Framework for path finding in multi-layer transport networks

Dijkstra, F.

Publication date
2009

Document Version
Final published version

Chapter 2

Optical Exchanges

> This chapter is based on *Optical Exchanges* by F. Dijkstra and C.T.A.M. de Laat [a1] and on *A Terminology for Control Models at Optical and Internet Exchanges* by F. Dijkstra, B. van Oudenaarde, B. Andree, L. Gommans, P. Grosso, J. van der Ham, K. Koymans and C. de Laat [a4]

In the introduction chapter, we have seen that national research and education networks (NRENs) responded to the demand for dedicated network connections by building hybrid networks, which provide both packet switched IP connections as well as circuit switched point-to-point connections.

This chapter dives into the design of these networks. In doing so, it establishes a model and terminology that is later used in chapter 5. Section 2.1 of this chapter introduces the concept of optical and hybrid networks.

The rest of this chapter is focussed on the interconnection points between existing hybrid networks, the network exchanges. By comparing the functions and services between Internet exchanges on one hand and so-called optical exchanges that are deployed in emerging transport networks, the differences and commonalities will become clear. During this comparison, this chapter answers the following side question: **Can optical exchanges, just like Internet exchanges, be completely transparent to a path finding algorithm in circuit switched networks?**.

The layout of this chapter is as follows. Section 2.2 examines the functional differences between exchanges for optical networks, and existing Internet exchanges. Section 2.3 discusses how technological incompatibilities affects functions and services of an optical exchange. Section 2.4 introduces a terminology for the owner and operator of exchanges, which facilitate descriptions of exchanges. Section 2.5 answers the research question and section 2.6 ends with conclusions.

2.1 Network Terminology

2.1.1 Photonic Networks

Originally, the term *optical networks* or *photonic networks* was used to denote fibre based networks, with laser light as the transmission technology. Examples of an optical transport technology are optical carriers (OC) in synchronous optical network (SONET) [s1] and wavelengths in wavelength division multiplexing (WDM). Articles dealing with optical transmission as well as some network design articles such as Saleh *et al.* [p35] use the term *optical* in its original sense, photonics. Others use *optical* to not only describe optical transmission such as optical carriers in synchronous optical network (SONET) [s1], but also use optical to describe the electrical components in otherwise optical networks, such as the SONET switches, resulting in some confusion.

2.1.2 Optical Networks and Transport Networks

Because of the wavelength division multiplexing (WDM) capabilities of optical networks, the term *optical network* was also used to describe networks that could provision dedicated network circuits. The term *lambda* (after the symbol λ for wavelength), which originally meant a single wavelength on a fibre [p5], started to be used for any link segment, and the term *lightpath* was used to refer to end-to-end circuits over these networks. This idea exists at least since 1992 [p7]. Optical BGP as defined by St. Arnaud *et al.* [s34] clearly uses *optical* in a non-photonic sense. Nowadays, the scope of the term has so broadened that it also includes non-optical technologies such as Ethernet. Authors such as DeFanti, De Laat and Smarr use the term *optical networks* to refer to networks providing lightpaths, circuit-switched based services [p37, p23, t16, p10, t15]. This is consistent with the use of word *optical*, *lambda* and *light* in projects like OptIPuter, Global Lambda Integrated Facility (GLIF), NetherLight, StarLight, etc. [p36, u8, u11, u15].

To avoid confusion with photonic networks and actual optical transmission technologies, we will use the term *transport networks* to refer to networks providing circuit-switched based services, regardless of the use of optical or other transmission technologies.

Transport networks can be created with any circuit-switched technology. Example technologies included wavelength division multiplexing (WDM), time division multiplexing (TDM) technologies such as synchronous digital hierarchy (SDH) [s40] and synchronous optical network (SONET) [s1], Optical Transport Network (OTN) [s44, s39], Asynchronous transfer mode (ATM) [s48]. In addition, the connection oriented properties of some packet switched technologies can also be used to create virtual circuits. Examples of the later technologies include MPLS [s15], Ethernet with the use of IEEE 802.1Q tags (VLAN tagged links) [s3] and Ethernet with Provider Backbone Bridge Traffic Engineering (PBB-TE) (the addition of quality of service to VLAN tagged links) [s5].

