

# Space-Division Multiplexing in Data Center Networks: On Multi-Core Fiber Solutions and Crosstalk-Suppressed Resource Allocation

Hui Yuan, Marija Furdek, Ajmal Muhammad, Arsalan Saljoghei, Lena Wosinska, and Georgios Zervas

**Abstract**—The rapid growth of traffic inside data centers caused by the increasing adoption of cloud services necessitates a scalable and cost-efficient networking infrastructure. Space-division multiplexing (SDM) is considered as a promising solution to overcome the optical network capacity crunch and support cost-effective network capacity scaling. Multi-core fiber (MCF) is regarded as the most feasible and efficient way to realize SDM networks, and its deployment inside data centers seems very likely as the issue of inter-core crosstalk (XT) is not severe over short link spans (<1 km) compared to that in long-haul transmission. However, XT can still have a considerable effect in MCF over short distances, which can limit the transmission reach and in turn the data center's size. XT can be further reduced by bi-directional transmission of optical signals in adjacent MCF cores. This paper evaluates the benefits of MCF-based SDM solutions in terms of maximizing the capacity and spatial efficiency of data center networks. To this end, we present an analytical model for XT in bi-directional normal step-index and trench-assisted MCFs and propose corresponding XT-aware core prioritization schemes. We further develop XT-aware spectrum resource allocation strategies aimed at relieving the complexity of online XT computation. These strategies divide the available spectrum into disjoint bands and incrementally add them to the pool of accessible resources based on the network conditions. Several combinations of core mapping and spectrum resource allocation algorithms are investigated for eight types of homogeneous MCFs comprising 7–61 cores, three different multiplexing schemes, and three data center network topologies with two traffic scenarios. Extensive simulation results show that combining bi-directional transmission in dense core fibers with tailored resource allocation schemes significantly increases the network capacity. Moreover, a multiplexing scheme that combines SDM and WDM can achieve up to 33 times higher link spatial efficiency and up to 300 times greater capacity compared to a WDM solution.

Published by The Optical Society under the terms of the [Creative Commons Attribution 4.0 License](#). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Manuscript received March 30, 2017; revised December 22, 2017; accepted January 3, 2018; published March 5, 2018 (Doc. ID 291823).

H. Yuan (e-mail: h.yuan@ucl.ac.uk), A. Saljoghei, and G. Zervas are with the Optical Networks Group, University College London, London WC1E 7JE, UK.

M. Furdek, A. Muhammad, and L. Wosinska are with KTH Royal Institute of Technology, School of ICT, Stockholm, Sweden.

<https://doi.org/10.1364/JOCN.10.000272>

**Index Terms**—Bi-directional transmission; Data center; Inter-core crosstalk; Link spatial efficiency; Multi-core fiber; Resource allocation algorithm.

## I. INTRODUCTION

The past few decades have witnessed the rapid development of optical communications. In traditional terrestrial networks, wavelength division multiplexing (WDM) and advanced modulation formats were utilized to stretch the capacity limit of single-core single-mode fiber (SMF). However, conventional SMF solutions based on WDM may fall short in satisfying the capacity, spatial efficiency, power consumption, and cost requirements of high-performance data center networks (DCNs) [1,2]. DCNs are considered as the first potential candidate for introducing homogeneous multi-core fiber (MCF), where all the cores in the MCF are made of the same material due to their high capacity requirements, short link spans (<1 km), and the relative ease of network infrastructure implementation (new data centers are deployed as green-field systems and the existing ones are upgraded every three to eight years) without requiring any trenching [1,3,4]. This type of fiber has been identified as the key technology enabler for space-division multiplexing (SDM) systems [5]. Recently, it has been shown that silicon photonic (SiP) on-board transceivers coupled on MCF [6] can be used to support MCF transmission without requiring any fan-in/out or core pitch conversion devices, which can increase the front panel density while offering better panel space management. Moreover, recently showcased optical switches can support purely MCF-based DCN links with SDM switching [7]. Self-homodyne transmission systems and detection methods have shown that MCF-based networks can reduce the complexity, cost, and power consumption of transceiver digital signal processing (DSP) [8]. Based on these advances and by considering the gradual enhancement to the technological road map for SiP on-board transceivers and optical switches, high-core-count MCF-based solutions can be envisioned for DCNs. Our previous work has shown the potential benefits of MCF-based DCNs in terms of cost and power savings compared to

SMF-based DCNs through proper architectural design [9]. However, such networks are also vulnerable to inter-core crosstalk (XT) between optical signals at adjacent cores [10].

To alleviate this problem for long-haul networks, routing and spectrum allocation (RSA) algorithms have been proposed for uni-directional (1di) MCF-based networks that guarantee signal quality by properly allocating spectral resources [11]. Significant reduction in XT, i.e., at least 20 dB, has also been achieved by transmitting optical signals in opposite directions on adjacent cores of the MCF link [12]. However, the benefits of bi-directional (2di) transmission at the network level still remain unexplored. This requires the development of a new analytical model for XT in bi-directional MCF and devising efficient resource allocation techniques using the new model. In short-reach networks inside a DC, the crosstalk suppression due to bi-directional transmission would allow for more densely populated MCF with a smaller core pitch, which can maximize the link spatial efficiency (defined as the capacity divided by the cross-sectional area of MCF measured in bits/s/ $\mu\text{m}^2$ ). Besides, this approach can aid in alleviating the fiber complexity compared to uni-directional transmission per fiber.

This paper aims to develop high-capacity and highly spatially efficient optical DCN solutions by exploiting the benefits offered by bi-directional transmission in MCF [13]. Research has demonstrated the use of directly coupled vertical-cavity surface-emitting lasers (VCSELs) in MCFs to minimize footprint, maximize density, and minimize cost [4]. However, the existing MCFs for terrestrial long-haul networks have sparse core density (e.g., 35–50  $\mu\text{m}$  core pitch) and large cladding diameters [14,15], which limit the spatial efficiency of the SDM system. Therefore, multiple new MCF layouts with high core density (25–30  $\mu\text{m}$  core pitch) are proposed in this paper, which aim to maintain high mechanical reliability for bending [16] and further reduction of front-panel and link density, which can lead to both cost and space savings for DCNs.

To the best of the authors' knowledge, none of XT models capture bi-directional transmission along a fiber. Thus, we develop new wavelength-dependent crosstalk formulations in this paper that consider both bi-directionality and uniform core pitch between adjacent cores for homogeneous MCFs, including normal step-index MCF (SI-MCF) and trench-assisted MCF (TA-MCF). To mitigate XT and reduce the associated computational complexity, RSAs are proposed based on three mechanisms: bi-directional core prioritization, core switching, and spectrum splitting. Bi-directional core priority mapping is developed to alleviate XT through judicious core selection. Core prioritization considers two different strategies denoted as *start1* and *start2* that perform core sequencing by starting from one fiber and two fibers, respectively. Core switching on path links is exploited to mitigate the wavelength continuity constraint [11]. Moreover, to alleviate the computational complexity of spectrum assignment for MCF, which is an NP-hard problem even in SMF based networks [17], spectrum splitting follows two different strategies of defining the part of the spectrum to be checked, denoted as soft- and hard-splitting strategy. Splitting the spectrum limits the search for the portion of the spectrum that satisfies the XT requirements and reduces the execution time for XT calculations.

The proposed RSA algorithms are first evaluated in terms of network utilization, blocking probability (BP, which is the ratio between the number of blocked requests and the total number of requests), and computational time for the SDM-WDM scheme using Spine-Leaf topology. Then, the best algorithm is selected for further network and fiber technology analysis in three topologies (Spine-Leaf, Facebook, and Three-Tier Fat Tree). We consider two types of requests: a) mixed-rate (10, 100, 110, 300 Gb/s) and b) single-rate (300 Gb/s), as well as all three types of multiplexing schemes (WDM, SDM, and SDM-WDM) for capacity and link spatial efficiency analysis. The authors in Ref. [18] state that the crosstalk in intra-DC interconnects could be negligible. However, we find that crosstalk may limit the transmission reach, and the limitation extent depends on the fiber type. Thus, several homogeneous hexagonal SI/TA-MCFs (7-core, 19-core, 37-core, 61-core) are evaluated in terms of capacity and link spatial efficiency. Different XT thresholds that can support different modulation format requirements (i.e., on-off keying, PAM4, etc.) are taken into consideration for the investigations. Simulation results indicate that the bi-directional XT-aware algorithm with core prioritization, soft-spectrum splitting, and core-switching mechanisms performs the best among all investigated algorithms. Moreover, the Spine-Leaf topology outperforms the others in terms of blocking probability. However, the two three-tier topologies are considerably more modular and scalable. In particular, the Facebook topology delivers the highest link spatial efficiency. When different multiplexing techniques are compared (WDM, SDM, and SDM-WDM), SDM networks outperform WDM networks in terms of both capacity and link spatial efficiency. The combined SDM-WDM solution also considerably improves the performance compared to the WDM solution using SMF.

The rest of the paper is structured as follows. A brief overview of the key enabling technologies and approaches used in the paper, including fiber technologies and RSA algorithms, is presented in Section II. Section III presents the new crosstalk model, bi-directional core prioritization, and RSA algorithms. The simulation environments and simulation assumptions are explained in Section IV. The numerical results are analyzed in Section V. Section VI provides the concluding remarks.

## II. ENABLING TECHNOLOGIES AND CONCEPTS

### A. Multi-Core Fiber

Homogenous MCF with a hexagonal layout is commonly used in research experimentation and trials. The basic features of homogenous MCF are that the core pitch (distance) between any two neighboring cores is identical following a triangle lattice and all core structures are identical (Fig. 1) [14,15,19].

1) Inter-Core Crosstalk (XT): This unwanted interference in MCF is generated by the power leak between neighboring cores. As shown in Fig. 2, crosstalk is dominantly generated when signals with the same wavelengths are transmitted in adjacent cores [20–22].

The statistical mean crosstalk of a homogenous MCF per meter, which can also be considered as the power leakage

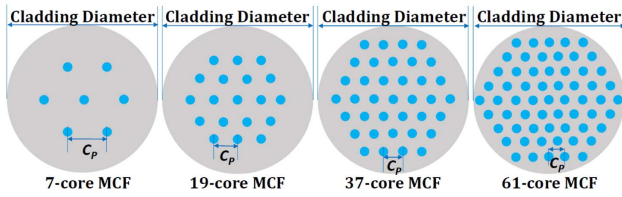


Fig. 1. Examples of homogenous MCFs ( $C_p$  represents core pitch).

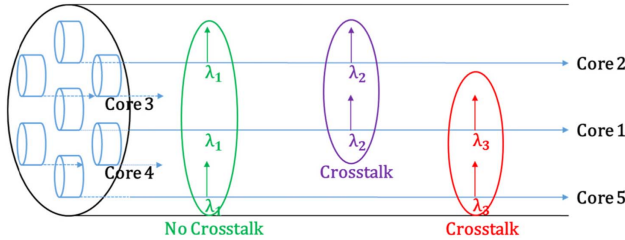


Fig. 2. Example of crosstalk in MCF.

TABLE I  
PARAMETERS USED IN ALL FORMULAS [4]

$n$	Number of adjacent cores
$\kappa$	Coupling coefficient
$\beta(\text{m}^{-1})$	Propagation constant (constant)
$R$ (m)	Bending radius (constant)
$C_p$ (m)	Core pitch
$L$ (m)	Distance of link

from one core to another, is expressed in Eq. (1) [23,24]:

$$h(\kappa, C_p) = 2 * \frac{\kappa^2 * R}{\beta * C_p}. \quad (1)$$

The parameters used in Eqs. (1) and (2) are shown in Table I.

Furthermore, by considering the coupled-power theory, the crosstalk of a homogeneous MCF can be expressed as in Eq. (2), which is used in uni-directional transmission [4]:

$$\text{XT} = \frac{n - ne^{-(n+1)*2*h*L}}{1 + ne^{-(n+1)*2*h*L}}. \quad (2)$$

The numerator and denominator represent the signal power of the neighboring cores and the target core, respectively, with the consideration of the energy leakage among them [25].

