Abstract: In this paper we consider the aspect of choosing an efficient and cost-effective technology to connect edge processing Central Offices to core and metro data-centers for the specific purpose of meeting OTT traffic requirements. To this end, we introduce a demarcation device with coherent optical DWDM interfaces – a Data Center Interconnect (DCI) Xponder. which has the dual function of transponding and muxponding of data as a plausible technology choice in the virtualized OTT traffic environment. Direct DWDM wavelength interconnect via DCI-Xponders (X for both transponder and muxponder functionality) is hedged against more traditional packet solutions such as IP and MPLS networking. To this end, we analyze through an optimization model the impact of the DCI-Xponder for meeting OTT requirements in virtualized metro networks. Simulations show sizable cost benefit in using DCI-Xponders with comparable performance to an IP/MPLS network. Further we consider the impact of a packet switch as data-center gateway combined with a DCI-Xponder, and observe that the performance of such a solution is comparable to an IP/MPLS network but with the advantage of severe cost reduction.

I. INTRODUCTION

Application traffic, especially from Over-The-Top (OTT) players, has witnessed an increase of 2x every other year. Much of this traffic begins at an end-user device, and terminates in one or more data-centers (DCs), which are usually owned by OTT operators. Much of the communication cost goes in transporting the traffic from the user to the data-center and back. Minimizing this communication cost is of interest to all the three parties involved: the OTT player (for whom this communication cost is an operational expenditure part of his business cost); the ISP (who does not make much revenue from OTT services); and the user (who has to pay for the access bandwidth). There are two models that have come about for interconnecting the user to the DC. In the first model, the OTT player leases bandwidth from the ISP into its DC, and it is the job of the ISP towards ensuring that users are connected to the DC. In this model, the ISP obtains a fixed revenue from the user for last mile connectivity, and usually a bandwidth usage-based revenue from the OTT towards the DC connectivity. The ISP has to invest in large parts of the network, and usually due to the uncertainty involved in figuring out which customer requires what type of service, and that to at a particular location, the ISP ends up over-provisioning the network with an alphabet soup of protocols and a variety of equipment. This means that for much of the OTT traffic, from the user to the OTT DC, there are functions of aggregation, disaggregation, switching, routing and transport that are thrust upon the traffic, when in actuality, if properly engineered, only transport of the packet stream is needed. In the second, and more recent model, the OTT player builds multiple large and small (core and metro) DCs, which are used to store its application data and the problem then becomes connecting the user to the nearest DC. The DCs (metro and core), are further interconnected across information highways, which are large pipes of bandwidth, usually at 100Gbps or more that periodically exchange content thus enabling and facilitating the end user to be connected to the right VM in the nearest DC. In this model as well, the ISP has to build and manage these high-speed pipes between DCs using traditional equipment of switches, routers and ROADMs implying that the cost per bit is high. Hence it is of interest to the entire OTT ecosystem (of all the 3 players), to reduce the cost-per-bit of this transport of information.

A variation to this model, is the use of the Central Office (CO), which now becomes a small DC that houses information and VMs. In this case, the CO is used as an edge computing platform that houses VNFs and regularly caches used user data, which makes user experience significantly better than any of the previous deployed mechanisms. The key then is to enable rapid connectivity between the CO and the DC, which is now of paramount importance from the perspective of seamless user experience across multiple data and video applications that reside in either of the CO or DC. We desire to then transport huge chunks of data between the CO and the DC, using low-cost and efficient equipment.

There are multiple ways to connect the CO to the DC, and it is a well-established fact [3, 4] that keeping the data as low as possible in the stack ensures minimal cost, minimal latency and free from vulnerabilities. To this end, direct data-center interconnect (DCI) [1-2] is a recently proposed transport artifact that facilitates communication between the Internet and the metro data-center (DC). This approach involves multiplexing across channels as well as consolidation of data. The DCI has gained significant traction in recent times with major deployments for both operators and OTT players. A particular implementation of interest to us, which we shall term the DCI-Xponder includes a multiplexing device and a transponder (transport) device combined and then tipped with coherent optics. Such a DCI-Xponder has the capability of taking in multiple packet streams using grey optics and mapping these to appropriate ports using coherent optics [10,
The DCI-Xponder is ideally suited to connect top-of-the-rack servers to aggregation routers in a DC.

Within the DC typically is a leaf-and-spine topology, whereby nodes in a rack are connected to top-of-the-rack switches that act as leaves, which are further connected to larger port-count switches (spines). Such a leaf and spine arrangement is then backhaul-connected to content delivery networks or the Internet through a DCI network across a border/ gateway router and firewall. The border router and firewall safeguard the DC against cyber-attacks and provisions connections between legitimate external users and VMs in the DC. Such connections also facilitate Service Function Chains (SFCs) [5, 6] within the DC across multiple servers.

