# Sustaining a Vertically Disintegrated Network through a Bearer Service Market

Petros Kavassalis[1], Thomas Y. Lee[2], Joseph P. Bailey[3]

[1] *Ecole polytechnique, Centre de Recherche en Gestion*
*<petros@rpcp.mit.edu>*
[2] *Massachusetts Institute of Technology, Context Interchange*
*<tlee@mit.edu>*
[3] *Massachusetts Institute of Technology, Technology, Management and Policy*
*<bailey@rpcp.mit.edu>*

Abstract

Based upon the Internet perspective, this chapter will attempt to clarify and revise several ideas about the separation between infrastructure facilities and service offerings in digital communications networks. The key notions that we will focus on in this paper are: i) the bearer service as a technology-independent interface which exports *blind* network functionality to applications development; ii) the sustainability of an independent market for bearer service and the organizational consequences associated with such a market.

## 1. Introduction

During the past two years, applications like email and the World Wide Web have combined with evolving network protocols to propel the Internet into the heart of a computer and communications convergence. Central to the Internet's immersion into digital convergence has been the effectiveness with which the Internet Protocol (IP) has played the role of "spanning layer." [1]

The IP abstraction enables applications to request network services independent of underlying, physical network technologies. Moreover, new underlying network technologies may either substitute for or co-exist with existing network technologies without significantly affecting the broader system. Based on this abstraction, the National Research Council recently articulated the "Open Data Network (ODN)" as an architecture for the networks of the future that generalizes the principle of separating service offerings from infrastructure facilities as demonstrated in the Internet [2]. In the same way that IP serves the Internet, the ODN relies upon a "bearer service" to function as a *technology-independent network layer* that resides above the technology substrate and enables interoperation between diverse, high-level applications and various underlying network infrastructures (figure 1).

The NRC report describes an Open Digital Network (ODN) as a four-level layered architecture: "i) at the lowest level is an abstract bit-level service, the bearer service, which is realized out of the lines, switches, and other elements of networking technology; ii) above this level is the transport level, with functionality that transforms the basic bearer service into the proper infrastructure for higher-level applications (as is done in today's Internet by the TCP protocol) and with coding formats to support various kinds of traffic (e.g., voice, video, fax); iii) above the transport level is the middleware, with commonly used functions (e.g., file system support, privacy assurance, billing and collection, and network directory services); and iv) at the upper level are the applications with which users interact directly.

This layered approach with well-defined boundaries permits fair and open competition among providers of all sorts at each of the layers". [2]

Certainly, the Internet demonstrates the technical and functional robustness of a technology-independent bearer service abstraction [3]. The bearer service is intended to support requests for service from all applications and to recognize all substrates, but as both application and infrastructure innovations turn increasingly towards user-oriented models of network architecture, technology and policy considerations related to the generalization of this abstraction should be carefully studied. Such a service blurs the boundaries of telecommunication markets.

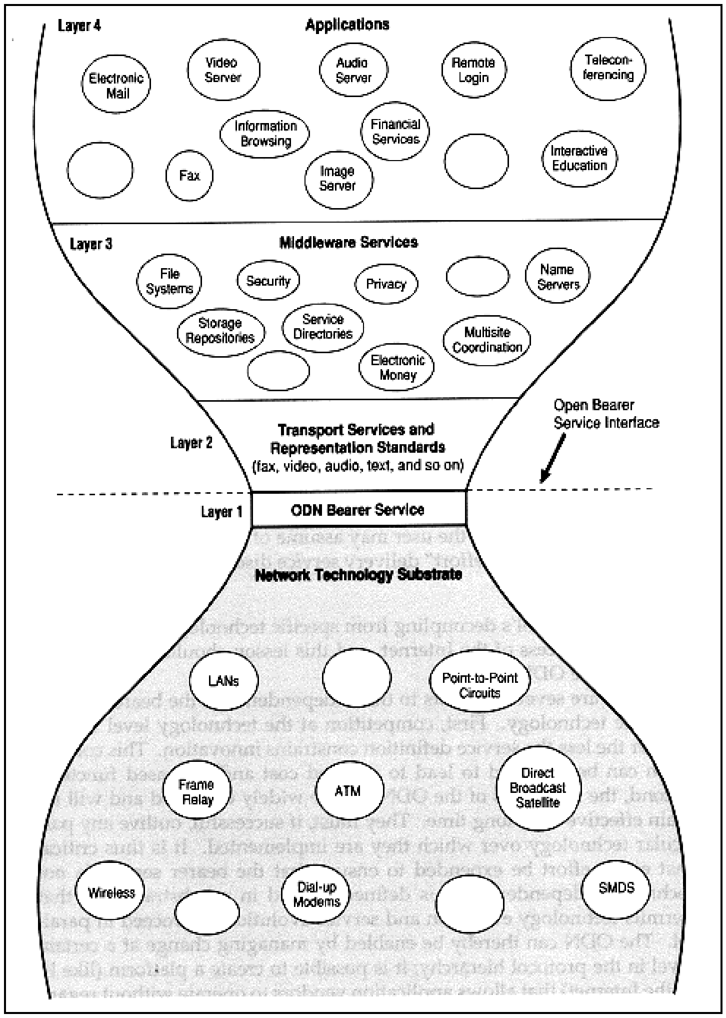



**Figure 1: The Bearer Service Concept**

For example, the promises of Internet telephony to combine the benefits of the public switched telephone network (PSTN) and the Internet would be possible through a bearer service – even though Internet telephony would weaken market boundaries and challenge the regulatory environment [4]. Regulators have a difficult time categorizing bearer service providers as belonging to any one existing market because such providers can offer services that cut across many existing telecommunications markets. Businesses that develop ubiquitous services or tailor applications for customers may be threatened by market entrants who have competitive products built around a new technology architecture that is able to provide great flexibility in applications design — the bearer service. Finally, customers can benefit from an integrated services environment because their data and voice communications can be transmitted across multiple telecommunications infrastructures. A new era of interoperability [5] is possible through the bearer service.

However, questions for this new market abound. Not only is the business model in question, but the technology is also in flux. In this chapter, we hope to shed some light on the bearer service issue by answering some of the questions arising from this new bearer service market. First, what does a technology independent bearer service look like? What are the technical characteristics of a bearer service that satisfies the requirements of both elastic and real time applications? Second, is a segregated network that separates infrastructure facilities from service offerings by a bearer service an economically viable market organization? This chapter spans both technical and economic issues to ascertain the potential of a market for bearer services. Specifically, this chapter will address the following subset of bearer service market questions:

i) What comparative advantage does a technology-independent bearer service provide over homogeneous infrastructures that support multiple service classes (e.g. ATM)? This question is especially important given the increasing interest in *Integrated Networks* solutions,.

ii) In a recent paper Gong and Srinagesh [6, 7] suggest that layered network architectures may inherently engender vertically integrated markets for network services. Is this a general statement or a phenomenological one? And if the phenomenological reading is correct, what factors tend to suggest the viability of disaggregated markets for network services both at and above the spanning layer as described above?

The remainder of the chapter is organized as follows. Section 2 employs a comparative analysis to define concepts of the bearer service and the *layered network architecture* concepts. Considering IP as a *bifurcation point* in the evolution of network design, the bearer services of traditional communications infrastructures and the Internet (both the current *best-effort* and the future *Integrated Services* Internet) are surveyed to elicit design characteristics and functional differences between a technology-dependent and independent bearer service. We suggest that in a network design with a technology-independent bearer service, the communications network supports a *flexible organizational* model capable of dealing with unanticipated (applications') variability.