### B. Resource Allocation Algorithms

A wide variety of RSA algorithms exist for SMF-based long-haul networks. Recently, a few studies proposed resource allocation schemes that utilize the additional flexibility offered by the spatial domain [11,22,26]. The introduction of a spatial dimension increases the complexity of the RSA process [27,28] as additional constraints such as XT need to be considered while allocating network resources. Existing uni-directional RSA mechanisms apply different mechanisms to address these constraints.

Core prioritization was proposed as a policy to reduce the crosstalk between adjacent cores by predefining the sequence of core usage for uni-directional transmission [22,29]. The principle of core switching is employed in an effort to alleviate the impact of spectrum continuity constraint by allowing connections to use different cores on each link along the path while still using the same wavelength, thus increasing the freedom in frequency-slot allocation as opposed to limiting each connection to always use the same core [11]. The spectrum contiguity constraint enforces the assignment of contiguous spectrum slots to each request [28], which reduces the flexibility of spectrum-slot selection and may lead to solutions with higher XT. Such increases of XT can have a significant effect on requests using higher-order modulation formats due to their sensitivity to distortion and noise. This constraint can be alleviated using the Slot Split Algorithm, which divides a request requiring a large chunk of spectral bandwidth into several requests with smaller spectral bandwidth requirements. The optical bandwidth assigned to each split request should not be smaller than the minimum allowable bandwidth allocation in the network [30].

### III. PROPOSED CROSSTALK FORMULATIONS AND ALGORITHMS

In this section, we formulate a new wavelength-dependent crosstalk model for bi-directional transmission in both SI-MCFs and TA-MCFs. We then present our XT-aware RSA algorithms for all bi-directional MCFs. We define two new core prioritization strategies for bi-directional transmission as well as a spectrum-splitting approach aimed at speeding up our XT-aware RSA approach.

#### A. Crosstalk Formulation for Bi-Directional Homogeneous SI-MCF With Uniform Core Pitch

In order to model the XT suppression ( $\Delta\text{XT}_{-dB}$ ) in bi-directional transmission, a power reduction coefficient  $P_r$  is defined to account for the crosstalk contribution from signals propagating in the opposite direction along the neighboring cores ( $P_r = 10^{(-\Delta\text{XT}_{-dB}/10)}$ ). This contribution can be modeled by multiplying the numerator in the existing XT formula from Eq. (2) by  $P_r$ . The resulting formula is shown in Eq. (3), and  $P_r$  can be calculated by Eq. (4) [12]:

$$\text{XT} = \frac{P_r * (n - ne^{-(n+1)*2*h*L})}{1 + ne^{-(n+1)*2*h*L}}, \quad (3)$$

$$P_r = \frac{S\alpha_R}{2\alpha} \left[ \frac{e^{\alpha L} - e^{-\alpha L}}{\alpha L} - 2e^{-\alpha L} \right]. \quad (4)$$

In the above equations,  $n$  represents the number of adjacent cores that carry signals in the opposite direction from the considered core.  $S$ ,  $\alpha_R$ , and  $\alpha$  stand for the recapture factor of the Rayleigh scattering component into the backward direction, attenuation coefficient results from Rayleigh scattering, and the fiber attenuation coefficient, respectively. As more than one core in an MCF can be assigned

to carry signals in any direction, a generalized form of the equation is necessary to consider crosstalk from adjacent cores carrying data in the same [Eq. (5)] and the opposite directions [Eq. (6)]:

$$\text{XT}_{\text{same}} = \frac{n_1 - n_1 e^{-(n+1)*2*h*L}}{1 + n e^{-(n+1)*2*h*L}}, \quad (5)$$

$$\text{XT}_{\text{opposite}} = \frac{P_r * (n_2 - n_2 e^{-(n+1)*2*h*L})}{1 + n e^{-(n+1)*2*h*L}}. \quad (6)$$

In the equations,  $n_1$  and  $n_2$  represent the number of adjacent cores carrying signals in the same and the opposite directions, respectively ( $n = n_1 + n_2$ ). Ultimately, Eq. (7) can be derived for bi-directional MCFs:

$$\begin{aligned} \text{XT}_{\text{hex}} &= \text{XT}_{\text{same}} + \text{XT}_{\text{opposite}} \\ &= \frac{n_1 - n_1 e^{-(n+1)*2*h*L} + P_r * n_2 - P_r * n_2 e^{-(n+1)*2*h*L}}{1 + n_1 e^{-(n+1)*2*h*L} + n_2 e^{-(n+1)*2*h*L}}. \end{aligned} \quad (7)$$

For a constant core pitch and number of adjacent cores, the crosstalk is proportional to the link distance. This indicates that crosstalk reduction due to bi-directionality can increase the supported fiber length.

### B. Wavelength-Dependent Crosstalk Formulations for SI-MCF

The transmission wavelengths for signals in two adjacent cores may affect the level of XT between them [31]. This additional wavelength-dependent XT contribution increases for wavelengths at the lower end of the C-band ( $\lambda_0 = 1530$  nm) and can be expressed as

$$\begin{aligned} \Delta \text{XT}_{dB} &= 10 * \log_{10}(1 - 0.001256 * \Delta\lambda)^4 \\ &+ 19.85\pi * r_1 * \sqrt{2\Delta_1} * \frac{\Delta\lambda * C_P}{\lambda * \lambda_0}. \end{aligned} \quad (8)$$

In the equation,  $\Delta\lambda = \lambda - \lambda_0$ , where  $\lambda$  represents the transmission wavelength, while  $r_0$ ,  $r_1$ , and  $\Delta_1$  stand for the refractive coefficient of cladding, the refractive coefficient of core, and the refractive coefficient difference between core and cladding, respectively [31]. The additional crosstalk  $\Delta \text{XT}_{dB}$  can be transformed into a power coefficient  $P_i$  ( $P_i = 10^{(\Delta \text{XT}_{dB}/10)}$ ). The  $P_r$  for bi-directional transmission also needs to be modified to account for the wavelength-dependent XT. The modified power coefficient  $P'_r$  can be expressed as

$$P'_r = 10^{\frac{\Delta \text{XT}_{dB} - \Delta \text{XT}_{dB}}{10}}. \quad (9)$$

Thus, after inserting  $P_i$  and  $P'_r$  in Eq. (7), the XT in bi-directional hexagonal MCF can be calculated using Eq. (10):

$$\begin{aligned} \text{XT}_{\text{wd-hex}} &= \frac{P_i * n_1 - P_i * n_1 e^{-(n+1)*2*h*L} + P'_r * n_2 - P'_r * n_2 e^{-(n+1)*2*h*L}}{1 + P_i * n_1 e^{-(n+1)*2*h*L} + P'_r * n_2 e^{-(n+1)*2*h*L}}. \end{aligned} \quad (10)$$

### C. Wavelength-Dependent Crosstalk Formulation for Bi-Directional Homogeneous TA-MCF With Uniform Core Pitch

The inherent features of TA-MCF facilitate lower XT levels than the normal SI-MCFs. According to Ref. [31], the mean crosstalk between two adjacent cores of TA-MCF can be expressed as

$$h'(\kappa, C_P) = h * \frac{W_1}{\left[ W_1 + (W_2 - W_1) * \frac{w_t}{C_P} \right]} * e^{[-4(W_2 - W_1) * w_t/a]}, \quad (11)$$

where  $W_1 = 1.1428V_1 - 0.996$ , and  $W_2 = (V_2^2 + W_1^2)^{1/2}$ .  $V_1(1.5 \sim 2.5)$  denotes the V number determining the modes propagating in a fiber, and  $V_2 = 2\pi a_0 \sqrt{2|\Delta_2|}/\lambda$ , where  $\Delta_2$  is the refractive index difference between trench and cladding. Moreover,  $w_t$  and  $a$  are the trench width and the core radius, respectively [31]. Thus, by the same principles that were presented in the previous, the following equation is derived for determining the XT in bi-directional TA-MCFs:

$$\begin{aligned} \text{XT}_{\text{TA-wd-hex}} &= \frac{P_i * n_1 - P_i * n_1 e^{-(n+1)*2*h'*L} + P'_r * n_2 - P'_r * n_2 e^{-(n+1)*2*h'*L}}{1 + P_i * n_1 e^{-(n+1)*2*h'*L} + P'_r * n_2 e^{-(n+1)*2*h'*L}}. \end{aligned} \quad (12)$$

In this paper, Eqs. (10) and (12) are used to calculate the XT for the links of the considered networks for all MCF types. The total XT of a path with several links connecting the source and destination nodes in DCN can be calculated using the following equation:

$$\text{XT}_{\text{path}} = \sum_{i=1}^{i=LN} \text{XT}_{\text{link}i}, \quad \text{LN: link number.} \quad (13)$$

### D. Proposed Resource Allocation Algorithms

1) *Bi-Directional Core Priority Mapping*: The bi-directional core priority policy reduces the crosstalk by defining a core usage sequence that avoids assigning contiguous blocks of adjacent cores for transmission in the same direction. The set core priorities also need to stimulate fairness, i.e., support equal capability of communication in each direction. We propose two strategies for core prioritization in a bi-directional link model with two fibers per link. The first one, denoted as *start1*, begins by using only a single fiber in the pair (Fig. 3), while the other one, denoted as *start2*, considers both fibers (Fig. 4). Figure 3 illustrates an example of the core priority map obtained by core prioritization process for a 7-core hexagonal MCF pair, where the propagation direction of each core is pre-assigned to ensure that each core transmits (optical signals) in the direction opposite to its neighboring cores, i.e., the green (dotted) cores carry signals in the direction opposite to the orange (solid) cores. Cores in one direction are assigned priorities independently of the other

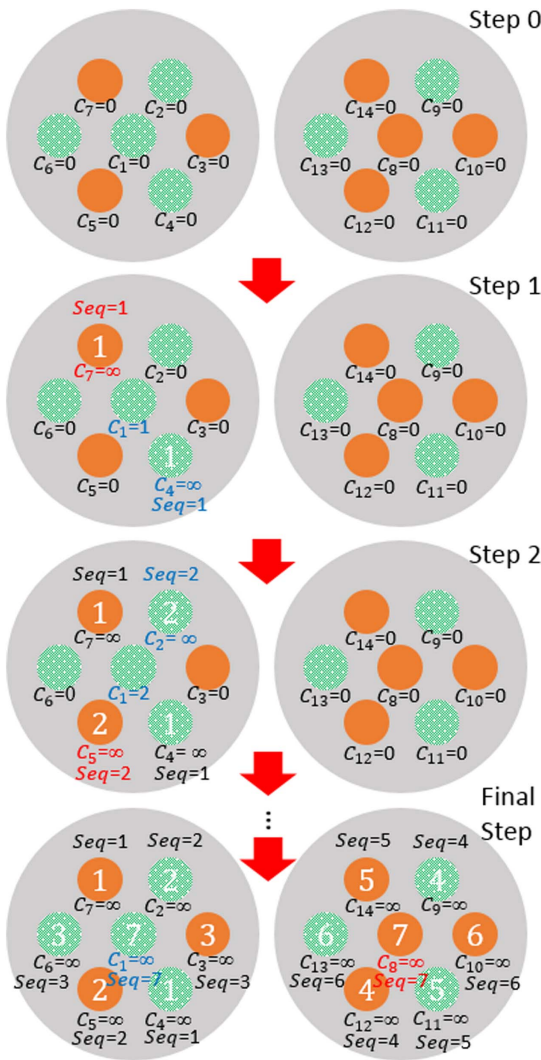


Fig. 3. Priority mapping starting from one MCF (*start1*).

direction. The numbers inside the core are the core sequence numbers (*Seq*) in each direction set by the process, while *i* is the core index, and  $C_i$  is the core cost computed during the process, initialized to 0.