The DCI-Xponder plays a pivotal role in facilitating communication to the DC: (1) the DCI-Xponder facilitates uninterrupted communication between users and servers in the DC through aggregated bandwidth pipes, which are usually provisioned wavelengths; (2) the DCI-Xponder facilitates multiple service providers to connect to the DC using wavelength and fiber diversity; and, (3) the DCI becomes a single point of interconnection to multiple providers.

A key goal of this paper is to evaluate whether the DCI-Xponder as a technology has been fully exploited and its role vis-à-vis the rest of technology elements such as switches, routers and ROADM. Our specific case study is an open network based virtualized Central Office (CO), where computing at the edge includes virtualized appliances [8, 9] and some cached data, from which the CO is directly connected to the metro DC using a DCI network. We evaluate the performance of such a network in which, the Central Office (DC) at the edge is connected directly or through a minimal number of hops, to a metro DC for OTT traffic applications.

In Section II we describe the network architecture that encompasses a typical OTT provider in access and metro environments. Section III presents a constrained optimization and abstraction of the topology design problem. Section IV provides for numerical results comparing DCI-Xponder with IP and MPLS technology from cost and performance perspectives. Section V concludes the paper.

II. THE NETWORK ARCHITECTURE

Shown in Fig. 1 is an example of a classic metropolitan network that consists of multiple DCs, provider routers, multi-degree ROADMs, OTN cross-connects and Ethernet switches along with COs connected to users. Shown in Fig. 2 is a schematic of a typical DCI Xponder (transponder/muxponder) inserted in the network of Fig. 1, with minimal use of routers etc. Fig. 2 is our target design and the larger goal of this paper is to compare cost and performance of the network in Fig. 2 to the network in Fig. 1. The DCI-Xponder is introduced in Fig. 2. The DCI-Xponder architecture is also shown in Fig. 2. Fig. 2 shows how a DCI-Xponder can be used at the COs as well as DCs in an underlay model, whereby, routers at intermediate nodes can potentially be avoided. A DCI-Xponder consists of client-side ports that supports grey-optics and network-side ports that support colored optics connected to long-reach fibers. The DCI-Xponder multiplexes several packet streams and maps these to metro-reach capable wavelengths. As such, the DCI-Xponder works entirely in layer 0-1 domain implies that the cost-per-bit is much lower than a corresponding switch or router. Naturally, both the equipment types (switch/router and DCI-Xponder) have a paradigmatic difference in functionality, it is still worth comparing the use of these in different network topologies. Our argument is that switches and routers, though they provide fine grained packet processing, have limited use for the particular case of OTT traffic in operator networks. This argument is specifically valid when there is some degree of edge-processing (such as in a CORD environment [3]). Intuitively, we argue and will show through simulation and analysis that once data is processed at the edge of the network on to appropriate wavelengths, then there is not much use of further packet processing. The network architecture that we consider, consists of COs that have open networking functionality [12] and DCs. In the core of the network, the DCs are architected such that DCI-Xponders serve as interface points between the DC and the network. The DCI-Xponders are backhaul-connected to spine switches, which are further connected to servers through leaf switches. Essentially, there are optical circuits (wavelengths) provisioned directly from the CO to the DCI. For redundancy, each CO is assumed to be connected to at least 2 DCs.

In comparison, our reference design (the classical solution in Fig. 1) consists of nodes in the metro network that have P-routers for IP routing functionality. A P-router at the metro/core node also connects to the DC. Further, in this architecture, COs have a variety of processing equipment and are connected to a PE router, which is then backhauled onto a metro network through ROADMs and P-routers, with the application data multihopping and reaching the DC. In the past, this architecture has often been chosen due to operational issues (such as variations across deployment workforce implying the need for various equipment types). This architecture also supports statistical multiplexing. These features are now mitigated due to edge processing through the use of Virtual Network Functions (VNFs) [7, 8] at COs and DCs.

