Resource Allocation in User-Controlled Circuit-Switched Optical Networks*

Wojciech M. Golab and Raouf Boutaba

School of Computer Science, University of Waterloo
200 University Avenue West, N2L 3G1 Waterloo, Ontario, Canada
{wgolab,rboutaba}@bbcr.uwaterloo.ca

Abstract. The concept of a user-controlled circuit-switched optical network is gaining popularity in an effort to fulfill the insatiable data transport needs of the scientific community. We consider the resource allocation challenges that arise in such a network, based on prior experience in developing a user-controlled lightpath management system. In particular, we examine problems related to partitioning of lightpaths and construction of end-to-end lightpaths in support of large data transfers. Through a simulation study, we show that bandwidth fragmentation is an important issue in partitioning of SONET circuits, and we demonstrate an effective countermeasure. We also explore novel optimization criteria for routing of end-to-end lightpaths, and present problem formulations with efficient polynomial time solutions.

1 Introduction

The Internet, while interconnecting a large part of the globe and supporting key services such as the World Wide Web and electronic mail, is not meeting the needs of emerging data-intensive scientific applications. Simply put, it is a vehicle too slow and unreliable to carry the massive volumes of data involved in complex distributed computations, such as weather modelling, and in collecting data from novel sensor-rich experiments, such as particle colliders and radio telescopes. These data-intensive applications require aggregate capacity in the range of Gigabits or even Terabits per second, which cannot be achieved using best effort delivery over a shared IP layer due to technological and economic limitations.

Circuit switching technologies offer scientists an attractive alternative to the Internet as they intrinsically provide guaranteed bandwidth and minimal delay, while avoiding the costly electronics associated with high-speed queuing and scheduling hardware. Optical circuits are particularly promising thanks to their superior capacity, low error rate, and low attenuation characteristics, as well as their favourable cost following a period of massive over-provisioning in the late nineties. In fact, there is a strong trend for research institutions, schools, and

* This research is partially supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

Z. Mammeri and P. Lorenz (Eds.): HSNMC 2004, LNCS 3079, pp. 776–787, 2004.

large enterprises to purchase optical wavelengths or entire strands of optical fibre in order to connect to each other, or to Internet service providers, using their own switching equipment. Such user-owned optical networks offer significant cost savings over conventional carrier-managed services, and make it possible for users to flexibly control and manage their infrastructure according to their needs. In addition, the ability to peer directly with other networks without the involvement of a service provider promotes the development of novel high-performance network applications that establish circuits on demand. For further information on user-owned and user-controlled optical networks the reader is referred to [1].

The hardware deployed in user-owned optical networks comprises a variety of optical and electronic devices. The most popular switching elements are digital cross-connect systems (DCSs) implementing the Synchronous Optical Network (SONET) or Synchronous Digital Hierarchy (SDH) standards, which provide powerful traffic grooming, performance monitoring, and protection switching capabilities. Digital switches are commonly used in conjunction with point-to-point Wavelength Division Multiplexing (WDM), which entails concurrent transmission of multiple optical carriers (i.e. wavelengths) over a common fibre in order to increase aggregate capacity. However, emerging all-optical switching devices based on micro-electro-mechanical systems (MEMS), for example arrays of tiny movable mirrors, are gaining popularity as they offer unprecedented switching capability at lower cost than equivalent high-speed electronics. Optical circuits provisioned over such devices are referred to as *lightpaths* since they represent continuous optical data paths. However, due to the popularity of legacy digital technologies we shall relax the definition of this term to include SONET streams, as well as circuits emulated over packet switching technologies.

In this paper, we consider the resource allocation problems associated with user-controlled circuit-switched networks, in which resources from multiple management domains are shared in a collaborative manner. Through shared control over switching nodes, referred to as lightpath cross-connect devices, users are able to combine resources obtained from multiple contributors into end-to-end lightpaths, and peer dynamically without relying on network service providers to configure core network equipment. We envision that each resource is a lightpath with a specific bandwidth, owner, and lifespan defined in terms of a finite duration ownership transfer mechanism. We consider the following operations on lightpaths:

- *advertisement* of a lightpath, signifying that it is available for lease by other users for a limited period of time
- *lease* of an advertised lightpath, which entails transfer of ownership for a limited period of time such that the lease terminates before the end of the advertisement period
- *partitioning* of a parent lightpath into a set of children having smaller bandwidth and spanning the same endpoints as the parent
- *concatenation* of lightpaths, whereby a sequence of shorter constituent lightpaths of common bandwidth is composed into a longer end-to-end lightpath.

