第 42 卷 第 16 期/2022 年 8 月/光学学报 研究论文

光学学报

基于 Ring-Clos的全光交换架构

杨晓雪[1]，胡冰[1,2*]

[1]浙江大学信息与电子工程学院，浙江 杭州 310027；

[2]之江实验室智能网络研究院，浙江 杭州 310027

摘要 数据中心面临带宽与能耗的双重挑战，光交换具有高带宽、低功耗和透明传输等优势，是一种极具前景的解决方案。针对Clos架构中使用电缓存所引起的高能耗与高延迟问题，提出了一种基于Ring-Clos的全光交换架构，通过级内连接与可调波长转换器为光分组提供相邻中间级路由，解决了部分输入级或中间级输出端口冲突的问题。利用并行匹配调度算法为光分组分配路径，算法复杂度低，硬件实现简单。仿真结果表明，所提架构的丢包率仅为Clos架构的48.81%，有效提高了网络性能。

关键词 光通信；数据中心网络；光交换；Clos架构；竞争解决；丢包率

中图分类号 TP308 **文献标志码** A **DOI**：10.3788/AOS202242.1606004

All-Optical Switching Architecture Based on Ring-Clos

Yang Xiaoxue[1], Hu Bing[1,2*]

[1]*College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou* 310027, *Zhejiang, China*;

[2]*Intelligent Network Research Institute, Zhejiang Lab, Hangzhou* 310027, *Zhejiang, China*

Abstract Data centers are facing dual challenges from bandwidth and energy consumption, for which optical switching is a promising solution owing to its advantages of high bandwidth, low power consumption, and transparent transmission. Considering the high energy consumption and long delay caused by the electronic buffers used in the Clos architecture, an all-optical switching architecture based on Ring-Clos is proposed. Routing between adjacent central modules is provided for optical packets by employing intra-stage connection and tunable wavelength converters, thereby solving the output port conflict problem at certain input or central modules. Meanwhile, the concurrent dispatching for Ring-Clos switch scheduling algorithm is applied to distribute routes for optical packets as it has low complexity and simple hardware implementation. The simulation results show that the packet loss rate of the proposed architecture is as low as 48.81% of that of the Clos architecture, which means the proposed architecture effectively improves the network performance.

Key words optical communications; data center network; optical switching; Clos architecture; contention resolution; packet loss rate

1 引 言

数据中心是一种连接成千上万台高性能服务器的网络，作为数字化世界的信息支柱，其承载着多种存储计算密集型服务，如云计算、大数据和搜索引擎等。随着这些应用的发展，指数型爆炸的数据流量给数据中心带来了巨大的挑战[1-4]。

Clos交换架构由多个交叉开关矩阵结构级连组成，以减少大规模网络的交叉点数目，并解决在电话网络中使用机电开关所引起的性能与成本问题。当前，Clos网络因复杂度低、易扩展等多种优点被广泛应用于交换网络中。大多数架构仍采取电交换技术，如：Joshi等[5]提出了一种用于片上通信的Clos网络，其在各种流量模式下均提供一致的低延迟与高带宽；Schröder等[6]将两张独立的中央交换卡分别连接至16张外围线卡处，这种基于Clos的无源双星互连架构具

收稿日期：2022-01-21；**修回日期**：2022-03-06；**录用日期**：2022-03-14

基金项目：国家重点研发计划(2019YFB1802905)、国家自然科学基金面上项目(61971377)、浙江省自然科学基金重点项目(LZ22F010008)

通信作者：*binghu@zju.edu.cn

有高冗余度，可确保在一张交叉板故障时，网络仍可正常工作。电交换技术需要大量光-电-光（O/E/O）转换，这带来了高功耗与高延迟，而光交换技术具有高带宽与透明传输等优势，有源光纤速率可达 400 Gb/s[7]，因此研究人员设计了多种光交换架构，如：Xi 等[8]以三级阵列波导光栅路由器（AWGR）为核心，构建了基于 Clos 的光分组交换架构，具有高吞吐量与扩展性；Lea[9]为减少可调波长转换器（TWC）的使用，将中间级更换为空间交换机，大幅降低了架构的功耗；邓鸿胜等[10]利用 16 个波长实现了机架内 64 个服务器的光互连，在提高数据中心吞吐量的同时，降低了设备成本。

在基于可重排无阻塞 Clos 拓扑的光分组交换架构中，若采用并行式调度算法[11-17]，分组将竞争相同端口，研究者们通常在输入线卡处设置电缓存以解决竞争问题[8]，但 O/E/O 转换将引起高功耗与高延迟。本文针对 Clos 架构中的分组竞争问题展开研究，提出了无缓存的 Ring-Clos 架构，其通过相邻中间级互连，为输入、输出端口提供了更多的冗余链路。同时，提出了一种并行启发式调度算法，用于为光分组分配路由路径。

2　Ring-Clos 架构设计

2.1　Ring-Clos 系统架构

在对 Clos 架构中的光分组进行路径分配时，若链路选择不当，则可能会出现如下情况。在图 1 中，实线表示此链路已被占用，虚线表示此链路未被占用。此时，输入为 3、输出为 3 的分组无法分配到空闲链路，即分组间发生竞争。