Section 3 associates the technology-independent bearer service with the Internet organizational model: a flexible system of regional or more extended backbones and access links (or access networks) to these backbones, managed by the Internet Service Providers (ISPs). Specifically, we address the question of whether the ISP model, which is characterized as a model where network operators exhibiting varying degrees of vertical integration compete in an open market, can sustain itself, or if one monolithic, integrated firm will emerge from mega-mergers?

To answer this question, we begin by considering the work of Gong and Srinagesh [6], who argue that a stable and sustainable equilibrium for healthy bearer service market growth might not be possible. Our analysis closely follows the definition for a bearer service formulated in section 2. Arguing that the bearer service is not a commodity product, we identify differentiable service attributes upon which an independent bearer service market could form. Through differentiation, the bearer service market can avoid a Bertrand equilibrium (i.e. pure price competition). Furthermore, we challenge assumptions about perceived trends towards vertical integration, by noting the relative independence of assets between bearer service providers and potential buyers of this service (i.e. connectivity providers who are situated at the edge of the network versus other kinds of service providers).

The chapter concludes in section 4 by underlying the role of the bearer service market in creating the potential for competition between independent and integrated providers. However, as revealed in section 3, while a market for bearer services may be economically viable, it will also likely be fragile. The conditions for sustaining a bearer service that exhibits "spanning" capabilities and promotes efficient interoperability (as in the Internet), requires an alignment of both technical and economic considerations. In addition, new institutions, not necessarily government-dependent, may be required to enable the market to form.

## 2. From the railroad gauge to information bitways: the evolution of the bearer service functionality

Given a pre-specified set of applications and a physical network which may include more than one substrate technology (e.g. a Ethernet-based LAN connected to a Frame relay network), the bearer service (BS) constitutes those common[1] functions which are implemented throughout the network rather than in the network end nodes and are necessary for pairing each application's communication requirements with the performance characteristics of all components of the heterogeneous network.

For example, consider voice and fax data services transmitted via traditional, analog telephony between the United States and Canada. Neither within nor between each country do the customers need to share the consumer premises equipment (CPE). Likewise, at the level of network technologies, materials like cables, switches, and even local numbering conventions might vary. However, since all CPE and network subsystems observe a common, international numbering convention and all networks and CPE transmit, carry, and receive signals in the 4 kHz frequency spectrum, voice conversations are conducted and faxes are conveyed. For voice and fax data services, the international numbering scheme, combined with the 4 kHz channel constitute the bearer service.

More abstractly, computer and communications technologies may be separated into three layers. The physical infrastructure (e.g. wires, switches, etc.) resides at the lowest layer. At the top lies the set of applications and service offerings supported by the underlying infrastructure. A *spanning layer*[2] bridges the two [1]. An application requests network services through the spanning layer to the substrate technologies.

In the early public switched telephone network (PSTN), telephony was tightly coupled to a specific infrastructure so the spanning layer supported only a single application and one technology substrate (typically, the spanning layer was located *in the wires).* The development of newer applications for computers and communications as well as advances in substrate technologies, prompted a refinement of the spanning concept. In the presence of a diverse suite of applications and a heterogeneous network, the *bearer service (BS)* constitutes a spanning layer which *escapes from the wires* and thus supports all applications over the entire network. This section uses Piore's model of organizational flexibility and production system transformation as a methodological framework for tracing the evolution of the spanning layer towards a technology-independent bearer service. Advances in shipping and the transport of physical goods are used as a metaphor for the transformation of yesterday's PSTN into tomorrow's ODN.

### *2.1. Production technologies and flexible specialization*

Piore [8, 10] describes the on-going transformation of industrial production systems towards greater variety and flexibility as a four-stage evolution[3]. The products in such systems are comprised of both independent and interdependent design features; changes in design features mark the different evolutionary stages. Independent design features "can be varied in isolation without complementary changes in other features of the design" while interdependent features "require a number of complementary adjustments" [8].

The initial stage, *mass production* is characterized by a production system tailored to a single product. There is no room for variation. In *Mass production with cosmetic variation*,

---

[1] If every application uses the function, then it is certainly a function in common and unambiguously a component of the BS. If only one application uses the function, then perhaps it is more appropriately considered part of the application (or of the transport layer). If two or more applications utilize the function but not all applications in the set use the function, then we need to question whether the function belongs in the BS. Recall also that a separate metric for distinguishing BS functionality is whether that function can be implemented in an endnode. BS functions include only those functions which cannot be implemented in an endnode.

[2] As suggested by Clark, "a spanning layer is characterized by three key parameters that characterize the sort of interoperation it supports: i) the span of infrastructure options over which the interoperation occurs, ii) the range of higher level applications the spanning layer supports, iii) the degree of symmetry in the services of the spanning layer and its interfaces, and in the higher layers up to the application definitions". [1]

[3] This work draws on *The Second Industrial Divide* [9]

product design may slightly vary existing or may introduce new independent design features. "The notion of cosmetic variation seems to imply a sharp dichotomy between design changes which are easy to make and those that are not" [8].

*Flexible mass production* extends cosmetic variation by introducing the potential for change in interdependent design features. Flexible mass production explicitly identifies, a priori, both the set of product design features which is subject to change and the set of values which each design feature may take. Therefore, the flexible mass production system represents a finite number of products, which vary in more than simple cosmetics.

Diametrically opposite mass production is *flexible specialization* where variation is virtually infinite. However, *closed flexible specialization* includes those systems where the set of design features that varies is defined a priori, but the domain over which each varying feature ranges is unknown. By contrast, *open flexible specialization* where both the set of variable design features and the domain over which each variable ranges is potentially infinite.

A different cognitive model applies in each stage, with flexible specialization involving a balance between "a deepening of understanding within a given cognitive frame and the pull to reintegrate back (in the production process) to a different frame in order to produce a sellable commodity" [11]. Similarly, a technology-independent bearer service offers more than a predesigned set of services. Rather, it supports an "*application-blind interface*" that enables the introduction of new applications independent of the initial strategies and service offerings of Telecommunicationos Operators (TOs). given (ex-ante designed) set of services: it provides a basic functionality, an *application-blind interface* in the network jargon. This functionality can be easily reintegrated to the application vendors and users' cognitive frames, thus allowing them to introduce new applications and operate independent of the strategies and the service offerings of the Telecommunications Operators (TOs). To illustrate the continuum that spans well-defined, mass production systems and flexible, application-blind interfaces, we will discuss two contrasting metaphors, the *gauge* and the *container.*

### *2.2. The gauge metaphor*

The gauge of a railroad is defined as the distance between rails or between the flanges of the wheel sets on a railroad car. The gauge determines the tracks upon which a given railroad car may travel. By extension, the gauge therefore also determines which railroad companies may exchange rolling stock and the transparency with which a customer may transport freight across boundaries between different railroad companies. Accordingly, diversity in gauge standardization implies transaction costs and other inefficiencies as customers and freight traverse rails of different gauges. Thus, for reducing technical complexity and transaction costs to internalizing mutual network externalities, railroads have been progressively converging towards a gauge standard [12, 13].

The emergence of the spanning layer concept may be derived directly from the convergence towards a rail gauge standard. Railroad tracks comprise the physical network. Differences in rail cars represent distinct applications from which a customer might choose. Gauge dimensionally is therefore a layer which resides between the physical track and the applications. Gauge standardization expanded the services that a particular rail system could offer by extending the reach of the rail network and expanding the scope of traffic (the kinds of cars) that could be carried. Standardization reflected a shift towards some flexibility where rail car design could vary (as long as it conformed to the gauge standard) and traffic could move across network boundaries.

### *2.3. The spanning layer in traditional communications networks*

Communications networks have traditionally been vertically bundled. Whether for telephony, radio, broadcast television, or community access television (cable), infrastructures have long been closely coupled to service provision. As Tennenhouse et al. notes [14], "telephone and cable services are each carried over their own wired systems.

