

面向新型DC互联的全光交叉传输系统功能需求研究

崔素峰 梁伟军 辜彪

河南省信息咨询设计研究有限公司 河南 郑州 450003

摘要：面对涌现的新型DC和不断DC化的传统通信机房，波分传输系统需要适应新的机房环境并探索新的组网方式、硬件形态和管控方式，打造绿色低碳、高效承载的智算网络光底座。通过分析新型DC业务需求、传统波分系统在DC场景中存在的问题，提出新型DC互联全光交叉传输系统在技术选择、功能设计、硬件设计和管控方式的需求，并给出组网优化建议。

关键词：智算；DCI；OXC；WSS；模块化

引言

新型DC包括智算中心和新型互联网交换中心（IXP）。智算中心部署的AI大模型通过不断增加参数规模，涌现智能性，AI小模型推动了行业效率。新型IXP优化数据中心跨网、跨地域数据交互，提升电信运营商和用户互联互通质量。

网络内容、网络硬件的不断云化，导致通信机房逐步DC化，对硬件设计、业务快速提供、灵活性等方面都提出了更高的要求。面向新型DC互联场景，如何降低建

网成本及运行能耗，配合打造统一管控的智算网络，最终实现开放解耦，是本场景需考虑的问题。

1 新型DC业务需求分析

1.1 算力新基建对传输的需求

“东数西算”是国家根据不同地区资源禀赋，采取的数字新基建与清洁能源统筹工程。相关政策根据业务时延要求，对“东数西算”业务进行区分，给出了不同业务的部署建议，详见表1。

表1 “东数西算”工程业务分类

业务分类	时延	业务部署建议	典型应用/服务及算力需求占比	
热业务	<10ms	城域部署	5%-10%	金融交易、AI推理、直播、游戏、工业控制、低时延物联网、车联网等
温业务	<30ms	区域部署	55%-60%	政务网站、智慧城市、协同办公、部分工业互联网等
温冷业务	<100ms	东数西算	20%-30%	异地容灾、视频转播、医疗影像、基因测序、大数据、云会议等
冷业务	>100ms	东数西算	10%	数据备份、归档、门户浏览、社交、邮件、电商、AI训练等

1.2 分布式智算网间需求

智算中心是集公共算力服务、数据开发共享、产业创新于一体的基础设施，未来将统筹运营。可通过将多个智算节点互联，形成智算资源的统一调度，达到算力能力提升的目的。同时，智算中心在万卡集群后，维护成本提升，电力供给成为算力提升的制约因素，区域内多个智算中心互联，形成统一算力调度，可以分摊算力集中的风险。

1.3 新型IXP建设需求

新型IXP建设是以省会节点为核心、区域城市为汇聚节点、区域热点为接入点的省内城市群网络，该网络不仅提升业务感知，也是未来区域分布式算力资源调度的潜在承载网络。

1.4 低时延需求

智算中心的业务时延与大模型训练效率之间存在密

切关系，对于温、热业务，传送网被要求提供更低时延，降低时延成为网络建设的新诉求。

1.5 低功耗需求

随着数据中心建设标准的提高，对PUE值的要求越来越低，针对DC机房环境，传统传输系统应从硬件设计、产品功能、风道设计等适应新环境^[1]。

2 新型DCI全光交叉传输系统技术选择

2.1 线路侧技术选择

2.1.1 码型技术选择

按照“东数西算”规划预测，算网流量预计新增3000T，是现有运营商骨干网带宽的3倍左右，现波分系统以单波100G为主，为承载新增流量，需提升单波速率。根据技术演进、成本因素，下一代大规模部署的波分技术应向400G演进。

$$\text{线路速率} = \text{复用偏振数} * \text{子载波数} * \text{Log}_2(\text{调制阶数}) * \text{波特率} \quad (1)$$

根据公式(1),提高调制阶数能提升线路速率,但传输距离会大幅下降。在给定的频谱宽度下,提升波特率是提高系统容量的常用方法,通过芯片的改进和算法的优化,波特率可有效提高。

400G技术目前有PM-16QAM、PCS-16QAM、PM-QPSK三种主流调制方案。16QAM有60和90Gbaud两种,前者采用75GHz间隔方案,成本低、传输距离较近,适用于点到点的传输,后者采用100GHz间隔方案,最远可传输1000km左右,可适用本地传输或小型干线系统。QPSK方案需将光电器件提升至128+Gbaud,采用150GHz间隔