This thesis will not use the terms 'lightpath' and 'lambda' but use the terms 'dedicated network connection', or simply 'network connection' for end-to-end dedicated network connection and 'link connection' for a segment of a dedicated network connection. This terminology is in line with the one used in the ITU-T G.805 recommendation [s42] (see section 4.3.3 for the precise definitions).

Whereas a lightpath refers to a single network connection, a collection of lightpaths for the same user or organisation is referred to as an *optical private network* (OPN) for that organisation. Just like a virtual private network (VPN) is an overlay network over the regular Internet, an optical private network is an overlay network over a transport network.

2.1.3 Hybrid Networks

National research and education networks (NRENs) such as SURFnet (Netherlands) and Canarie (Canada) have responded to the need for dedicated network connections by building hybrid networks [p25, p13]. *Hybrid networks* in this context are networks that provide both regular Internet as well as point-to-point connections using the same physical infrastructure. Figure 2.1 shows a basic layer stack for such hybrid networks.

Example implementations of hybrid networks include SURFnet6 (Netherlands), CAnet4 (Canada) and National Lambda Rail (USA) [t4, u10, u6].

This use of the term *hybrid network* as a network providing both packet and circuit based services, should not be confused with the same term meaning partial wired and wireless networks, as can be commonly found in the literature.

Hybrid networks utilise the property that most computer networks are *layered*, where services on a lower layer are offered to a client on a higher layer. For example, in figure 2.1 the fibre infrastructure provides services to

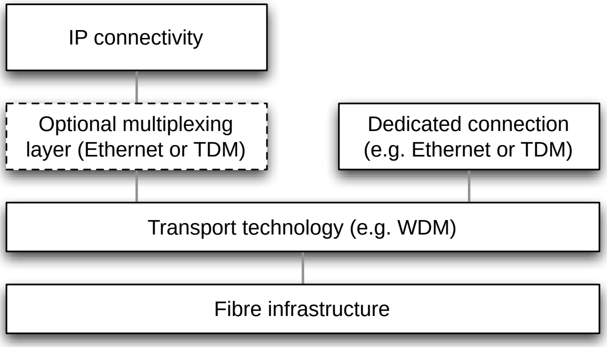


Figure 2.1: *Dual stack for a hybrid network example: the underlying infrastructure provides services at multiple layers.*

the Wavelenght Division Multiplexing (WDM) layer. The OSI model [s53] is a multi-layer model that is commonly referred to in network modelling, although the actual layer stack that it defines has hardly gained any use. The TCP-IP network stack has been more common since the rise of the Internet. Unless otherwise noted, the term *multi-layer network* in this thesis refers to a network that is *configurable at multiple layers*. Hybrid networks are an example of a network which can be reconfigured both at a circuit switched layer and the IP layer.

2.2 Exchanges

The Internet is literally a collection of interconnected networks. Due to economic advantages the interconnections between networks often cluster at specific locations [p6, p18, p28]. This argument holds for all interconnections between networks, not just the regular Internet.

This section examines the commonalities and differences between interconnection points in the Internet and in transport networks.

2.2.1 Peering, Exchanges and Members

Peering, in most literature, is limited to providing connectivity to each other's networks and to the customers of the other network, but not to other destinations. *Transit* on the other hand implies that the traffic is handled for all destinations, usually for a fee [p28]. In this thesis, we do not differentiate between these relations, and simply speak of *peering* to refer to the exchange of data between two networks.

The most trivial interconnection point is a co-location that provides no other functionality than rack space and power. This already gives the networks at the co-location the ability to initiate bilateral peerings with other networks at the same facility.

Our research focus is not these simple co-location facilities, but on *exchanges*, networks dedicated to facilitate the interconnections between different networks. We use the term *member* for the networks connected to an exchange. In our terminology *member* does not imply an organisational structure of the exchange. The term is common in the Internet exchange community. The GLIF community has no particular term, although *autonomous optical domains* is used in some descriptions.

2.2.2 Classification

Based on the functions, rather than the technical implementation, we observe four types of interconnection points:

- Internet exchange points (IXP),
- Mobile roaming exchanges (MRX),
- GLIF Open Lambda Exchange (GOLE), and
- Points of presence (POP).