Although radio and television share the airwaves, for practical purposes, they are discrete distribution vectors, since separate portions of the spectrum have been allocated to each type of service" . With one wired network for any given application, each service resembles a mass production system. *Spanning* is trivial.

Even within a single service offering, however, the utility of a spanning layer is apparent. Telecommunications heterogeneity is subject to the same economic forces promulgating rail gauge convergence. Monopoly structures and the political environment contributed to both the rapid acceptance of the 4 kHz circuit as a worldwide standard for voice communications[4]. Subsequently, that standard emerged as a spanning layer providing transparent support for multiple applications (for instance fax and data transmission via modems) over technologically evolving analog and digital telephony[5].

The advent of digital computer and communications technologies, in addition to introducing the prospect of new applications and services, began to extend the scope of the spanning layer concept. For digital telephony, the 4MHz channel gave way to a transfer rate of 64 Kbps using pulse code modulation (PCM). While corresponding to a single voice conversation, the data rate of 64 Kbps also introduced a new category of *value-added* applications and services (e.g. credit-card calls, file transfer, audio-conferencing, etc.).

For value-added services, however, a raw 64 Kbps bit pipe proved inadequate. Specifications for data format, routing, signaling, and many other parameters of transmission are required to enable high speed data services. Such services include packet switching, Frame Relay, intelligent network applications, and Integrated Services Digital Network (ISDN). Consequently, value-added services led to a new network model which extended the 64 Kbps spanning layer with delay, loss, and other data characteristics. Rather than cosmetic, these additions drove the system towards flexible mass production where the network supported a discrete set of delay values, etc. to support a limited number of new, pre-specified services.

The additional switching and control functions for these new services are *logically* implemented on top of the elementary 64 Kbps spanning layer yet *physically* implemented elsewhere than the Central Office Equipment. As examples of the distinction between the logical and the physical, consider X25 and ISDN. The transmission of packets between nodes attached to the same transmission link is assured by a specific *connection-oriented* protocol, the X25[6]; the *data link* layer of the X25 converts an *unreliable* for packet transmission bit pipe to a packet-link, and, in collaboration with the *network* layer, supervises naming, addressing, routing and congestion control[7] [16, 17]. In a similar way, the ISDN (2B+D) provides a customer with two B channels of 64Kbps[8] and one X25 full duplex D channel at 16Kbps for "out-of-band" signaling; furthermore, under the *Intelligent Network* model, ISDN separates the flow of control information from user information[9] [18, 19].

Despite the new functionality and extensions to functionality supported by the spanning layer, the spanning capability was essentially limited. While a single infrastructure may have supported more than one application, service offerings continued to be tied to the infrastructure in a largely many-to-one correspondence. Infrastructures were not application blind. For example, changes in the X.25 protocol would necessitate change both within the network and at the end nodes. Overall, the public network resembled the shape of a

---

4 As in the case of the railroads, however, those forces may sometimes be unable to achieve complete standardization. In broadcast television, *spatial lock-in* contributed to the emergence of three different standards (NTSC, PAL, and SECAM).

5 The development of applications other than those selected by the network operator, though technically feasible from the beginning, became however effective only following a change in political climate as signaled by the FCC's *Carterfone* decision [15].

6 Other technologies for *connection oriented* packet-switching include SNA of IBM and DECNET of Digital Equipment.

7 We refer to OSI terminology

8 These channels can be indifferently used for voice and data transmission

9 Even though theoretically possible and very appealing to separate user and flow information, this separation *(unbundling* in the telecommunications jargon) is actually difficult to implement, for reasons relating to the particular strategic interests of the Telecommunications Operators [20, 21].

"patchwork" of vertically integrated, flexible mass production systems where each network supported a different set of applications. Within each of these production systems, one can still distinguish between the conception of a service and its delivery, because of the spanning layer is *being embedded* in the substrate technology. Variation exists but only the variation which has been ex-ante conceived and designed by the engineers of the infrastructure-facilities owners.

### *2.4. The new metaphor: the container*

Early transportation systems provide examples of vertically integrated systems. For the purpose of transporting goods over land, dedicated technologies including railroad tracks, cars, and rail yards were created. Likewise, air and water transportation services warranted similarly single-purpose infrastructures for ships and airplanes. Over time, however, it became common for freight transporters to link different media, such as water, air, or land. Intermodal freight transportation refers to the linear combination of two or more transportation services. For example, moving a product from a warehouse in Hong Kong to a retailer in the American city of Chicago might entail trucking the product to the harbor, shipping the product to a port on the West Coast of the United States, and then conveying it by rail to Chicago.

Contrasting vertically integrated architectures, where a dedicated infrastructure is used to realize separate applications, is the *horizontal integration* model [22],where diverse applications are developed independent of any underlying architecture. Instead, developers build to a common *abstraction*[10]. The common abstraction, a spanning layer now located *on top* of the transportation (or communication) modes, projects a virtual infrastructure which itself comprises one or more interoperable physical networks. The spanning layer pairs applications to infrastructure facilities because applications request network service[11] from the spanning layer not the underlying infrastructure, and the spanning layer translates a request for network service to the infrastructure protocols.

For the transportation industry, horizontal integration, which emerged only in the 1950's, was made possible by the development of and agreement upon a standard freight container. Muller noted that "when it was publicly demonstrated in 1956 that standard containers could move successfully on a land-sea intermodal journey, a commercial revolution was started… *It was the container's unique role as common denominator among modes that was revolutionary* (we underline)" [23].

### *2.5. The Internet Protocol: a robust technology-independent spanning layer*

In the communications industry, diversity and, consequently, requirements for a common denominator among different technological media, also become increasingly important. Local and Metropolitan Networks providing inexpensive and fast interconnections for personal computers and workstations proliferated inside and outside firms and among research institutions and universities. But, successive technological waves have dispersed the computer networking environment in different architectures (LANs: Ethernet, Token Ring, Apple Talk, MANs: Frame Relay, SMDS etc.). Interconnection

---

[10] Messerschmitt [22] defines *abstraction* as "the conscious hiding of unnecessary implementation of functional details while making visible behavioral properties that are essential and important to other modules. Abstraction helps to ensure the independence of the architectural modules, as one module cannot be dependent on the hidden properties of another module".

[11] The term "network service" indicates here the actions and responses (i.e. functionality) provided by underlying infrastructures to applications' requirements. Not confuse this *network service* with that we usually call *services (or service offerings),* i.e. the customer services provided by a network infrastructure to final users.

between these architectures, as well as connections between these private networks and the public telecommunications infrastructure, appeared as the natural priority. In the 80s, three technologies were compete to capture the interconnection market (SNA, B-ISBN, ARPA's Internet protocols), with the *Internet protocol suite* becoming more recently the solution most often adopted [24].

The Internet approach to interconnection involves the separation of infrastructure bitways from applications by defining an interface (Internet Protocol) to the basic infrastructure facilities and then exporting that interface for application development. Formally, IP provides a uniform method of addressing (which is independent of physical hardware addressing) and a variable size datagram *(i.e. a standard data container)*[12]. From this interface any number of diverse applications (e.g. the World Wide Web as an example of new, unanticipated applications) may be constructed [25]. Likewise, principles of datagram encapsulation enables any number of substrate technologies to transmit IP (e.g. Ethernet, token ring, X25, ISDN, Frame Relay, even ATM) [26].

IP takes an additional step towards supporting flexibility by disassociating *transport* from *network* functions (respectively performed by TCP and IP, TCP being over IP) [25]. As a result, IP incorporates only that functionality which must be implemented within the network as opposed to at the periphery or end-user nodes. In this way, IP minimizes the possibility that changes in applications (including transport protocols) or infrastructure components will require changes in all of the routers, hubs, and gateways throughout the network substrate.