当光分组在无缓存的 Clos 网络中发生竞争时，一般的处理方式是在 Clos 的入口线卡处设置电缓存，本

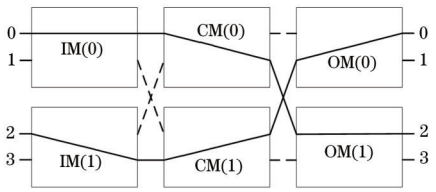

图 1　中间级分配中的竞争问题

Fig. 1　Contention in central module allocation

时隙中竞争失败的光分组将等待之后的时隙并与其他光分组重新竞争，但 O/E/O 转换不可避免地增加了网络时延与功耗。为避免冲突，本文提出了 Ring-Clos 架构，包括输入级（IM）、中间级（CM）和输出级（OM），其可以同时路由 $M(M-2)$ 个光分组，其中 M 为每个交换单元的端口数。IM 与 OM 的每个交换单元可任意选择一种配置时间在纳秒级的结构，如基于 TWC 与 AWGR 的波长交换结构或基于半导体光放大器（SOA）的广播与选择结构等。CM 由互连的交换单元组成，每个交换模块为由 TWC 与 AWGR 组成的光波长交换结构。如图 2 所示，黑线表示 IM/OM 与 CM 的级间连接，灰线表示 CM 的级内连接，IP 代表输入端，OP 代表输出端。本文用 R(M) 代表一个 Ring-Clos 架构，其中 IM、OM 级由 $M-2$ 个交换单元组成，CM 级由 M 个交换单元组成，每个交换单元的大小均为 $M\times M$。每个 CM 级交换单元的 $M-2$ 组端口用于与 IM 或 OM 互连，另外 2 组端口用于与相邻 CM 互连，由于 AWGR 的循环波长路由特性，不同波长的 M 个光分组可能会同时从一个 CM 路由至相邻的多个 CM 中，因此在这些端口处放置带有复用器（MUX）和解复用器（DEMUX）的 M 个 TWC 用于 CM 级内连接。

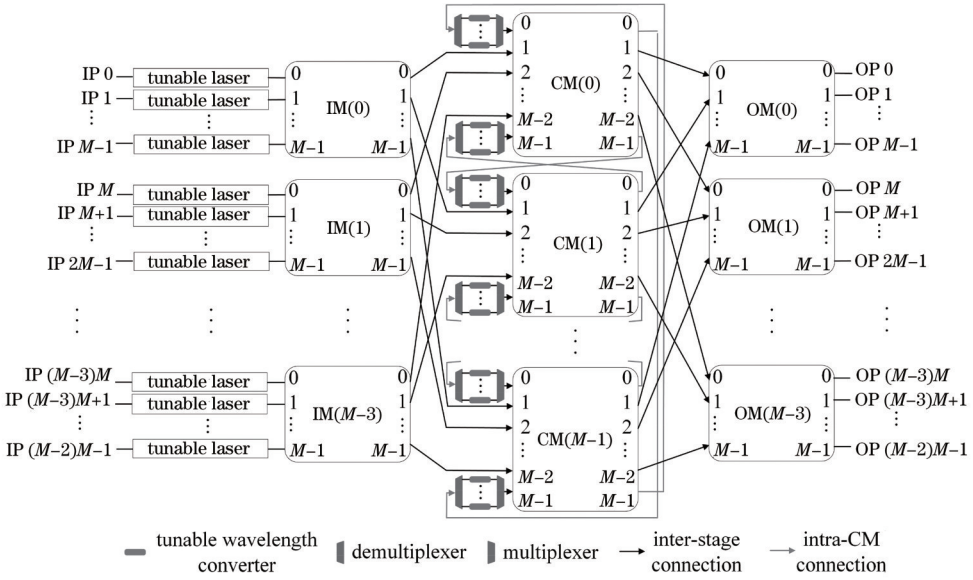

图 2　Ring-Clos 架构示意图

Fig. 2　Schematic diagram of Ring-Clos architecture

2.1　Ring-Clos 路由

本文将以具有 8 个节点的 Ring-Clos 架构[R(8)]为例，具体描述其路由方式。在某一时隙中，已为 5 个光分组分配路径，现有连接状态如图 3 所示，粗线表示链路已被占用，细线代表链路空闲。