For conducting a thorough comparison between the two architectures we show in Table 2, a cost comparison of DCI-Xponder transport vis-à-vis other technologies on a cost-per-bit basis. Our eventual goal remains to compute which architecture leads to a lower cost, and measure the performance of both solutions from a provisioning perspective. To this end, an optimization model that computes the cost of a DCI+CO based network is presented.
III. ABSTRACTION AND OPTIMIZATION MODEL

The use of open DCI-Xponders in networks is essentially a topology design problem, where DCI-Xponders become a logical underlay that feed to a layer 0-1 (ROADM based) network while providing express pathway capability to an overlay of IP routers. The idea is to map the aggregated data at a CO on to a correct wavelength, such that this wavelength will guide the data stream to the appropriate DC. The OTT players’ VNFs in the DC then terminates the service. In this process, we absolve the need for IP routing, Ethernet switching and OTN switching, at intermediate nodes, while providing end-to-end optical path from the CO to the DC. The goal in our optimization formulation is to consider the costs involved in (1) a DCI-Xponder network; (2) a DCI-Xponder+packet switch network and (3) an IP/MPLS network. The ROADMs continue to be common elements across all the three network architectures. To this end, we consider the virtual topology problem with parameters shown in Table 1.

| $G(V,E)$ | Physical graph of $|V|$ nodes and $|E|$ edges |
| $T_{abkm}$ | The $m^{th}$ traffic connection between nodes $V_a$ and $V_b$ on the $k^{th}$ service type, |
| $d_{abkm}$ | =1, if traffic connection $T_{abkm}$ is routed through DC $d$, = 0 otherwise. |
| $c_{dci}$, $c_{IP}$, $c_{OTN}$, $c_{SW}$ | Cost per bit for DCI, IP, OTN and Switch respectively |
| $\beta_{IP}$, $\beta_{SW}$, $\beta_{DCI}$ | =1, if IP router or Ethernet switch or DCI at node $V_j$, = 0 otherwise |
| $\{R_{ab}\}$ | Set of nodes along the shortest route from $V_a$ to $V_b$ |
| $\gamma_{abkm}$ | = 1, if the traffic $T_{abkm}$ has to terminate at DC $d$, = 0 otherwise. |
The objective function is: \( \forall \theta_{abkm} > 0 \),

\[
\min \sum_{a,b,k,m} T_{abkm} \cdot \left[ \beta_{DCI}^j \cdot c_{DCI} + \beta_{IP}^j \cdot c_{IP} + \beta_{SW}^j \cdot c_{SW} \right. \\
+ \left. \beta_{OPT}^j \cdot c_{OPT} \right]
\]

The objective function minimizes cost of the entire network, inclusive of all equipment that is possible.

Subject to the following constraints:

**Delay constraint:** \( \forall \theta_{abkm} > 0 \)

\[
\delta_{CO(j)} \cdot \beta_{DCI}^j + \delta_{SW} \cdot \beta_{SW}^j + \delta_{IP} \cdot \beta_{IP}^j \leq \Delta_{abkm}
\]

The delay constraint ensures that the total delay for each service is less than a given threshold.

**CO size constraint:**

\[
\sum_{a,b,k,m} \frac{T_{abkm} \cdot \beta_{DCI}^j}{n} \leq n \cdot S_{CO}
\]

The above constraint ensures that the amount of data allocated to a CO is no more than the maximum size of the CO.

**Switch constraint:**

\[
\sum_{a,b,k,m} T_{abkm} \cdot \beta_{IP}^j \leq n \cdot S_{SW}
\]

The above constraint ensures that an Ethernet switch at node \( j \) can accommodate transit + add/drop traffic.

**ROADM Capacity constraint:**

\[
\sum_{a,b,k,m} T_{abkm} \cdot \beta_{OPT}^j \leq n \cdot S_{OPT}
\]

The above constraint ensures that a ROADM at node \( j \) can accommodate transit + add/drop traffic.

**IP Router size constraint:**

\[
\sum_{a,b,k,m} T_{abkm} \cdot \beta_{IP}^j \leq n \cdot S_{IP}
\]

The above constraint ensures that an IP router at node \( j \) can accommodate transit + add/drop traffic.

**Routing and wavelength assignment:**

\[
\forall \theta_{abkm}, \quad \alpha_{abkm}^p = 1, \quad p \in \{R_{ab}\}
\]

If the \( p^{th} \) wavelength is used for connection \( T_{abkm} \), leading to the constraint: \( \sum_{a,b,k,m} T_{abkm} \cdot \alpha_{abkm}^p \leq C_{i(p)} \), where, \( C_{i(p)} \) is the capacity of the \( p^{th} \) wavelength.

**DCI-Xponder constraint at a CO.**

We say, \( \forall j \in \{ CO \} \) and \( \theta_{abkm} = 1 \)

\[
\sum_{a,b,k,m} T_{abkm} \cdot \beta_{DCI}^j \leq n \cdot S_{DCI}
\]

where \( \{ CO \} \) is the set of all COs.

