# How to Buy a Network: Trading of Resources in the Physical Layer

*Vassilis Prevelakis, Drexel University*
*Admela Jukan, INRS-University of Quebec*

## ABSTRACT

Recently, a number of new research initiatives, most notably UCLPv2 and GENI, have promoted the dynamic partition of physical network resources (infrastructure) as the means to operate the network, and to implement new protocols and services. This has led to a number of open issues such as resource discovery, implementation of resource partitioning, and the aggregation of resources to create arbitrary network topologies. To us, the key issue is the design of a mechanism to trade, acquire, and control network resources, given a choice of providers of physical resources (infrastructure providers). In this article we present an architecture that allows physical resources to be traded, while granting users controlled access to the acquired resources via a policy enforcement mechanism. In addition, it allows resource provider domains to be linked via configurable, provider-neutral resource exchange points that are the physical resource equivalents of the *pooling point*, or Internet Exchange Point (IXP). We demonstrate how our trading system will operate by presenting a use case in which a network topology is constructed using resources from multiple providers, be it Internet Service Providers (ISPs), or National Research Experimental Network (NREN) providers. The use case also shows how a dynamic reconfiguration can be effected by the customer though the use of simple access control policies, without involving the provider.

## INTRODUCTION

The demand for applications using guaranteed bandwidth such as video-on-demand, multiplayer gaming, and grid computing is reviving interest in advance and on-the-fly resource reservation schemes across large geographically distributed networks. Earlier attempts did not catch on for a number of reasons, notably lack of business incentives on the part of the service providers. Most notably, providers lack efficient accounting and charging architectures to be assured that a bandwidth service (transaction) is actually *paid* by the customer, be it a research organization or a home user. The dynamic resource management required to facilitate efficient service trading in a large-scale environment presents a great challenge and will be the critical issue before it is deployed into commercial operation.

Recently, however, the user's interest in dynamic bandwidth provision has evolved into a more generic need for dynamic provision of physical network resources, that is, raw infrastructure, including links, switches, and IP router equipment. Two observable trends have emerged. First, large international corporations as well as municipal offices are acquiring or leasing fiber and switches to build their own networks, similar to the L1VPN concept [1]. By acquiring their own physical infrastructure, the customers become empowered to create their specialized network infrastructures, with customized service provision decoupled from the physical topology. Second, various government supported research initiatives, such as UCLPv2, are based on the premise that for networking research to advance it is important that researchers be able to create multiple parallel topologies over the same physical infrastructure and experiment with packet delivery systems beyond the current Internet [2, 3]. To reach the vision of the next-generation Internet, the first step is to provide mechanisms to dynamically partition the physical network infrastructure into parallel networks, each running their own protocols and services.

With increasing numbers and types of resources included in the process of dynamic partition and acquisition, questions of resource discovery and trading as well as user authorization for resource usage has become a challenge. In this article we address this challenge and propose a resource trading architecture applicable in both connection-oriented and connection-less networks. In our approach, users are allowed to purchase resources using an open market where providers, such as ISP or NREN, advertise resources and users bid for the resources. To this aim, we introduce the configurable, carrier-neutral resource exchange points (RXPs), which

0163-6804/06/$20.00 © 2006 IEEE

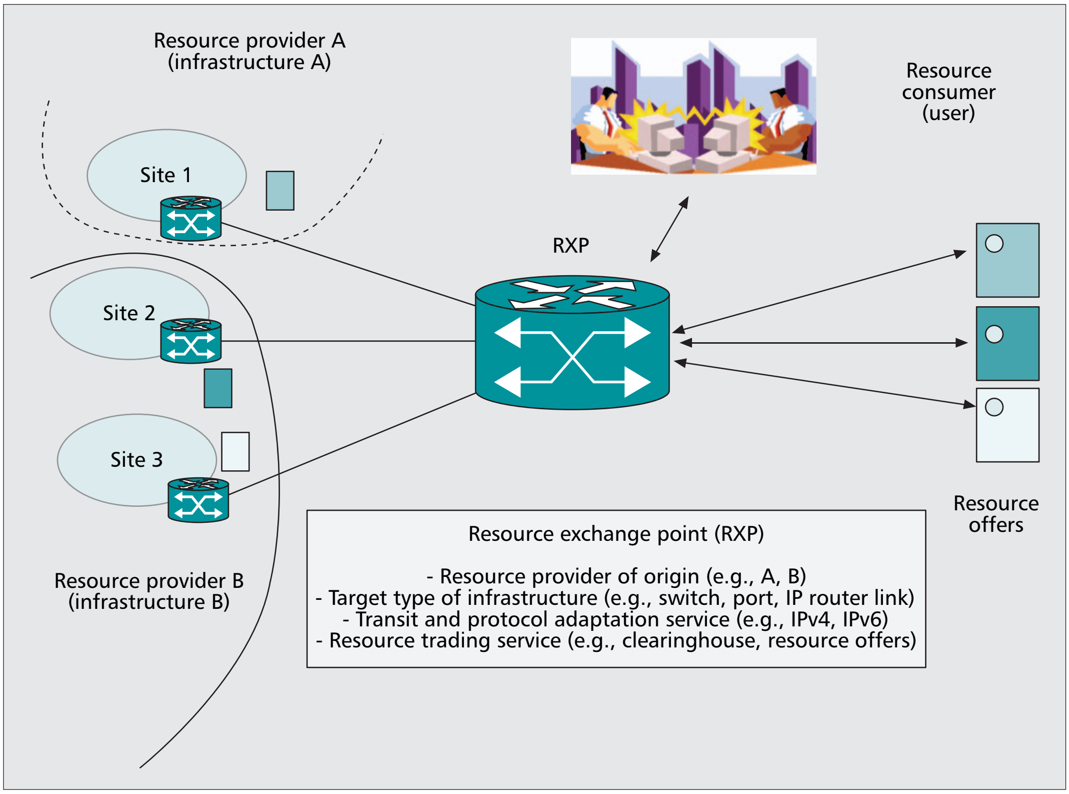

Resource provider A
(infrastructure A)
Resource consumer (user)
Site 1
RXP
Site 2
Site 3
Resource offers
Resource exchange point (RXP)
- Resource provider of origin (e.g., A, B)
- Target type of infrastructure (e.g., switch, port, IP router link)
- Transit and protocol adaptation service (e.g., IPv4, IPv6)
- Resource trading service (e.g., clearinghouse, resource offers)
Resource provider B
(infrastructure B)