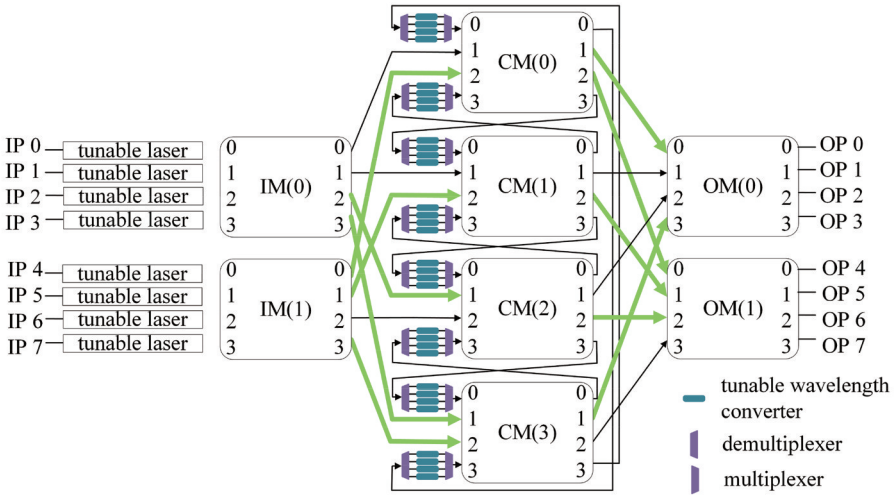

图 3　R(8)架构的初始连接状态

Fig. 3　Initial connection state of R(8) architecture

用 $P(p,q)$ 表示输入端口为 p，输出端口为 q 的光分组。当为光分组 $P(0,0)$ 分配路径时，首先要找到输入端口所在的 IM 与输出端口所在的 OM，然后确定是否存在空闲 CM。IM(0)的可用 CM 为 CM(0)与 CM(1)，OM(0)的可用 CM 为 CM(1)与 CM(2)。当限制光分组最多只能通过一个 CM 时，仅有一条路径可以选择，即经过 CM(1)的路径，如图 4 所示，并将这种经过一个 CM 交换单元的路径称之为 1 跳。在这种情况下，竞争分别发生在 IM 输出端口与 CM 输出端口处。

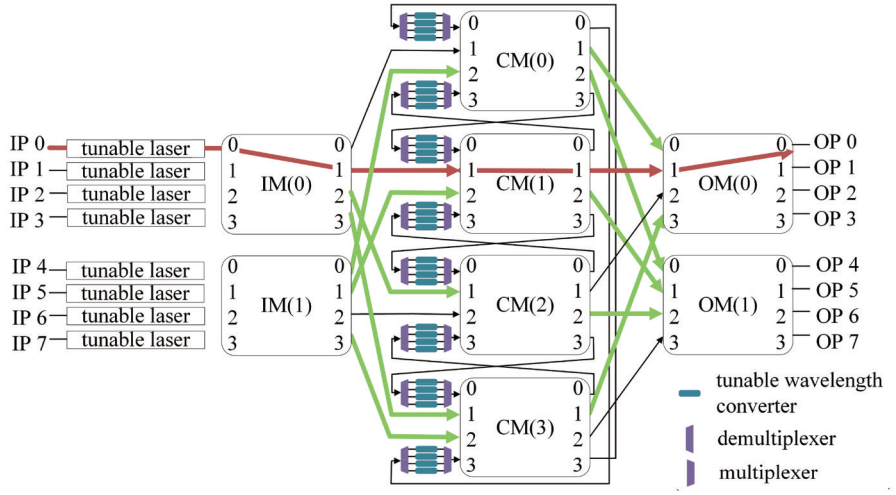

图 4　R(8)架构中光分组 $P(0,0)$ 的路径分配

Fig. 4　Path allocation of optical packet $P(0,0)$ in R(8) architecture

当为光分组 $P(4,4)$ 分配路径时，通过观察 IM(1)和 OM(1)，发现 IM(1)的可用 CM 为 CM(2)，OM(1)的可用 CM 为 CM(3)，即当光分组限制为最多经过 1 跳时，没有可用的 CM。因此，光分组可以通过途径多个 CM 的方式到达目的 OM。例如，光分组可以先后通过 CM(2)与 CM(3)到达目的 OM(1)，如图 5 所示。将这种经过两个 CM 交换单元的路径称为 2 跳。同时，利用波分复用(WDM)技术，CM 级内连接链路可以同时通过多个具有不同波长的光分组，故当第一跳 CM 的输入端口，第二跳 CM 的输出端口可用时，此分组必定不会与其他分组发生竞争。在这种情况下，竞争分别发生在 IM 输出端口与第二跳 CM 输出端口处。

当为光分组 $P(1,1)$ 分配路径时，IM(0)的可用 CM 仅有 CM(0)，而 CM(0)的输出端口 2 被占用，代表其无法通过 0 跳到达目的 OM(0)。同时，CM(1)、CM(3)的输出端口 2 也被占用，代表 $P(1,1)$ 也无法通过 1 跳到达目的 OM(0)。CM(2)具有唯一空闲的输出端口 2，则光分组可以先后经过 CM(0)，CM(3)，CM(2)到达目的 OM(0)，如图 6 所示，并称这种经过三个 CM 交换单元的路径为 3 跳。在这种情况下，分别考虑三个 CM 交换单元的竞争情况：在第一跳 CM 中，竞争只发生在输入端口处；在第二跳 CM 中，从输入端口 3 到输出端口 0 仅能经过一个光分组；在第三跳 CM 中，竞