Further, \( \forall j \in \{ D \} \) and \( \theta_{abkm} = 1 \),

\[
\sum_{a,b,k,m} T_{abkm} \cdot \alpha_{abkm}^d \leq n \cdot S_{DC}
\]

where \( \{ D \} \) is the set of all DCs.

The size and number of DCI-Xponder at an ingress CO is upper bounded to the size of the ROADM

\[
\frac{1}{c} \sum_{a,b,k,m} T_{abkm} \cdot \alpha_{abkm}^d \leq 5000
\]

Where, \( S_c \) is the capacity of a server.

### IV. RESULTS FROM NUMERICAL EVALUATION

To evaluate our DCI-enabled virtualization hypothesis and compare with current IP-based non-virtualized architecture, we developed a Matlab based optimization model that was called by a Python-based event simulator. We assumed a two-city network with each metropolitan city having 10 and 14 nodes, 2 and 3 DCs. Each DC was modeled to have 5000 servers. In addition to the DCs, we also modeled edge COs, such that a typical CO had up to 30 VNF types [3]. For comparison, modern virtualized COs were assumed with up to 50 servers of 2x10Gbps IOs. Other metro equipment is modeled as follows: P-routers are assumed to have non-blocking cross-connects of size up to 6.4Tbps with 10 and 100Gbps IOs. ROADM are assumed to be CDC (colorless directionless, contentionless) with up to 192 wavelengths and support for up to 6-degrees of fiber directions. The PE routers have a non-blocking cross connect size of 1 Tbps with 10Gbps and 100Gbps IOs. OTN switches that can process from 1Gbps and 10Gbps ports. At the intersection of metro and long-haul network, we assume OTN switches that can process from 2.88Tbps to 25.6 Tbps of traffic. The DCI-Xponder is assumed to have inputs of up to 16-64 wavelengths per card with an output that consists of colored lambdas at rates 10Gbps to 400Gbps (coherent).

The optimization model runs on a i7 Core 8th gen. machine and takes on average 45 seconds to produce a single iterative solution.

Load is computed as the average occupancy in the network compared to the sum of the capacity of all the edge

<table>
<thead>
<tr>
<th>Table 2: Cost Assumptions.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>10Gbps per port cost</td>
</tr>
<tr>
<td>100Gbps per port cost</td>
</tr>
<tr>
<td>200Gbps per port cost</td>
</tr>
<tr>
<td>400Gbps per port cost</td>
</tr>
<tr>
<td>IP/MPLS router</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1.5</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>OTN switch</td>
</tr>
<tr>
<td>0.25</td>
</tr>
<tr>
<td>0.35</td>
</tr>
<tr>
<td>0.5</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>DCI-Xponder</td>
</tr>
<tr>
<td>0.1</td>
</tr>
<tr>
<td>0.2</td>
</tr>
<tr>
<td>0.3</td>
</tr>
<tr>
<td>0.5</td>
</tr>
<tr>
<td>Ethernet Packet switch</td>
</tr>
<tr>
<td>0.4</td>
</tr>
<tr>
<td>0.75</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1.5</td>
</tr>
</tbody>
</table>

models.
interfaces divided by the average hop count. Load is hence normalized in the range [0,1].

Fig. 3. Percentage savings.

Fig. 4. Per-node cost comparisons.

Fig. 5. Drop percentage of traffic.

Traffic is assumed to be only OTT type traffic (no enterprise traffic), since we want to evaluate the limited question as to which architecture suits OTT traffic better. Though the traffic is only OTT type – i.e. beginning from a user in the access network and terminating in the metro DC or in 15% of the cases within the COs, the traffic is further subdivided into multiple traffic types: video, data, text messaging, and wireless backhaul. The traffic differentiation is twofold: bandwidth and delay. 70% of the traffic is video, 30% is data and other services. Video traffic has a maximum delay tolerance of 5 ms exclusive of propagation delay, while data traffic has a delay tolerance of 100ms.