方案,预计可传输2000km左右。

高调制增大设备复杂度和造价成本,需将光电/电光转换器和FEC芯片从16nm提升到7nm或5nm,也会降低信号OSNR容限、缩短传输距离。

2.1.2 频谱选择

不同的调制方式意味着不同的波特率和不同的频谱宽度,最终决定不同的频谱效率。

$$\text{单波频谱效率} = \text{单波容量} / \text{单波频谱宽度} \quad (2)$$

根据公式(2),可计算出不同调制方案下的频谱效率。

表2 不同调制下频谱效率及系统容量(C4波段)

调制方式	波特率 (GBd)	波道间隔 (GHz)	无电中继传输距离 (km)	频谱效率	C4波段系统容量 (T)
PM-16QAM	60	75	小于600	5.3	21.2
PCS-16QAM	90	100	小于1000	4	16
PM-QPSK	130	150	小于2000	2.9	10.4

100G的频谱效率为2,从表2可以看出,单波速率提升4倍,频谱效率非成比例提升。为提升400G系统容量,需采用扩频技术拓展谱宽,增加波长数量。目前L及L+波段的放大器可实现增益大于20dB,且NF小于8,初步具备部署条件,不过C和L分立放大器不利于维护习惯、设备配置。

覆盖C+L波段超带宽一体化放大器的商用,是400G长距离场景规模部署的必备条件。目前,实验室环境下,共掺铋钕的新型放大器实现C+L宽带覆盖,增益大于20dB,NF最低为3.8dB,在解决成本问题后,相信在不久的将来,C+L一体放大器可实现规模部署。

2.1.3 线路侧技术方案建议

(1)对于大型城市内两个或多个DC互联的场景,码型方案可选择成本较低的PM-16QAM,频谱可仅采用C波段,系统容量足够满足该场景DCI流量需求。

(2)对于业务密集省内城市群DC互联场景,从容量、C+L一体化等因素考虑,码型方案选择PCS-16QAM,频谱应支持C6T+L6T方案,前期可仅使用C波段。

(3)对于数据中心集群互联的省际传输场景,码型方案应选择PM-QPSK,频谱应支持C6T+L6T方案。

2.2 全光交叉技术

2.2.1 一体化波长选择开关(WSS)

目前WSS采用LCOS实现,技术已经非常成熟,谱宽拓展至L波段后,WSS器件也需要一体化,分设WSS不利于产品的集成度。

2.2.2 光交叉连接技术(OXC)

新型DCI传输网具有MESH化、跨平面立体化、安装空间省、线缆少等趋势。OXC创新在于引入了预制光背

板,由传统PCB到新型POB,通过把光纤预制在POB上实现光支路单元、光线路单元全互联的能力,实现光信号任意方向的交叉调度。

2.2.3 全光交叉技术方案建议

通过提高光芯片分辨率、采用光学超表面等技术,支持L波段的光信号处理,实现器件C+L一体化,成本会比C和L分立WSS方案降低30%。新型DC节点光方向较多,光预制背板是必选项,可有效减少机房内连纤。

2.3 系统与DC场景匹配

2.3.1 功能单元模块化

传统设备具有集成度低、能力强、槽位多、功耗大等特点,通信设备云化是不可阻挡的趋势,云化后利于实现开放解耦、功能堆叠、功耗降低等目标。

2.3.2 匹配DC环境的风道设计

相关政策提出东部枢纽电能利用效率指标(PUE)控制在1.25以内,西部为1.2以内,波分设备在传统机房的PUE为1.8-2.0。DC机房的气流组织由传统机房的“洪灌风”改为“封闭冷通道”为主,波分设备也应改善气流组织。

2.3.3 系统功能DC化建议

(1) 模块化设计

DC之间电路以100G为主,辅以少量10G,对于DCI传输系统,可优化掉电交叉单元,使余下光层、电层功能实现模块化设计,集成度更高,设备功耗更低,可按以下方案划分功能区:

功能区1:线路主用信号放大器、线路保护单元、线路侧WSS(1:N);

功能区2:线路备用信号放大器;

功能区3: 预制光连接背板(维度可拓展);
 功能区4: 本地侧WSS矩阵(N:M)、波长转换单元;
 功能区1和2按实际光方向扩容; 功能区3采用独立式设计, 通过多芯光缆与功能区1和4预连接; 功能区4按需堆叠扩容。

(2) 气流组织选择

系统硬件设计应充分利用DC机柜纵深, 槽位横向设计。设备由“前进前出(上出)”优化为“前进后出”。引电支持220V交流或高压直流, 避免外接AC/DC模块。

2.4 开放解耦

当前传统波分设备管控设计属于烟囱式架构, 无法实现异厂家共同组网, 不利于网络的开放性和延展性。“全国一体化大数据”对传输的一个关键需求是推动数据中心与网络高效对接与协同发展。在此背景下, 开放和解耦是面向新型DCI传输系统的发展趋势, 建设方选择最佳组件、灵活扩容、自主可控, 利于降低单位数据承载造价。