By separating spanning layer development from both infrastructure facilities (and applications), IP characterizes a different technological trajectory[13] that stems more immediately from the container metaphor. Essential characteristics of this trajectory are that applications may work over multiple substrates (e.g. network technologies) and that these substrates do not pre-specify the development of new applications. The separation both above and below, move IP further along Piore's framework towards flexible specialization. Certainly IP offers flexibility[14] and features variability both above and below the spanning layer.

### *2.6. Enhanced IP with RSVP and ATM: different approaches to horizontal integration*

The Internet protocol suite (TCP/IP) provides the foundation for the current data communications infrastructure in the United States, the Northern Europe and much of the rest of the world. However, the current architecture provides only a *best-effort* delivery service. Consequently, it can not deal with applications requiring service guarantees, such as real-time applications (telephony, video etc.). To respond to this objective, new service models are being defined in the Internet together with protocols to reserve capacity according to applications' requirements. Essentially, the goal is to bring out an architecture for Integrated Services Packet Networks -- ISPN [29].

New concepts for spanning, therefore, propose to complement IP datagrams with functions to provide not only best-effort delivery but also variable Quality of Service (QoS). With the major part of the research effort devoted now, to create a new "building block" able to specify QoS parameters (as bandwidth, delay and loss characteristics), the search for a new Internet architecture seems to cognitively approach the ODN model. Both approaches impose a narrow point in the protocol suite, "isolating the application builder from the range of underlying network facilities, and the technology builder from the range of applications" [3].

---

[12] To use Rose's terms [26], *IP has no wires associated with it.*

[13] in the sense of Dosi [27]

[14] Borrowing from Bar [28], "... it is the flexibility of its network resources will determine a company's ability to experiment with telecommunications technologies, learn from the experimentation, and repeatedly reorganize itself to capture cumulative benefits. True network flexibility is not only the ability to support a range of applications over a given network configuration but also, and perhaps more importantly, the simultaneous ability to re-configure a network in order to provide various applications mixes and to design new applications that take advantage of new configuration possibilities".

This newer approach for internetworking goes beyond networks interconnection, to horizontal integration. It adopts one of the most successful architectural elements of the current Internet – the existence of a spanning layer located *outside of the wires*, provides, on top of the IP, *additional functionality* for applications supporting services with *multicast* and *bounded delays* constraints. At the time this chapter was written, the Internet Engineering Task Force (IETF) is working towards standardization of a set of extensions to the datagram functions. Essentially, their effort will standardize the protocol which will generate a *flow* from the source to the destination (i.e. a sequence of packets associated with a single application and sharing common requirements for bandwidth allocation) and then, give to the network elements (routers and switches) the ability to directly recognize the existence of this sequence [30]. Stated in another way, this approach aims to define a *service model* which describes[15], in terms of Quality of Service, the set of network services offered to the applications and then, establish and maintain at each network element a *reservation state* to provide the requested by the application service.

Key mechanisms for the implementation of that scenario include [31, 32]: i) an application-initiated *flowspec* (corresponding to a set of information), signaling to the network the type of service this application requires; ii) a process of *admission control* decides whether the network can provide the requested service; iii) a *resource reservation* process which establishes and maintains state[16] to all nodes along the path(s) of the flows, thus providing the requested service[17]; iv) a *packet service algorithm* to schedule packet transmission which lets the network switches and routers meet the flow's requirements. A specific protocol, the RSVP, has been standardized (and now is being implemented) for these purposes. Even though RSVP resides on the top of the datagram, it does not belong to the transport protocols family (usually implemented in the end nodes), since RSVP does not transport applications data but only accommodates flow control functions. The new Internet spanning layer can therefore be defined as the *reservation paths*[18] the applications "follow", in order to operate under their own requirements in Quality-of-Service (QoS). Apparently, as RSVP over IP is only concerned with the QoS of the forwarded packets[19], this bearer service can work with any substrate technology, including ATM virtual paths.

ATM proposes a different approach to the horizontal integration of the existing patchwork of vertically integrated communications services, through a single substrate which has been designed to accommodate a diverse type mix of traffic (voice, data, video etc.) [22]. In such a model, the virtual network consists of a single technology. ATM's container, corresponds to the 4 kHz circuit or the 64 Kbps channel, is a 53-byte cell [33]. But unlike the PSTN, ATM constructs virtual paths of varying bandwidth out of 53-byte cells. As alluded to above, a second characteristic of ATM model is the architectural integration of bitway technologies (figure 2).

---

[15] Shenker [29] proposes such a service model with two kinds of real-time service, several classes of best-effort services and a modified virtual leased-line service (for "bulk" virtual capacity).

[16] Clark [25] calls that a *soft-state*, in the sense that reservation is maintained at each network element but this is periodically refreshed by the end points.

[17] This is called *signaling* in the telecommunications world.

[18] We can imagine individual resource reservation but also reservation for groups of "calls", or multicast data distribution and multipoint-to-multipoint paths.

[19] As Braden et al. [31] point out, the RSVP it is not itself a routing protocol: "RSVP is designed to operate with current and future unicast and multicast routing protocols. An RSVP process consults the local routing database(s) to obtain routes".

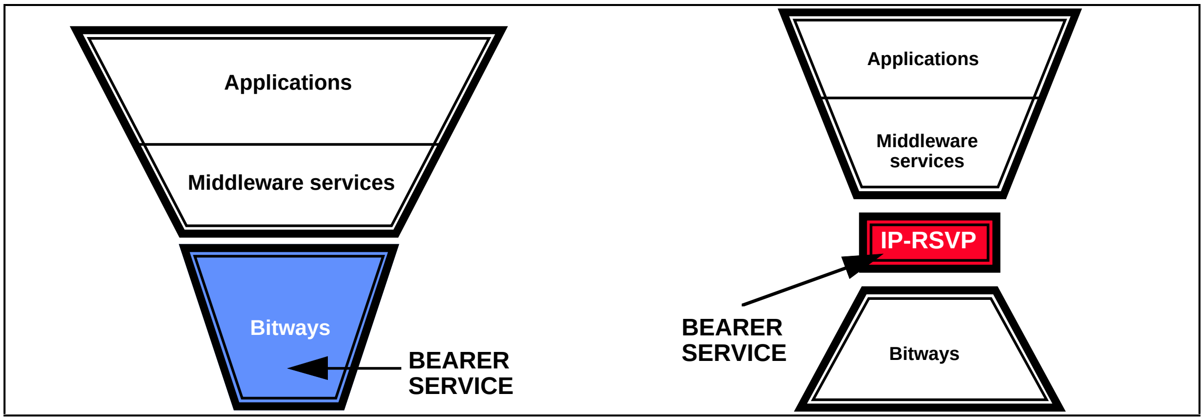



**Figure 2: ATM versus RSVP over IP**

In more general terms, the consideration of the ATM layer as the "spanning layer" raises some problems. One is that ATM is simply one technology among many, and given the rapid evolution of LAN technologies (fast Ethernet etc.), it would be hard to imagine ATM becoming the universal technology basis for applications development[20]. A second problem relates to implementation issues. Telecommunications Operators have adopted a "command and control" approach for signaling (SS7) between ATM switches, which has been criticized for promoting centralized control [22]. Consequently, if ATM should go "to the desktop", specific regulation for "unbundling" and more symmetric implementation might be defined. Still, the efficiency of such policies is open for debate, given the relative success of the Internet which is not founded on a "command and control" philosophy and is governed with minimal regulation.