An **Internet exchange point** (IXP), or simply *Internet exchange* serves as an interconnection point to exchange packet data between individual members. The members have one or a few physical connections to a central core infrastructure. The core network can be Ethernet LAN, SONET, ATM, or MPLS-based. The Ethernet variant is by far the most common variant [t2] and is stateless, while the other three are stateful and require that the individual members set up a path between them. Such a path is a channel in the physical connection. Internet exchanges are known in the USA as *Network Access Points* (NAP).

Mobile roaming exchanges (MRX), such as GPRS roaming exchanges [p4] and UMTS exchanges, exchange packet data for respectively 2.5th and 3rd (3G) generation mobile telephony. In telecommunications, however, the term *exchange* is different from the usage in this thesis and refers to a transit provider rather than an interconnection point. An exchange point between mobile roaming exchanges is technically not different from a packet-based[1] Internet exchange.

GLIF Open Lambda Exchanges are interconnection points where members exchange traffic at OSI layer 1. The switched circuits through the exchanges are part of larger circuits – the lightpaths. Most GLIF Open Lambda Exchanges (GOLEs) use SONET/SDH technology to provide the circuits, although some create pseudo-circuits using Ethernet VLANs. The term *GOLE* is used within the GLIF community since 2005. Examples of GOLEs are: NetherLight in Amsterdam, StarLight in Chicago, MAN LAN in New York, T-LEX in Tokyo, HKOEP in Hong Kong, UKLight in London, NorthernLight and Pacific Wave [u8]. The later two exchanges are geographically distributed (along Scandinavia and the East coast of the USA respectively).

In this thesis, we use the term *optical exchanges* [a1] instead of GOLE since that term will also be recognised outside the GLIF community. Optical refers to optical networks, and does not require that the switching technology is purely photonic. Optical exchanges are also known as grid exchange points (GXP), since they focus on switching circuits on transport networks for use in Grid-based applications. GMPLS Internet exchanges (GMPLS-IX) as defined by Tomic and Jukan [p39] share the concept of circuit-switched interconnection points, but have not been implemented yet.

Points of presence (POP) are interconnection points where access networks connect with a transit network provider. While the transit network provider may connect to multiple networks, we do not regard it as an exchange, since the function of the transit network is not primarily to facilitate data exchange between the members, but between a member and itself.

We will examine Internet exchanges and optical exchanges in more detail, and see where they differ in function and technology.

2.2.3 Internet Exchanges

There are currently three types of Internet exchanges [p18, p27] defined:

LAN-based Internet exchanges: The most common exchange [t2], creating a Local Area Network (LAN) that connects the routers of the member networks with each other. The network is usually a star topology with a layer 2 Ethernet switch at the core, though earlier variants had a ring topology (based on Fiber Distributed Data Interface). This is the only stateless exchange. Congestion is possible if multiple members want to send traffic to the same member at the same time.

[1] GPRS and UMTS are packet based. The older CSD system is circuit switched.

ATM-based Internet exchanges: A stateful exchange, with an asynchronous transfer mode (ATM) network. The member networks are connected with each other with permanent virtual circuits (PVCs) through the ATM network. PVCs in an ATM network can either be variable bit rate or constant bit rate [s49]. If variable bit rate (VBR) circuits are used, there is no guaranteed congestion-free transmission. Usage of constant bit rate (CBR) on the other hand, results in very inefficient usage of resources and poor scalability.

MPLS-based Internet exchanges: Stateful packet based exchange, based on Multiprotocol Label Switching (MPLS) [s15]. The destination of incoming packets is determined by a lookup at the IP routing table at the ingress edge label switching router (LSR). This is a relative expensive operation for very-high bandwidth data streams.

Given their properties, these types of exchanges are not designed to support few high bandwidth flows that do not require routing. Either it is technically impossible (for LAN-based Internet exchanges) or it yields unnecessary and costly overhead (for ATM- and MPLS-based Internet exchanges).

2.2.4 Internet versus Optical Exchanges

Table 2.1 highlights the differences between Internet exchanges and optical exchanges.

Members at an Internet exchange interconnect to exchange IP traffic. This is shown at the first two rows in the Internet exchange column in table 2.1. An Internet exchange contains exactly one circuit per peering relation between two members. In contrast, a transport network supports circuits between end-users, so at an optical exchange there is a circuit between members for each end-to-end connection that goes through the exchange. The table further emphasises that for an optical exchange, these circuits can carry any layer 1 or layer 2 traffic. Differences in function and purpose lead to different choices in technology between Internet exchanges and optical exchanges. Finally, the table highlights that an optical exchange may offer more advanced services than an Internet exchange.