In each step, the algorithm selects the core with the lowest cost as the next core in the priority sequence. The cost of each unnumbered core will increase when an adjacent core in same direction is assigned priority. Based on the aforementioned rules, any core other than the central one can be

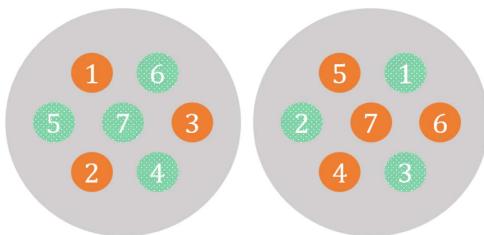


Fig. 4. Priority mapping starting from different MCFs (*start2*).

chosen as the first core for either direction at step 1. In the example shown in Fig. 3, the orange core with index 7 and the green core with index 4 are selected first for each propagation direction. Consequently, their costs are set to infinity ( $C_7, C_4 = \infty$ ) to avoid reassigning their priorities. Simultaneously, the cost of their adjacent cores increases. As there are no orange cores adjacent to the core with index 7, only the cost of the green core with index 1 ( $C_1$ ) is incremented. In step 2, the orange core with index 5 and green core with index 2 have the lowest costs in the same fiber and are thus assigned the second-highest priorities in each direction. In this step,  $C_1$  is incremented again. If two cores in different fibers have the same cost, the core located in the same fiber as the core with higher priority is selected first. The process starts considering the other fiber when the cost of its cores becomes the lowest or if all cores in the first fiber are processed. Subsequently, both green and orange cores 3, 4, 5, 6, and 7 will be added to the priority sequences. The complete core priority map for the *start1* strategy is shown in the final step of Fig. 3. In this case, the fairness is ensured, as the same number of cores is assigned in both directions.

In the *start1* strategy (Fig. 3), transmission in both directions begins from the cores in the same fiber. In the alternative, *start2* strategy, the cores with the highest priority are located in different fibers, as shown in Fig. 4. In practice, the *start1* approach provides modularity and scalability, as the independence of core sequence among fibers enables increased capacity by directly adding new fibers. The *start2* approach may offer better performance since there is no XT at lower traffic loads, but may require increased capital expenditures (CAPEX).

2) *Spectrum-Splitting Scheme*: To alleviate XT, our RSA algorithm needs to evaluate XT for each new request using Eq. (9), which may be computationally demanding. In an effort to relieve this complexity, we apply a spectrum-splitting scheme that divides the whole spectrum (i.e., the C-band) into two bands. This allows for a reduction of algorithm computation time since it scans only half of the resource pool at a time in search of free slots. Moreover, arranging the bands on adjacent cores in a non-overlapping fashion can further mitigate crosstalk. The proposed algorithm is fully compatible with the bi-directional model. Two spectrum-splitting approaches are considered: soft split and hard split. The main steps of the soft-split scheme for the SDM-WDM case for a given core priority map and with core switching enabled are as follows:

- i. Divide the spectrum (100 slots) into two divisions of 50 slots. For each core, the first 50 slots are denoted as Division 1 (D1) and the other half as Division 2 (D2).
- ii. According to the priority map, define the initial allocation division for each core in both directions. The general defining rule is shown in Table II.

In Table II, *W* equals the total number of cores in each direction. According to the core priority map, Core 1 to Core *V* are not adjacent in each direction, and they will be allocated first. On the other hand, Core (*V* + 1) to Core *W* are adjacent to previous cores, and they will be utilized later.

TABLE II  
PREDEFINED SPECTRUM DIVISION FOR BI-DIRECTIONAL MODEL

Defined Sequence (One Direction)	First Division for Request Allocation	Defined Sequence (Opposite Direction)	First Division for Request Allocation
Core 1	D1	Core 1	D2
Core 2	D1	Core 2	D2
...	...	...	...
Core V	D1	Core V	D2
Core (V + 1)	D2	Core (V + 1)	D1
...	...	...	...
Core W	D2	Core W	D1

Figure 5 shows an example of the final map for the 19-core MCF. The green (dotted) cores represent one direction, and the orange (solid) ones the other. The numbers inside the cores denote the core usage sequence per direction, where the black digits imply that spectrum division D1 is used first, and the white ones mean that D2 is used first. Thus,  $V = 13$  and  $W = 19$  in this case.

iii. For each connection request, the following steps are applied: 1) check the predefined division slot by slot and test crosstalk when enough available slots for the request are found, 2) move to the next core in the priority map if there are no sufficient slots in the current core and current division, and 3) repeat the previous steps until suitable slots are found or all cores and all slots in the current division have been checked. This process is depicted in Fig. 6, where the numbers inside the slots are the spectrum slot indices. The green slots carry signals in one direction and the orange ones in the other. The white slots represent the unused spectrums.

iv. If there are not enough slots available on all the cores for either direction in the original division, then the two divisions (in all cores) will be swapped, and the algorithm starts checking the other division. In the example above, it means that the orange cores 1 to V are allowed to scan and utilize the slots of D2, while the requests in green core 1 to V can use D1. As shown in Fig. 7, the unused slots in D2 of the first orange core are the next division to be scanned and utilized after the original division (D2) of the last orange core is fully occupied and the divisions swapped.

v. When the next request comes, the Soft-Split Algorithm will still check the spectrum slots from the original division and switch to the other division if no

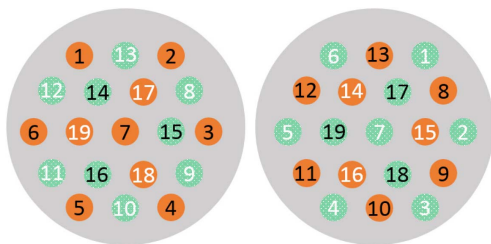


Fig. 5. Core priority map with defined spectrum division for 19-core MCF.

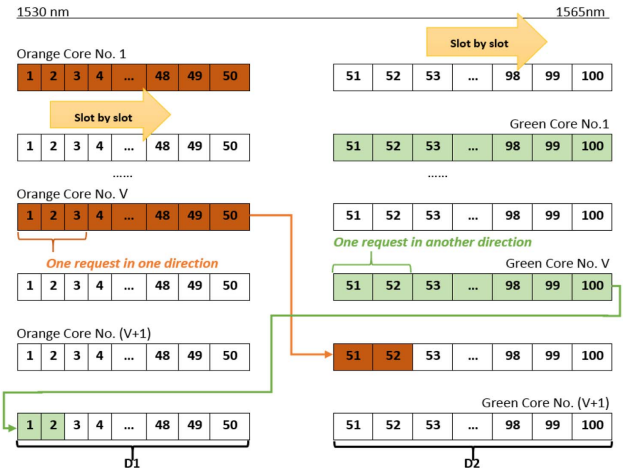


Fig. 6. Resource checking in original division based on priority map.

sufficient slots are available. If neither of the two divisions can provide sufficient slots, then the request is blocked.

Unlike the soft-split approach, the hard-split approach will block the current request directly when there are no sufficient slots available in the original division without checking the slots in the other division. It will also update the BP value upon processing each request. Once the BP in the original division reaches a predefined threshold (e.g., 10%, 1%, or 0.1%), the two divisions will be swapped permanently for the following requests. Obviously, a higher BP threshold means that the spectrum slots in the original division are more likely to be fully occupied before the divisions swap permanently. In other words, more spectrum slots in the first division could be utilized compared to the case with a lower threshold. As a consequence, a higher threshold may contribute to a reduction of the final overall BP and an increase in resource utilization at the expense of temporary intermediate blocking. The flow chart of the hard-spectrum-split scheme is shown in Fig. 8.

The soft split seems more flexible, and the first blocking of requests occur only after the whole spectrum has been checked. However, this principle increases computational complexity and processing time. In other words, when the original division is fully occupied, for the next coming

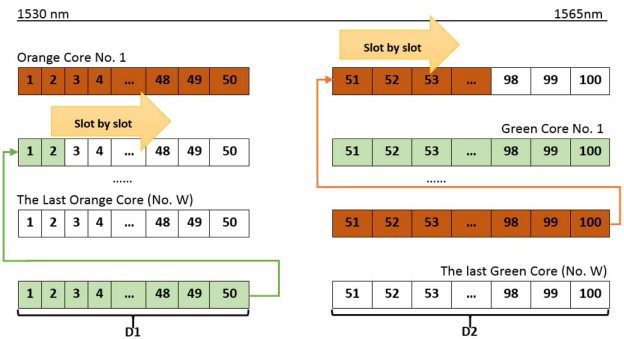


Fig. 7. Spectrum division swap after a lack of slots in the original division has occurred.

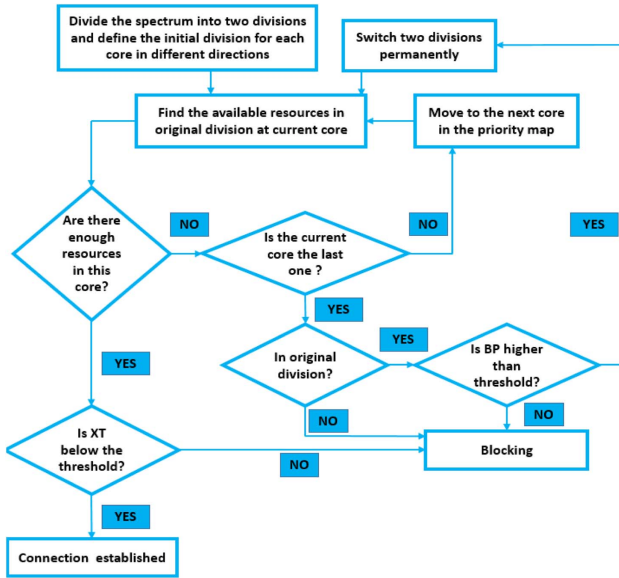


Fig. 8. Flowchart of the hard-spectrum-split approach.

requests, the system still spends time on checking the original division before attempting to look for slots in the other. As for the hard split, though it will block requests early on, the system will not check the original division after the BP threshold in the original division (e.g., 10%) has been reached. For all subsequent requests, only one division will be searched, which translates to a shorter execution time than for the soft split. In other words, the soft-split scheme can provide lower BP at the cost of increased time, while the hard-split scheme can provide better processing efficiency and lower total BP (up to a certain network utilization level) but may reject more requests under low network utilization.

The pseudo-code of the XT-aware RSA algorithm with core prioritization, spectrum splitting, and core switching is shown in Table III.

TABLE III  
SYMBOLS USED

Symbol	Description
LN	Number of links for end-to-end path
$S_i$	Source/ingress node of $i$ th link
$T_i$	Destination/egress node of $i$ th link
BW	Required bandwidth for the connection request
D	Spectrum division
D1	Division 1 of available spectrum resources (first 50 slots; 1-50)
D2	Division 2 of available spectrum resources (last 50 slots; 51-100)
CPM	Core priority map (specifying usage sequence)
CPM <sub>1</sub>	Core priority map for direction 1
CPM <sub>2</sub>	Core priority map for direction 2
SA	Available spectrum slots
XT	Total crosstalk for the current request (initially 0)
XT <sub><math>i</math></sub>	Crosstalk value for the current request on the $i$ th link (XT <sub><math>i</math></sub> = -1 indicates spectrum slot unavailability)

(Table continued)

Symbol	Description
CS	Status of division switch (different between soft and hard); 1 means true and 0 means false
$t$	Indicator of the division being checked (0 is initialized and 1 indicates division switched)
W	Total number of cores
V	Index of the highest used core without adjacent cores carrying signals in the same direction

---

Procedure of generic spectrum split for each request

**Input:** Routing result (LN,  $S_i$ ,  $T_i$ ,  $i = 1, \dots, LN$ ), core priority map (CPM<sub>1</sub> & CPM<sub>2</sub>)

```

1  for i = 1 to LN
2  XT $i$  ← 0, t ← 0 # initialization of parameter
3  if  $S_i$  less than  $T_i$  in the node index value
4  D ← D1, CPM ← CPM1
5  else
6  D ← D2, CPM ← CPM2
7  endif
8  for k = 1 to W
9  if k == V + 1 && D == D1
10 D ← D2
11 endif
12 if k == V + 1 && D == D2
13 D ← D1
14 endif # define the division for allocation
15 Check the whole D in kth core to find SA
16 if SA == BW
17 Calculate XT $i$ 
18 break
19 else
20 XT $i$  ← -1
21 endif
22 endfor # get the crosstalk for the current request
23 if XT $i$  == -1 # spectrum resources are not available
24 on current D
25 if t == 0 && CS == 1
26 t ← 1
27 if D == D1
28 D ← D2
29 else
30 D ← D1
31 endif # division swapping
32 Turn to line 8 # search spectrum slots on the new D
33 endif
34 Block the request # blocked due to resource
35 unavailability
36 break
37 elseif XT $i$  != -1 && XT $i$  >= threshold # Check the
38 crosstalk for current link
39 Block the request # blocked due to high XT
40 break
41 else # XT $i$  is lower than threshold
42 XT = XT + XT $i$  # Calculate the total crosstalk of
43 link 1 to link i
44 if XT >= threshold
45 Block the request # blocked due to high XT
46 break
47 else # XT is lower than threshold
48 if i == LN # the crosstalk for the whole path (LN
49 links) has been checked
50 Connection established
51 endif
52 endif
53 endif
54 endfor

```

## IV. SIMULATION ENVIRONMENT AND ASSUMPTIONS

## A. Topologies Used for Simulations

Unlike the conventional long-haul optical networks, data center networks operate over a short range, with link span generally less than 1 km. One of the common investigated indirect topologies for DCNs is Spine-Leaf, which consists of the top of rack (ToR) or leaf nodes and spine nodes. It provides both scalability and flexibility since the number of paths can be increased by adding more spine nodes, which can also facilitate connectivity to other parts of the DCN. In fact, it can provide almost an ideal DCN with a large non-blocking switch where all servers are directly connected [32].