### *2.7. From spanning layer to bearer service*

This section began by considering the traditional telecommunication infrastructure and its successive evolutions in light of Piore's perception of the transformation of the industrial system. The organizational archetype is the *Bell System* design, a ubiquitous *mass production* system for telephony, where application and substrate are not only tightly coupled but in fact developed in concert. In time, this model has produced a slight differentiation, by offering a number of additional services (fax and data transmission via modems) over the initial spanning layer (i.e. a 4 kHz point-to-point channel, the telecommunication equivalent of the *railroad gauge).* With the introduction of digital technologies, we observed a more fundamental transformation in network production towards increasing variety at the level of service offerings. However, the application and service design are still designed into the system ex ante. Therefore, flexibility is *formally* developed within a "closed set with a finite number of elements" *(flexible mass production).*

The section then explored newer network models. Interconnection of heterogeneous networks (as exemplified by the Internet) and, now, horizontal integration, mark the continuing shift from *cosmetic variations* in mass production to more *flexible mass production* or possibly *flexible specialization.* In these models, applications may be able to go over multiple network technologies (substrates) without the need to design applications and services in advance, as a function of particular substrate constraints.

By fitting the evolution of network architectures to Piore's framework (figure 3), we were able to extract functions of the *spanning layer* and predict how the *bearer service* may evolve. A trivial spanning layer, embedded in the substrate technology, seems to characterize early railroads and telecommunications mass-production networks. Increasing service variety, due to digital transformation of the network, required to develop extensions to the functionality supported by the spanning layer. However, the spanning layer itself

---

[20] *Network World,* The shrinking world of ATM, 30 June 1997

remained tied to the infrastructure *wires*. It escapes from the wires only with the Internet, which proposes a network model that might be described by the *container metaphor:* a specific protocol (IP) plays the role of a common denominator among different substrate technologies. Because IP is a technology-independent spanning layer which enables independent variation of applications above it as well as substrate technologies below it, it is at a layer above the bitways which enables more design flexibility and innovation.

In the near future, integrating flow control mechanisms (assuring Quality of Service) with IP should extend Internet's application variability to include services with real-time constraints. Again, rather than implementing this functionality within the bitways, an enhanced IP embeds the additional functionality into the spanning layer. In this way, a new "building block" (RSVP over IP) appears for supporting a network composed from Integrated Services Packet Networks (ISPN). These characteristics of the new Internet "building block" (the span of underlying infrastructure facilities and the range of applications that can be served) in combination with the existence of *symmetrical* properties (the type of delivered service is enforced by the end points) make it look as a sort of implementation in the Internet real-world of the ODN *bearer service* concept for two reasons. In the ODN network model, the discussion on the bearer service is in abstract terms and there is no provision for implementation. But the position of the "spanning layer" is exactly the same as the bearer service, it is close to the network technology substrate. Throughout, the whole system may scale over number of networks and different substrates. And second, the ODN bearer service layer is clearly defined as providing quantified Quality-of-Service measures for band-width, delay and loss characteristics. Simultaneously, this layer must also provide mechanisms for accounting at a sequence of packets level and incorporate feedback loops to utilize accounting statistics in capacities such as evaluating QoS commitments and traffic management.

| **Piore transformation** | | **Freight transport** | **Comm networks** | |
|---|---|---|---|---|
| **Mass production** | **Tech-embedded** | **Early railway** | **Early voice** | **Spanning** |
| **Mass prdctn w/ cosmetic variation** | | | **Fax and data 4 KHz analog** | |
| **Flexible mass production** | | | **Digital voice** | |
| **More flexible mass production** | | | **ATM substrate-bound** | |
| **Flexible specialization** | **Tech indpt** | **Container based multi-modal** | **Internet + multiple apps/subs** | **Bearer service** |

**Figure 3. Trajectories of technical change**

This approach contrasts with an "all ATM network" strategy for horizontal integration. Though ATM uses a common container (a fixed size cell), in at least one respect, ATM appears to evolve directly from the *railway gauge* metaphor: while ATM is application blind, it does not offer independence from the substrate technologies. Essentially, ATM tends more towards a flexible mass production system (or a *flexible* *specialization with large and discontinuous change* system[21]) by pre-specifying the basic commodity products for meeting customer needs or constructing higher level applications (ATM was designed as an applications independent platform with three general applications at its center: telephony, high speed data communication and video delivery).

To conclude: A spanning layer lies between a set of network substrates and higher level applications. *As a special case of a spanning layer, the bearer service is both substrate independent and application blind.* Only those functions which cannot be implemented at the periphery of the network substrate are included in the bearer service. Because not all substrate technologies can support all application requirements, application requests must be matched to particular substrate characteristics. To preserve transparency, the bearer service conducts the pairing rather than having applications explicitly request particular underlying technologies.

## 3. The economics of the bearer service

The evolution of communications spanning layers from homogenous substrate/application pairings to the bearer service may change the economics of this industry. While section 2 concentrated on how spanning layers have evolved to incorporate heterogeneous applications and substrates, this section will explore the economics of this evolution and describe the eventual structure of the bearer service market. We describe why the conclusion of vertical integration described by Gong and Srinagesh [6, 7] may not be the only way to sustain a bearer service market. Essentially, we will raise questions about possible *governance structures* in networks with layered architectures.

### *3.1. Institutional forms associated with the "container" network model*

The argument we developed earlier built around a distinction between two different trajectories of technical change. The first derives from traditional telecommunications *mass production* organizational models and involves a pattern of innovation consisting of anticipating user needs in communications services and designing the corresponding underlying network architecture. In contrast, the Internet and newer approaches for horizontal integration of an infrastructure becoming increasingly heterogeneous, and separate bitways from applications (and service offerings). Technically, this is possible by defining an interface (IP) to the basic infrastructure facilities and then building on top of that interface the general functionality for applications. We argued that this particular approach to technological change leads to a new technological trajectory which we may metaphorically call *container-based interoperability*. Its essential difference from previous organizational models is that applications may work over multiple substrates (i.e. network technologies) and that these substrates do not determine the development of new applications. It is increasingly clear that this new trajectory brought up considerable flexibility in the design of applications the customers use (as exemplified by the extra-ordinary development of the World Wide Web). Because it "blindly" uses network resources, system production (applications and service offerings) takes the characteristics of an open-ended product line able to generate (possibly infinite) varieties.

The *container approach* appears to be different than mass production since it is associated with institutional innovations which have begun to crystallize into a *new specialization*, as is apparent from the market of *Internet Service Providers (ISPs).* To explain ISPs sudden appearance, one may argue that the adoption of a design imposing a technology-independent bearer service, introduced incentives to explore new forms of organization which had been previously rejected [24]. If IP involved a separation of the

---

[21] Piore [8] defines *flexible specialization with large and discontinuous change* as follows: "In any system, some moves will be incremental and some will be large and discontinuous... An open (flexible system) for which a change is large and discontinuous is in some respect similar to a system of flexible mass production".

applications from the infrastructure facilities, efficient utilization of the resources would not require a multilevel switching system (and the corollary huge investments raising powerful entry barriers), as in the case of traditional telecommunications networks, but only simple forwarding of "IP packets" (i.e., the network is *not cognizant* about the applications they carry). The question then is to organize a flexible system of regional or more extended backbones and access links to these backbones. In order for this to occur, entrepreneurial, flexible, and small organizations have emerged in the Internet Services Provider market. They offer a bundle of services to their customers (such as Internet Access Providers, individual organizations, firms, and academic establishments), mainly for IP transport, but also may provide the necessary hardware and software, and customer support [34].

Although still a somewhat embryonic sector, the ISP proliferation challenges the huge classic organizational model of the Telecommunications Operators (TOs). However, a number of colleagues suggest that ISPs might not be able to stay as unintegrated firms and will move downwards, trying the acquisition of their own infrastructure facilities. And, on the other side, backbone providers (mostly, traditional Telecommunications Operators) will integrate upwards, since, according to Gong and Srinagesh [6, 7], the economics of the *layered networks* would not tolerate niche players leveraging off of the investments of the infrastructure owners[22]. Essentially, as the same authors also point out, the question is to know whether ISPs and TOs with varying degrees of integration can coexist or if the ISP model based on external provisions of network capacity is only a transitory one? In a large part, the response of this question depends on the existence of an efficient market for technology-independent bearer service.