**Figure 1.** *At the resource exchange point (RXP) infrastructure providers offer resources and customers (users) bid for the resources.*

*The environment in which our system is expected to operate is extremely varied. For example, some resource providers may offer dark fiber, while others optical switches, or IP routers, or more generally, a whole network domain running a certain protocol stack, such as Ipv6.*

facilitate the trading of physical resources between infrastructure providers and users. This facility allows purchasing resources in advance (effectively creating a "futures" market for resources) as well as on the "spot" market. Users can select from a range of offerings by various resource providers to create network topologies or simply end-to-end pipes piece-meal, or can choose to purchase a complete package from a single provider (or consortium of providers), where available.

The article is organized as follows. We first describe the architecture with the RXPs, which includes a sample usage scenario and issues of usage policies and access to the physical resources. We address the resource trading, including advance and on-the-fly reservation, as well as a brief overview of the work related to exchange points. We describe various components of our resource trading prototype system, along with an analysis of credentials processing. Finally, we conclude the article.

## ARCHITECTURE

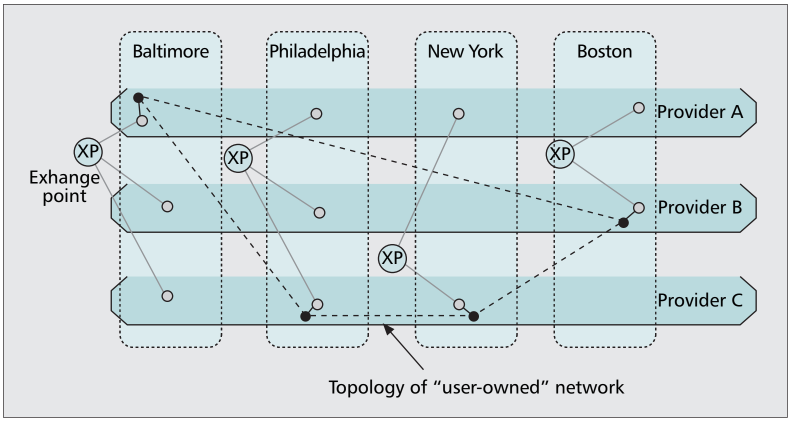
Traditionally the role of network providers was to move data through their network. The management of the network elements (NEs) was the responsibility of the provider, while users who are purchasing services would have a rather passive role akin to passengers in a train. Recent trends in high-speed networking have been leading the way away from this model; efforts in experimental networking, such as UCLP, have shown that it is not only possible but desirable to allow users direct access to the network infrastructure (optical paths, parts of switches and routers, etc.).

The operation of such a system poses a number of questions, such as:

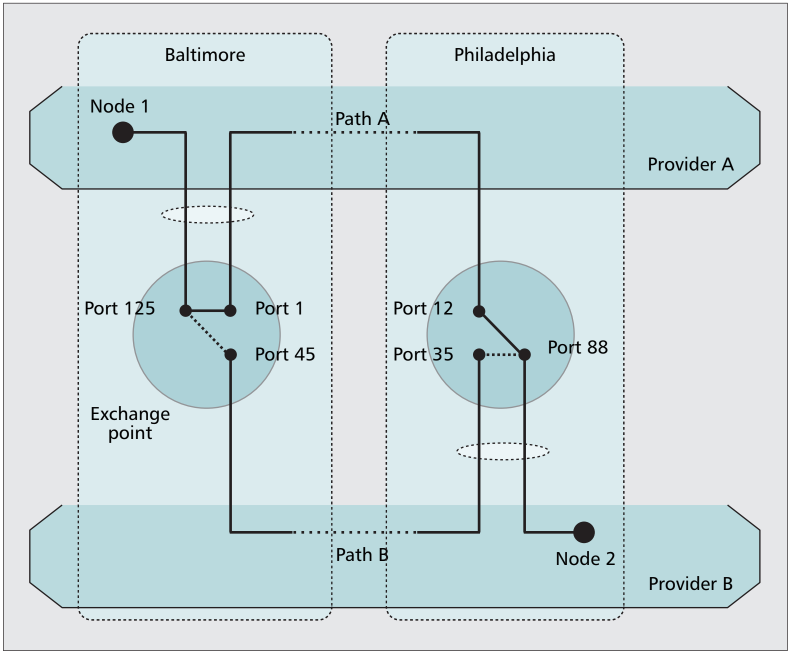
- How can a physical infrastructure be traded?
- How do we allow users access to the configuration of the NEs while ensuring that they are only allowed to perform approved actions?
- How can such a framework be deployed in an existing network (i.e., how do we handle integration with the existing NEs)?