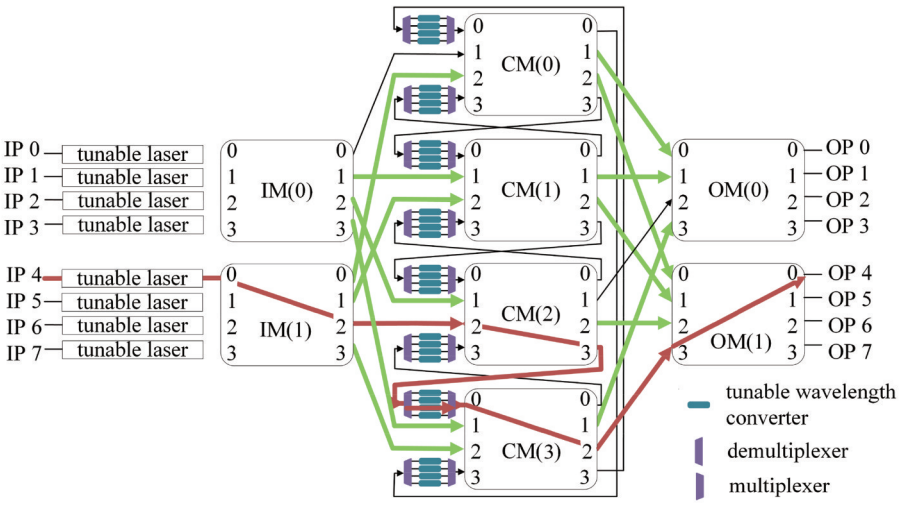

图 5　R(8)架构中光分组 $P(4,4)$的路径分配

Fig. 5　Path allocation of optical packet $P(4,4)$ in R(8) architecture

争仅发生在输出端口处。对比 1 跳与 2 跳的路由路径，3 跳路由路径的竞争条件增加。

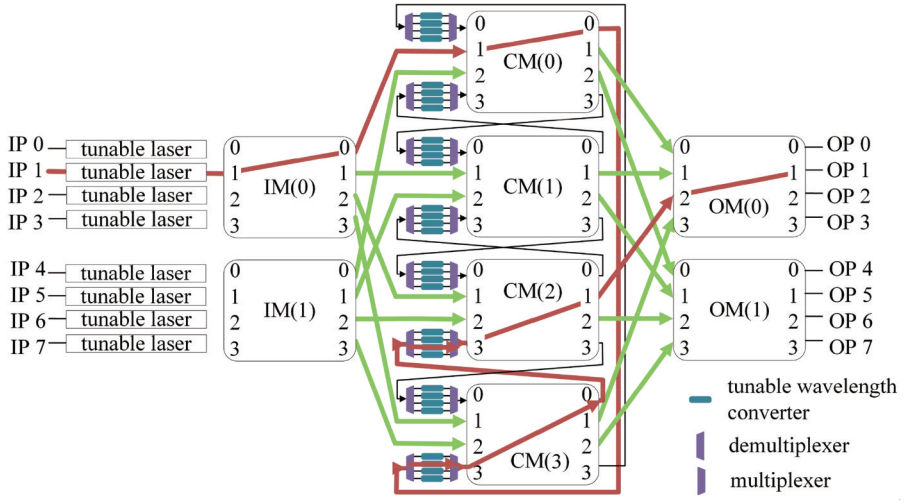

图 6　R(8)架构中光分组 $P(1,1)$的路径分配

Fig. 6　Path allocation of optical packet $P(1,1)$ in R(8) architecture

综上所述，对 Ring-Clos 架构而言，若光分组可以通过无数跳 CM 到达目的 OM，则来自相同 IM 的多个光分组仍可以通过相同 CM 到达目的地址，只是去往 CM 的路径不同。

3　Ring-Clos 调度算法

若不对光分组的跳数作任何限制，则当各个光分组的输入、输出端口均不相同时，此架构的吞吐量可达近 100%。然而，当经过多级级连的 TWC 时，信号的功率与质量降低不可避免，故本文限制光分组最多仅经过 2 跳，若在此条件下，仍然无法找到可行路径，则发生分组丢失。在此前提下，本文提出了一种并行启发式调度算法——并行匹配调度算法(CDRC)，其每个阶段的算法过程分别如图 7～9 所示，每个阶段均迭代 5 次。

3.1　第一个阶段(1 跳)

1 跳的具体过程如下。

1) 每个非空且未在之前几次迭代中匹配的 $\mathrm{IIP}(i,v)$向该 IM 内的所有输出链路发送请求，$\mathrm{IOP}(i,k)$将所有请求发送至对应 CM，其中 $\mathrm{IIP}(i,v)$表示第 i 个 IM 交换单元的第 v 个输入端口的仲裁器，$\mathrm{IOP}(i,k)$表示第 i 个 IM 交换单元的第 k 个输出端口的仲裁器。

2) 当 $\mathrm{CIP}(k,i)$收到请求时，根据请求的目的 OM 将请求发送至对应的 $\mathrm{COP}(k,j)$，其中 $\mathrm{CIP}(k,i)$表示第 k 个 CM 交换单元的第 i 个输入端口的仲裁器，$\mathrm{COP}(k,j)$表示第 k 个 CM 交换单元的第 j 个输出端口的仲裁器。