### *3.2. Unbundled bearer service?*

This market may not be sustainable, according to Gong and Srinagesh [6, 7]. Their argument consists of a series of propositions which we summarize here:

i) In networks with layered architectures, competition at the bottom layer (infrastructures facilities) of the network hierarchy, is unsustainable. Markets for raw transport under conditions of excess capacity and oligopolistic competition for a homogeneous good, turn easily to “destructive competition” arising from Bertrand equilibrium[23] (prices decrease to marginal costs following the example of the leased line market). Consequently, facilities-based carriers competing for raw transport essentially on price, may not cover their sunk and fixed costs and fail to afford competition with non-facilities based resellers and ISPs.

iii) One way to avoid Bertrand competition is through bundling of bottom-layer transport with higher services, closer to final customer (vertical integration). With bundling, carriers are able to differentiate their services, segment the market and price accordingly.

iii) As facilities-based companies integrate with others at higher layers, variable costs rise significantly along with the minimum efficient size.

iv) Given the above hypothesis, “policies promoting competition in the provision of unbundled bearer service among owners of physical networks may ultimately fall”.

Gong and Srinagesh extend their analysis to other forms of bundling and horizontal mergers to demonstrate that integration may be a natural outgrowth of competition in the convergence spurred by technological advancement (layered digital architectures) and deregulation. Albeit increasing diversity in service offerings, *economies of scale* and *sub-additivity* seem still powerful and responsible for generating growth, as in the old trajectory

---

[22] As P. Srinagesh kindly reminded us, recent mergers between backbone providers and ISPs (GTE-BBN, MFS-WorldCom-Alternet) seem to give empirical evidence to this statement. We note however, that in many industries, waves of mergers coexist with, or are punctuated by other less centralized forms of organization.

[23] Gong and Srinagesh describe raw transport as a service that is like a commodity. A commodity is a product that is so standardized and sold by enough firms that no firm can set the price — all firms are price takers. The Bertrand competition model in economics holds for such a market. Assuming that all firms have a different marginal cost and they will exit the market if the price is set below or equal to their marginal cost, the only sustainable market price is equal to the marginal cost of the second-lowest firm.

of *mass production.* However, counter examples from the Internet come easily to mind. While it may still be too early to tell whether the Internet Service provision is "sustainable", it is true that there is a mix of integrated and non-integrated ISPs, with market competition rewarding more and more product differentiation[24] and competence accumulation (i.e. learning) in improving offerings[25].

In analyzing the economics of layered networks, one should probably make a distinction between economies of scale associated with fixed investments in facilities and other forms of growth (new applications generating markets and user-oriented learning). Some firms may grow even though vertically integrated suppliers representing closed systems will dominate the market. Stated in another way, the question is to know whether the classical Chandlerian hypothesis [35] indicating that in infrastructures, merger-generated megafirms yield a clear advantage in the competition, is a general statement or a phenomenological one[26]?

This chapter does not suggest answers to all these questions. It simply addresses the topic of an efficient market for bearer service which seems to be a necessary condition (but not sufficient) for allowing different organizational structures to coexist. Refocusing into our primary concern, the remainder of this section will explain the characteristics of the "bearer service" *product* (very different from a commodity product). Next, we will explore opportunities for a sustainable bearer service market under two different perspectives, i) *product differentiation* and ii) allocation of *rights of control.*

### *3.3. What exactly is the "bearer-service" product?*

Defined as a network substrate-independent interface with Quality of Service (QoS) characteristics (section 2), the bearer service *as a product* looks like *sheer bandwidth with QoS attributes*, available over wholly owned networks as a standard (platform-independent) good with quality options ($QoS^1$, $QoS^2$, ... $QoS^n$). From this definition it should be inferred that bearer service is different from a commodity product and, consequently, the bearer service market would be different from raw transport markets.

A commodity product has the characteristic that the demand elasticity of substitution is infinite — if the price of the commodity goes up, then the demand goes to zero. If the price goes down, the vendor receives all the demand[27]. This particular characteristic of a commodity results in the service providers (or vendors) to set their price at the same level[28]. The leased line market may not have an infinite demand elasticity of substitution, but it is much larger than the elasticity for substitution of a bearer service product. The reason comes from the ability of the leased line vendor to fulfill the service request it receives from the customer versus the bearer service provider to do the same. In the leased line case, customers specify the bandwidth of a connection often days or weeks in advance of using that bandwidth. There may be some specification about the reliability of that link, but, in general, the leased line is always there for the customer. In a bearer service market, the customer demands service in a dynamic manner dependent upon the application's service requirements — perhaps only a few milliseconds before the service is delivered. Therefore, the bearer service provider must choose technology substrates to service application requests based upon the expected demand. This not only results in a usage sensitive pricing policy to limit the use, but it also encourages the service provider to make the exact provision in network capacity to minimize costs (taking advantage of statistical sharing to

---

[24] Varian in [30] recognizes an evolution of the Internet with ISPs now differentiating their offerings without necessarily getting into the content business. Options for differentiation include fast-packet switching technique, proprietary routing techniques, development of tacit complementary assets for security and network management and, probably, capacity management techniques such as yield management.

[25] A quick examination of *The List* (http://www.thelist.com), a listing of Internet Service Providers, offers many examples of non-integrated ISPs with good local or more extended implementation.

[26] On this interesting theoretical question, see in particular Dosi [36].

[27] This is true as long as we live in a world with weak *increasing returns.*

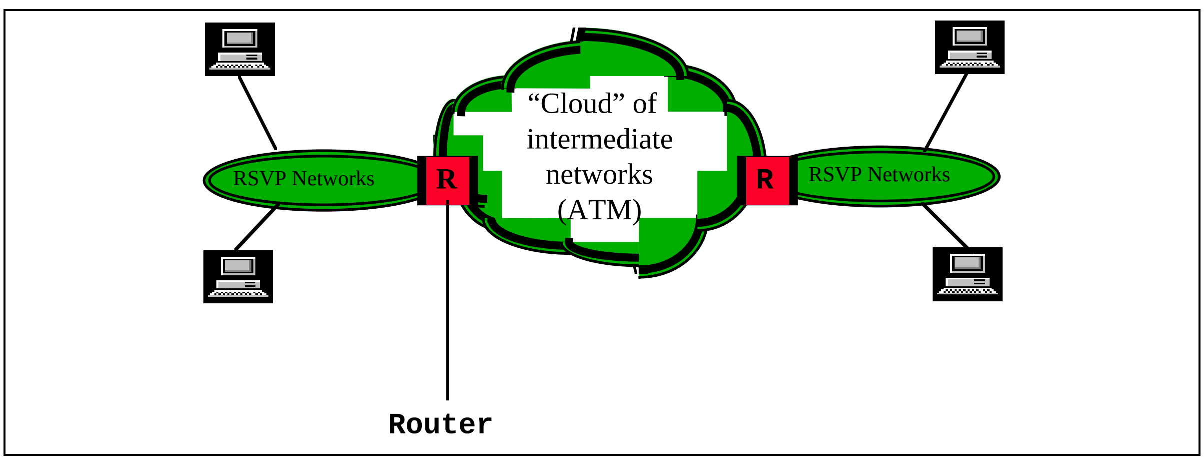
[28] Assuming slight product differentiation, the market would be shared by a number of companies and an equilibrium price would be established – given different product characteristics and availability.

service their users requests). The result is a probability of blocked service request that does not exist in the leased line market. Customers then see a noticeable difference between bearer service providers resulting in a price versus performance decision on their part. Because one bearer service provider underprovisions its network less than its competitor, it will also increase its performance by refusing fewer service requests, thereby commanding a greater price for its service. The result is a demand elasticity of substitution that is less than the leased line market indicative of a non-commodity product.