Our architecture addresses these questions. We assume a number of network infrastructure providers, offering different kind of resources (Fig. 1). The environment in which our system is expected to operate is extremely varied. For example, some resource providers may offer dark fiber, while others optical switches, or IP routers, or more generally, a whole network domain running a certain protocol stack, such as Ipv6. There may be large number of nodes, providers, or consumers and arbitrary heterogeneity of resources offered. In addition, and possibly separately, specific protocols as well as services may be offered, such as multiplayer gaming, video, Ipsec, and VPN. Similarly, resource consumers (users, or applications) will have different and varied requirements (e.g., QoS, cost, security, reliability). For the purposes of this article we do not distinguish between "complex" and "simple" users, that is, we do not distinguish between the home user requesting a bandwidth pipe for video-on-demand service, and a campus network operator requesting a set of resources to run a "user-owned" network akin to UCLP. We will in general assume that users will want to minimize expenses, and maximize performance, within specific time-frames. As for the resource providers, we assume that their primary interest is to maximize resource yields. In that way, resource providers can benefit from efficient resource utilization.

In this setting, which conceptually may not be new, we now introduce the novel concept of *resource trading* and argue that this functionality can be best implemented at provider-neutral RXPs. At the exchange point, infrastructure providers will offer their physical resources and users will bid for the resources (Fig. 1). Users eventually select a network topology, including bandwidth, switching technology, and control mechanism, possibly purchasing different pieces of infrastructure from different providers. Various QoS criteria may be considered for trading. The configuration may involve "protocol adaptation features" such as combination of the distributed and centralized routing control. The flexibility and potential for low-cost configuration is attained through the transparency of the entire mechanism and the ease of collecting all the relevant information about resources. Further optimizations would be possible, such as the case where resource providers allowed for full visibility of resources, so that users could release or sublet unused resources. In fact, the implementation of a RXP is not only concerned with the network control and protocols in support of resource trading, but it may also include a multiservice switching node that provides additional functions, such as topology discovery, routing service, and dynamic service-level agreements [4]. This is an equivalent of the concepts of the Internet exchange point or pooling point [5], but applied on physical network infrastructure.



■ **Figure 2.** *Assuming four cities (each with its own RXP) and three providers, there are many ways we can we construct a new network infrastructure to link POPs in all four cities.*



■ **Figure 3.** *By providing the end user controlled access to the switches within the exchange points, we allow fast and efficient reconfiguration of user controlled resources.*

## Use Case

To illustrate the way users can utilize exchange points to create privately managed networks, let us consider the following example of use. A regional broadcaster (RB) would like to link all its offices in the North East region of the United States via an optical network. In Fig. 2 we see the layout of the RB optical network (dotted black lines) consisting of a number of intercity optical links purchased from various network providers linking RB's points of presence (POPs, indicated by solid black circles) in each city.

Each POP reaches the local RXP via a link (solid black line) provided by a local network provider. The POP contacts the RXP via User Network Interface (UNI) signaling [6]. Our model assumes the physical links have been purchased and installed ahead of time. The dotted lines shown in Fig. 2 indicate the desired topology of the purchased network infrastructure (i.e., the connectivity graph that the RB wishes to create), which may be implemented by connecting the various RXPs via links from the three provider networks (A, B, C). For example, starting from the Baltimore POP, we can reach the Philadelphia POP using any provider, but the Philadelphia POP can reach the New York POP only via providers A and C, since provider B does not link with the New York XP. Providers A, B, and C established their business relationship over RXPs by implementing, for example, External Network-Network Interface (E-NNI) signaling at their interconnections.

Searching the offers made by the different providers, we end up with a set of intercity links described as <*provider*, *source*, *destination*, *cost*> and a set of exchange ports described as <*XP*, *port*, *cost*>. With this data on hand, we can run our analysis to create a list of the desired resources we wish to use for the implementation of the new network. For the purposes of this discussion let us assume that provider A gives the best offers, so we implement the optical network using intercity links only from provider A. However, to provide a backup path, the RB can reserve an additional intercity link between Baltimore and Philadelphia from provider B and the corresponding ports on the two RXPs.

The layout is shown in Fig. 3. As we can see, the connections within the two RXPs are arranged so as to use Path A (from provider A). If there is a problem with that link, the RB can switch to the back-up link via provider B by simply reconfiguring the port connections in the two RXPs.

### On Usage Policy

The key characteristic of the presented use case is that, since both the links and the resources have been allocated to us, we have administrative control over these resources and hence can implement the change without contacting the network providers or the managers of the RXPs; instead, we simply send the appropriate reconfiguration commands directly to the resources (e.g., via UNI connecting to RXP). Some systems provide the same functionality via the network management system of the provider, but we believe that by handing over management of the leased parts of the infrastructure to the customer, we reduce the overhead (both in terms of costs and delays). This implies that the customer is able to use control-plane signaling not only to establish connections, but also to send configuration commands to the NEs directly, which in turn poses the problem of authentication (who is sending the command) and access control (whether that entity is authorized to issue that particular command).

We use the term *policy* to describe the actions that can be taken by various actors (or principals) and require that any request made be consistent with the policy that is currently active for that particular equipment (referred to as network element, or NE). Policy can be defined as the commands that may be entered by a given principal (i.e., by associating ASCII strings that correspond to CLI commands with a particular principal). However, since the syntax of the commands may depend on the particular version of software running on the element (e.g., IOS for Cisco NEs), we need to express our policy in a more abstract manner, so that it is independent of platform-specific peculiarities. Policies are evaluated at the point of policy enforcement, which in our case is the resource or the NE we are trying to control and, eventually, reconfigure.