3) $\mathrm{COP}(k,j)$根据其指针 $P_{\mathrm{COP}}(k,j)$选择一个请求，并向相应 $\mathrm{CIP}(k,i)$返回应答信息。同时，将指针更新至下一个位置。

4) CIP(k,i)根据其指针$P_{\mathrm{CIP}}(k,i)$选择一个应答信息，并发送给其对应IM的IOP(i,k)。

5) IOP(i,k)收到应答信息后，获取此应答信息对应的OM，根据其指针$P_{\mathrm{IOP}}(i,k)$选择一个去往此OM的IIP(i,v)，并发送此应答信息。

6) IIP(i,v)根据其指针$P_{\mathrm{IIP}}(i,v)$选择一个应答信息，发送确认信息，并更新相应的$P_{\mathrm{IIP}}(i,v)$、$P_{\mathrm{IOP}}(i,k)$、$P_{\mathrm{CIP}}(k,i)$指针至下一个位置。

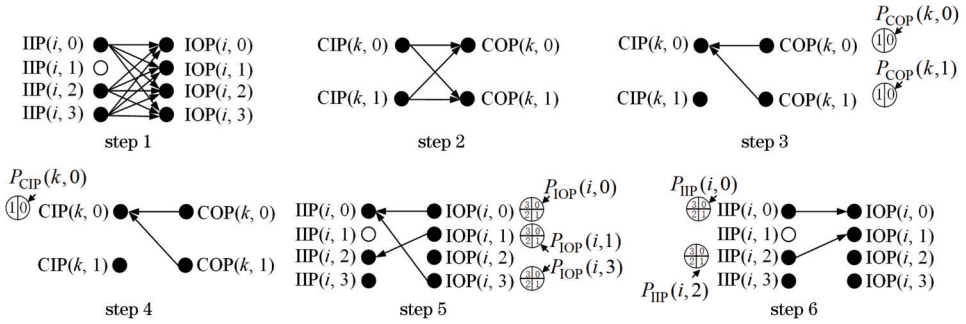

图7　CDRC算法第一个阶段示意图

Fig. 7　Schematic diagram of CDRC algorithm in first stage

3.2　第二个阶段(2跳)

2跳的具体过程如下。

1) 与第一个阶段的步骤1)相同。

2) 与第一个阶段的步骤2)相同。

3) 第一跳CM(t_1)将请求转发至其相邻的上下两个CM的CIN($t_1-1,1$)与CIN($t_1+1,0$)，将这些第二跳CM统称为CM(t_2)，其中CM(t_1)表示第t_1个CM交换单元，亦即第一跳CM交换单元，CIN($t_1-1,1$)表示第t_1-1个CM交换单元上与相邻CM相连的第1个输入端口的仲裁器。

4) 第二跳CM(t_2)的CIN根据请求的目的OM，将请求发送至对应的COP(t_2,j)。

5) 第二跳CM(t_2)的COP(t_2,j)根据其指针$P_{\mathrm{COP}}(t_2,j)$选择一个请求。当多个请求具有相同IM时，选择来自CM$\left[(t_2+1)\bmod(r)\right]$的请求，向对应CIN($t_2,u$)发送应答信息，并将指针更新至下一个位置，其中CIN(t_2,u)表示在第t_2个CM交换单元(第二跳CM交换单元)上第u个输入端口的仲裁器，mod(•)为取余函数，r为CM级交换单元的数量。

6) CIN(t_2,u)向此请求对应的第一跳CM(t_1)返回应答信息。此处注意到，从相同CIN去向不同COP的光分组必定具有不同波长，且多个不同波长的光分组可以同时经过CM级内连接，故CIN可以同时向相邻CM的多个CIP(t_1,i)返回应答。

7) 第一跳CM(t_1)的CIP(t_1,i)根据其指针$P_{\mathrm{CIP}}(t_1,i)$选择一个应答信息。此处需注意的是，当CIP(t_1,i)向上下两个相邻CM的请求均获得了去往相同OM的应答时，选择来自CM$\left[(t_2-1)\bmod(r)\right]$的请求，并发送给其对应IM的IOP($i,t_1$)。

8) 与第一个阶段的步骤5)相同。

9) 与第一个阶段的步骤6)相同。

3.3　第三个阶段(3跳)

3跳的具体过程如下。

1) 与第一个阶段的步骤1)相同。

2) 与第一个阶段的步骤2)相同。

3) 第一跳CM(t_1)将请求转发给第三跳CM$\left[(t_1-2)\bmod(r)\right]$和CM$\left[(t_1+2)\bmod(r)\right]$的CIN。

4) 第三跳CM(t_3)的CIN(t_3,u)根据请求的目的OM，将请求发送至对应的COP(t_3,j)，其中CM(t_3)表示第t_3个CM交换单元，亦即第三跳CM交换单元。

5) 第三跳CM(t_3)的COP(t_3,j)根据其指针$P_{\mathrm{COP}}(t_3,j)$选择一个请求。此处需注意的是，当两个请求具有相同IM时，选择来自CM$\left[(t_3+2)\bmod(r)\right]$的请求，向对应CIN($t_3,u$)发送应答信息，并将指针更新至下一个位置。