Another commodity characteristic that leased lines have and the bearer service product does not, is transport between two points as opposed to between a point and a shared "cloud" (figure 4). The leased line customer specifies the points they wish to connect, not only their bandwidth requirement. It is a point-to-point connection. In the Internet, however, the service between two points traverses through the cloud and the full point-to-point performance is dependent on the sender's access link *and* the recipients. Likewise, the bearer service does not specify two communication points a priori, but it only gives access to the shared cloud. While the bearer service cloud does more than the Internet cloud (e.g., provides a guaranteed quality of service), the ultimate functionality depends upon the access link of bearer service provider connecting the sender to the cloud and the quality of the access link connecting the recipient of the cloud. If the same bearer service provider connects both the sender and the recipient, there may be no reason to traverse the cloud and the provider can handle the transmission internally. The result is a provider that is able to offer better service because it has a large network of customers. The economics of *increasing returns* apply here giving customers greater value if there are more customers connected to their bearer service provider's network.

The "cloud" metaphor is used to describe the statistically shared wide-area transport of the Internet. In this chapter, we argue that the "cloud" metaphor can also be used to describe the bearer service network topology. More precisely, we assume a network topology with a fast-packet (for example ATM) "cloud" and RSVP connections to Users' Networks.

Overall, the leased line product does look more like a commodity than a bearer service because the offering is relatively homogenous and static. The service demands of customers in the leased line market do not change dramatically from day-to-day because the service has been established a priori for a fixed time period (usually months or days). The bearer service provider must service requests very dynamically and it must service requests for an unknown or short period of time (perhaps seconds or minutes). Furthermore, the leased line product that one provider can offer versus another provider is not very different. In the bearer service market, there are distinguishing characteristics about the provision of bearer service that can differentiate the providers, thus making competition less susceptible to "destructive prices"[29]. In addition, bearer service provision requires complex investments in different families of assets ranging from network hardware and management tools to link capacity.




29 Besides, even globally decreasing prices in leased line market are not so "destructible". Big discounts and "back of the book" tariffs for large volume and long term contracts, seem to correspond to *non-linear* pricing strategies and other *pricing discrimination* techniques.

**Figure 4: The next Internet**

### *3.4. Bearer Service differentiation*

Since the bearer service is not a commodity, it is important to explore ways that it can be differentiated when provided by competitors. Independent bearer service providers can: i) choose different substrate technologies, ii) design a different network topology than their competitor and, iii) design a different pricing policy for their service. Each of these three methods will be explored in detail below and compared to the leased line market.

The choice of substrate technology is the first way that bearer service providers can differentiate themselves from a competitor. As discussed earlier, while the bearer service can support all applications and all substrates, it is not possible for all application/substrate pairings to work together. But, by choosing a particular substrate technology, the application space can be constrained. For example, if the bearer service provider chooses Ethernet as its substrate technology, it cannot offer a guaranteed quality of service to the customer while their competitor with an ATM substrate can offer this service. The subset of applications offered by the Ethernet-substrate provider is smaller than the ATM substrate provider, but their cost structure is also different. By choosing Ethernet, their costs are lower and, therefore, they can pass on some savings to the customer. In summary, the application space may be constrained by choosing a particular substrate technology.

The choice of network technology is another degree of differentiation between the bearer service providers. As outlined above, connection to the *cloud* is very different than connection between two points. Therefore, the way a bearer service provider connects to the cloud and how dependent the provider is on the cloud for service are differentiating factors. The better the connection to the bearer service cloud and the less dependent the provider is on the cloud, the lower the probability of a denied service request. To use a metaphor from transportation, the bearer service provider can provide transport to a "hub" or exchange point where you can hop on another link. If this exchange point is in a remote location, or experiences heavy congestion, then the service is providing you less value than if it takes you to hub with greater service.

Pricing policies can be a powerful service differentiation mechanism. Since service requests for bearer service provision must be answered dynamically, it is likely that pricing will also be set in a dynamic manner (for example, as function of the requested QoS). Pricing policies for the Internet to give a quality of service better than "best effort" are discussed in McKnight and Bailey [37]. How a provider chooses to make capacity provisions and how defines prices to maximize its revenues, should distinguish one bearer service provider from another[30].

### *3.5. Shared "residual control rights"*

Perhaps another way to evaluate the sustainability of a bearer service market is to consider the physical assets, in the post-investment stage, which make the relationship between suppliers (S) and buyers (B) of bearer service – B may correspond to a non-facilities based ISP. Then, following Hart [38], we can assess the economic implications of different governance structures.

One critical issue for governance structure is what Grossman and Hart [39] call the *rights of residual control*. These are the rights to make decisions about the usage of an asset

---

[30] Other ways the bearer service providers could differentiate themselves include marketing and customer service. While these differentiating factors have nothing to do with the transport of bits or the ability for an application to access the bearer service, it does influence the customers' decision when they trust their provider or achieve benefits from better customer support. Brand recognition that results from advertising and marketing can influence a decision maker and also lead to differentiation between competing service providers. These differentiation tools are understood by the marketing literature and are beyond the scope of this chapter for adequate analysis.

which is not anticipated in a contract between the parties involved (incomplete contracts). There are different ways to allocate these rights, integration (in both directions, from S to B or B to S), or non-integration, with integration involving the ownership of an asset (which goes together with the residual rights of control over the asset). The choice of the ownership structure depends upon the specifics of the relationship investments and the distribution of the information required to make decisions. Grossman and Hart [39] develop a theory of integration based on the above principles which allows to formally evaluate the costs and the benefits of integration (or non-integration).

The allocation of the residual rights of control is obvious in the mass production model. As long as the bearer service resided *within the wires,* the rights of control over bearer service assets (i.e. network resources offered under *Open Network Architecture* provision) are naturally held by the infrastructure providers. Resellers and other *third parties* are bound by a contract which gives them access to network resources but it is understood that the infrastructure owner has the right to use the asset any way that is not inconsistent with the law. The infrastructure provider (S) would have the right to define production capacity and decide on technological investments, assuming that the initial contract is silent about these. Because the operating firm possess the residual rights, it will receive a greater fraction of the *ex post* surplus created by its investments, so these rights work as incentives to make investments and expand the relationship with the third parties.

But with bearer service *escaping from the wires,* flexibility is increasing because the *openness* of the IP protocol makes the assets less complementary, i.e. less specific to the relationship. Under the assumption of a network with *fast-packet cloud* and RSVP routers to connect to that cloud (figure 4), investment decisions at both the *cloud* and the *periphery,* so non-integration is likely to be a serious hypothesis[31]. As Piore [8] concludes in more general terms, when the rights of control are held by organizational components upstream in the production process, the organization may be more decentralized[32].

---

[31] Hart [38] considers non-integration as optimal, as long as assets are non-complementary (independent). In this case, the effect of the full control transfer to S (a backbone operator) is to reduce B's relationship-specific investments, which may reduce openness and, corollary, total surplus.

[32] The thoughts outlined above are at a preliminary stage and one should pursue them further to come up with firm conclusions.

## 4. Conclusion

What we suggest is that the impressive diffusion of the Internet may signify a transition towards flexible models of organization in the communications industry. Flexible organization involves independence of applications from the investment cycles of the infrastructure facilities and it significantly depends upon the efficient existence of a bearer service market. Such a market, if implemented, may establish the *symbiosis* of backbone suppliers with independent ISPs.