Many existing NEs include Unix- or Linux-based control processors, and these can be adapted to include the policy-enforcement framework within the NE itself. However, many NEs cannot be adapted to our system so easily, due to either memory limitations or proprietary software. In such cases, the policy enforcement can run on a single board control computer that can be located next to the NE and communicate with it via either a dedicated network port or via a serial port. If a private network link can be set up, the control computer can use control plane signaling commands (e.g., UNI), SNMP or TL1 commands to configure the NE, or otherwise command line interpreter (CLI) [6, 7]. Such control computers can be small in size (a couple of square inches) and low cost (less than $400).

### Access Control at the NE Level

Our model allows the role of the network provider to be redefined, where a network provider can act as either infrastructure or service provider or both, shifting more power and flexibility to the end user. The key property of our approach is that the *policy credentials contain the usage policy*, that is, what can be done by the entity that owns the key to the credential (and is, thus, able to sign the requests authorized by the credential). We apply this mechanism to the control and management of any NE, thus allowing the owner of the network infrastructure (network provider) to create authorizing credentials granting the customer limited access to some NE. The customer may then send a signed configuration request to that NE along with the network provider credential *directly* to the NE. The NE can determine whether the configuration request is acceptable by comparing the request against the policy contained in the credential. The credential, in effect, completes a chain of trust, linking the customer to the provider's piece of infrastructure (NE) via the network provider (see the example shown in Fig. 4).

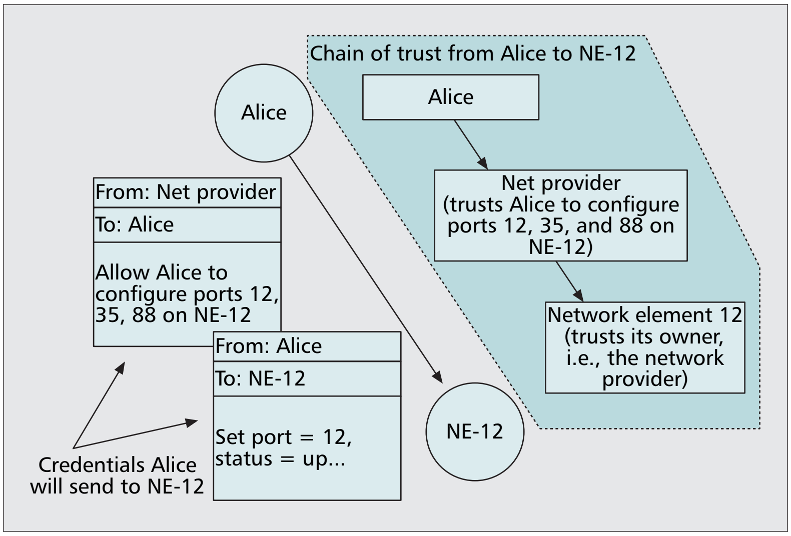

Chain of trust from Alice to NE-12
Alice
Alice
Net provider (trusts Alice to configure ports 12, 35, and 88 on NE-12)
Network element 12 (trusts its owner, i.e., the network provider)
From: Net provider
To: Alice
Allow Alice to configure ports 12, 35, 88 on NE-12
From: Alice
To: NE-12
Set port = 12, status = up...
Credentials Alice will send to NE-12
NE-12

**Figure 4.** *Credentials (left) form a chain of trust (right) from Alice to the network element via the network infrastructure provider. This allows Alice to submit configuration commands directly to the NE, provided they comply with the policy defined by the network provider.*

The benefit of this approach is that configuration requests can be checked locally without having to consult a global (centralized) database, or having to go via the provider's Network Management System (NMS). This approach is far more scalable than the centralized approach, while granting a lot of flexibility to the customer, especially in cases where changes in the configuration need to be made at short notice (e.g., to compensate for unforeseen circumstances or some network disruption). As such, it can be easily implemented in the exiting control-plane frameworks, such as GMPLS, and the RSVP-TE protocol within.

A key characteristic of our architecture is that customers acquire limited control of providers' infrastructure. They are given the ability to instruct NEs to perform specific tasks, such as "connect two ports of an switch forming a path through that switch." The customer is allowed to carry out any desired command, provided it is allowed by the acquired capability. In the example above, we discussed how an optical switch could be reconfigured to divert traffic from a primary link to a secondary without the intervention of the owner of the equipment. This is made possible by the fact that the offer credentials contain the actual policy directives that will be used to vet the instructions provided by the customer. This is similar to what has been shown for IPSEC VPN where policy credentials can be used to convey VPN configuration information to VPN gateways construction [8].

Under our system, network providers post their available resources on notice-boards hosted by various participating sites. The system can accommodate one or more such sites, since they merely announce the offers. Apart from that, the notice-board is not involved in the actual resource reservation.

## RESOURCE TRADING

Under our system, network providers post their available resources on notice-boards hosted by various participating sites. The system can accommodate one or more such sites, since they merely announce the offers. Apart from that, the notice-board is not involved in the actual resource reservation. The postings are in the form of credentials that describe the identity of the resource provider and promise to abide by a set of service specifications which may be processor resources, storage, a path between two points in the provider's network, and so on. The credential may also contain the time period that the offer is valid (which may be different from the expiration of the credential), the price of the concession, and additional information related to the resource provider. For example, in the case of network segments, this can be the path that should be taken between two points. While offer credentials are signed by the resource provider who issues them, the user can have a contract (legal) relationship to either the RXPs, or infrastructure providers directly (e.g., ISP). Users contact one or more notice boards to collect offers. For complex tasks, an end-user may need to combine multiple offers.