In this paper, we use Spine-Leaf, shown in Fig. 9(a), as a small-size topology for algorithms and network analysis to benchmark a range of algorithms for test cases reported in Subsections V.A and V.B. The most promising scheme is then tested for another two three-tier topologies, i.e., Facebook Data Center topology [Fig. 9(b)] and Three-Tier Fat Tree topology [Fig. 9(c)], and the extensive analysis results are reported in Subsection V.C. We assume all topologies support 20 racks and each server is interconnected to the ToR with a single channel (in the SDM or the WDM scheme) or a single core carrying several channels (in the SDM-WDM scheme). We consider two fibers per link with different lengths (per link) for evaluation purposes.

## B. Simulation Setup

In order to investigate the performance of the proposed algorithms for the three interconnect topologies, a simulator in Matlab is created. It supports the SDM-WDM network by implementing the homogenous hexagonal (7-core, 19-core, 37-core, and 61-core) MCFs with 100 of 25 GHz spectrum slots in the C-band (WDM), which can be realized by either a passive arrayed-waveguide grating (AWG) [33] or an active but more flexible bandwidth-variable wavelength selective switch (BV-WSS) [34].

Figure 10 displays the overview of the simulation process divided into five main steps. The XT-aware RSA algorithms simulated in step 3 are listed in Table IV with a summary of mechanisms they apply. All algorithms consider core

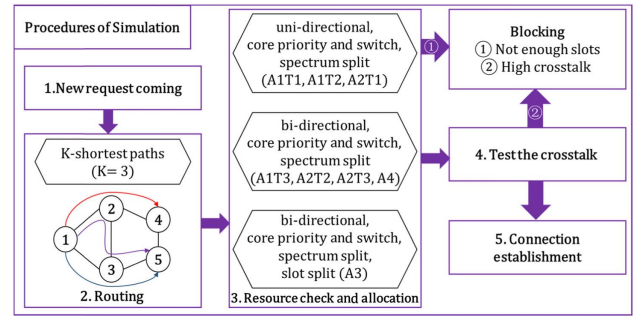


Fig. 10. Procedure of simulation for each request.

TABLE IV  
ALGORITHMS USED IN THE SIMULATIONS<sup>a</sup>

Algorithm and Type	Direction (1di/2di)	Core Priority ( <i>start1/start2</i> )	Spectrum Split (Soft/Hard/N)	Slot Split (Y/N)
A1T1	1di	<i>Start1</i>	N	N
A1T2	1di	<i>Start2</i>	N	N
A1T3	2di	<i>Start1</i>	N	N
A2T1	1di	<i>Start2</i>	Soft	N
A2T2	2di	<i>Start1</i>	Soft	N
A2T3	2di	<i>Start2</i>	Soft	N
A3	2di	<i>Start2</i>	Soft	Y
A4	2di	<i>Start2</i>	Hard	N

<sup>a</sup>A/T, Algorithm/Type; Y, with; N, without.

switching, and each of them is run in independent simulation sequences. Algorithm 1-Type 1 (A1T1) is from [29], and it considers uni-directional transmission with core prioritization using the *start1* scheme. It is used as the benchmarking approach for the approaches proposed in this paper. Compared to A1T1, A1T2 utilizes the *start2* scheme, while A2T1 combines the soft-spectrum-split scheme presented in Section III with A1T2. A1T3 considers the bi-directional transmission with the proposed core prioritization using the *start1* scheme. A2T2 adds the soft-spectrum-split scheme to A1T3, while A2T3 replaces the *start2* scheme with the *start1* scheme in A2T2. The combination of A2T3 and the slot-split scheme, which has been described in Subsection II.B, is denoted as A3. A4 is the only one that considers the hard-spectrum-split scheme. During this process, after checking the  $K$  ( $K = 3$ ) shortest paths individually, the first path that satisfies the spectrum resource and crosstalk level (below the threshold) requirements of a request will be selected for

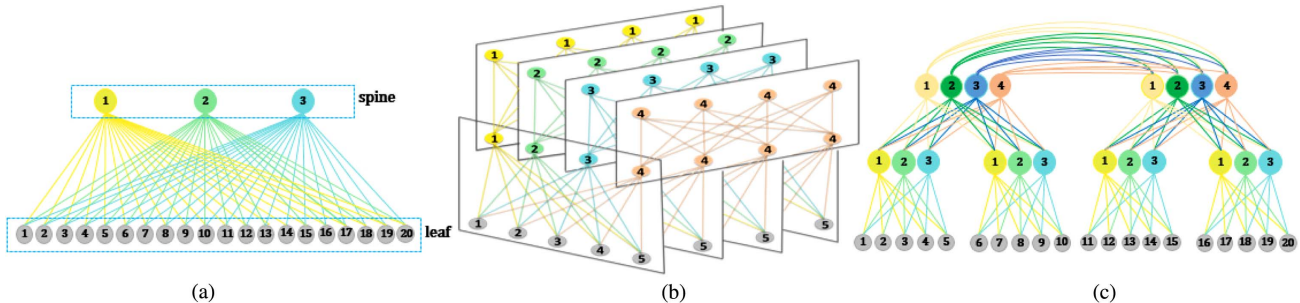


Fig. 9. Topologies used in the simulations: (a) Spine-Leaf topology, (b) Facebook Data Center topology, and (c) Three-Tier Fat Tree topology.



connection provisioning. If none of the  $K$  paths meets the requirements, the request will be blocked.

### C. Traffic Characteristics

In data centers, the traffic characteristics vary depending on the specific applications being run on the network. For instance, the traffic for a DC network that supports video streaming will be different from a DC network used by a research institution for carrying out complex computations. Moreover, apart from a few theoretical studies analyzing DCN traffic at the packet level in Refs. [2,35], there is a lack of literature on intra-DC traffic from DC operators, making it difficult to model such traffic. Therefore, in this work that focuses on optical circuit switched (OCS) networks, we assume lightpath requests associated with virtual machines (VMs)/virtual tenant interconnection arrive following a Poisson distribution, with an average inter-arrival time of 10 time units. A similar approach has been considered in other related works, such as in Refs. [36,37]. We consider an incremental traffic scenario where each request has a holding time of 200,000 time units. This study allows us to assess the impact of different existing and new fiber types as well as new allocation schemes on the attainable core packing density and the achievable capacity and reach.

We consider two types of requests with the assumed modulation and multiplexing schemes shown in Table V. Uniform distribution of the number of requested frequency slots (bandwidth) is assumed for the first type requests. Different bandwidth corresponds to different data rates. This scenario can apply to a DCN that goes through phased migration, resulting in some servers having 10 Gb/s transceivers and others with either 100 Gb/s or 300 Gb/s. Particularly, the case study of 110 Gb/s reflects the sum of 10 Gb/s and 100 Gb/s client rates in two channels to represent DCN evolution where multiple rates could coexist. The second type refers to a maximum-capacity single-rate green-field DCN realization that has only 300 Gb/s transceivers. The XT threshold for each type of modulation corresponds to the XT power at which it induces 1 dB of penalty at bit error rate (BER) of  $10^{-3}$  [44,45]. Based on Ref. [46], we assume a 6 dB lower threshold for the PAM8 format compared to the PAM4 scheme.

The simulations also consider three multiplexing cases: WDM, SDM, and their combination. WDM assumes SMF

links with either AWGs or WSSs. SDM considers MCF links and fiber switches. SDM-WDM can be supported by MCF links, AWG/WSS, and fiber switches.

### D. Fiber Type Characteristics

The assumptions on characteristics of different fiber types are listed in Table VI, while Table VII lists parameters used for XT calculation. To the best of our knowledge, none of the existing studies provide the value of  $\kappa$  for a 25  $\mu\text{m}$  core pitch. Thus, we predict it by interpolation as shown in Fig. 11.

### E. Crosstalk Reduction Due to Bi-Directional Transmission

The exploited bi-directional scheme can offer around 20 dB crosstalk suppression per core pair based on

TABLE VI  
MCF PARAMETERS

Type of MCF	Core Pitch ( $\mu\text{m}$ )	Fiber Diameter ( $\mu\text{m}$ )	Area ( $\mu\text{m}^2$ )
7-core	30	140	15,393.80
19-core	30	200	31,415.92
37-core	30	260	53,092.91
61-core	25	260	53,092.91
SMF [4]		125	12,271.85

TABLE VII  
PARAMETERS AND VALUES USED FOR XT CALCULATION  
[4,16,31,44]

Symbol	Description (Core Pitch)	Value (Unit)
$\kappa_1$	Coupling coefficient (40 $\mu\text{m}$ )	$4 * 10^{-4}$
$\kappa_2$	Coupling coefficient (30 $\mu\text{m}$ )	$6 * 10^{-2}$
$\kappa_3$	Coupling coefficient (25 $\mu\text{m}$ )	$7 * 10^{-1}$
$\beta$	Propagation constant (constant)	$4 * 10^6 (\text{m}^{-1})$
$R$	Bending radius (constant)	$50 * 10^{-3} (\text{m})$
$r_0$	Refractive coefficient of cladding	1.45
$\Delta_1, \Delta_2$	Refractive coefficient differences	0.35, -0.35 (%)
$w_t/a$	Trench width/core radius	1
$\lambda$	Transmission wavelength	1530–1570 (nm)

TABLE V  
ASSUMED MODULATION AND MULTIPLEXING SCHEMES OF THE REQUESTS

Bandwidth	Capacity (Gb/s)	Possible Modulation	Threshold (dB)
25 GHz	10	On-off keying [38]	-14
50 GHz	100	DP-QPSK [39], PAM4 [40,41]	-18
75 (50 + 25) GHz	110 (100 + 10)	2 $\lambda$ : OOK + PAM4	-18/-14
100 (50 + 50) GHz	300	PAM8 [42,43]	-24

Type 1 request: Combination of 10, 100, 110, 300 Gb/s

Type 2 request: Fixed data rate, bandwidth (300 Gb/s, 100 GHz)

Type of multiplexing used on networks considered (fiber type, technique) a) SDM (using MCFs, fiber switch)  
b) WDM (using SMF, AWG/WSS)  
c) SDM-WDM (using MCFs, AWG/WSS and fiber switch)

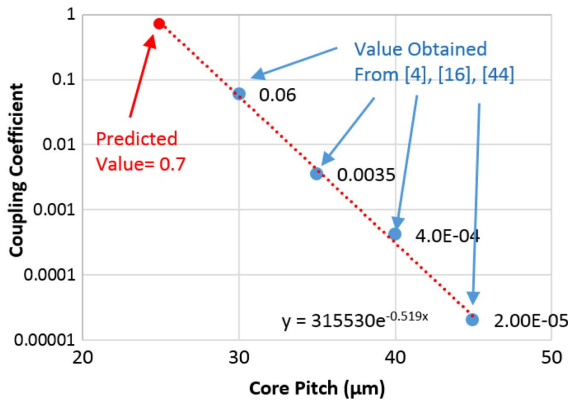


Fig. 11. Coupling coefficient versus core pitch values.

experimental measurements [12]. Note that the results in Ref. [12] were obtained for a long transmission link, i.e., 100 km. For short-reach transmission in DCNs, the effect of some parameters may be negligible. Therefore, the practical XT reduction could be higher than 20 dB ( $P_r < 0.01$ ) for a short-distance connection. Based on Eqs. (2), (10), and (12), Fig. 12 shows the impacts of the bi-directional transmission and trench-assisted technique on the value of total XT of the central core.