One of the main conclusions of this chapter is that the bearer service may be sustainable in the absence of vertical integration. It is the flexibility of the bearer service itself which helps the bearer service providers to differentiate themselves from their competition and set prices based upon varying degrees of value to their customers. The differentiation of the service can be realized in many forms. The bearer service provider can choose a substrate technology to better support a set of applications, they can design a network topology to minimize the number of denied service requests, and they can formulate a pricing policy different than their competition. The different offerings can appeal to different classes of users and provide a mechanism for price discrimination and service flexibility so both the provider and the customer benefit.

However, the ability to vertically integrate to absorb the bearer service is also possible. We realize that there are benefits to vertical integration from a substrate provider to a bearer service provider, or from a bearer service provider to a middleware provider. While this chapter does explore a rationale behind an independent bearer service market, it does not adequately explore the comparative benefits of vertical integration and disintegration (neither the conditions favoring the one or the other approach). These are important questions which would be a valuable extension to this chapter. To respond to these questions, it appears that one must combine elements from the theory of *contracts* [38, 39] with the *competencies-based* analyze of organizations [40].

Overall, the bearer service appears as an opportunity which might prevent "closure" and collusive price-fixing in future integrated services networks and, more generally, should reinforce Schumpeterian *creative destruction* forces. But, while a market for bearer services may be economically viable, that market is also fragile. New institutions may be required to enable the market to form. For example, to check short-term influences for horizontal and vertical market integration in a digital network with a layered architecture, regulation mechanisms, although not necessarily government intervention, might be appropriate for defining risk sharing agreements between providers of services above and below the bearer layer. May the stock-market create these mechanisms?

## References


[1] D. Clark, Interoperation, Open Interfaces, and Protocol Architecture. Laboratory for Computer Science, Massachusetts Institute of Technology, mimeo, 1995.

[2] Computer Science and Telecommunications Board (CSTB), National Research Council, Realizing the Information Future: The Internet and Beyond. National Academy Press, Washington, D.C., 1994.

[3] Computer Science and Telecommunications Board (CSTB), National Research Council, The Unpredictable Certainty: Information Infrastructure Through 2000. National Academy Press, Washington, D.C., 1996.

[4] Internet Telephony Interoperability Consortium (ITC), Workplan 1997, MIT.

[5] J. Bailey et al., The Economics of Advanced Services in an Open Communications Infrastructure: Transaction Costs, Production Costs, and Network Externalities, *Information Infrastructure and Policy* **4** (1995) 255-278.

[6] J. Gong and P. Srinagesh, The Economics of Layered Networks. In: L. McKnight and J. Bailey (eds.), Internet Economics. MIT Press, Cambridge, MA, 1997, pp. 63-77.

[7] J. Gong and P. Srinagesh, Network Competition and Industry Structure, *Industrial and Corporate Change* **5** (1996) 1231-1241.

[8] M. Piore, Corporate Reform in American Manufacturing and the Challenges to Economic Theory. Department of Economics, Massachusetts Institute of Technology, WP #90-079, 1989.

[9] M. Piore and C. Sabel, The Second Industrial Divide. Basic Books, USA, 1984.

[10] M. Piore, Beyond Individualism. Harvard University Press, Cambridge, MA, 1995.

[11] M. Piore, Fragments of a Cognitive Theory of Technological Change and Organizational Structure. In: N. Nohria and R. Eccles (eds.), Networks and Organizations: Structure, Firm and Action. Harvard Business School Press, Boston, MA, 1995.

[12] C. Kindleberger, Standards as Public, Collective and Private Goods, *Kyklos* **36** (1983) 377-396.

[13] D. Puffert, The Economics of Spatial Network Externalities and the Dynamics of Railway Gauge Standardization. Stanford University, Dissertation Thesis, 1991.

[14] D. Tennenhouse et *al.*, Virtual Infrastructure: Putting Information Infrastructure on the Technology Curve. Laboratory for Computer Science, Massachusetts Institute of Technology, TM-523, 1995.

[15] G. Brock, The Telecommunications Industry: The Dynamics of Market Structure. Harvard University Press, Cambridge, MA, 1981.

[16] J. Walrand, Communication Networks. Irvin, Boston, MA, 1991.

[17] D. Bertsekas and R. Gallager, Data Networks. Prentice-Hall, Englewood Cliffs, N.J, 1992.

[18] G. Dicenet, Le RNIS: Techniques et Atoûts. Masson, Paris, 1988.

[19] M. Elton (ed.), Integrated Broadband Networks: The Public Policy Issues. Elsevier Science Pub., North Holland, 1991.

[20] R. Mansell, The New Telecommunications. Sage Pub., London, 1994.

[21] P. Kavassalis et *al.*, The Internet: A Paradigmatic Rupture in Cumulative Telecom Evolution, *Industrial and Corporate Change* **5** (1996) 1097-1126.

[22] D. Messerschmitt, The Convergence of telecommunications and computing: What are the implications today?, *Proceedings of the IEEE* **8** (1996) 1167-1185.

[23] G. Muller, International Freight Transportation. Eno Transportation Foundation, USA, 1995.

[24] P. Kavassalis and R. Solomon, Mr. Schumpeter on the Telephone: Patterns of Technical Change in the Telecommunications Industry before and after the Internet, *Communications and Strategies* **26** (1997) 371-408.

[25] D. Clark, The Design Philosophy of the DARPA Internet Protocols, *Computer Communication review* **18** (1988) 106-114.

[26] M. Rose, The Simple Book: An Introduction to Management of TCP/IP - based internets. Prentice-Hall, Englewood Cliffs, N.J, 1991.

[27] G. Dosi, Technological paradigms and technological trajectories, *Research Policy* **11** (1982) 147-162.

[28] F. Bar, Network flexibility: a new challenge for telecom policy, *Communications & Strategies* **2** (1991) 113-123

[29] S. Shenker, Service Models and Pricing Policies for an Integrated Services Internet. In: B. Kahin and J. Abbate (eds.), Standards Policy for Information Infrastructure. MIT Press, Cambridge, MA, 1995, pp. 315-337.

[30] R. Braden et *al.*, Resource Reservation Protocol (RSVP) - Version 1 Functional Specification. Internet draft, 1996 (ftp.isi.edu/internet-drafts/draft-ietf-rsvp-spec-14.ps).

[31] R. Braden et *al.*, Integrated Services in the Internet Architecture: an Overview. Internet Engineering Task Force, RFC 1633, 1994.

[32] D. Mitzel et *al.*, A Study of Reservation Dynamics in Integrated Services Packet Networks. Proceedings of Infocom, 1996 (ftp://catarina.usc.edu/pub/mitzel/Infocom96/infocom96.ps.Z).

[33] S. Downs, From Narrowband to Broadband: ATM and the Transformation of the Telephone Network. Massachusetts Institute of Technology, Master's Thesis, 1992.

[34] P. Srinagesh, Internet Cost Structures and Interconnection Agreements. In: L. McKnight and J. Bailey (eds.), Internet Economics. MIT Press, Cambridge, MA, 1997, pp. 121-154.

[35] A. Chandler, The Visible Hand. The Belknap Press of Harvard University Press, Cambridge, MA, 1977.

[36] G. Dosi, Organizational competencies, size and the wealth of nations: some comments from a comparative perspective. Presented to XI International Economic History Congress, Milan, Sept. 1994.

[37] L. McKnight and J. Bailey (eds.), Internet Economics. MIT Press, Cambridge, MA, 1997.

[38] O. Hart, Firms, Contracts and Financial Structure. Oxford University Press, Oxford, 1995.

[39] S. Grossman and O. Hart, The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration, *Journal Of Political Economy* **94** (1986) 691-719.

[40] D. Teece et *al.*, Understanding corporate coherence, *Journal of Economic Behavior and Organization* **23** (1994) 1-30.