In an environment with a single notice board, the end-user can issue queries to get lists of offers matching user's requirements. If there are many notice boards, the end-user may employ various strategies such as dispatching an intelligent agent to collect the offers and come back with a recommendation that meets preassigned constraints (price, reliability of the provider, etc.), query each notice board independently, or use a meta-search engine [9]. At the end of the search, the end-user will hold one or more offer credentials that describe the desired resources and related specifications. At this point, the end-user has not actually purchased the resources. In order to issue payment and reserve the resources, a number of steps have to be taken. The end-user contacts a host offering a resource (we call such a host a *resource holder* or RH) and activates the reservation protocol, for example, via specific extensions of UNI. The RH issues a challenge, which is then returned signed by the end-user. This response also contains the offer credentials collected by the end-user and a credit-worthiness credential issued by the end-user's credit institution. This exchange accomplishes the following:

- Identifies the end-user (the key that has signed the RH challenge)
- Provides proof of good standing
- Limits payment only to the offer credentials provided
- Can be used only for that particular transaction, since it depends on the challenge issued by the RH

Our system allows individual resources to be reserved separately, regardless of whether they exist within a single or multiple provider domains. In general, however, there is no need for the provider's offers to match exactly the requirements of the end-user. For example, if user Alice requires a GigE link from Atlanta to Dublin, she may use an offer for a OC-48 connection, but purchase only 1 Gb/s. Alternatively, Alice may purchase the entire block and resell the portion that she does not need (by posting an offer on an RXP site). On the other hand, if Alice is allowed to resell the portion of OC-48 in one domain, this may not be the case in another domain. The providers in every domain may include clauses in their offer credentials allowing, or prohibiting, such unbundling. The flexibility of the policy language used in RXP allows many such special considerations to be encoded within the offer credentials. The advantage of having these restrictions expressed as policy is that they can be used directly by the provider's infrastructure without any need for conversion. Moreover, the end-user cannot alter these restrictions, since they are an integral part of the credential (and are protected by the resource provider's digital signature on the offer credentials).

### ADVANCE VS. ON-THE-FLY RESOURCE RESERVATION

On-the-fly (or just-in-time) resource reservations occur just before the resource is required. In contrast, advance reservations book resources well before they are used. In the case of on-the-fly reservations, the user collects the offers and proceeds to access the resources in short order, because the offers are effective immediately and have a short lifetime. There is no need to negotiate with the resource provider before accessing the resources. In the case of advance reservations the situation is different and more complex, since the providers need to know which offers will be exercised to plan their resource allocation. Once the end-user collects the offers, a notional reservation negotiation will be initiated. From the end-user perspective, the process is identical to an immediate reservation, but no resources are actually allocated. The resource providers involved note the reservations and issue receipt credentials to the user. Under pay-per-use schemes, payment is made at this time because the providers commit themselves to make the paid resources available at the requested time frame. In other words, the receipts serve as service-level agreements (SLAs). The receipt credentials are then used in the same manner as the offer credential was used in the on-the-fly scenario. When the resources are actually required, the end-user initiates a reservation negotiation, but sends only the reservation credential (instead of the offer and payment credentials). The allocation of reserved resources is handled in the same way as in the earlier case. Most large-scale experimental networks today do not support advance reservations, on the assumption that supply will outstrip demand in the near future. Our framework supports both types of infrastructure creation (advance and on-demand) leaving such decisions to the providers and users.

### Connecting Users with Infrastructure Providers

The requirement that end-users should be able to reserve resources from multiple providers without having to register with them implies that the system includes an entity that is trusted by everybody involved (usually referred to as a trusted third party, or TTP). Our model includes one or more credit institutions (CIs) that would normally handle payments from users to providers. In cases where we have collaborative sharing, the CI is still there, but the transactions need not involve real money. The CI issues users with credentials expressing their credit-worthiness (called credit-worthiness credentials or CWCS, see below), in other words, how much the user is allowed to spend. A credit-worthiness credential (CWC) is signed by a CI and associates a user key with a spending amount. Users are expected to supply a CWC with their reservation requests, thus providing the holder of the resource with the means to establish that the request is valid and that payment will be made. Thus, the CI establishes a connection of trust between users and resource providers. Like the offer notice boards, there is no requirement to have a single CI. It is, however, important that the resource providers have a way of confirming the keys of the various CIs. This is because the CWCs issued by the CIs to their customers will have to be verified by each provider. If a provider cannot verify a CWC, then it may be fake; trusting it may result in the equivalent of a bounced check.

### Exchange Points Revisited

In the current Internet, the Internet exchange (IX) performs managed interconnections of every autonomous system (AS). Whereas one AS advertises its local routes to the others, the dynamically changing *price of reachability* ("trading") cannot be established at IXs. In [10], a framework has been proposed that addresses the issue of optimal location of peer points for data exchange between ASs. We believe that this framework can be combined with the resource trading architecture proposed here, since the peering objective proposed in [10] was to minimize the cost of peering. Most recently, a few networks around the globe, such as KAREN (the Kiwi Advanced Research Education Network), BCnet, Amsterdam Internet Exchange (A-IX), and FirstMile, promoted the use of concept of carrier-neutral "peering exchanges" [11]. Some of them are interconnecting more than 200+ ISPs per exchange point, with the claim that "all data packets were created equal," or ISP-neutral.

In mobile networks, the concept of the resource exchange is implemented in the so-called GRX exchange points [12]. At every GRX, providers can dynamically negotiate resource exchange. For example, if one mobile provider lacks capacity, it can request its peer at GRX to provide it. As such, GRX plays the crucial role not only for users' roaming, but also for the proliferation of the new service providers that can provide service without owning any infrastructure network.

In the optical domain, the "distributed exchange" concept based on the optical BGP [13] has been first proposed in the CA*net4 research network. In the CA*net4, users create their own network by using the UCLP tool that allows for automatic interconnection of user-owned, dark fiber network with the optical core of the CA*net4 network. The optical core acts as a reconfigurable distributed exchange point for resource reservation; every cross-connect acts an infrastructure exchange point ("distributed"). When one ISP (or user) wants to establish a direct peering session, an IP BGP session is initiated and the lightpath is established under full control of the user. In addition to lightpaths, users can also allocate other resources such as optical switches or their partitions.

