systems.

4.2 The MSF Architecture

Figure 4-1 illustrates two practical and one generic example based on the MSF architecture. Common to all examples is a natural decomposition of a system into components. For intra-system interfaces, the MSF is endorsing a small set of protocols (currently GSMPv3 and MEGACO/H.248) to be used for realising these interfaces across a broad set of services. For inter-system interfaces, a wider set of industry standard protocols have been used. A merit of this architecture is that the components can be developed, deployed and replaced independently of each other.

Starting from the left-hand side of Figure 4-1, a separate Media Gateway Controller (MGC) controls one or more Media Gateways (MG) through the MEGACO/H.248 protocol. In the middle of the figure, a separate switch controller controls and manages one or more switches via GSMPv3.

The right-hand side of the figure is the generic example of other controllers using extension of GSMP or other protocols to remotely manage a forwarding, switching, or routing device. However, the principal focus of Release 1 is on the MEGACO/H.248 and GSMP protocols.

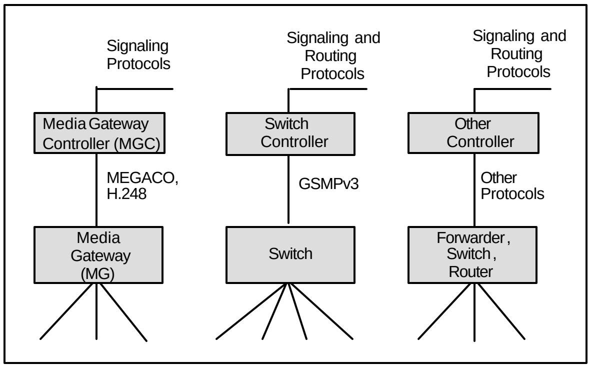
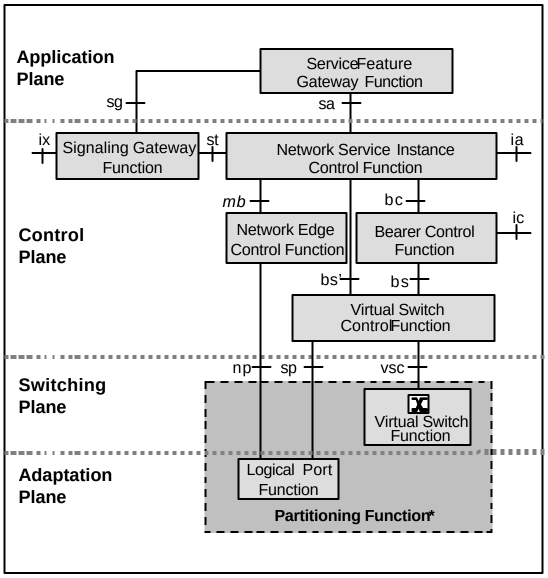


Figure 4-1 Examples of interfaces between control and switching in the MSF architecture

How has the MSF identified the small set of interfaces that can support a wide variety of network services? The approach taken by MSF was to define an architecture of finer granularity than what is expected to be implemented as individual components. Figure 4-2 shows all the functions and reference points that the MSF Architecture comprises.



Notes:
- Italicized reference points are not considered open reference points for release 1.
- Bearer transport reference points are not shown
▨ - Management functions overlaid on functional architecture
* The Partitioning Function maintains partition integrity between partitions of a partitioned entity.

Figure 4-2 The MSF Reference Architecture

The architecture has been verified by bundling the functions into sets of functions that would be natural in a physical implementation. Section 5 illustrates three such physical realization examples from the MSF System Architecture Implementation Agreement.

Figure 4-2 also shows the multi-planar system model chosen. Starting from the bottom of Figure 4-2, the Adaptation Plane supports the physical interface to a user or another network element. The Switching Plane supports the actual switching fabric by which physical interfaces are connected. The Control Plane provides the generic capability to manage network service events and provides control over both the adaptation and Switching Planes. Standard protocols are used in communicating between the Control Plane and the Switching/Adaptation Planes. The Application Plane provides services that use the capabilities of the Control Plane and it also provides enhanced services which control the services realised within the Control Plane.

The planes aggregate functions that interact to realise the generic behavioural model of an MSF compliant system. Short hand definitions of the functions are provided below:

Logical Port Function (LPF)
Each Logical Port Function (LPF) provides the necessary media mapping and service specific adaptation functions related to the incoming media stream.