Within this framework, we examine resource allocation issues from the perspective of resource efficiency and computational complexity.

We begin by reviewing relevant software technologies and mathematical theory in Section 2. Next, we discuss lightpath partitioning in the context of SONET circuits in Section 3, and propose an effective countermeasure to the problem of bandwidth fragmentation. Then, in Section 4, we explore the routing of end-to-end lightpaths based on novel optimization criteria, and propose efficient polynomial time computation techniques. Section 5 concludes the paper, reviewing open research problems and natural extensions to the work presented herein.

2 Related Work

2.1 Management and Control of Optical Transport Networks

Traditional optical transport networks are managed through a centralized operations and support system (OSS), where provisioning of connections is a slow, error-prone, and costly manual process. Consequently, the emerging approach is to incorporate signalling functionality that makes it possible to create connections on-demand through a user-network interface (UNI). This new functionality is typically placed in a distributed control plane, which in addition provides automated configuration and resource discovery [2]. A similar concept is the network-network interface (NNI), which allows signalling between control domains in support of end-to-end service provisioning [3].

Current standardization activities related to optical control planes include the GMPLS signalling protocols defined by the IETF [4], the Automatic Switched Optical Network (ASON) proposed by the ITU-T [5], as well as the UNI and NNI signalling specifications of the OIF [6, 7]. The NNI specification is particularly relevant to this paper as it addresses the exchange of information across control domain boundaries, though it is the least advanced of the above. Presently, only the intra-carrier scenario has been considered, where a single management domain is partitioned into multiple control domains, for example due to lack of interoperability between vendor-specific control planes. The problem of flexible user control is not considered.

The most direct efforts to realize the possibility of provisioning end-to-end lightpaths across multiple management domains were pioneered by CANARIE Inc., Canada's advanced Internet development organization. An initial proposal in 2001 considered extending the Border Gateway Protocol (BGP) [8] in order to solve the problems of disseminating lightpath availability and allocating inter-domain lightpaths along the Autonomous System (AS) path from a source to a destination. However, user-controlled lightpath provisioning (UCLP) software did not appear until 2003, when CANARIE Inc. and Cisco Canada Inc. launched a shared-cost Directed Research Program. The goal of the program is to grant users shared control over lightpath cross-connect devices (i.e. optical switches) in support of cross-connecting lightpaths, partitioning cross-connections, and transferring control of such partitions to other users. These low-level capabilities, in turn, make it possible to realize the four high-level operations defined in

Section 1, though our definition is also based on recommendation Y.1312 of the ITU-T [9]. The latter discusses similar functionality in the context of layer 1 virtual private networks (VPNs).

A variety of service-oriented software architectures and implementation technologies were proposed by the participants of the Directed Research Program [10]. The software architectures share a number of common elements including a module that interfaces with network elements (i.e. lightpath cross-connect devices), a service access point that exposes lightpath operations for machine-to-machine interaction, and a graphical user interface that invokes these operations on behalf of a human user. The solutions also differ in terms of offered functionality and data management. In particular, the University of Waterloo's system [11, 12] empowers users with the ability to control lightpaths as per all four of the operations defined in Section 1, while other solutions sacrifice the ability to partition lightpaths but use a distributed data model where each domain maintains information concerning its own resources. The latter affords greater administrative autonomy, but complicates user control over end-to-end lightpaths provisioned across multiple management domains.

2.2 Routing

Graph theory underlies the concept of routing, which plays a role not only in forwarding of Internet traffic at layer 3 of the OSI model, but also in the establishment of circuits. The fundamental approach is to compute a least cost path on a graph G with edge set E and vertex set V based on some cost metric (weight function) $W : E \rightarrow \mathbb{R}$. Centralized computations using information from link state routing algorithms can be performed using Dijkstra's algorithm, which assumes nonnegative link costs and runs in $\mathcal{O}(|E|+|V|\log|V|)$ time given an efficient Fibonacci heap data structure [13]. In the case of distance vector routing, a distributed form of the Bellman-Ford algorithm is used instead.