The GMPLS-based optical exchange architecture (GMPLS-XP) was first proposed in [4], where the architectural difference between the GMPLS-XP and multiple UNI was discussed and evaluated. The optical exchange points were then implemented by the mechanisms of the optical control plane and demonstrated in [14]. Exchange points have been also proposed for the MPLS networks [15]. In Grid computing, significant efforts have been made to advance the issues of economics in computation; however, the network as a Grid resource has not been considered for trading as yet [16].

What is still missing in the current, related research is a network architecture that can support generic multiprovider trading of the network infrastructure and its pieces, with selective capabilities of users to control the NEs physically, which can bring benefits to both resource providers and consumers.

*What is still missing in the current, related research is a network architecture that can support generic multiprovider trading of the network infrastructure and its pieces, with selective capabilities of users to control the NEs physically, which can bring benefits to both resource providers and consumers.*

## Proof of Concept

We created a proof-of-concept test setup to demonstrate the sample scenario discussed in the beginning of the Architecture section above. The two objectives that drove the basic implementation were the need to:

- Determine the cost and latency imposed by the credential evaluation
- Allow us to investigate how the command structure of a network element can be mapped into a the credential-based framework

First, latency in the execution of commands is crucial to the use of the system in an environment where rapid reconfiguration is required in order to minimize service disruption. Our concern was that as the chain of credentials increases, so too does the cost of evaluating a request in terms of computing resources. This, in turn, delays the execution of the command itself (i.e., the access control overhead is added to the command execution overhead). Later in this section we include measurements that demonstrate that techniques such as credential caching have substantially reduced overheads.

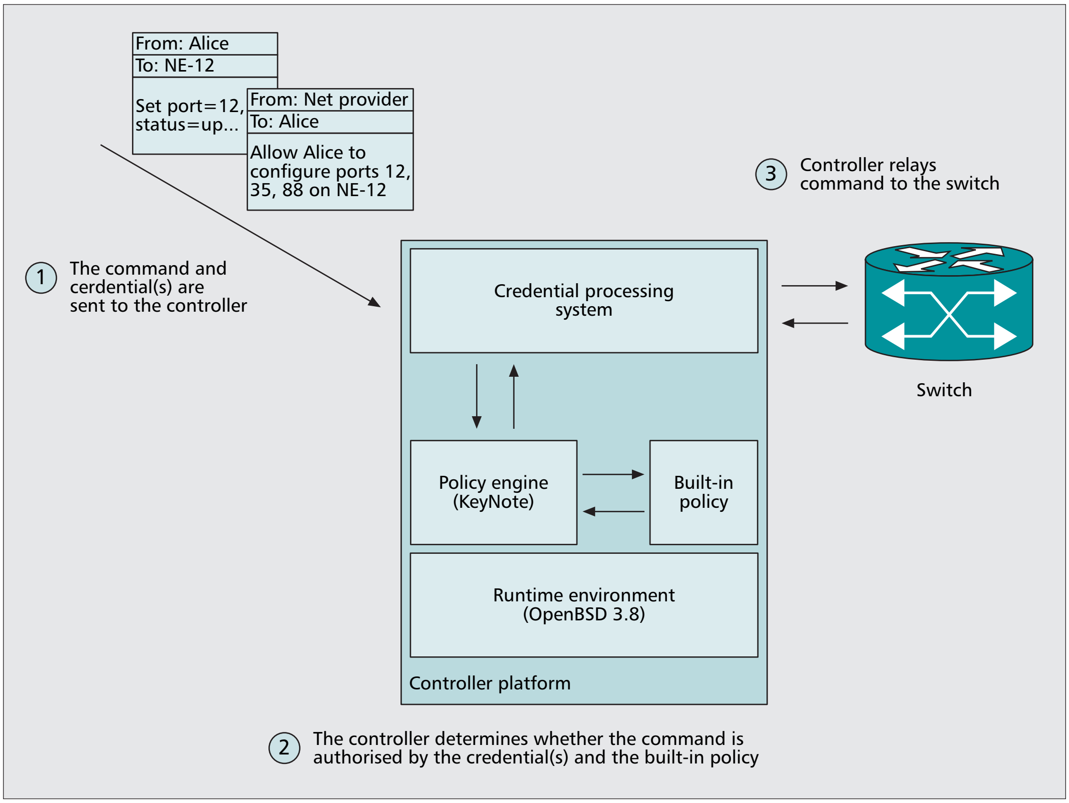

From: Alice
To: NE-12
Set port=12, status=up...
From: Net provider
To: Alice
Allow Alice to configure ports 12, 35, 88 on NE-12
1 The command and cerdential(s) are sent to the controller
3 Controller relays command to the switch
Credential processing system
Switch
Policy engine (KeyNote)
Built-in policy
Runtime environment (OpenBSD 3.8)
Controller platform
2 The controller determines whether the command is authorised by the credential(s) and the built-in policy

**Figure 5.** *The software architecture of the controller in the prototype system.*

| Number of credentials | Total time with policy evaluation (ms) | Total time without policy evaluation (ms) | Average time difference (ms) |
|---|---|---|---|
| 3 | 2793.14 | 248.96 | 2544.18 |
| 4 | 3479.87 | 261.30 | 3218.57 |
| 5 | 4753.54 | 304.14 | 4449.40 |
| 6 | 5518.06 | 321.46 | 5196.60 |
| 7 | 6721.52 | 337.72 | 6383.80 |