Virtual Switch Function (VSF)
A partition of switch resources is called a Virtual Switch Function (VSF), and is an arbitrary subset of switch resources that can be controlled as a unit..

Switching resources are responsible for the switching of media streams from one (logical) port to another or to a functional entity. The switching resources may provide packet switching, frame switching, cell switching etc.

Network Edge Control Function (NECF)
The NECF is responsible for sending and receiving control information to and from a LPF regarding the media streams and other services (e.g., encryption) on those streams that it supports.

Virtual Switch Control Function (VSCF)
The VSCF controls and monitors the VSF (Virtual Switch Function) and LPFs (Logical Port Functions) within a Partition. The VSCF provides the required cross-connect information, including traffic and QoS information, across the VSF from one LPF to one or more LPF. It receives information about the switching function and propagates this information to other functions that may change state as a result (e.g., BCF).

Bearer Control Function (BCF)
The bearer control function for establishes, modifies and releases end-to-end bearers between end point(s) of a bearer connection. The BCF interacts with the appropriate instance of the NSICF and receives the information required to setup a bearer connection.

Network Service Instance Ctrl Function (NSICF)
The NSICF establishes, maintains, modifies, and releases network service instances. Some examples of network service instances include a context, circuit switched calls (e.g., PSTN connection) and an inter-domain route (e.g. BGP-4 route).

Signaling Gateway Function (SGF)
The Signaling Gateway Function is the functional block that processes signaling.. Mandatory is to map, relay or tunnel the signaling. The SGF may also screen or terminate the associated signaling.

Service Feature Gateway Function (SFGF)
The SFGF is a functional block allowing access to intelligent network services and other network provided applications. It also allows directly signaled services in the Application Plane to access the Control Plane functionality.

4.3 Switch Partitioning

Within the vision of Multi-Service Switching, the functions in the Control Plane can share the resources of the Switching Plane and of the Adaptation Plane. A partition of switch resources is called a Virtual Switch Function, and is an arbitrary subset of switch resources that can be controlled as a unit.

Management creates a Virtual Switch Function by specifying the switch resources that are to make up the partition. These resources include physical port resources such as bandwidth and buffer space, and physical switch resources such as forwarding table space.

A partitioned switch entity may be a single physical switch, or a group of inter-connected physical switches, or other VSFs.

4.4 Deterministic and statistical sharing of resources

It shall be possible to combine Virtual Switches in such a way that they interact with each other only in a predictable and controlled manner. The sharing of resources between Virtual Switches may be either deterministic or statistical.

The first application of partitioning is expected to provide multiple instances of services to separate groups of users. These user groups would not expect interference from each others' activities. At the same time, the users of each partition are interested in efficient utilization of the resources allocated to their Virtual Switch. This implies that the principle of resource sharing <u>between partitions</u> should be **deterministic**. Note that the control logic of a partition still requires an ability to use statistical allocation principles, for the resources <u>within</u> its partition.

More efficient utilization may be achieved if Virtual Switches statistically share pooled resources among several, otherwise non-cooperating control logic instances, as illustrated in Figure 2-3. As maybe obvious, the concept of **statistical** resource sharing between partitions is more difficult to define stringently than a deterministic one. The challenge lies as previously (c.f. ATM Forum service classes, IETF Diff-Serv/Int-Serv etc) in:

- Defining a control mechanism together with a behavioral conformance definition to maintain a utilization contract with each instance of partition control logic.
- Defining a performance target that governs the allocation of partitioned resources in such a way that additional requests for allocation are blocked when the probability of failing to maintain the performance is unacceptably high.

These terms, deterministic and statistical, are not to be confused with static and dynamic configuration. The term static configuration means that the Virtual Switch partitions are established prior to the switching system as a whole being brought in service. The static configuration can be changed first after the switching system as a whole is brought out of service. The dynamic configuration principle allows for changes without bringing the switching system as a whole out of service.

