東北大学機関リポジトリ
TOUR
Tohoku University Repository

Design and Analysis of Optical Router Architectures

著者	王 暁亮
号	54
学位授与機関	Tohoku University
学位授与番号	工博第4201号
URL	http://hdl.handle.net/10097/61640

氏名	おう しょ りょう 王 暁 亮
授与学位	博士（工学）
学位授与年月日	平成２１年９月９日
学位授与の根拠法規	学位規則第４条第１項
研究科，専攻の名称	東北大学大学院工学研究科（博士課程）電気・通信工学専攻
学位論文題目	Design and Analysis of Optical Router Architectures （光ルータ構成方式の設計と解析）
指導教員	東北大学教授 亀山 充隆
論文審査委員	主査 東北大学教授 亀山 充隆 東北大学教授 山田 博仁 東北大学教授 徳山 豪 東北大学准教授 姜 暁鴻

論文内容要旨

The network router, which serves as the most crucial equipment in a network node for packets processing, switching and buffering, is now performing its three main functions electronically. Hence, the growth in router capacity is closely linked to Moore`s law and to memory bandwidth growth, which are historically slower than the growth in optical link capacities. This is why routers have increasingly become a bottleneck for Internet communications. To keep up with incoming fiber traffic, network operators now need to piece together many routers per backbone node. Further, routers need to be packed more densely. In the past years, the power consumption per core router rack has exponentially increased and already exceeds standards. Clearly, this trend cannot continue indefinitely. Some future alternative solutions for router design are mandatory.

A natural approach is to fully harness the power of optics by using all-optical router, where its three main functions are all performed optically. For this purpose, it is necessary to explore the effective router architectures which do not need complex switch fabric scheduling, scale to larger number of interfaces and is suitable for high speed optical transmission. In this thesis, the router architectures to be investigated include the load-balanced architecture, the output queued (OQ) architecture and the shared buffer architecture. We will study how to emulate the above architectures based on the available optical switches and buffering elements.

We first explore the load-balanced router architecture that is based on load-balancing packets uniformly inside the router before forwarding them to their correct destination. Recent researches have shown that a load-balanced router does not require any scheduler and that it can guarantee 100% throughput for a broad class of traffic. Besides, as both switches follow the same predetermined circulation reconfiguration, they are

suitable to be realized by using optical switch fabric like Arrayed Waveguide Grating (AWG) or Micro Electro Mechanical Systems (MEMS) to achieve high switching speed, large switch size and almost zero power consumption. This makes the load-balanced router an appealing architecture to study. However, the main problem of the load-balanced router is that packets can be mis-orded. Chapter 2 provides a solution to this problem which not only keeps packet in-order but also guarantees the good performance as that of the basic architecture. We proposed a revised load-balanced router architecture by introducing re-sequence buffer at the departure ports. The mis-ordered packets can be re-ordered in the re-sequence buffer. To bound the re-sequence buffer size, we further apply limited central buffer such that the amount of mis-ordered packets at the output port is finite. Through theoretical analysis, we show that our proposed architecture requires only N^2 (N is the port number) re-sequence buffer size to avoid packet mis-order while guarantee 100% throughput as well. In order to evaluate the average packet delay performance in the proposed load-balanced router, we found that the proposed architecture can provide prior delay guarantees under broad traffic patterns.

We then study the output queued router architecture where the memory is set at the output port of switch fabric. Under this architecture, packets are immediately forwarded to the destined output ports once they arrive at the inputs. The OQ router architecture benefits from 100% throughput and the best delay performance as the destining packets are always available at the output ports. However, this architecture suffers from memory speed constraint. This is even a major impediment in the optical domain due to the lack of optical random access memory. Instead, the only feasible way of realizing optical buffers is to use fiber delay lines (FDL) in which packets can propagate for a fixed time interval. For the implementation of OQ router architecture, we investigate both the feedback-based and feedforward-based designs, since both kinds of designs have their advantages and disadvantages (e.g., feedback architecture is cost effective while the feedforward architecture is easy to control).

In chapter 3, the feedback OQ router architecture was studied which first proposed by using the loop fiber delay lines for buffering packets. This architecture has multiple feedback FDLs of appropriately-selected length that are shared by all input and output ports. The optical switching fabric used is memoryless, rearrangeable non-blocking. Any two delay lines may have the same or different delay values, and delay lines of length greater than one packet duration reduce the number of recirculation loops needed. The corresponding scheduling algorithm alleviates the recirculation problem by the implementation of FDLs and output port reservation and thus provides efficient utilization of loop fiber buffers. However, this algorithm does not attempt to keep packets in their proper first-in-first-out sequence. To solve this problem

but do not degrade the packet loss and delay performance, we have provided a new scheduling algorithm. By first revising the fiber delay lines setting from exponential distribution to linear distribution, we alleviate the possibility of packet mis-order. Then, we proposed a new algorithm by introducing time label to completely avoid the packet mis-order. The simulation results concerning the previous algorithm and our new algorithm have shown that the new algorithm achieves an similar packet loss rate and delay performance as the available reservation algorithm but keeps packets in-order. Finally, we further extended our work to support variable-length burst swiches.