Since we considered homogeneous MCFs with small core pitches, the XT for uni-directional cases is higher than the ones reported in other similar studies, such as [47,48], which employ a larger core pitch. In the bi-directional cases, we measure the worst-case crosstalk on a central core that has six adjacent cores, shown in Fig. 4. Half of them are carrying signals in the same direction as the central core, while the others have the opposite direction. As illustrated in Fig. 12, when we assume a 20 dB XT reduction ( $P_r = 0.01$ ) from bi-directional core pair transmission for two extreme bi-directional cases, a) 7-core with largest core pitch and b) 61-core with smallest core-pitch, there is a

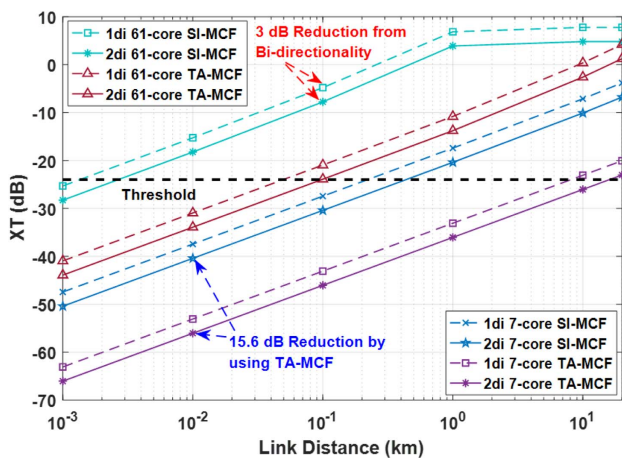


Fig. 12. XT reduction in the central core due to bi-directional transmission and trench-assisted technique.

3 dB total XT suppression compared to uni-directional transmission. A 3 dB reduction means that XT is reduced by half, which means adjacent cores on the same direction have major XT contributions. This is also clear by analyzing Eq. (7). When  $P_r = < 0.01$ ,  $XT_{\text{opposite}}$  contributes to  $XT_{\text{hex}}$  only marginally. Thus, all the results in the following section are obtained by assuming  $P_r = 0.01$ . Furthermore, as shown in Fig. 12, compared to SI-MCFs the TA-MCFs can provide 15.6 dB of reduction in the total observed XT for both the uni-directional and bi-directional cases. This reduction allows the fiber length with tolerable XT (below the threshold) to increase from 1.2 m to 50 m for a 61-core MCF and from 200 m to 8 km for a 7-core MCF. It should be mentioned that these two values can be doubled after introducing bi-directionality to the TA-MCFs.

### V. SIMULATION RESULTS

In this section, we first evaluate the benefits of bi-directional transmission and compare a range of algorithms in terms of network behavior (Subsection V.A) and execution time (Subsection V.B) on the Spine-Leaf topology. After identifying the most promising algorithm, we use it to investigate the network behavior (Subsection V.C) and network capacity as well as link spatial efficiency (Subsection V.D) for all topologies.

#### A. Algorithm Comparison in Terms of Network Behavior

Figure 13 shows the comparison between the proposed approach with bi-directional MCF, core prioritization, and soft-spectrum splitting and the uni-directional benchmarking algorithm. The results were obtained for the 7-core MCF with 30  $\mu\text{m}$  core pitch on the Spine-Leaf topology [Fig. 9(a)] with a 250 m link span. All results

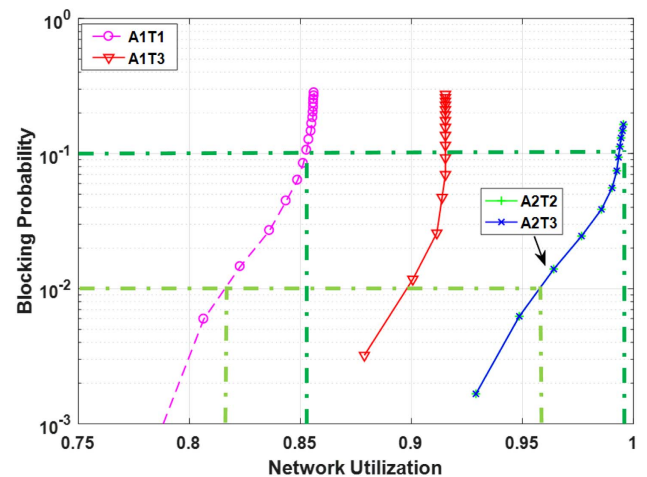


Fig. 13. Network behavior for uni-directional and bi-directional transmission in the Spine-Leaf topology.

shown in Figs. 13–18 consider type 1 requests and the SDM-WDM scheme (Table V).

Compared to the benchmark (A1T1, purple line), the equivalent bi-directional scheme (A1T3, red line) is able to significantly enhance the performance in terms of the overall BP. The BP is further reduced when the soft-spectrum-split approach is incorporated (green/blue line) even when the XT of MCF is already at a low level. Under BP in the order of 0.01 and 0.1, the proposed methods (green/blue line) offer 17% and 16% higher network utilization, respectively. This proves that the proposed resource allocation scheme is beneficial even for DCNs with low-XT MCFs. The *start2* approach will be used in the following sections.

After demonstrating the ability of the proposed bi-directional model with spectrum splitting to drastically mitigate XT effects compared to the benchmark uni-directional approach in the previous figure, we evaluate all variations of our proposed algorithms in order to select the best scheme for further DCN investigation. Figure 14 shows the network performance obtained by all three algorithms with 7-core bi-directional MCF (A2T3, A3, and A4) and two algorithms with 7-core uni-directional MCF (A1T2 and A2T1) in the Spine-Leaf topology.

According to Fig. 14, the hard-split method used in A4 produces blocking at an early time since only one slot division can be allocated at the beginning and a request is blocked if there are no sufficient slots in that division. When BP meets the threshold (0.1%, 1%, and 10%), two spectrum divisions will be swapped, and slots from the second division become available. Thus, the subsequent requests will not be blocked due to lack of slots, and the total BP decreases. However, as new requests arrive, slots in the new division will become insufficient as well and cause the BP to rise again. Eventually, the final network utilization will not be able to reach the same level as the soft-split approach A2T3, which translates to lower spectral efficiency of the hard split. The BPs of the remaining three approaches A1T2, A2T1, A2T3, shown in dashed purple, dashed blue, and solid blue lines, increase steeply.

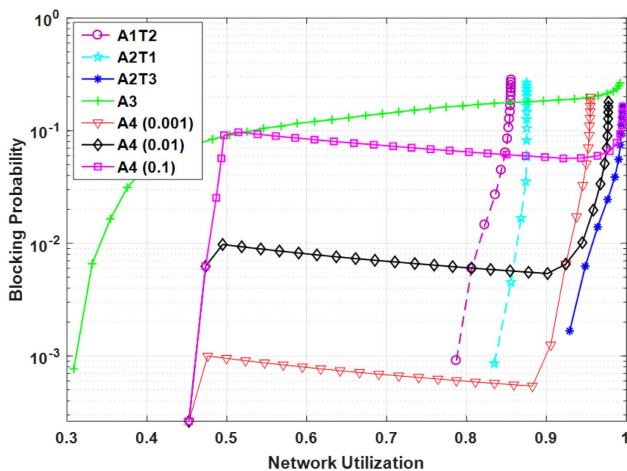


Fig. 14. Algorithm comparison for 7-core hexagonal MCF in the Spine-Leaf topology.

Comparing to the hard-split cases, these approaches check the whole spectrum rather than only one division. Thus, the blocking predominately occurs due to resource unavailability when the network becomes saturated, causing a sharp increase in the BP.

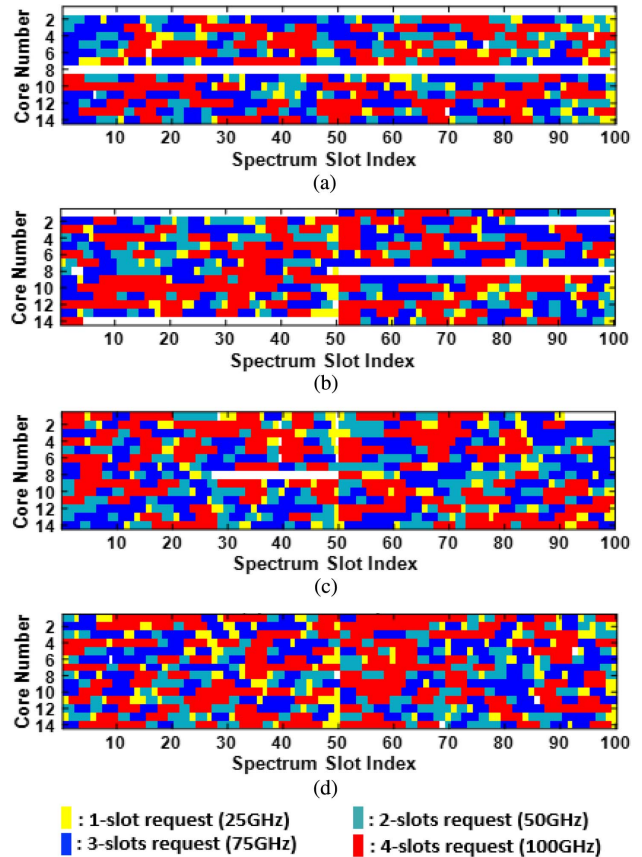


Fig. 15. Spectrum fragmentation for algorithms in 7-core hexagonal MCF: (a) A1T1, benchmark; (b) A4, hard split; (c) A2T3, soft split; and (d) A3, soft split and slot split.

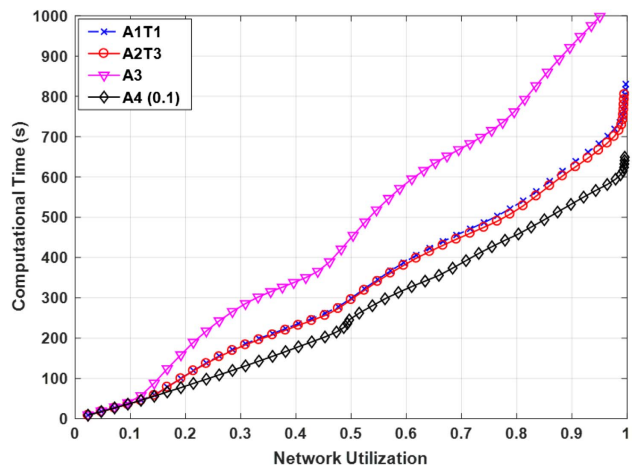


Fig. 16. Computational time for A1T1, A2T3, A3, and A4.

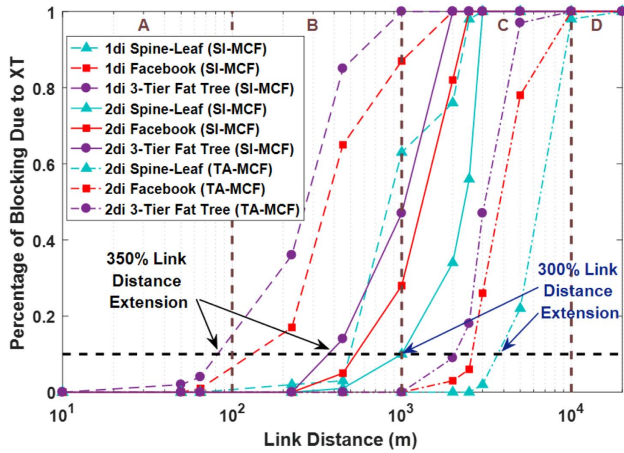


Fig. 17. Percentage of blocking due to XT with 7-core MCF.