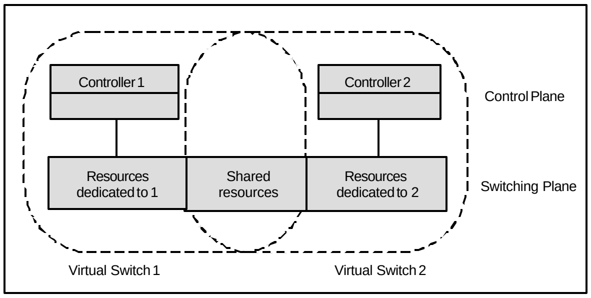


Figure 4-3 Statistical sharing of switching and adaptation resources between two controllers

4.5 Implications on management

The MSF Management Plane complements existing standards with functions, MIBs and interfaces that are required by the MSF Architecture. These new requirements stem mainly from the fact that the MSF reference architecture:

- Allows components to be ‘mixed and matched’
- Defines Virtual Switch Function, responsible for partitioning of switching resources into Virtual Switches.

The first bullet leads to a requirement for autoconfiguration. Comparing again with the PC industry, the concept of ‘Plug&Play’ has been maybe the largest challenge but also a major strength in the architecture of PCs.

The second bullet allows Virtual Switches to be created on some set of switches in the network to form a virtual network. Each virtual network (the VSs that make up the network) can be controlled by an independent Control Plane that should not be aware that it is controlling virtual switches rather than physical ones. Similarly, each virtual network can be managed by a separate network manager that need not be aware of the fact that it is managing a virtual network and not a physical one. In this environment, management functions are required on two levels:

- Super-ordinate management of the Switch Function and the VSF
- Sub-ordinate management of individual VSs

Figure 4-4 illustrates that each Sub-ordinate manager has access to its Virtual Switch MIB, but only the Super-Ordinate Manager interact with the Partitioning MIB.

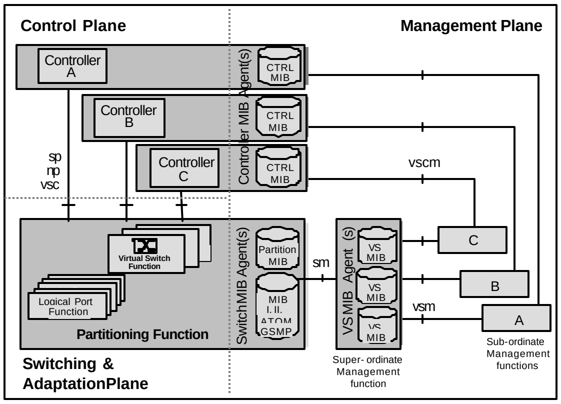


Figure 4-4 The MSF management architecture

5 Realisation examples

5.1 Classical IP router

Figure 5-1 describes the application of the MSF architecture to a representative best effort IP forwarding capability as defined in RFC 1812, requirements for IPv4 routers. The purpose is to illustrate how a router is optimally decomposed into two separate ‘boxes’.

An MSF compatible ‘box’ is composed of some or all functions described in the MSF Architecture. The ‘boxes’ are connected over an interface traversing one or more MSF defined reference points. To emphasise that the examples represent a physical implementation, and not just a set of logical relationships, all traffic enters/leaves the ‘boxes’ through physical ports.

The bundling applied in this example has as its purpose to separate the IP routing control component and the IP forwarding component. This is viewed to support a natural specialization and the independent scaling of each component. The upper control part shall maximise reachability while maintaining stability. The scaling requirement is moderate since the number of IP prefixes to be handled currently grows linearly in the Internet. The processing intense computations in the control part boil down to simple forwarding tables in the forwarding component. With the volume of traffic in the Internet growing exponentially the key feature of the forwarding system is the ability to forward an ever increasing volume of traffic.

The example illustrates a minimal split into two physical elements, thus requiring the smallest number of open interfaces. In the future it may be feasible/required to introduce a finer modularity in the physical model.