There is no clear boundary between the different interconnection points since each interconnection point may take multiple roles. We expect that the differences listed in table 2.1 will change over time, as new technologies become available and are implemented. For example, customers at a POP may also directly peer with each other, a function typically seen at exchanges. Circuit

	Internet Exchange	Optical Exchange
OSI Layer	Transports traffic at layer 2, members connect with layer 3 devices	Transports traffic at layer 1 or layer 2, members connect at that same layer.
Traffic type	IP traffic only	Any packet data or any data at a specific framing or bit rate
End-points	Connection between two member networks	Connections are part of a larger circuit between two end-hosts
Dynamics	Stateless, or state changes only when (peering) relation between members changes	State changes for each data transport
Technology	Often packet switched, sometimes label-switched (with virtual circuits like MPLS and ATM)	Circuit or virtual-circuit switched (e.g. using SONET or VLANs)
Services	Only data transport	Data transport and other services, like the conversion of data between different formats and layers

Table 2.1: *Functional differences between Internet exchanges and current optical exchanges. These characteristics will change over time, as new technologies become available and are implemented.*

switching is typically associated with optical exchanges, but not a technical necessity: ATM- and MPLS-based Internet exchanges are also circuit switched and it might be possible to create a non-circuit switched optical exchange using optical burst switching (OBS) [p32].

In fact, we have reason to believe that the distinction between Internet exchanges and optical exchanges will disappear as time progresses. First of all, it is economically beneficial to build optical exchanges at a location where a large number of member networks are already present. The location where most networks are present is at Internet exchanges, since networks that connect to optical exchanges are often hybrid networks that are also connected to Internet exchanges. Secondly, Internet exchanges also tend to offer more services that are now regarded as optical exchange functions, like private circuits between two members. For example, the Amsterdam Internet Exchange (AMS-IX) already provides private interconnects and closed user groups [t1].

Third, optical exchanges must often provide all sorts of services, as we will see in section 2.3.3. Multiparty peering like in a LAN-based Internet exchange, is just another service.

2.3 Incompatibilities

2.3.1 Progressing Technology

The interfaces in a network may be of different type. For example, one interface may be used to carry undefined traffic over 32 wavelengths using dense wavelength division multiplexing (DWDM) [s36], while an other interface may only carry one signal at 1310 nanometre, which carries SDH-framed traffic [s40]. A third interface may be LAN-PHY based Ethernet [s6], where traffic must reside in the 192.0.2.0/24 subnet.

Most of these types map to different network layers, and some devices may deal with layer 1 functions, while another device may deal with layer 2 functions. A device that deals with higher layer functions will nevertheless make use of the lower layer functions to transport data on its interfaces. In the above example the SDH framed traffic was carried over a 1310 nanometre wavelength.

This variety of encodings at different network layers would not be a problem if everyone would use the same encoding for each layer. Unfortunately, there are many, possible incompatible ways to transport data. There are multiple ways to carry signals in one fibre using DWDM, due to the variety of wavelength bands and channel spacings. Similarly not all devices encapsulate Ethernet in the same way over a SONET/SDH connection.

Networks can only communicate with each other if they agree on a specific encoding of their data. A common interface may be agreed upon if one technology has been proven to be robust and sufficient for all network operators. If a new network is built, the operator will choose that particular technology. For example, the de-facto standard on the Internet is IP, and alternatives such as the OSI protocols [s53] are no longer used.

If technologies are still in development and new possibilities are introduced on the market as time progresses, network operators may choose to use the new technology instead, which may be incompatible with the technologies from networks that have been built while that technology was not available. This has a severe impact on the design of networks, as networks using two different technologies can not communicate by default, and the data must be translated to an agreed upon encoding.

Technology advances are rather common. For example, during the time this thesis was written, multiple technologies have been used to create dynamic circuits: MPLS [s15], GMPLS [s20], Ethernet VLANs[s3], SONET/SDH [s1, s40], and Ethernet PBB-TE [s5]. Also technologies that were considered stable for decades still change. For IP there is a distinction between IPv4 and IPv6 [s12], and recently the size of AS-identifiers has changed from 2 bytes to 4 bytes [s29], causing potential incompatibilities.