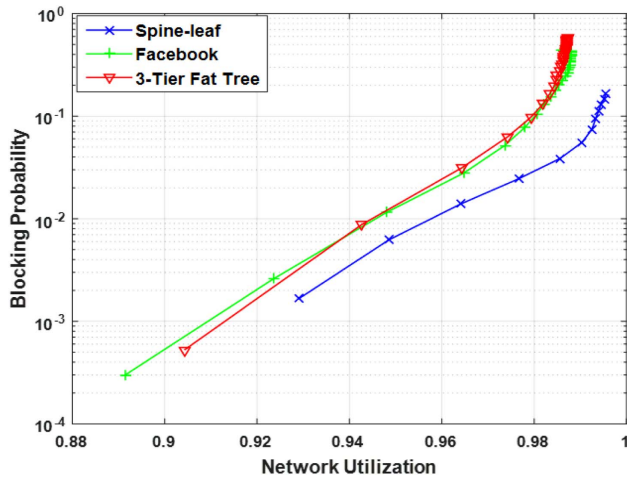


Fig. 18. Blocking probability as a function of network utilization obtained by A2T3 for the considered topologies.

In terms of blocking probability versus network utilization, A2T1 (light blue line) shows significant improvement compared to A1T2 (purple line) that does not split the spectrum. Moreover, the improvement can be enhanced when bi-directional transmission is adopted (A2T3, blue line) as expected. A2T3 is the best scheme, while A3 seems as the poorest choice in this case. This can be explained by the fact that the narrowest band request in the split-spectrum set controls the path selection for all other requests, depriving A3 of the benefits offered by the relaxation of spectrum contiguity constraint. In other words, if a path in the K-shortest path set satisfies the demands (spectral resources and XT level) for the narrowest band request, then that path is selected for all the other requests in the split set irrespective of whether or not it has enough free spectral resources to meet the requirements of these other requests in the set. Furthermore, the whole request is blocked when the requirements for any split request are not fulfilled. As a consequence, the first request blocking occurs already at around 30% utilization.

Figure 15 snapshots the final states of the highest loaded link for each scheme after processing 20,000 requests. For each subplot, the vertical axis represents the core index whereas the horizontal axis stands for the frequency slots. Unused slots are shown in white. The results show that due to high XT, two central cores are left completely empty by the benchmarking scheme [Fig. 15(a)]. When the Bi-Directional MCF and Spectrum-Split Algorithms are introduced [Figs. 15(b) and 15(c)], a boundary between the first 50 slots (D1) and the last 50 slots (D2) clearly emerges in each subplot. As explained in Subsection III.D, the two spectrum divisions will be swapped when the original division starts lacking free frequency slots. While both spectrum-splitting schemes reduce fragmentation drastically, the soft-split approach (A2T3) performs better. It leaves only 3% slots unused, while this number is 12.5% for the hard-split approach (A4). These results clearly illustrate why the proposed schemes obtain higher final network utilization than the benchmark in Fig. 14. Although A3 [Fig. 15(d)] has the lowest fragmentation, it will not be utilized for further analysis due to the poor blocking performance shown in Fig. 14.

### B. Algorithm Comparison in Terms of Execution Time

Figure 16 shows the computational time for each algorithm with 20,000 requests. A2T3 shows similar performance to the benchmark (A1T1). A4 (black line) runs faster than the benchmark for the same network utilization, demonstrating how limiting the search on a part of the spectrum speeds up calculations. On the contrary, A3 (the top line) requires longer running time as it divides one request into several requests. Although A4 runs faster, A2T3 is considered for the next capacity and link spatial efficiency evaluations since it offers the best blocking probability while having similar computational complexity to the benchmark approach.

### C. Comparison for Different Topologies and Fibers

Figure 17 shows a comparison of XT-induced blocking and transmission reach values obtained by the bi-directional and the benchmarking uni-directional scheme in the three topologies. Four scales of DCNs are classified by link distance as declared in Fig. 17: intra-cluster network (<100 m; type A), larger multi-cluster (100–1000 m; type B), building-to-building data center farm (1–10 km; type C) and metro-to-metro DCN (>10 km; type D).

As it can be seen in the figure, the link distance that can be traversed by signals before crosstalk blocking occurs is longer in the Spine-Leaf topology than in the Facebook topology and Three-Tier Fat Tree topology in both the uni-directional (dotted lines) and bi-directional (solid lines) scenarios. This can be explained by the fact that in Spine-Leaf topology there are only two hops between any node pair, while the average number of hops in Facebook and

Three-Tier Fat Tree topologies equals 3.6 and 4.1, respectively. Thus, the performance of the two three-tier topologies is better when the path distance is considered. The above chart clearly indicates that the bi-directional MCF transmission is capable of drastically extending the XT-dependent transmission distance in all three topologies. In the Spine-Leaf topology, the bi-directional approach allows for transmission over 1000 m with 10% blocking due to XT for a 7-core SI-MCF, while the value reduces to 500 m with the benchmark algorithm, yielding a 100% increase of the transmission reach. In the Facebook and the Three-Tier Fat Tree topology, this enhancement reaches approximately 320% and 350%. This indicates that the bi-directional transmission has greater impact in the multi-tier topologies. However, as it can be seen in Fig. 17, the use of bi-directionality in SI-MCFs can only allow for the realization of data center types A and B without any blocking caused by XT (considering a Spine-Leaf topology). Nevertheless, by using bi-directional TA-MCF, the transmission distance can be further extended (e.g., 300% distance extension for a Spine-Leaf topology) enabling the realization of data center type C.

Figure 18 shows the BP for different utilization levels obtained by A2T3 for all three topologies. The Facebook and the Three-Tier Fat Tree topology have similar performance, both reaching 98% utilization at 10% BP level. The Spine-Leaf topology outperforms them, achieving 99.5% utilization at the same blocking level. This slight enhancement (1.5%) owes to the reduction of the tier number (two-tier). However, this tier suppression limits the connectivity and scalability of the Spine-Leaf topology.

#### D. Comparison of Network Capacity and Link Spatial Efficiency

Figures 19 and 20 provide the overall comparison of the maximum network capacity and link spatial efficiency for the three topologies using various fibers with different types of requests and multiplexing techniques, as summarized in Table V. The network capacity is the sum of the capacity used on every link by all accepted requests at 10% blocking probability. For a SDM scheme, only a type 2 request can be utilized since we consider only one channel per core. This scheme is studied to prove that MCF-based SDM can provide benefits in DCNs over SMF-based WDM,

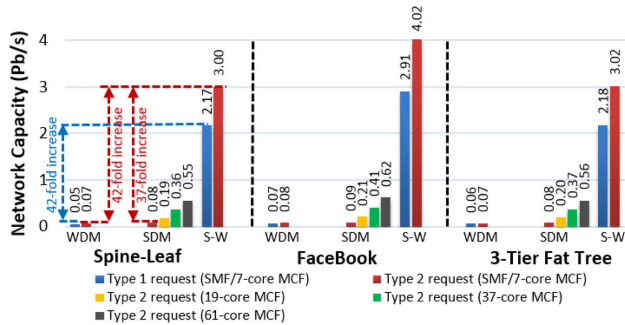


Fig. 19. Total network capacity obtained by A2T3 in different topologies and for different schemes. S-W, SDM-WDM.

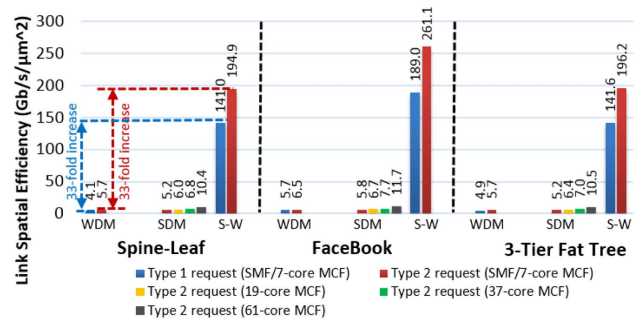


Fig. 20. Spatial efficiency obtained by A2T3 in different topologies and for different schemes. S-W, SDM-WDM.

which is the baseline for investigation of the SDM-WDM scheme. The Facebook topology can achieve the highest capacity and link spatial efficiency among all topologies since it has more links connected to each end node than the other two topologies. The Spine-Leaf topology and Three-Tier Fat Tree topology both have three links connected to each end node and perform similarly.

For the Spine-Leaf topology (Fig. 19), the SDM scheme with the 7-core MCF and type 2 requests obtains 0.08 Pb/s of total capacity, which is 14% greater than the WDM scheme (0.07 Pb/s, red bars). It is important to note here that the WDM case assumes 250% more channels per link compared to the 7-core MCF SDM case. This is 25 channels per fiber link direction (each request occupies four frequency slots while having 100 in total) whereas in the SDM case there are only as many channels as there are cores, i.e., seven in this case. Moreover, when the core (channel) number increases, the 19-core (0.19 Pb/s), 37-core (0.36 Pb/s), and 61-core (0.55 Pb/s) MCFs using pure SDM offer 1.7-fold, 4.1-fold, and 6.8-fold capacity increases compared to WDM, respectively.

For the SDM-WDM scheme, the number of available channels per core is 25, while each channel comprises four spectrum slots. The type 1 requests will not always occupy all four slots of the available channels due to their diverse data rates (from 10 Gb/s to 300 Gb/s). On the other hand, type 2 requests with a fixed data rate of 300 GB/s will always need all four slots to satisfy the capacity demand. In a 7-core MCF with Spine-Leaf topology, the SDM-WDM (S-W) scheme shows 42-fold and 37-fold increases in capacity compared to WDM and SDM, respectively, when serving type 2 requests (red bars). Not only that, the SDM-WDM scheme also attains over 42-fold capacity enhancement compared to a WDM scheme with type 1 requests (blue bars). For the Facebook topology, the SDM-WDM scheme enhances the DCN capacity 49-fold (i.e., from 0.08 to 4.02 Pb/s) and 43-fold (i.e., from 0.09 to 4.02 Pb/s) for the type 2 requests compared to the pure WDM and SDM schemes, respectively. This capacity enhancement is 41-fold related to the WDM scheme (i.e., from 0.07 to 2.91 Pb/s) for the type 1 requests.

Apart from the total data rate, the spatial efficiency of MCF is also a crucial performance indicator for data center deployment. The WDM scheme offers better spatial efficiency than the 7-core MCF SDM scheme for

all topologies in Fig. 20, since MCF has a higher cross-sectional area than SMF (Table VI) but a considerably reduced capacity. On the contrary, the 19-core MCF compensates the spatial issue by raising the core number. Because of this, the SDM scheme improves the spatial efficiency by 5%, 3%, and 12% compared to the WDM scheme for Spine-Leaf (i.e., from 5.7 to 6.0 Gb/s/ $\mu\text{m}^2$ ), Facebook (i.e., from 6.5 to 6.7 Gb/s/ $\mu\text{m}^2$ ), and Three-Tier Fat Tree topologies (i.e., from 5.7 to 6.4 Gb/s/ $\mu\text{m}^2$ ), respectively. These respective enhancements increase to 19%, 18%, and 23% for the 37-core MCF case, and reach as high as 82%, 80%, and 84% when a 61-core MCF is considered. The SDM-WDM scheme in MCF proves itself as the best one, offering over 33-fold (7-core MCF) higher spatial efficiency than the WDM scheme with SMF. Although the type 2 requests yield better performance than type 1, the type 1 requests are more suitable for further network investigation considering the practical ability of enabling different transmission capabilities in DCN.