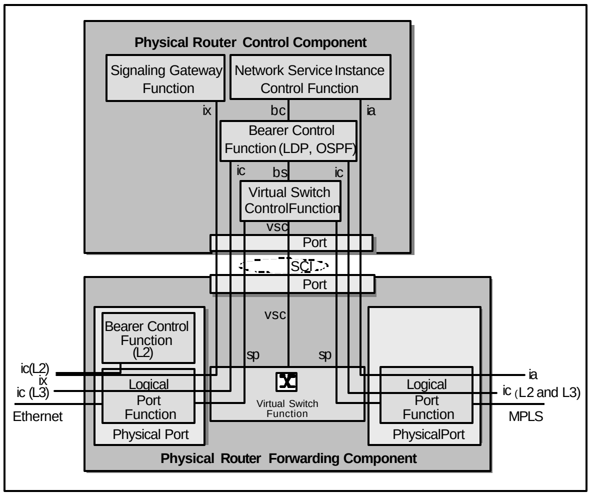


Figure 5-1 Physical model of a Classical IP router

5.2 Telephony Gateway (MGC/MG)

Like the previous example, the next example decomposes two distinctly different functional domains into two separate implementations: A Media Gateway component and a Media Gateway Controller component. This system provides conversion between the information carried on telephone circuits and data packets carried over the Internet or over other IP networks.

The MGC component is responsible for protocol interactions with the service logic, eventually leading to establishment of a call. It is important to note that for the telephony service the operator profile is mainly established through specific implementation of the service logic exercising this protocol interaction.

The MG component is responsible for adaptation of the media streams as packaged on either side of the network boundary. The issue here is more interoperability and compatibility rather than service profile building.

This decomposition 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 forwarding should allow more rapid and cost effective introduction of new services and features in comparison with the traditional monolithic approach.

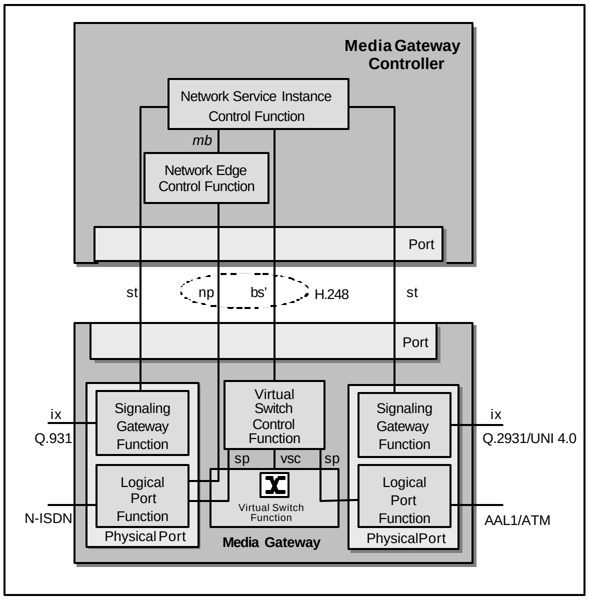


Figure 5-2 Physical model of a decomposed Media Gateway

5.3 Partioned Label Switch Router

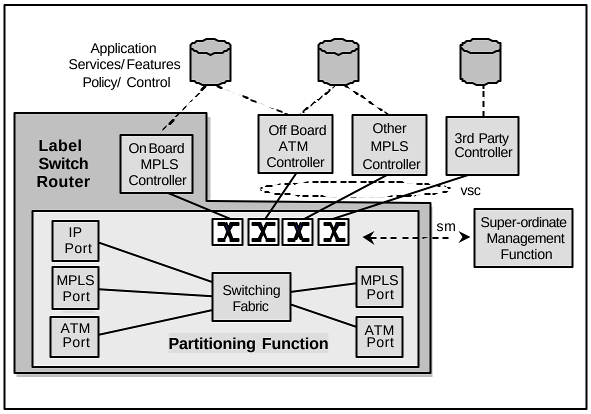


Figure 5-3 A partioned Label Switch Router

The concept of partitioning enables the operation of multiple protocols. Figure 5-3 illustrates this concept for an MPLS Label Switch Router (LSR). Here, the assumption is that the vendor has implemented an on-board MPLS controller. Using the switch Partitioning Function, several other controllers can share switch resources. In the example, there may be two versions of MPLS to provide interworking with different vendor implementations of pre-standard MPLS signaling and routing information interchange. Partitioning the switch so that an off-board ATM controller can provide an ATM capability is a logical design choice. Additionally, the use of 3rd party controllers that allow wholesale resale of resources, or else empower the deployment of non-standard, innovative application are also enabled via this architecture.