6) 第三跳CM(t_3)的CIN(t_3,u)根据其指针$P_{\mathrm{CIN}}(t_3,u)$选择一个应答，并向第一跳CM(t_1)的CIP(t_1,i)返回应答。

7) 第一跳CM(t_1)的CIP(t_1,i)根据其指针$P_{\mathrm{CIP}}(t_1,i)$选择一个应答信息。此处需注意的是，当CIP(t_1,i)向两个第三跳CM的请求均获得了去往相同OM的应答时，选择来自CM$\left[(t_1-2)\bmod(r)\right]$的请求，并发送给其对应IM的IOP($i,t_1$)。

8) 与第二个阶段的步骤8)相同。

9) 与第二个阶段的步骤9)相同。

在上述调度算法中，每个指针代表一个轮询仲裁器。算法的复杂度取决于仲裁器的规模，一个规模为M的轮询仲裁器执行一次仲裁操作的时间复杂度为$O(\log M)$。若CDRC算法在每个阶段均迭代i次，则时间复杂度为$O(i\log M)$。

4　仿真实验和结果

仿真采用OMNeT++软件，一种离散事件网络模拟器，构建了规模为$R(32)$的Ring-Clos光分组交换架构，此交换架构具有960个端口，并采用均匀非重复的伯努利业务源模型对网络进行仿真分析。在该模型下，每个输入端口在每个时隙中以P_{load}的概率发送一个光分组，此光分组去往各输出端口的概率为

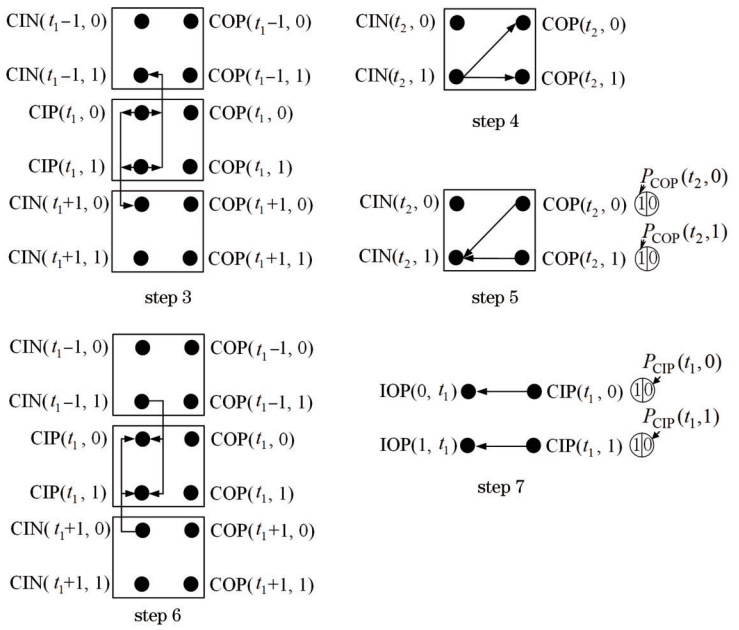

图 8　CDRC 算法第二个阶段示意图

Fig. 8　Schematic diagram of CDRC algorithm in second stage

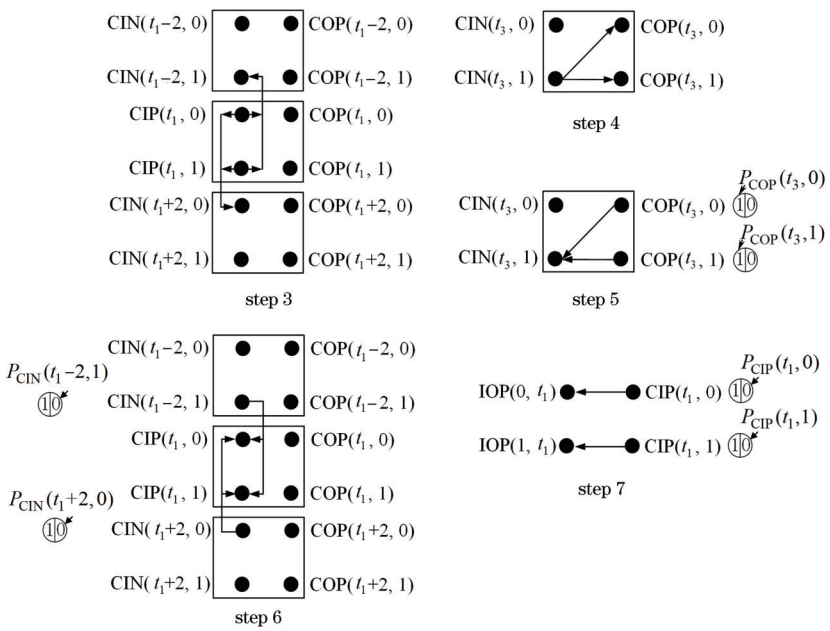

图 9　CDRC 算法第三个阶段示意图

Fig. 9　Schematic diagram of CDRC algorithm in third stage

$P_{load}/960$，且各输入端口的目的输出相异，以排除因输出端口冲突所产生的丢包情况。CDRC 在每个阶段的迭代次数均为 5 次，仿真持续 2×10^6 个时隙，每个时隙长度为 1 μs。同时，在光分组发送的间隔中插入保护时间，使得之前时隙光分组的端口占用情况不影响本组光分组的路由。