Figure 21 depicts the front panel fiber core density that can be achieved by using standard commercial connectors to accommodate the conventional SMF, SMF-based fiber ribbon, and single-mode MCFs on a typical 1U rack mount shelf. Note that the results in Fig. 21 assume a 70% front panel dedicated to connectors and 30% to the ventilation and other devices. To compare the SMF solutions with MCF solutions, a commercial lucent connector (LC) connector [49] is considered, where we assume that the MCF connectors have the same surrounding area as that of the SMF connector. As the results indicate, the achievable core density for the front panel is linearly proportional to the core number of a fiber (yellow bars in Fig. 21), i.e., the 61-core MCF can achieve 61 times the density of the SMF. This can be explained by the fact that the cladding area has only a limited effect compared to the surrounding area, which dominates the connector size. To explore the effect of the cladding diameter of the MCF on the front panel density, a highly dense 72-fiber multi-fiber push-on/ multi-fiber termination push-on (MTP/MPO) connector [50] is employed for comparison between fiber ribbon and the MCF solutions. In this case, we assume the same (multi-fiber) ferrule size, in which fiber alignment is dependent on the eccentricity and pitch of the fiber and the alignment of pin holes for all fiber types to ensure the same connector size for all types of fiber. Thus, the number of fibers in a connector depends on the cladding area of the fiber and can be calculated by Eq. (14):

$$\text{Fiber Number per Connector} = \frac{72 \times \text{Cladding Area of SMF}}{\text{Cladding Area of the Fiber}} \quad (14)$$

Based on Eq. (14), the connector area per core and the core number per front panel can be calculated by Eqs. (15) and (16), respectively:

$$\text{Connector Area per Core} = \frac{\text{Connector Size}}{\text{Fiber Number} \times \text{Core Number per Fiber}}, \quad (15)$$

$$\text{Core Number per Front Panel} = \frac{\text{Front Panel Area} \times \text{Percentage of the Area Dedicated to Connectors}}{\text{Connector Area per Core}}. \quad (16)$$

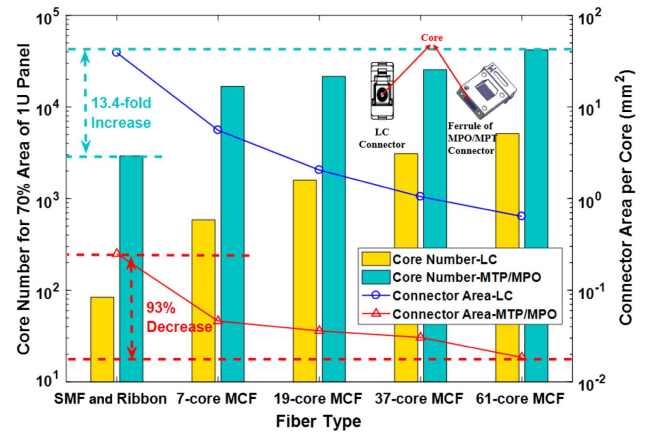


Fig. 21. Comparison of front panel core density.

The results show that the MTP/MPO can considerably improve the front panel density compared to the LC connector. Moreover, the investigated MCFs achieve a maximum 13.4-fold front panel core density increase for the 61-core MCF compared to the fiber ribbon. The results of the connector area per core can be used for calculating the core density for any other type of panels based on Eq. (16).

Figure 22 shows the total network capacity for different link distances with various types of SI-MCFs. The results were obtained by A2T3 on four hexagonal SI-MCFs (all homogeneous) in three topologies and compared to the uni-directional 7-core SI-MCF benchmarking approach. As shown in the figure, the maximum capacity of the same layout SI-MCFs in each topology is proportional to the number of cores. Compared to the single-core fiber with a WDM scheme in Fig. 19, all the schemes offer considerable improvements. In particular, the 61-core SI-MCF shows the highest, i.e., 379-, 361-, and 317-fold increases in Spine-Leaf, Facebook, and Three-Tier Fat Tree topology, respectively. The 7-core bi-directional SI-MCF obtains the same maximum capacity as the benchmarking approach. The reason for this is that the max capacity is achieved when there is no blocking due to XT (i.e., XT is below threshold), which means that the XT suppression provided by bi-directional transmission does not affect the network capacity. However, XT reduction contributes to protecting the requests from being blocked due to excessive XT values. Thus, bi-directional SI-MCF can maintain the maximum capacity for longer distances without XT blocking (e.g., 225 m for 1di transmission, and 425 m for 2di transmission with 7-core SI-MCF in the Spine-Leaf topology).

As depicted in Fig. 23, compared to the network capacities that are achievable by SI-MCFs, each type of

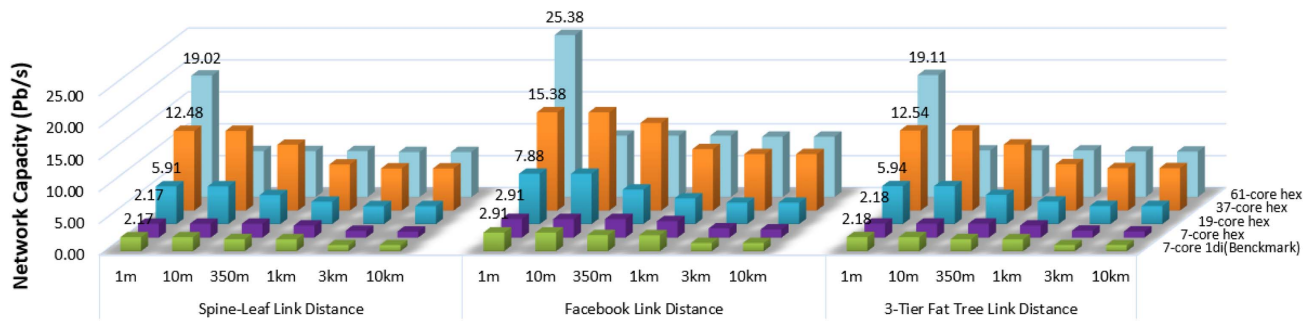


Fig. 22. Total network capacity as a function of link distance for different normal step-index fiber types.

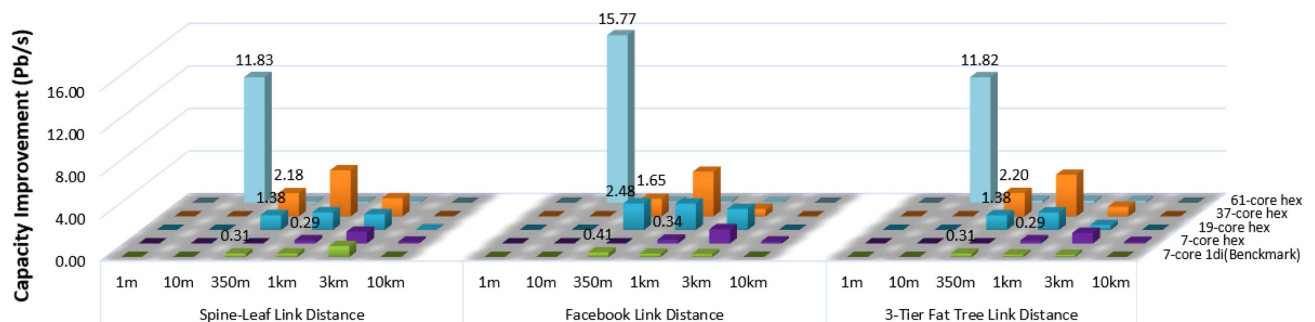


Fig. 23. Improvement of the total network capacity by using TA-MCF over using SI-MCF.

the corresponding TA-MCF provides the same maximum capacity. However, all TA-MCFs enlarge the link distances where the maximum capacities can be achieved (without any blocking caused by XT). For instance, the 61-core TA-MCF can support a DCN with  $>10$  m link distance with maximum capacity, yielding an over 730% link distance extension compared to that of the SI-MCF for each of the topologies (around 1.2 m). Moreover, by using the TA-MCFs, different levels of capacity improvements can be achieved for DCNs with 10 to 3,000 m link distances. Particularly, the 7-core TA-MCF obtains a 103% capacity increase over the 7-core SI-MCF in the Spine-Leaf-topology-based DCN with 1000 m link distance. However, when the link distance goes to 10 km, SI-MCFs and TA-MCFs perform the same since all of the blocking obtained for both of types of MCFs results from severe XT. The results obtained in Figs. 22 and 23 could be a useful reference for MCF selection when designing, evaluating, and deploying MCFs for different applications (i.e., intra-cluster DCN with  $<100$  m link distances, data center farms with up to 10 km end-to-end distances), which depend on the requirements for both resource efficiency and the physical scale.

## VI. CONCLUSION

In this paper, we propose a bi-directional data center networking solution using single MCF for crosstalk reduction between adjacent cores. We mathematically derived new wavelength-dependent crosstalk formulations for homogeneous MCFs including SI-MCF and TA-MCF, which for the first time considers bi-directionality and a uniform

pitch between adjacent cores (hexagonal MCF). The derived model is based on several experimentally proven analytical results for MCF systems, validating its application to real SDM network systems. A bi-directional core priority map and two spectrum split algorithms (hard and soft) are introduced to improve the DCN performance, which is thoroughly examined in terms of blocking probability, network utilization, network capacity, and link spatial efficiency. Several homogenous SI-MCFs and TA-MCFs are investigated for three different topologies with the objective to maximize capacity and spatial efficiency, and find the best fit between fiber type and data center environment. Simulation results demonstrate that the bi-directional model is able to extend the transmission range of MCF (100%, 320%, and 350% increase on transmission reach with 10% XT-caused blocking compared to uni-directional benchmark in the Spine-Leaf, Facebook, and Three-Tier Fat Tree topology respectively) under the same conditions. The spectrum split algorithms mitigate XT by dividing the whole available spectrum into two disjoint bands. Among the proposed strategies, the bi-directional XT-aware algorithm with core prioritization, soft-spectrum-splitting, and core-switching mechanisms outperforms all investigated algorithms. In terms of topology, the Spine-Leaf topology shows a slight advantage in network behavior, while the Facebook topology provides the highest network capacity and link spatial efficiency. In the multiplexing schemes, the experimental results support the superiority of SDM over WDM networks. For SDM-WDM networks with MCF, more than 33 times higher link spatial efficiency (7-core MCF) and up to over 300 times increased capacity (61-core MCF) are attained compared with SMF-based

WDMs for all three topologies. This indicates the potential to support future high-capacity DCNs. Eventually, the studies carried out here on network capacity with respect to the link distance clearly highlight that crosstalk suppression enables highly dense MCFs (i.e., 61-core SI/TA-MCF) to be the best candidate for small-scale data centers [i.e., within a) a rack, b) data center in a box, c) points of delivery (PODs)] with link spans ranging from few meters and metro-to-metro data centers with >10 km link distance. On the other hand, it is found that the 37-core hexagonal TA-MCF outperforms other MCF types in the larger multi-cluster DCNs with link spans of up to 100 s of meters and building-to-building data center farms with link spans up to 1000 s of meters.

#### ACKNOWLEDGMENT

This work was supported by the European Commission H2020 dRedBox project, by a SONATAS project funded by EPSRC (UK), and by EPSRC EP/L026155/1.