**Table 1.** *Cost of processing requests vs. number of credentials used.*

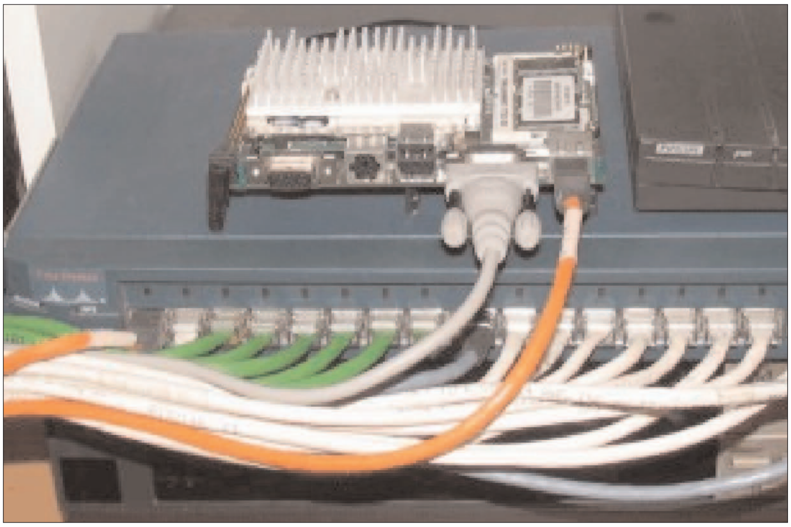
With this basic prototype, we plan to move up and integrate this architecture to a more complex network control and management platform in collaboration with the UCLPv2 team at the University of Quebec in Montreal (UQAM). First, we plan to include the demonstration of the RXP architecture in presence of multiple network infrastructures. Another important experimentation will be to analyze the side-effects, which can interfere with parts of the NE that are not under the control of the supplied credentials. Note that the specific implementation of the user-network interface or standard network management interface is out of scope here. The goal of our joint effort with UQAM team will be to experiment with the UCLPv2 mechanisms for admission, authentication, and access control with the credential-based system presented in this article.

The proof-of-concept prototype described here is based on a managed Ethernet switch, which supports VLANs. We used this switch because it was available and easy to integrate with our controller platform, and its command interface was well known. In this prototype, if we wish to link port 12 of the switch to port 88, we create a VLAN which includes only ports 12 and 88. The software architecture of the controller is shown in Fig. 5, while the switch and associated controller are shown in Fig. 6.

We assume that the switch receives commands only via the controller. Configuration requests and the appropriate authorization credentials are sent to the controller (step 1 in Fig. 5), which then uses the Keynote policy engine to determine whether the request should be carried out (step 2). Finally, assuming that the request is acceptable, the controller passes the command to the NE via the console interface (step 3). While for the purposes of the prototype we had to use an external computer (essentially an embedded version of a standard PC), we envisage that this functionality will eventually be integrated in the management software of the NE (the Ethernet switch in this case), thus removing the need for the external controller.

The credential illustrated below (expressed in the Keynote policy definition language, RFC-2704) has been issued by the network provider (`NP_KEY`) to Alice (`ALICE_KEY`) allowing her to send to network element 1234 "`set vlan`" commands involving ports 12, 35, or 88 (note that the keys have been truncated to reduce clutter). In the case of the use case described in the Architecture section, the network provider can give Alice a similar credential to allow her to reconfigure the right-hand NE to use the alternate path via provider B. A detailed description of the protocol and the structure of the credentials may be found in [17].

```
Keynote-Version: 2
Local-Constants:
ALICE KEY = “rsa-base64:MCgCIQ…
NP_KEY = “rsa-base64:MIGJAo…”
Authorizer: NP_KEY
Licensees: ALICE KEY
    Conditions: app domain == “Resource-XP”
    && date < “20060924”
    && command == “set vlan”
    && &element == 1234
    && (&port == 12 || &port == 35 || &port
    == 88) -> “true”;
Signature: “sig-rsa-sha1-base64:QU6SZ…”
```



■ **Figure 6.** *Controller connected to an Ethernet switch. The controller is a single board computer with a 266 MHz Pentium processor, 256 Mbytes RAM, running OpenBSD 3.7. The orange Ethernet cable connects the controller to port 1 of the switch, which is the only port assigned to the administrative VLAN (VLAN 1); the serial cable (gray DB-9) also connects the controller to the console port of the switch, allowing the controller to download configurations via SNMP (over the Ethernet port) or CLI commands over the serial console port.*

The dominant cost of this scheme is that of authentication and authorization, that is, the evaluation of the credentials to determine whether the request is consistent with the policy. While in the examples shown above we demonstrate requests involving two parties, this is not always the case. For example, a provider may resell resources leased from another provider, or a user may sublet resources to other users. For this reason, NEs may receive requests containing chains of two or more credentials and hence expend more computational power in order to evaluate them. Table 1 shows how the addition of credentials affects the overall processing of requests (the numbers originated from tests run on a Dell PowerEdge 1550 and represent average times over 100 trials). We show the time taken to process requests which include three to seven credentials with the policy engine active (column 2) and with the policy engine replaced with a routine that always returns true (column 3). The difference between the two measurements (column 4) shows the cost of credential evaluation, which is significant. The request processing time increases even if the policy engine is disabled because the system must still parse and process the larger requests.