将 Ring-Clos 与 Clos 架构作对比，其中最大跳数限制（HL）是指路径经过 CM 交换单元的次数。当 HL 为 1 时，光分组仅能通过一个 CM 交换单元到达目的 OM，此时的 Ring-Clos 架构等同于相同规模的 Clos 架构。若仅采用 CDRC 算法的第一个阶段，则其同样适用于 Clos 架构。从图 10 中可以发现，当 HL 增大时，丢包率大幅降低。在负载等于 0.5 时，Ring-Clos 架构下 HL 为 2 和 3 的丢包率分别为 Clos 架构下 HL 为 1 的丢包率的 1.29% 和 1.10%。在负载等于 1 时，相应比例分别为 62.09% 和 48.81%。可以发现，丢包率在不同负载下均有不同程度的降低，当负载较小时，丢包率降低幅度极大。这是因为负载小时，CM 端口占用概率较低，光分组通过相邻 CM 到达目的 OM 的概率极高。对 Ring-Clos 架构而言，其时延极低，几乎等于在内部链路的传播时延。当跳数增加时，二次路由导致的时延非常小，为光分组经过光纤、TWC 和 AWGR 器

件的传播时延，约为几十纳秒，而光分组的时隙大小为1 μs，相比而言，二次路由的开销非常小，故而不会对架构的性能造成影响。在Ring-Clos架构中，所有光分组经过IM与OM交换单元的时间相同，其区别仅在于经过几跳CM，进而用平均跳数（AH）来代指时延。从仿真结果中可以看出，HL为2和3的平均跳数相较HL为1的平均跳数的增幅不大，分别在4.05%和6.80%。由上述分析可知，与Clos架构相比，Ring-Clos架构在大幅降低丢包率的同时，几乎不影响时延性能。

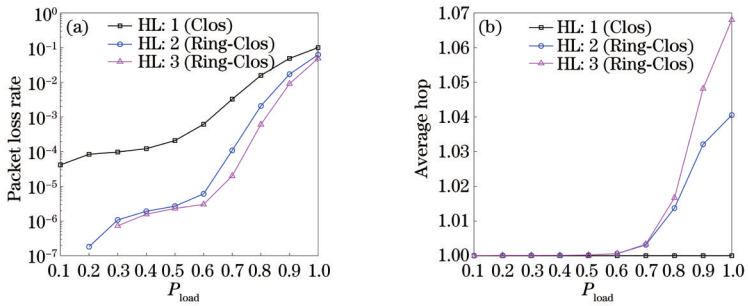

图10　Ring-Clos与Clos的网络性能表现。(a)丢包率；(b)平均跳数

Fig. 10　Network performance of Ring-Clos and Clos. (a) Packet loss rate; (b) average hop

迭代次数越多，算法的复杂度越高，但网络的性能会更佳，故需要在两者之间进行权衡。当HL为3时，对CDRC算法进行2～7次迭代时的性能进行对比分析，如图11所示。仿真结果表明，当迭代次数增加时，丢包率与平均跳数均有降低趋势，且降幅均逐渐减小，这是因为迭代次数增加后，更多请求将有机会获得确认信息，且更容易在先前的阶段获得匹配。同时，待匹配的请求随着迭代次数的增多而越来越少，且这些请求有极大概率是无法被匹配的，所以降幅会越来越小。随着迭代次数的增加，性能收益会越来越小，而迭代次数增加至一定阈值后，几乎所有有可行路径的光分组都已匹配到相应路径，进而此时迭代次数对性能产生的影响非常小。从图11(a)可以发现，迭代次数为5～7次的平均跳数曲线在负载不小于0.7时的丢包率曲线重合。在算法复杂度与性能间进行权衡后，5次迭代为一个较优的选择。

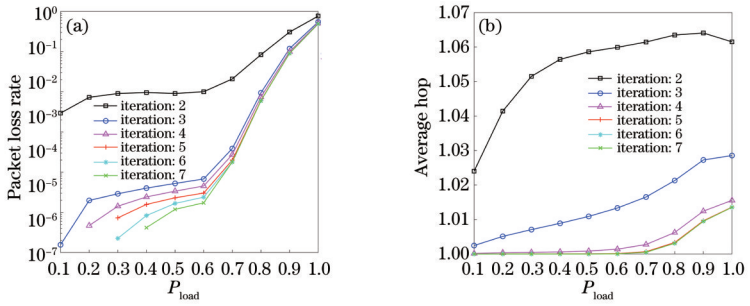

图11　Ring-Clos在不同迭代次数下的网络性能表现。(a)丢包率；(b)平均跳数

Fig. 11　Network performance of Ring-Clos under different iterations. (a) Packet loss rate; (b) average hop

5　结　　论

对Clos架构中产生竞争的原因进行分析，在此基础上提出了一种全新的Ring-Clos架构。通过CM级内连接，光分组可经多跳CM到达目的OM，有效降低了分组争用概率。CDRC算法通过多阶段的多次迭代以$O(i\log M)$的复杂度实现了光分组的路径分配。仿真结果表明，与Clos架构相比，Ring-Clos架构的丢包率大幅降低。所提架构中可以进一步引入缓存或丢包重传机制，以便更好地满足实际应用中的通信质量需求。

参　考　文　献

[1] Guo L, Congdon P. IEEE 802 Nendica report: intelligent lossless data center networks[J]. IEEE SA Industry Connections, 2021: 1-44.