The basic least cost path problem can be extended to accommodate more sophisticated optimization criteria and QoS constraints [14, 15]. Of particular relevance to this paper is the problem of computing the widest path (i.e. one with greatest bottleneck bandwidth), which is a natural optimization criterion when creating a lightpath in support of a file transfer. The solution can be obtained using a slightly modified version of Dijkstra's or the Bellman-Ford algorithm, where the + operator is replaced with the *max* operator, and the cost metric is taken to be the inverse of bandwidth. The time complexity of the computation is not affected as the additional cost is merely $\mathcal{O}(|E|)$. The problem of finding the widest cost-constrained path can also be solved in polynomial time, although the solution in that case entails performing $\mathcal{O}(|E|)$ ordinary least cost path computations (one for each distinct bandwidth exhibited by some link in the network).

3 Partitioning of SONET Lightpaths

Partitioning of lightpaths is a fundamental operation that enables flexible sharing of resources. It allows the owner of a lightpath to separate the needed bandwidth from the unused or excess bandwidth, and make the latter available to another user. In this section we delve into the details of partitioning lightpaths realized as SONET circuits.

3.1 Challenges and Designs

Ideally, the partitioning operation would allow the bandwidth B of a lightpath to be distributed among partitions $P_1, P_2, \ldots, P_n$ of arbitrary size, provided that $\sum_{i=1}^{n} P_i = B$ and $P_i > 0$ for $1 \leq i \leq n$. However, in the case of SONET one must contend with constraints related to bandwidth granularity and placement of partitions within the parent circuit. The basic unit of bandwidth from which SONET circuits are formed is the Synchronous Transport Signal level 1 (STS-1), which has a raw bit rate of 51.840 Mbps. Frames carrying multiple STS-1 streams are multiplexed for transmission over the same optical carrier (OC) signal through byte interleaving, which gives rise to STS levels 3, 12, 48, and 192. This multiplexing process imposes an order on the STS-1 streams, which in an STS-N frame are numbered from 1 to N, the latter numbers being referred to as STS channels. Data-bearing circuits can be formed by grouping multiple STS-N streams, according to certain constraints imposed by the particular type of hardware used.

The most popular technique for forming circuits in SONET hardware is contiguous concatenation [16], which involves grouping the small (783 byte) payloads of N consecutive STS-1 streams into a larger concatenated payload. The result is an STS-Nc circuit, where "c" denotes contiguous concatenation, as illustrated in Fig. 1. In order to simplify hardware implementation, N is typically restricted to the set $\{3, 6, 9, 12, 24, 48, 192\}$, and the channel number of the first STS-1 in a contiguously concatenated group is also constrained [17]. As an example of the latter, the Cisco ONS 15454 platform [17] behaves as follows. If $N \in \{3, 12, 48, 192\}$, then an STS-Nc circuit can begin at any channel number of the form $1 + Nk$ where k is a nonnegative integer, provided that $N(k+1)$ is not greater than the total number of channels. For $N \in \{6, 9, 24\}$, an STS-Nc circuit can only begin at a channel number of the form $1 + 3k$, but not all such combinations are valid and the exact constraints do not follow a simple mathematical formula.

<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td></tr>
<tr><td colspan="3">STS-3c</td><td colspan="6">STS-6c</td><td colspan="3">STS-3c</td></tr>
</table>

Fig. 1. Grouping of STS-1 streams inside an STS-12 frame through contiguous concatenation.

A more flexible circuit formation technique that is gaining popularity is virtual concatenation (VCAT) [18]. VCAT enables formation of a virtual concatenation group (VCG), denoted as STS-Nc-Mv, by combining M different STS-Nc circuits. For example, Gigabit Ethernet traffic can be efficiently carried using STS-3c-7v or STS-1-21v circuits, instead of underutilizing a larger STS-24c. Although each member of a VCG is governed by the constraints of contiguous concatenation, the members need not be placed back-to-back in a single frame. Another benefit is that the additional processing is limited to the source and destination network interface cards, which makes it possible to deploy VCAT in legacy networks that only support contiguous concatenation in core switching nodes.

Contiguous concatenation has important implications on the design of the partitioning operation. Consider an STS-24c lightpath occupying channels 1 to 24. Suppose that a request arrives to create an STS-9c child lightpath, and is serviced by allocating channels 1 to 9. The remaining bandwidth constitutes 15 channels, but cannot be accessed as a single unit because STS-15c is not a supported circuit size. Thus, it is not intuitive to represent it as a single partition, i.e. as a single child lightpath. However, in considering multiple partitions of channels 10 to 24, one finds that there are multiple solutions: STS-3c and STS-12c, or STS-9c and STS-6c. It is not clear which is the better choice as each solution prohibits certain configurations of child lightpaths that are supported by the other.