Summarising, we can conclude that **the use of multiple technologies causes incompatibilities**. In addition **technologies evolve over time**, thus **incompatibilities will continue to exist**.

2.3.2 Impact on Optical Exchanges

While technologies evolve, this gives rise to possible incompatibilities, and the design of an exchange should take that into account. Either the exchange should decide on a common technology, or only members with the same technology can communicate, or the exchange must contain a service to enable the interworking between different technologies.

Network engineers avoid technology incompatibilities as much as possible. Indeed, it is very uncommon to see incompatibilities within a single domain, but it is very common to see them across domains. Networks are not built at the same time, and network owners may have different opinions about the best technology. This means that **exchanges must take potential incompatibilities into account to avoid non-working network connections**.

If an exchange must perform one of the services mentioned above, it must also be aware about the possible incompatibilities between the different interfaces, as explained in section 2.3. At a minimum, it must know about the layer and framing of the interface. For example, it must know the difference between 1 Gbit/s and 10 Gbit/s Ethernet links, and must know how to extract the Ethernet frames from the carrier. A more advanced incompatibility is the different ways in which Ethernet can be encapsulated in SDH and SONET channels.

Exchanges try to avoid incompatibilities by standardising the interfaces to the members. For GLIF Open Lambda Exchanges (GOLEs), the optical exchanges in the GLIF community, exchanges either switch individual VC-4 timeslots (for SDH), or use Ethernet VLANs to switch. The embedding of Ethernet channels in VC-4 timeslots is standardised using Framed Generic Framing Procedure (GFP-F) [s45], virtual concatenation (VCAT) [s47] and Link Capacity Adjustment Schema (LCAS) [s46]. The effect of all these standards is that egress timeslots can be chosen independently of the ingress timeslots, and thus are less possible incompatibilities. Nevertheless, some exchanges standardise on Ethernet, others on SDH, and others on GMPLS. Thus technology conversions are still required for end-to-end connections.

2.3.3 Services

Since it is undesirable that only members of an exchange with the same technology can communicate, and because technology conversions are still required for end-to-end connections, exchanges may offer interworking services. The following list gives a partial overview of the different services that may exist for optical exchanges.

Cross connect: The basic functionality of any exchange is to transport traffic from one member to another member. For interfaces of equal type, this can be accomplished using a simple cross connect between the two interfaces.

Adaptation: If a signal goes in at one layer and goes out at a different layer, the optical exchange effectively acts as an ‘elevator’, lowering or raising traffic to different layers. For example if a signal comes in as a wavelength using DWDM, and is injected in a VLAN, the signal is elevated from layer 1 to layer 2.

Multiplexing: Combining different signals into a single carrier is called multiplexing. The extraction is called de-multiplexing. For example, combining multiple wavelengths into a single fibre.

Interworking: Conversion between different incompatible technologies, for example, conversion between WAN PHY Ethernet and LAN PHY Ethernet or between a wavelength with 50 GHz and 100 GHz spacing.

Wavelength conversion: A special case of Interworking. Changing the colour (wavelength) of a lambda

Signal regeneration: Either simple amplification or attenuation of the power levels to match a certain output power level, or full regeneration consisting of reamplification, reshaping and retiming (3R) of the signal.

Optical multicast: The ability to duplicate an optical signal as-is. This can only be done one-way, not for bidirectional connections (at least not without merging the return signals, or ignoring all but one).

Traffic aggregation: Traffic may be aggregated as packets into a single data stream. This is a special case of multiplexing. Aggregation with (IP or Ethernet) packets, (ATM) cells or optical burst switching may lead to congestion and packet loss.

Store-and-forward: One way to reduce blocking chances is to transport large chunks of data on a hop-by-hop basis, rather than reserving the entire end-to-end path. The data may first be transported along a first segment of the end-to-end path, stored at an intermediate location, and transported along the rest of the path as soon as that section of the path becomes available for the data transport.

2.3.4 Control Plane Services

So far we only discussed data plane services.

The *data plane* of a network is where all data transport functions take place. The control plane of a network is where the operational control takes plane, and externally influenced state changes of the network are processed. The control plane includes the routing and signalling protocols that react to state changes in the network (e.g. links coming up or down, external connectivity announcements), and to connectivity requests by users or applications (e.g. path requests and tear down messages).