[2] 王斌锋, 苏金树, 陈琳. 云计算数据中心网络设计综述[J]. 计算机研究与发展, 2016, 53(9): 2085-2106.
Wang B F, Su J S, Chen L. Review of the design of data center network for cloud computing[J]. Journal of Computer Research and Development, 2016, 53(9): 2085-2106.

[3] 魏祥麟, 陈鸣, 范建华, 等. 数据中心网络的体系结构[J]. 软件学报, 2013, 24(2): 295-316.
Wei X L, Chen M, Fan J H, et al. Architecture of the data center network[J]. Journal of Software, 2013, 24(2): 295-316.

[4] 李韦萍, 孔淼, 余建军. 基于偏振复用光调制器产生PDM-16QAM射频信号[J]. 光学学报, 2020, 40(23): 2306002.
Li W P, Kong M, Yu J J. Generation of PDM-16QAM

radio frequency signal based on a polarization multiplexing optical modulator[J]. Acta Optica Sinica, 2020, 40(23): 2306002.

[5] Joshi A, Batten C, Kwon Y J, et al. Silicon-photonic clos networks for global on-chip communication[C]//3rd ACM/IEEE International Symposium on Networks-on-Chip, May 10-13, 2009, La Jolla, CA, USA. New York: IEEE Press, 2009: 124-133.

[6] Schröder H, Neitz M, Whalley S, et al. Multi-layer electro-optical circuit board fabrication on large panel [C]//IEEE 66th Electronic Components and Technology Conference, May 31-June 3, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 468-476.

[7] 皮顿, 单子豪, 吴兴坤. 光纤通信波段微光学件的抗反射纳米结构[J]. 光学学报, 2020, 40(6): 0622002.
Pi D, Shan Z H, Wu X K. Nanostructured antireflection micro-optics in the optical fiber communication band[J]. Acta Optica Sinica, 2020, 40(6): 0622002.

[8] Xi K, Kao Y H, Chao H J, et al. A petabit optical switch for data center networks[M]//Kachris C, Bergman K, Tomkos I. Optical interconnects for future data center networks. New York: Springer, 2010: 135-154.

[9] Lea C T. A scalable AWGR-based optical switch[J]. Journal of Lightwave Technology, 2015, 33(22): 4612-4621.

[10] 邓鸿胜, 卢旸, 曹露芳, 等. 基于NRZ+Manchester信号和偏振复用的无源光互连数据中心[J]. 光学学报, 2021, 41(15): 1506001.
Deng H S, Lu Y, Cao L F, et al. Passive optical interconnection data center based on NRZ+Manchester signal and polarization multiplexing[J]. Acta Optica Sinica, 2021, 41(15): 1506001.

[11] Chao H J, Lam C H, Oki E. Broadband packet switching technologies: a practical guide to ATM switches and IP routers[M]. New Jersey: John Wiley & Sons, 2001: 279-335.

[12] Chao H J, Deng K L, Jing Z. A Petabit Photonic Packet Switch (P3S) [C]//IEEE International Conference on Computer Communications, March 30-April 3, 2003, San Francisco, CA, USA. New York: IEEE, 2003: 775-785.

[13] Chao H J, Deng K L, Jing Z G. PetaStar: a petabit photonic packet switch[J]. IEEE Journal on Selected Areas in Communications, 2003, 21(7): 1096-1112.

[14] Chao H J, Liew S Y, Jing Z. A dual-level matching algorithm for 3-stage Clos-network packet switches[C]// 11th Symposium on High Performance Interconnects, August 20-22, 2003, Stanford, CA, USA. New York: IEEE Press, 2003: 38-43.

[15] Oki E, Jing Z G, Rojas-Cessa R, et al. Concurrent round-robin dispatching scheme in a clos-network switch [C]//IEEE International Conference on Communications, June 11-14, 2001, Helsinki, Finland. New York: IEEE Press, 2001: 107-111.

[16] Oki E, Jing Z G, Rojas-Cessa R, et al. Concurrent round-robin-based dispatching schemes for Clos-network switches[J]. IEEE/ACM Transactions on Networking, 2002, 10(6): 830-844.

[17] Pun K, Hamdi M. Static round-robin dispatching schemes for Clos-network switches[C]//Workshop on High Performance Switching and Routing, Merging Optical and IP Technologies, May 29-29, 2002, Kobe, Japan. New York: IEEE Press, 2002: 329-333.