Another important issue in partitioning is bandwidth fragmentation, which occurs as partitions are allocated and released. In the establishment of end-to-end connections, it is often necessary to partition some of the available lightpaths (i.e. those owned by the calling user, or advertised for lease by others) before the appropriate sequence of constituents can be concatenated. The excess bandwidth can be returned to the pool of available lightpaths, and used to service subsequent requests. Thus, repeated partitioning occurs, and larger blocks of unused bandwidth become fragmented, which makes it more difficult to form circuits using contiguous concatenation. To counter this process, one can consider reversing the partitioning operation to the extent possible. However, just as in the case of partitioning, multiple combinations of child lightpaths can be considered (this time for merger), and it is not clear how or when this should be performed.

We consider several designs of the partitioning operation in the context of SONET hardware, varying in complexity, intelligence, and transparency to the user. We group these into two families, according to the strategy used to allocate the excess bandwidth of the parent given a request to create a single child.

Dynamic Partitioning. A single child lightpath is created in each partitioning operation, and the unused bandwidth of the parent lightpath is treated as a single pool. If VCAT is used, we assume that bandwidth is maintained in the form of STS-1 streams and an arbitrary group can be allocated to realize a particular child lightpath. In the case of concatenated framing, the child lightpath is placed within the parent according to one of two policies: *first-fit* or *smart*. In the former, the first valid starting position is chosen that does not conflict with existing children. In the latter, a position is chosen such that the number of contiguous blocks of unused channels (i.e. fragments of unused bandwidth) is reduced, or else stays the same. If this is not possible, the *first-fit* strategy is used. In either case, the user is not directly aware of the constraints imposed by hardware on allocation of the unused bandwidth of the parent lightpath.

Static Partitioning. This approach only applies in the case of contiguous concatenation. Multiple children are created in each partitioning operation, such that the parent lightpath retains no unused bandwidth. The requested child lightpath is allocated at the first valid starting channel within the parent, and the remaining bandwidth is partitioned by repeatedly allocating additional children of maximal size using the same algorithm. The partitioning of a parent lightpath is reversed when all of the children are released. An active defragmentation procedure can also be used, whereby a set of idle children of a common parent that occupy a contiguous range of channels can be replaced with a single larger child, provided that the size and position of the corresponding channels translate into a valid circuit. The search for groups of children that can be merged can be performed whenever a lightpath is released by a user.

3.2 Performance Comparison

In this section we compare the performance of three variants of dynamic partitioning (VCAT, *first-fit*, *smart*) and two variants of static partitioning (*defragmentation off*, *defragmentation on*), as defined in the previous section. Specifically, we examine repeated partitioning of a single parent lightpath to service a sequence of bandwidth requests, as might occur during repeated end-to-end lightpath establishment. In the case of dynamic partitioning, a new child is created for each request and destroyed after the corresponding holding time. In static partitioning, the idle lightpath (i.e. the parent itself, or its child, or a child of a child, etc.) is used whose size is closest to the requested amount but no smaller. The lightpath is allocated directly if it has the correct bandwidth, or else partitioned, with the excess bandwidth distributed among a set of unallocated children.

For performance evaluation we consider partitioning an STS-48c parent lightpath, subject to the constraints imposed by the Cisco ONS 15454 SONET platform [17]. Requests for child lightpaths arrive one at a time, the circuit size being selected either uniformly at random from the set of legal values (*flat distribution*), or with probability inversely proportional to the bandwidth (*skewed distribution*). Request inter-arrival time and holding time have independent exponential distributions. Performance results are presented in Fig. 2, each bar representing a mean over 1000 simulation runs, each run lasting 1000 time units. The mean request inter-arrival time and holding times are indicated in the figure respectively as IAT and HT. The standard error of the mean in all cases is approximately 0.1%.