#### REFERENCES

- [1] D. J. Richardson, J. M. Fini, and L. E. Nelson, "Space-division multiplexing in optical fibres," *Nat. Photonics*, vol. 7, pp. 354–362, 2013.
- [2] C. Kachris and I. Tomkos, "A survey on optical interconnects for data centers," *IEEE Commun. Surv. Tutorials*, vol. 14, no. 4, pp. 1021–1036, 2012.
- [3] N. Farrington, G. Porter, S. Radhakrishnan, H. H. Bazzaz, V. Subramanya, Y. Fainman, G. Papen, and A. Vahdat, "Helios: A hybrid electrical/optical switch architecture for modular data centers," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 41, pp. 339–350, 2011.
- [4] G. M. Saridis, D. Alexandropoulos, G. Zervas, and D. Simeonidou, "Survey and evaluation of space division multiplexing: From technologies to optical networks," *IEEE Commun. Surv. Tutorials*, vol. 17, no. 4, pp. 2136–2156, 2015.
- [5] J. Perelló, J. M. Gené, A. Pagès, J. A. Lazaro, and S. Spadaro, "Flex-grid/SDM backbone network design with inter-core XT-limited transmission reach," *J. Opt. Commun. Netw.*, vol. 8, no. 8, pp. 540–552, Aug. 2016.
- [6] T. Hayashi, A. Mekis, T. Nakanishi, M. Peterson, S. Sahni, P. Sun, S. Freyling, G. Armijo, C. Sohn, D. Foltz, T. Pinguet, M. Mack, Y. Kaneuchi, O. Shimakawa, T. Morishima, T. Sasaki, and P. D. Dobbelaere, "End-to-end multi-core fibre transmission link enabled by silicon photonics transceiver with grating coupler array," in *European Conf. and Exhibition on Optical Communication (ECOC)*, 2017, paper Th.2.A.
- [7] H. C. H. Mulvad, A. Parker, B. King, D. Smith, M. Kovacs, S. Jain, J. Hayes, M. Petrovich, D. J. Richardson, and N. Parsons, "Beam-steering all-optical switch for multi-core fibers," in *Optical Fiber Communications Conf. and Exhibition (OFC)*, Los Angeles, California, 2017, pp. 1–3.
- [8] W. Klaus, B. J. Puttnam, R. S. Luis, J. Sakaguchi, J. M. D. Mendinueta, Y. Awaji, and N. Wada, "Advanced space division multiplexing technologies for optical networks [Invited]," *J. Opt. Commun. Netw.*, vol. 9, no. 4, pp. C1–C11, Apr. 2017.
- [9] Y. Liu, H. Yuan, A. Peters, and G. Zervas, "Comparison of SDM and WDM on direct and indirect optical data center networks," in *42nd European Conf. on Optical Communication (ECOC)*, Dusseldorf, Germany, 2016, pp. 1–3.
- [10] I. Morita, K. Igarashi, and T. Tsuritani, "1 exabit/s-km transmission with multi-core fibre and spectral efficient modulation format," in *OptoElectronics and Communication Conf. and Australian Conf. on Optical Fibre Technology*, July 6–10, 2014, pp. 316–318.
- [11] A. Muhammad, G. Zervas, D. Simeonidou, and R. Forchheimer, "Routing, spectrum and core allocation in flex-grid SDM networks with multi-core fibers," in *Int. Conf. on Optical Network Design and Modeling*, Stockholm, Sweden, 2014, pp. 192–197.
- [12] F. Ye, K. Saitoh, H. Takara, R. Asif, and T. Morioka, "High-count multi-core fibers for space-division multiplexing with propagation-direction interleaving," in *Optical Fiber Communications Conf. and Exhibition (OFC)*, Los Angeles, California, 2015, pp. 1–3.
- [13] A. Sano, H. Takara, T. Kobayashi, H. Kawakami, H. Kishikawa, T. Nakagawa, Y. Miyamoto, Y. Abe, H. Ono, K. Shikama, M. Nagatani, T. Mori, Y. Sasaki, I. Ishida, K. Takenaga, S. Matsuo, K. Saitoh, M. Koshiba, M. Yamada, H. Masuda, and T. Morioka, "409-Tb/s + 409-Tb/s crosstalk suppressed bidirectional MCF transmission over 450 km using propagation-direction interleaving," *Opt. Express*, vol. 21, pp. 16777–16783, 2013.
- [14] J. Sakaguchi, B. J. Puttnam, W. Klaus, J. M. D. Mendinueta, Y. Awaji, N. Wada, A. Kanno, and T. Kawanishi, "Large-capacity transmission over a 19-core fiber," in *Optical Fiber Communication Conf. and Expo. and the Nat. Fiber Optic Engineers Conf. (OFC/NFOEC)*, Anaheim, California, 2013, pp. 1–3.
- [15] M. Koshiba, "Design aspects of multicore optical fibres for high-capacity long-haul transmission," in *Int. Topical Meeting on Microwave Photonics (MWP) and the 9th Asia-Pacific Microwave Photonics Conf. (APMP)*, Oct. 20–23, 2014, pp. 318–323.
- [16] K. Saitoh, M. Koshiba, K. Takenaga, and S. Matsuo, "Crosstalk and core density in uncoupled multicore fibers," *IEEE Photonics Technol. Lett.*, vol. 24, no. 21, pp. 1898–1901, Nov. 2012.
- [17] S. Talebi, E. Bampis, G. Lucarelli, I. Katib, and G. N. Rouskas, "Spectrum assignment in optical networks: A multiprocessor scheduling perspective," *J. Opt. Commun. Netw.*, vol. 6, no. 8, pp. 754–763, Aug. 2014.
- [18] M. Fiorani, M. Tornatore, J. Chen, L. Wosinska, and B. Mukherjee, "Spatial division multiplexing for high capacity optical interconnects in modular data centers," *J. Opt. Commun. Netw.*, vol. 9, pp. A143–A153, 2017.
- [19] P. Zhou, X. Xu, S. Guo, and Z. Liu, "Analysis on power scalability of multicore fiber laser," in *IEEE PhotonicsGlobal@Singapore (IPGC)*, Singapore, 2008, pp. 1–3.
- [20] E. L. Goldstein, L. Eskildsen, and A. F. Elrefaie, "Performance implications of component crosstalk in transparent lightwave networks," *IEEE Photonics Technol. Lett.*, vol. 6, no. 5, pp. 657–660, May 1994.
- [21] F. Ye, T. Morioka, J. Tu, and K. Saitoh, "Theoretical investigation of inter-core crosstalk properties in homogeneous trench-assisted multi-core fibres," in *IEEE Photonics Society Summer Topical Meeting Series*, July 14–16, 2014, pp. 180–181.
- [22] S. Fujii, Y. Hirota, H. Tode, and K. Murakami, "On-demand spectrum and core allocation for reducing crosstalk in multi-core fibers in elastic optical networks," *J. Opt. Commun. Netw.*, vol. 6, no. 12, pp. 1059–1071, Dec. 2014.
- [23] A. Sano, H. Takara, T. Kobayashi, and Y. Miyamoto, "Crosstalk-managed high capacity long haul multicore fibre



- transmission with propagation-direction interleaving," *J. Lightwave Technol.*, vol. 32, no. 16, pp. 2771–2779, 2014.
- [24] M. Koshiba, K. Saitoh, K. Takenaga, and S. Matsuo, "Analytical expression of average power-coupling coefficients for estimating intercore crosstalk in multicore fibres," *IEEE Photonics J.*, vol. 4, no. 5, pp. 1987–1995, Oct. 2012.
- [25] T. Hayashi, T. Taru, O. Shimakawa, T. Sasaki, and E. Sasaoka, "Design and fabrication of ultra-low crosstalk and low-loss multi-core fiber," *Opt. Express*, vol. 19, pp. 16576–16592, 2011.
- [26] K. Igarashi, T. Tsuritani, I. Morita, and M. Suzuki, "Ultra-long-haul high-capacity super-Nyquist-WDM transmission experiment using multi-core fibers," *J. Lightwave Technol.*, vol. 33, no. 5, pp. 1027–1036, Mar. 2015.
- [27] H. Zang and J. P. Jue, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks," *Opt. Networks Mag.*, vol. 1, no. 1, pp. 47–60, 2000.
- [28] L. Velasco, A. Castro, M. Ruiz, and G. Junyent, "Solving routing and spectrum allocation related optimization problems: From off-line to in-operation flexgrid network planning," *J. Lightwave Technol.*, vol. 32, no. 16, pp. 2780–2795, Aug. 2014.
- [29] H. Tode and Y. Hirota, "Routing, spectrum and core assignment for space division multiplexing elastic optical networks," in *16th Int. Telecommunications Network Strategy and Planning Symp. (Networks)*, Sept. 17–19, 2014, pp. 1–7.
- [30] A. Pagès, J. Perelló, S. Spadaro, and J. Comellas, "Optimal route, spectrum, and modulation level assignment in split-spectrum-enabled dynamic elastic optical networks," *J. Opt. Commun. Netw.*, vol. 6, no. 2, pp. 114–126, Feb. 2014.
- [31] F. Ye, J. Tu, K. Saitoh, K. Takenaga, S. Matsuo, H. Takara, and T. Morioka, "Wavelength-dependence of inter-core crosstalk in homogeneous multi-core fibers," *IEEE Photonics Technol. Lett.*, vol. 28, no. 1, pp. 27–30, Jan. 2016.
- [32] M. Alizadeh and T. Edsall, "On the data path performance of Leaf-Spine datacenter fabrics," in *IEEE 21st Annu. Symp. on High-Performance Interconnects*, San Jose, California, 2013, pp. 71–74.
- [33] "AWG multi/demultiplexer" [Online]. Available: [http://www.ntt-electronics.com/en/products/photonics/awg\\_mul\\_d.html](http://www.ntt-electronics.com/en/products/photonics/awg_mul_d.html).
- [34] "9/1 × 20 flexgrid wavelength selective switch" [Online]. Available: <https://www.finisar.com/roadms-wavelength-management/10wsaxxfl>.
- [35] T. Benson, A. Akella, and D. A. Maltz, "Network traffic characteristics of data centers in the wild," in *10th Annu. Conf. on Internet Measurement (IMC)*, 2010, pp. 267–280.
- [36] G. Zervas, F. Jiang, Q. Chen, V. Mishra, H. Yuan, K. Katrinis, D. Syrivelis, A. Reale, D. Pnevmatikatos, M. Enrico, and N. Parsons, "Disaggregated compute, memory and network systems: A new era for optical data centre architectures," in *Optical Fiber Communications Conf. and Exhibition (OFC)*, Los Angeles, California, 2017, pp. 1–3.
- [37] S. S. Chandrasekaran, "Understanding traffic characteristics in a server to server data center network," Master's thesis, Rochester Institute of Technology, Rochester, NY, 2017.
- [38] N. Amaya, M. Irfan, G. S. Zervas, K. Banias, M. Garrich, I. Henning, D. Simeonidou, Y. R. Zhou, A. Lord, K. Smith, V. Rancano, S. Liu, P. Petropoulos, and D. Richardson, "Gridless optical networking field trial: Flexible spectrum switching, defragmentation and transport of 10G/40G/100G/555G over 620-km field fiber," in *37th European Conf. and Expo. on Optical Communications*, 2011, paper Th.13.K.1.
- [39] O. Gerstel, M. Jinno, A. Lord, and S. J. B. Yoo, "Elastic optical networking: A new dawn for the optical layer?" *IEEE Commun. Mag.*, vol. 50, no. 2, pp. s12–s20, Feb. 2012.
- [40] "PAM modulation for 400G SMF," IEEE P802.3bs 400 Gb/s Ethernet Task Force, May 2014 [Online]. Available: [http://www.ieee802.org/3/bs/public/14\\_05/](http://www.ieee802.org/3/bs/public/14_05/).
- [41] "Opportunities for PAM-4 modulation," Huawei Technologies Co., Ltd. and IEEE 802.3 400 GbE Study Group, Jan 2014 [Online]. Available: [http://www.ieee802.org/3/400GSG/public/14\\_01/](http://www.ieee802.org/3/400GSG/public/14_01/).
- [42] M. A. Mestre, H. Mardoyan, A. Konczykowska, R. Rios-Müller, J. Renaudier, F. Jorge, B. Duval, J. Y. Dupuy, A. Ghazisaeidi, P. Jennevé, and S. Bigo, "Direct detection transceiver at 150-Gbit/s net data rate using PAM 8 for optical interconnects," in *42nd European Conf. on Optical Communication (ECOC)*, 2015, pp. 1–3.
- [43] "Update on technical feasibility for PAM modulation," IEEE 802.3 NG100GE PMD Study Group, Mar. 2012 [Online]. Available: <http://www.ieee802.org/3/100GNGOPTX/public/mar12/plenary/>.
- [44] A. Muhammad, G. Zervas, and R. Forchheimer, "Resource allocation for space-division multiplexing: Optical white box versus optical black box networking," *J. Lightwave Technol.*, vol. 33, no. 23, pp. 4928–4941, Dec. 2015.
- [45] P. J. Winzer, A. H. Gnauck, A. Konczykowska, F. Jorge, and J. Y. Dupuy, "Penalties from in-band crosstalk for advanced optical modulation formats," in *37th European Conf. and Exhibition on Optical Communication*, Geneva, Switzerland, 2011, pp. 1–3.
- [46] "PAM8 & FEC options," IEEE P802.3bm 40 Gb/s and 100 Gb/s Fiber Optic Task Force, Nov. 2012 [Online]. Available: [www.ieee802.org/3/bm/public/nov12](http://www.ieee802.org/3/bm/public/nov12).
- [47] F. Ye, J. Tu, K. Saitoh, K. Takenaga, S. Matsuo, and T. Morioka, "A new and simple method for crosstalk estimation in homogeneous trench-assisted multi-core fibers," in *Asia Communications and Photonics Conf.*, 2014, paper AW4C.3.
- [48] T. Hayashi, T. Taru, O. Shimakawa, T. Sasaki, and E. Sasaoka, "Ultra-low-crosstalk multi-core fiber feasible to ultra-long-haul transmission," in *Optical Fiber Communication Conf. and the Nat. Fiber Optic Engineers Conf. (OFC/NFOEC)*, 2011, paper PDP2.
- [49] LC Product Specification Outline [Online]. Available: <http://lcalliance.net/lcInterface/pdfs/LC-Product-Spec.pdf>.
- [50] USCONEC, "C9730 datasheet" [Online]. Available: <http://www.usconec.com/images/drawings/C9730.pdf>.