Even though these operations are relatively expensive, the impact of the overhead is minimal, since:

- Such operations occur only when a reconfiguration is desired (e.g., switching to a backup link, modifying some parameter in the NE such as traffic shaping, etc.)
- The cost is distributed among the NEs, so that if a new network configuration affects *n* NEs, the total time for the operation is that of the slowest NE
- The number of credentials in a chain depends on the parties involved (i.e., the number of “middlemen” in the transaction), so it is unlikely to be higher than four or five

## CONCLUSION

We have presented one of the first approaches to trading of physical network resources, which includes not only a multiprovider infrastructure setup, but it also supports resource management which allows users to access network elements (NEs) based on credentials that contain the usage policy. Our experiments demonstrate the ability to dynamically lease resources based on their quality, cost, and availability according to user requirements and expectations. Moreover, our framework foresees the free-market trading of resources between multiple parties without previous agreement or sharing of authentication information. Authorization is handled in a distributed manner at the level of the physical NEs, including fibers, links, switches, and routers.

In this article we have assumed that all infrastructure providers are interested in exploiting the opportunities of opening up their infrastructures as well as in profiting from the economic benefits of trading. For this development to happen, network providers (ISPs) need to have an economic incentive, and we believe that this incentive will both emerge with the user’s desire for universal connectivity and the providers’ desire to return their investments in the infrastructure, if portions of it are underutilized for a manageable period of time. While we are aware that the implementation of that vision carries challenges not entirely technical in nature, we believe that this development will happen and will soon become relevant not only in experimental networks, but also commercially.

We believe that the use of the infrastructure trading paradigm is essential for pushing networks into the next phase of evolution, towards the truly globally available connectivity of the 21st century.

## REFERENCES

[1] T. Takeds *et al.*, “Layer 1 Virtual Private Networks: Service Concepts, Architecture Requirements, and Related Advances,” *IEEE Commun. Mag.*, Mar. 2004.

[2] CA*net User Controlled Lightpah, UCLP v2, http://grid2.canarie.ca/wiki/index.php/UCLPv2

[3] NSF Initative GENI, http://www.nsf.gov/cise/geni

[4] S. Tomic and A. Jukan, "Policy-based Lightpath Provisioning over Federated WDM Network Domains," *Proc. OFC 2002*, 2002.

[5] C. Metz, "Interconnecting ISP networks," *IEEE Internet Computing*, vol. 5, no. 2 , Mar.–Apr. 2001, pp. 74–801.

[6] M. Z. Hasan and S. Kapoor, "UNI Manageability — Provisioning in Optical UNI Enabled Networks," oif2001-350.

[7] V. Prevelakis, "A Secure Station for Network Monitoring and Control," *8th USENIX Security Symp.*, Washington, DC, 1999.

[8] V. Prevelakis and A. Keromytis, "Designing an Embedded Firewall/VPN Gateway," *Proc. Int'l. Network Conf. 2002*, Plymouth, U.K., 2002.

[9] K. Krauter, R. Buyya, and M. Maheswaran, "A Taxonomy and Survey of Grid Resource Management Systems for Distributed Computing," *Software: Practice and Experience*, vol. 32, no. 2, 2001, pp. 135–64.

[10] D. Awduche, J. Agogbua, and J. McManus, "An Approach to Optimal Peering Between Autonomous Systems in the Internet," *Proc. Int'l. Conf. Comp. Commun. and Networks*, 1998.

[11] KAREN (http://www.reannz.co.nz/home/), BCnet (http://www.bc.net), Amsterdam Internet Exchange (http://www.ams-ix.net), FirstMile (http://www.firstmile.us).

[12] K. J. Blyth and A. R. J. Cook, "Designing a GPRS Roaming Exchange Service," *Conf. 3G Mobile Commun. Technologies*, 2001, pp. 201–05.

[13] B. St. Arnaud *et al.*, "BGP Optical Switches and Lightpath Route Arbiter," *Optical Networks Mag.*, Mar./Apr. 2001.

[14] F. Dijkstra and C. de Laat, "Optical Exchanges," *Proc. GRIDNETS 2004*, Oct. 2004.

[15] I. Nakagawa *et al.*, "A Design of a Next Generation IX using MPLS Technology," *Apps. and the Internet '02*, 2002, pp. 238–45.

[16] R. Buyya, D. Abramson, and J. Giddy, "A Computational Economy for Grid Computing and Its Implementation in the Nimrod-G Resource Broker," *Future Generation Comp. Sys. J.*, Elsevier Science, 2002.

[17] D. Turner, V. Prevelakis, and A. D. Keromytis, "The Bandwidth Exchange Architecture," Presented at the *10th IEEE Symp. Computers and Commun.*, La Manga del Mar Menor, Spain, June 27–30, 2005.

## BIOGRAPHIES

VASSILIS PREVELAKIS (vp@drexel.edu) is an assistant professor of computer science at Drexel University. He has worked in various areas of security in systems and networks both in his current academic capacity and as a freelance consultant. His current research involves issues related to automation network security, secure software design, and auto-configuration issues in secure VPNs. He has published numerous papers in these areas and is actively involved in standards bodies such as the IETF. He has received research funding from DARPA (CHATS) and NSF (CAREER). He received his Ph.D. from the University of Geneva, Switzerland, and his M.Sc. and B.Sc. from the University of Kent, United Kingdom.

ADMELA JUKAN (jukan@emt.inrs.ca) is an associate professor at INRS/University of Quebec, Montreal, Canada, and a visiting professor at the University of Illinois at Urbana-Champaign. Prior to that she served as program director in Network Systems Research at the National Science Foundation, Arlington, Virginia. She has authored numerous papers in the field of networking and is recognized for having introduced the concept of quality of service routing in optical networks. Her current research interests include network performance, control, and management, with applications in wireless cellular networks and optical networks. She received an M.Sc. degree from Polytechnic of Milan, Italy, and a Ph.D. degree (cum laude) from Vienna University of Technology, Austria.