6 New market opportunities

Efficient operation of a network requires economies of scale, i.e. standardised solutions offered by a multitude of vendors to all vendors. An operator's dilemma is that the standardised solutions erase the unique profile necessary to offer specialised networks for different markets. The Multiservice Switching Forum has adopted the concept of Virtual Switches; combining the economies of scale for bit transport and switching while allowing profiling of service providers.
The network control logic and the switching/forwarding functionality are allowed to evolve independently from each other – most important since there is no reason why a leading vendor of switching hardware would automatically also be a leading vendor of network control logic for differentiated service offerings.

Manufacturers sometimes argue that their software for e.g. routing or signalling is superior, so why would any 3rd party components be necessary? The answer is that a vendor with a superior component has an excellent opportunity to expand the market by selling to other less fortunate vendors. The potential for market growth is in fact tremendous for both vendors and operators through the realisation of the following steps:

The potential for market growth is in fact tremendous for both vendors and operators through the realisation of the following steps:

1. **Open** the intra-system interfaces to allow for integration of best-of-breed components from multiple vendors. Open interfaces enable more actors to participate in the development of future products, which in turn gives room for a larger variety of services –and network operator profiles.
2. **Widen** the geographical market for operators without physical presence by offering a logical partition of a physical switch. A network operator can in this way offer a service similar to Leased Line, now upgraded to Leased Virtual Switch.
3. **Lower** the threshold for introduction of new types of network control and at the same time allowing graciously phasing out network control with declining demand. No more fork lift replacement!
4. **Shorten** the time from the original idea to market for non-standardised services. The concept of Virtual Switches allow prestandard products to be introduced on a small scale while waiting for a standard to settle.

7 The MSF Vision

“The Multiservice Switching Forum (MSF) envisions a future of global connectivity in which network elements can be interconnected and intermixed with relative freedom. This next-generation network infrastructure will spawn new markets for entrepreneurial developers of network elements, spurring competition and accelerating the creation of innovative solutions for all facets of global communication.

Both large and small providers of telecommunications services and equipment are aligned behind this vision because it will enable rapid deployment of new and enhanced services—without requiring new investments in switching and transmission resources. Such an infrastructure will support quick deployment of innovative—even experimental—services, controlled by software that can be adapted to meet the new and evolving requirements of end users.

The MSF vision is shared by a broad range of service providers and equipment vendors, and will continue to gain support as the industry realizes the benefits of an integrated network and services architecture with interoperable elements. The MSF is dedicated to ensuring global acceptance of this vision, which will lead to a rapid proliferation of network services and lasting improvements in global communication.”

8 Acknowledgements

Thank the members of the MSF Technical Committee for their ground breaking contributions. The authors are especially indebted to Ms Irene Andersson for clarifying unification of all drawings in the MSF System Architecture IA and for redrawing some of them one more time for this paper.

9 Literature

[1] van der Merwe J.E. et al. *The Tempest – A Practical Framework for Network Programmability*. IEEE Networks (available from: http://www.cplane.com)

[2] Rooney S et al. *The Tempest a Framework for Safe, Resource Assured, Programmable Networks*, IEEE Communications, (available from: http://www.cplane.com)

[3] Angin O et al. *The Mobiware Toolkit: Programmable Support for Adaptive Mobile Networking*. IEEE Communications, Aug. 1998.

[4] Campell A et al. *OPEN SIGNALING FOR ATM, INTERNET AND MOBILE NETWORKS (OPENSIG’98)*. October 1998.(available from: http://comet.columbia.edu/opensig/activities)

[5] Newman P et al. *Ipsilon’s General Switch Management Protocol Specification*. IETF RFC 1987, Aug. 1996.

[6] Newman P et al, *Ipsilon’s General Switch Management Protocol Specification version 2.0*, IETF RFC 2297. Mar. 1998.

[7] GSMPv3: http://www.ietf.org/html.charters/gsmp-charter.html

[8] Bjorkman N, et al. *MSF System Architecture Implementation Agreement*. April. 2000. (Available from: http://www.msforum.org)

[9] “The Parlay Group, Introduction and Objectives”, (White paper on Parlay Group, available from: http://www.parlay.org)

[10] H.248/MEGACO: http://www.ietf.org/html.charters/megaco-charter.html

[11] Softswitch forum: http://www.softswitch.org

[12] IEEE P1520 Proposed IEEE Standard for Application Programming Interfaces for Networks http://www.ieee-pin.org/