The results show that dynamic partitioning under VCAT is the best performer. This is expected, since the VCAT scenario presents the most liberal constraints in the partitioning operation. The other dynamic approaches perform 6%

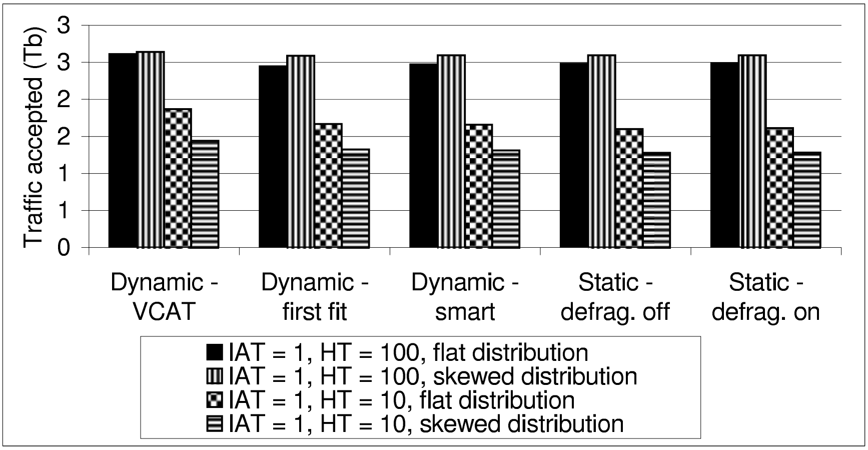

Fig. 2. Volume of traffic accepted during simulation period under various partitioning schemes.

worse when performance is averaged over the four request distributions tested, while static approaches perform 7% worse. The *smart* dynamic partitioning technique performs 0.2% better than *first-fit*. In static partitioning, the performance gain due to defragmentation is of the same magnitude. The simulation was repeated for other parent lightpath sizes, and similar relative performance results were observed in the sense that the non-VCAT algorithms achieved comparable performance levels. However, the performance difference between VCAT and the other approaches was amplified as the size of the parent lightpath was decreased. Overall, we conclude that both dynamic and static partitioning approaches work efficiently under contiguous concatenation.

4 End-to-End Lightpath Establishment

In this section we consider the computation of a path through a topology representing the set of lightpaths available to a particular user (i.e. those owned by the calling user, or advertised for lease by others) in order to establish an end-to-end lightpath between a particular pair of endpoints. We consider that a lightpath l has a bandwidth $B(l)$ and expiry date $T(l)$, the latter being defined in the context of a single request as

$$T(l) = \begin{cases} \text{lease expiry date of } l \text{ if calling user owns } l \\ \text{advertisement expiry date of } l \text{ otherwise} \end{cases} . \quad (1)$$

We define novel optimization criteria using the properties of bandwidth and expiry date, and propose efficient solutions based on Dijkstra's algorithm. Note that since the networks under consideration may contain multiple lightpaths between a particular pair of nodes (a, b), we consider only the minimum cost edge from a to b in a least cost path computation. Also note that given a bandwidth and duration requirement from the user, lightpaths having lesser bandwidth or expiry date are discarded when forming the input graph.

Minimizing the Hop Count, Breaking Ties. In a circuit-switched network it is natural to define hop count in terms of physical hops, i.e. fibre spans between physically adjacent cross-connect devices. Given such a metric, it is possible to use Dijkstra's algorithm directly to perform a shortest path computation. However, due to the fact that multiple parallel lightpaths can exist between a given pair of nodes, it is worthwhile to consider an additional criterion in order to break ties among paths of equal length. A sensible approach is to consider bandwidth, and favour those lightpaths that match the requirement of the user, which avoids unnecessary partitioning and fragmentation of bandwidth. A weight function that makes it possible to achieve this goal using a least cost path computation is

$$W(l) = H(l) + \frac{B(l)}{|V| \max_{k \in E} B(k)} \tag{2}$$

where H is the hop count, B is the bandwidth, V is the set of cross-connect devices, and E is the set of lightpaths. The correctness of W follows from the fact that it is positive-valued, in which case the least cost path is acyclic and contains fewer than $|V|$ edges. This implies that the contribution of the second term over a path is less than unity, which means that the least cost path has minimal hop count. Since the second term is proportional to $B(l)$, lightpaths that match the user's requirement are preferred when multiple paths exist with minimum hop count.

Maximizing the Expiry Date. The formation of an end-to-end path can also be optimized with respect to the expiry date, which for a compound lightpath is equal to the minimum expiry date of the constituent lightpaths, as per the definition above. This is useful when a user is unable to specify a minimum acceptable lease expiry date because of uncertainty, or because the end-to-end lightpath is needed for as long as possible, in which case the safest approach is to maximize the expiry date given the bandwidth requirement. Mathematically, the problem is analogous to finding the widest path, and can be solved using the modified Dijkstra or Bellman-Ford algorithms described in Section 2.2.