The control plane may also contain services for users or applications. These control plane services may include:

Provisioning service: set up or tear down of a network connection based on a user's request

Scheduling service: checking the availability of certain resources and making future reservations;

Authentication service: verifying the usage policy on a resource and authorising their usage;

Index server: listing the available resources;

Broker service: combining different resources together necessary to fulfil an entire request.

The resources in the list are typically network resources, but could be other types of resources, like storage or computational resources.

The services offered by the control plane are sometimes also referred to as the *service plane*.

2.4 Ownership

In this section we will introduce a concise terminology to describe exchanges, in particular the facilitating role of exchanges where the members rather than the exchange decide on the business policy. We need this terminology later in chapter 5.

One feature often quoted by exchanges is that they take a *neutral* role, and do not prefer one member over another. While neutrality is often seen as a commercial aspect, there are also technical aspects at stake. We need to define different roles to distinguish between the different aspects.

We relied as much as possible on existing terminology. In particular, the ownership terminology in section 2.4.1 builds upon the management layers in Telecommunication Management Network (TMN) [s50] and current practice in economic and legal communities [t10], while the term open exchange draws upon discussion in the GLIF community [t8].

2.4.1 Owner, Operator and Users

Network element is a generic term to include network devices, links, interfaces and hosts.

We distinguish between *legal owner*, *economic owner*, *operator* and *user(s)* for each network element. The *legal owner* of a network element is the entity that purchased the device and the *economic owner* is the entity that acquired the usage rights from the legal owner.

The *economic owner* determines its *policy* for the network, a set of rules to control access to the network resources [s16]. This entity carries the responsibility for the behaviour of a device and has the final responsibility in case of hazards and abuse. In addition, each network element can also have a separate *operator*, the person, organisation, or software component that configures and operates the device on behalf of the economic owner. The economic owner determines the policy for a network element; the operator enforces this policy. Finally, the *users* may use or invoke an element, if their request is in compliance with the active policy.

In this thesis, we use the terms *network administrator* and *network operator* interchangeably. This is in line with the Telecommunication Management Network (TMN) standard, which notes that "the expression 'Administration' is used for conciseness to indicate both a telecommunication administration and a recognized operating agency". Network operators in the GLIF community use the term *network administrator* to mean *network economic owner*. To avoid confusion, we prefer the term *network operator*.

We assume that each network element has exactly one legal owner, one economic owner, and one operator, but may have multiple users over time (though typically only one at a specific time).

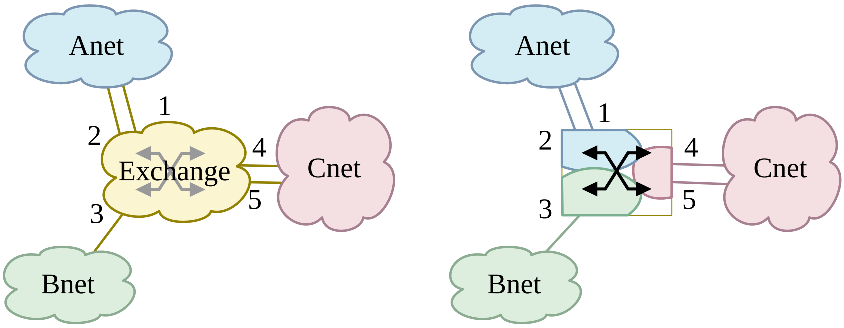
2.4.2 Open Control

Often the legal owner, economic owner, and operator of a network element are the same entity. For example, in the Internet, a transit provider is typically owner and operator of its network. But this is not always the case.

If an organisation leases a trans-oceanic fibre from a carrier to lease a fibre for a year, the carrier is the legal owner, while the other organisation is the economic owner.

If an organisation outsources the maintenance of its network, the economic owner and operator of this network are different entities.

An exchange may be neutral. Rather than applying its own policy, it outsources the policy decision to allow or deny a connection request between two interfaces to the domains that are connected to those interfaces. Therefore, members of an open exchange have the ability to configure 'their' interface in the exchange and thus can decide who connects to their network. We call this concept *open control*, and exchanges that apply such a policy are called *open exchanges*.