In Chapter 4, we study the feedforward OQ router architecture which adopts multistage of fibers delay lines connected by memoryless optical switches. The delay paths of packets contain FDLs from the first stage to the last stage. A good property of this architecture is that it enables packets self-routing to their output ports at right time. This is because the delay path can be expressed by the r-ary presentation of assigned delay which can be known in advance according to the exact emulation of OQ router. However, the packets may conflict with each other for the same delay line inside this architecture. In order to construct a non-conflicting architecture, the available work has already proposed a upper bound for the number of required delay lines in each bound. But this bound did not consider carefully the effect of packet size, and thus require more hardware cost than the real requirement. To improve the loose upper bound, we first derive a tighter upper bound for the number of required delay paths in the feedforward architecture. To know what`s the minimum hardware cost we can expected for a non-conflict feedforward OQ router architecture design, a new lower bound of required delay paths is further derived. By comparing the hardware costs of the previous and our new upper bounds, we found that our new upper bound can heavily reduce the number of fiber delay lines and corresponding switch fabric size. Then, we verified by computer simulation that the new lower bound, although require less hardware, can still achieve the similar packet loss rate as that of the upper bound under both uniform and un-uniform traffics. These results are very useful for router designer to evaluate the hardware costs of a non-conflicting feedforward OQ router.

Finally, in Chapter 5, we focus on the shared buffer router architecture where the buffers are sandwiched between two switch fabrics in a space-time-space model. Packet can be immediately written into an idle buffer and read from a non-idle buffer. It is thus said buffers are shared by all the input and output ports. It is capable to achieve the same performance as the OQ router architecture but has the maximum memory utilization. This is really promising router architecture, while until now there is no corresponding optical implementation. Our work constructs aims at realizing this router architecture by using optical switch and fiber delay lines. A straightforward idea is to follow its electronic space-time-space model, where the FDLs

are sandwiched between optical switch fabrics. Obviously, the single stage feedforward construction is easy to control but its complexity is high. We further explored the single stage feedback structure to build more effective shared buffer router architecture. We have found that by properly setting the length of feedback fiber delay lines (linear distribution here), and scheduling packets properly, it is possible to emulate the behaviors of the electronic shared buffer router. By theoretical analysis, we prove the proposed feedback construction can realize a shared buffer router. This result is the first attempt for optical shard buffer router design.

In summary, we studied the router architectures that can be implemented by using the available optical switch and buffer technologies. These architectures are potentially applied to future optical transmission network.

論文審査結果の要旨

近年の光ネットワークにおいては，パケットの中継を行うルータが通信性能向上のボトルネックになってきており，光信号を電気信号に変換することなく高速に中継が行える光ルータの開発が望まれている．著者は，光ルータとして実現が期待されているロードバランスルータ，出力キュー付きルータ，共有バッファルータに対して，ミスオーダと呼ばれるパケットの入出力順が変化する問題を解決し，さらに処理遅延とハードウェアコストを減少可能な新しい構成方式を提案し，その有効性を実証した．本論文はその成果を取りまとめたもので，全文６章よりなる．

第１章は緒言である．

第２章では，パケットのミスオーダが発生しないロードバランスルータの構成方式を提案している．本ルータでは，ミスオーダを解決するための再配列バッファと共に，入出力スイッチブロック間に再配列時の処理遅延を減少させるための二次元配列構造バッファを備えている．ミスオーダがなく 100%のスループットを達成できることを理論的に保証し，シミュレーションにより従来の構成法よりも平均処理遅延を大幅に短くできることを示している．これは，ミスオーダの解決と高速化を共に達成する有用な成果である．

第３章では，フィードバック形の出力キュー付きルータに対して，ミスオーダの発生を回避するパケットスケジューリング法を提案している．パケットのバッファリングに用いられるファイバ遅延ラインを整数倍の長さに設定し，タイムスタンプを用いたスケジューリング制約を加えることで，ミスオーダを回避できることを示している．これは，ハードウェアコストを増加させることなくミスオーダを解決できる重要な成果である．

第４章では，フィードフォワード形の出力キュー付きルータのハードウェアコストを削減するために，パケットの競合が起こらないことを保証するファイバ遅延ライン数の上界値を導出し，さらに下界値を求めるアルゴリズムを提案している．これらの上界値や下界値を用いれば，従来法と同程度のパケット損失率を保ちながら，ファイバ遅延ライン数を大幅に削減可能であることを，シミュレーションにより明らかにしている．これは実用上優れた成果である．

第５章では，共有バッファルータの実現に向けて，フィードフォワード形ルータの概念を応用した構成方式と共に，ファイバ遅延ライン長の設定法及びパケットスケジューリング法を提案している．これにより，共有バッファルータの機能を完全に実現可能であることを理論的に明らかにしている．これは，共有バッファルータ機能を実現した初めての構成方式であり，優れた成果である．

第６章は，結言である．

以上要するに本論文は，次世代の光パケット通信の実現に向けて，光スイッチとバッファ技術を用いた光ルータの構成方式を提案し，シミュレーションと理論的解析によりその有用性を明らかにしたものであり，情報通信工学の発展に寄与するところが少なくない．

よって，本論文は博士(工学)の学位論文として合格と認める．