Minimizing File Transfer Time. One of the key features of a user-controlled circuit-switched network is the ability to dynamically provision high performance data paths in support of individual file transfers. In that case, a useful optimization strategy is to select an end-to-end path such that file transfer time is minimized. This is a generalization of the widest path problem since one must take into consideration the additional constraint that the expiry date of the end-to-end path must be sufficiently far in the future to permit the transfer of the file in question at the optimal bandwidth.

Given the size S of a file, the above problem can be solved by performing multiple reachability computations to determine the maximum feasible bandwidth, followed by a least cost path computation to determine the optimal path at that bandwidth. The former phase is presented below as Algorithm 1. At a high level, the algorithm identifies the possible bandwidths of the end-to-end path, and for each such bandwidth B_i performs a search (i.e. Breadth First Search) to determine whether a suitable end-to-end path exists at that bandwidth. At each iteration, only those lightpaths are considered whose expiry date permits the transfer of S units of data at rate B_i.

Algorithm 1 *Maximum Transfer Rate*

Input: $G(E, V)$ – network topology (lightpaths and cross-connect devices)
v_s, v_d – source and destination vertices
S – size of file to be transferred
Output: $B_{\max}$ – maximum bandwidth of a path from v_s to v_d through G that can be formed sufficiently long to transmit a file of size S
Begin:
$T_{\text{now}} \longleftarrow$ current time + safety margin
$B \longleftarrow$ list $\langle B_1, B_2, \ldots, B_N \rangle$ of distinct elements of $\{B(e) \mid e \in E\}$ in descending order
for i from 1 to N do
  $E_i \longleftarrow \{e \mid e \in E$ AND $B(e) \geq B_i$ AND $B_i(T(e) - T_{\text{now}}) \geq S\}$
  if there is a path from v_s to v_d in $G_i(E_i, V)$ then return B_i
end for
return failure

If N is the number of distinct bandwidths in the input graph, the worst case time complexity of Algorithm 1 is $\mathcal{O}(N(|E| + |V|))$, since $\mathcal{O}(|E| \log N)$ steps are needed to create the list $\langle B_1, B_2, \ldots, B_N \rangle$, and there are N iterations of at most $|E| + |V|$ steps each. However, the cost of each iteration is typically smaller as the number of edges in G_i is typically less than $|E|$. In addition, if G_{i-1} is a subgraph of G_i, which is more likely to happen when S is small, then the results of the reachability computation from iteration $i-1$ can be reused in iteration i. Thus, the running time can be reduced, although the worst case time complexity remains the same.

Concerning the number of iterations in Algorithm 1, the factor N is a small integer for a SONET network since the set of possible bandwidths is limited by hardware. For example, $N \leq 8$ for a typical SONET switch equipped with OC-192 cards. If lightpaths are provisioned using technologies that do not impose such a coarse bandwidth hierarchy, for example ATM, then N can potentially be as high as $|E|$, and the worst case running time of Algorithm 1 is $\mathcal{O}(|E|^2 + |E||V|)$.

5 Conclusions and Future Research

In this paper we have discussed resource allocation problems that emerge in a user-controlled circuit-switched network. We focussed on the two fundamental operations in such a system, namely lightpath partitioning and concatenation. Although our examination of partitioning in Section 3 is based on development experience with a specific SONET platform, the constraints and optimizations discussed are relevant to SONET hardware in general as concatenated framing technology is widely deployed. Through a simulation study we have shown that the dominant issue in efficient partitioning is fragmentation of bandwidth, and that its effects can be effectively treated in software. The path computation techniques presented in Section 4 are not specific to any hardware technology. They provide efficient polynomial time solutions to novel optimization problems that arise in a user-controlled circuit-switched network, where each link is subject to an expiry date and multiple parallel links can exist between each pair of nodes.

We envision several directions for future research in the area of resource allocation in user-controlled circuit-switched networks. One of these is to investigate the behaviour of the partitioning operation under a mixture of hardware technologies, for example digital SONET or ATM switches interconnected using an all-optical WDM layer based on photonic cross-connects. The relevant hardware constraints must be identified and current techniques extended in order to ensure efficient allocation of resources. Another open research problem is related to scalable inter-domain routing. While the path computation techniques presented in Section 4 assume a global view of the inter-domain technology, it would be worthwhile to consider a more efficient path vector routing scheme, whereby each domain receives aggregated information from its neighbours. Finally, whereas we assume that each domain can cross-connect any pair of incident inter-domain paths in a nonblocking manner, one can consider the more general case where the latter process is constrained by resource availability.