OpenZR和OpenROADM定义了光模块的特性、板卡指标、管控接口, YANG模型、Netconf定义了管控标准, 无论是遵守以上协议, 还是自定义我国开放协议, 只有遵从相同标准, 才能实现同一系统的光电解耦、光光解耦。利用SDN的开放性和集中控制的特点, 实现对DCI系统、多域组网环境的统一管控, 实现网络动态智能控制, 提高业务开通速度^[2]。

3 网络建设优化

3.1 网络重构

传统DCI依托运营商传输网, 属于运营商网络末端的延申, 流量通过运营商节点中转、绕行, 增加了时延。

新型DCI网络建设应打破行政区划, 以各等级DC机房为核心、汇聚节点组网, 减少数据绕行、优化时延, DCI传输系统与原有传统传输系统在省核心节点打通, 用于承载DC至其他外网节点的时延不敏感流量, 新组网架构见图1。

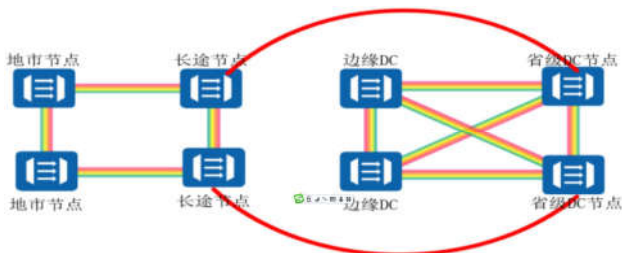


图1 新型DCI传输系统组网架构图

3.2 光缆优化

时延是影响用户对数据中心业务体验的关键因素, 不同类型的服务对时延的要求差异较大, 而业务时延主要由光缆距离决定, 传输设备侧的时延累计一般不超过1ms。现有长途光缆建设是从一个端站敷设至另一个端站, 中间设置若干中继机房, 中继机房一般是在市区附近的支局、基站机房, 以上节点至市区核心节点的光缆在市区绕行距离较长。

为缩短DC节点之间的光缆距离, 可在离市区最近的两个中继节点之间敷设直通光缆, 将这两个中继机房本应设置的OA替换为城郊调度OXC, 优化方案见图2。

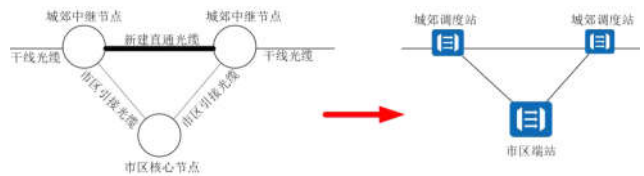


图2 面向DCI传输系统的长途光缆结构优化

本方案可用于智算节点与传统节点混合组网, 优点是可将DC之间业务通过两个城郊调度OXC站点转发, 不再经过市区端站, 每个城市可缩短至少30km的传输距离。但缺点也很明显, 两端OXC的设备造价要高于两端OA设备^[1]。

4 结语

算力网络作为新型数字信息基础设施, 已成为支撑数字经济持续发展的重要引擎, 随着“数字中国”战略的扎实推进, 产业数字化将加速发展, 催生千行百业数智赋能, 算力网络发展将面临千载难逢的好机遇。作为算力网络的光底座, DCI传输系统承担着基础管道的使命, 关键在于实现“光算协同”。

首先, 新型DCI传输系统在技术上应具备光谱超宽能力, 单波速率演进至400G, 具备支持C6T+L6T的能力。在光模块、WSS、信号放大器等光器件上, 也应具备C+L波一体化能力, 尽量避免光器件分设。

其次, 在功能设计上应匹配DC机房, 采用散热能力更优的风道设计, 实施节能技术, 打造超高能效的新型DCI传输平台, 实现绿色低碳、模块堆叠等目标。

最后, 在管控系统上应全网传输资源可视、可管、可控, 具备光电开放解耦能力, 打造资源一体化调度能力。

参考文献

[1]YD/T 4297《集成型可重构光分插复用(ROADM)设备技术要求》中华人民共和国工业和信息化部;
 [2]《全国一体化大数据中心协同创新算力枢纽实施方案》发改高技〔2021〕709号;
 [3]《算力时代网络运力研究报告(2022年)》中国信通院;