(a) *Closed control model* **(b)** *Open control model*

Figure 2.2: *Comparison between the closed control model and the open control in the same network.*

Figure 2.2 shows an example of a closed and an open control model. While control model on the left follows the physical topology of the operational domains, the control model on the right follows the *owner domains*. In the open control model, there is no owner domain for the exchange, because not the exchange, but its members decide on policies.

2.4.3 Domains

We define a *domain* as a set of network elements. A set of network elements may even be disjoint. An *owner domain* is a set of network elements with the same economic owner. An *operational domain* is a set of network elements with the same operator. Observe that this definition does not require that an operational domain contains *all* network elements with the same operator.

An *exchange network* is an operational domain within an interconnection point whose primary goal is to transport data between all its members. Exchanges tend to be neutral, and are policy-free. Exchange networks are of special interest throughout this chapter and we use the term *exchange* to refer to a core network and its operator.

2.5 Transparency

So far, we investigated the properties of Internet and optical exchanges in detail. It is time to answer our research question: **Can optical exchanges, just like Internet exchanges, be completely transparent to a path finding algorithm in circuit switched networks?**

The routing protocol on the Internet is the Border Gateway Protocol (BGP) [s24]. It works by distributing reachability paths, consisting of lists of individual networks, the autonomous systems (AS). These lists of AS numbers do only contain autonomous systems with routers. It does not included Internet exchanges, since Internet exchanges typically don't contain a router. Thus, Internet exchanges are transparent for the path finding algorithm in the Internet.

It is desirable that exchanges are neutral. The members of an exchange expect that an exchange has a facilitating role, and does not have its own policy.

An exchange is *transparent* to a path finding algorithm if the path finding algorithm needs no knowledge of the exchange. Thus, if the path finding algorithm regards two members as directly connected. This is true if two conditions are met. First of all, the members must be able to connect at will. Thus, the exchange must have a facilitating role, and not have a policy on its own. Secondly, the connections through the exchange must behave as a direct connection. So, the interfaces must be of the same type and the exchange does not offer any services. **Exchanges can only be ignored during path finding if the connections through an exchange are modelled as direct connections, and the exchange does not define a usage policy on its own**.

Optical exchange can, like Internet exchanges, be transparent. The condition is that the optical exchange does not provide a policy of its own. If the path finding algorithm finds a path between these two domains, then the domains are responsible to signal the exchange to provision a connection between the two members, just as the members are responsible for setting up the connection through their own network. If the optical exchange would have its own policy, then the members can not be sure that the exchange adheres to their request, and the exchange must be visible to the path finding algorithm. Given that exchanges have a facilitating role, this condition is typically true.

The second condition is that a path finding algorithm can treat the connection through the exchange as a direct connection between the members. This condition may not be true: if the two interfaces are of different type, or if the exchange offers services besides simple cross connects, then the exchange can not be regarded as a direct connection, and it is necessary that the path finding algorithm is aware of the exchange.

2.6 Conclusion

In this chapter, we investigated the fundamental differences between the Internet and transport networks. We created a definition framework that makes comparisons between exchanges possible, not only between Internet and optical exchanges, but also between different optical exchanges. Our terminology on circuits and lightpaths, the service overview in section 2.3.3 and the classification in section 2.1 helped the discussion in the community. The distinction between hybrid network and multi-layer network was not obvious before these definitions. Furthermore, We were the first to define the concept of 'open control' and relate that the difference between an owner and an operator of a network.

The most obvious distinction between the Internet and transport networks is that the Internet is packet based, while transport networks are circuit based. This difference also has implications for the exchanges in these networks. While Internet exchanges and optical exchanges may use the same technologies (e.g. Ethernet and SONET), an important distinction is that optical exchanges change state with each connection request, while the state of Internet exchanges only change when peering relations change.

Since technologies change over time, and so do networks, incompatibilities will continue to occur. Exchanges must take potential incompatibilities into account to avoid non-working network connections. Services like adaptation and interworking must be offered by exchanges or elsewhere in the network to solve incompatibilities. This last implication means that optical exchanges may offer services, such as interworking between different interfaces, and multiplexing of channels. If such services are offered, they are relevant to a path finding algorithm, and optical exchanges can not be transparent to a path finding algorithm, as Internet exchanges are.

It is possible for an optical exchange to retain its facilitating role, even when it is considered to be a regular domain by a path finding algorithm. In this case, the policy of the domain would *not* be to always grant access, but instead, to delegate the decision making to the affected neighbours.