We believe that user-controlled circuit switching is an important trend not only in connecting the scientific community but also in commercial networks. User control translates into faster and more flexible service provisioning, while optical circuits offer superior performance and security compared to higher layer connectivity or leased copper lines. We foresee much research activity in this area, especially as all-optical WDM switching technologies mature and agile optical networks are deployed.

Acknowledgement

Thanks to Youssef Iraqi for his valuable comments concerning the design of the partitioning and end-to-end lightpath establishment operations. The University of Waterloo development team [12] is acknowledged for the implementation of the Waterloo UCLP system.

References

1. St. Arnaud, B., Wu, J.: Customer-controlled and -managed optical networks. Journal of Lightwave Technology **21** (2003) 2804–2810

2. Alanqar, W., Jukan, A.: Extending end-to-end optical service provisioning and restoration in carrier networks: Opportunities, issues, and challenges. IEEE Commun. Mag. **42** (2004) 52–60
3. Saha, D., Rajagopalan, B., Bernstein, G.: The optical network control plane: State of the standards and deployment. IEEE Commun. Mag. **41** (2003) S29–S34
4. Internet Engineering Task Force: Generalized Multi-Protocol Label Switching architecture. IETF draft, draft-ietf-ccamp-gmpls-architecture-07.txt (2003) http://www.ietf.org/internet-drafts/draft-ietf-ccamp-gmpls-architecture-07.txt.
5. Telecommunication Standardization Sector of ITU: Recommendation G.8080/Y.1304: Architecture for automatic switched optical network (ASON). (2001)
6. Optical Internetworking Forum: OIF-UNI-01.0 - User Network Interface (UNI) 1.0 signaling specification. http://www.oiforum.com/public/documents/OIF-UNI-01.0.pdf (2001)
7. Optical Internetworking Forum: OIF-E-NNI-01.0 - intra-carrier E-NNI signaling specification. http://www.oiforum.com/public/documents/OIF-E-NNI-01.0.pdf (2004)
8. Blanchet, M., Parent, F., St. Arnaud, B.: Optical BGP (OBGP): InterAS lightpath provisioning. IETF draft, draft-parent-obgp-01.txt (2001) http://www.ietf.org/internet-drafts/draft-parent-obgp-01.txt.
9. Telecommunication Standardization Sector of ITU: Recommendation Y.1312: Layer 1 virtual private network generic requirements and architectures. http://www.itu.int/itudoc/itu-t/aap/sg13aap/recaap/y1312/ (2003)
10. CANARIE Inc.: CA*net4 directed research program – funded projects. http://www.canarie.ca/funding/research/projects.html (2003)
11. University of Waterloo UCLP Team: Statement of work: User controlled lightpaths project. http://bbcr.uwaterloo.ca/∼canarie/SOW.pdf (2003)
12. University of Waterloo UCLP Team: User Controlled Lightpaths Project. http://bcr2.uwaterloo.ca/∼canarie/ (2004)
13. Fredman, M., Tarjan, R.: Fibonacci heaps and their uses in improved network optimization algorithms. Journal of the ACM **34** (1987) 596–615
14. Younis, O., Fahmy, S.: Constraint-based routing in the Internet: Basic principles and recent research. IEEE Communications Surveys & Tutorials **5** (2003) http://www.comsoc.org/livepubs/surveys/public/2003/sep/pdf/fahmy.pdf.
15. Chen, S., Nahrstedt, K.: An overview of quality of service routing for next-generation high-speed networks: Problems and solutions. IEEE Network **12** (1998) 64–79
16. Cisco Systems Inc.: Understanding concatenated and channelized SONET interfaces on Cisco routers. http://www.cisco.com/warp/public/127/concat_16147.pdf (2002)
17. Cisco Systems Inc.: Cisco ONS 15454 reference manual, releases 4.1 and 4.5. http://www.cisco.com/univercd/cc/td/doc/product/ong/15400/r4145doc/r4145ref/index.htm (2003)
18. Cisco Systems Inc.: Leveraging transport for data services with Virtual Concatenation (VCAT) and Link Capacity Adjustment Scheme (LCAS). http://www.cisco.com/warp/public/cc/so/neso/meso/meipac/lcas_wp.pdf (2003)