**Main result:** Our main objective is to compare the DCI-Xponder solution to the IP router-based solution. Fig. 3 is a plot showing the capital cost of the DCI-based solution compared to a vanilla IP solution. Fig. 4 is based on the assumptions shown in Table 1 (where we compare the unit costs of the various technologies). To generate Fig. 3, we ran the optimization model on the topology 10-times for each load value, and averaged the results. The standard deviation is about 14%. We plot 3 sets of results for: a single CO directly connected to the DC, multihopping from one ingress CO to the DC through another CO (which we call as the 2-CO case) and finally multihopping through multiple COs. Intuitively the single CO (connected to a DC) will perform the best, followed by the 2 CO solution (in which we hop from one CO to another CO and then to a DC). The results in Fig 3 in fact exactly show exactly this behavior. We also modeled the >2CO solution and observed (though not plotted here) some cost savings compared to pure IP/MPLS due to virtualization of resources. The pertinent point to note beyond the intuitive results is that there is an increase in savings with load, followed by a saturation point – after which the benefits begin to recede (due to wastage of resources at high-loads – over-provisioning). The single CO-connected to the DC solution works best when the services have high-bandwidth (1Gbps and beyond), in which case, by simply putting a DCI-Xponder at a CO allows the application data to be put on an express wavelength towards the appropriate DC. For sub-1Gbps services, (average of about 25Mbps), the 2CO solution works best – there is enough multiplexing gain across the 2 multi-hopped COs and the average cost-benefit is about 34%.

In Fig. 4, we plot comparative per-site costs for the IP, MPLS and DCI-Xponder+packet switch solutions. A packet switch is assumed to provide packet aggregation at each CO so that the traffic is shuffled at the ingress CO and correctly placed on to a wavelength that is one-or-two CO hops away from the corresponding DC. This plot was constructed from the optimization algorithm tweaked for each individual solution. For the IP-only and MPLS-only solutions, we assumed, IP-over-ROADMs and MPLS-over-ROADMs, with colored optics as transponders plugged into IP and MPLS routers. Load at a metro site is varied from 1Tbps through to 25.6 Tbps, and costs are computed across the entire network, and then averaged for a single node. The node-costs are illustrated in Fig 4. The DCI-Xponder+packet switch cost is the lowest among the three curves. MPLS-only cost is about 30% more than the DCI-Xponder+packet switch cost. The IP-router-only solution is about 25% more expensive than the MPLS solution. The interesting aspect is that both the MPLS and IP solutions exhibit step-wise increases. For the MPLS solution, the steps may not appear as distinct, but that is due to the resolution of the figure. However, for the IP solution the steps are distinct. For both the cases of IP and MPLS, at load points of 0.2, 0.5 and 0.8 there are significant jumps in costs as the number of equipment chassis required at a node site changes. This behavior is interestingly not exhibited by the DCI-Xponder solution. The near constant gradient for the DCI-Xponder+packet solution is because of the low-cost of the solution per-se, which implies that churn in traffic has minimal impact over the solution.

In Fig. 5 we measure the amount of dropped traffic. Naturally the packet solutions perform better, but only slightly (and that too only at high loads). For this plot, the measurements were taken across 1 million services, with average granularity 100Mbps. Even at high loads and dynamic traffic (randomly
selected source-destination pairs), the loss in the DCI+CO (with IP router/packet switch as BNG) is 10E-6, which is quite acceptable for TCP or OTT traffic. More importantly for typical provider operating loads (<0.6), the loss is close to the more expensive non-virtualized pure IP network.

In Fig. 6 we plot the number of alarms that are raised in a DCI+2CO network comparing with a packetized IP network with the same service mix. Our objective was to get an indication on the operational characteristics of the different approaches. To generate this figure, we considered the connectivity matrices in the CO, IP, Ethernet, OTN and ROADM sizes and introduced an alarm rate of one-alarm per 3 months-per-card, with no direct relation with the number of services, but with the total amount of traffic. The DCI approach has on average 50% lesser alarms than IP/MPLS due to a smaller number of cards used to build the complete solution.

Fig. 7 is a plot of selecting an appropriate size of a packet switch at the CO. We consider 3 switch sizes: a 100Gbps, a 500Gbps and a 1Tbps non-blocking cross-connect. For the 100Gbps switch, we have 10x10Gbps IO ports at the client side multiplexed into a single 100Gbps backhaul port. For the second and third types, we have multiple 100Gbps backhauled ports connected to the DCI-Xponder. We compare the cost of this solution to the IP solution. Fig. 7 shows that the 1Tbps solution has the least improvement for our traffic types, essentially due to constant overprovisioning and high residual capacity. The 100Gbps solution does well at medium loads, but too many of those instantiations are required at high load, and hence the performance degrades. The 500Gbps solution is most suited for our traffic mix with an average 37.8% improvement in cost (lower cost) compared to the IP solution.

V. CONCLUSION

We study the problem of provisioning OTT traffic in metro and access networks, specifically taking into consideration edge processing at access Central Offices. We model the access and metro network for OTT providers with various technology mixes. Our study introduces DCI-Xponders as a coherent optical device that facilitates multiplexing of edge-processed packet streams appropriately mapping these to metro and core-centric wavelengths. We first show a network architecture that uses DCI-Xponders as a technology building block for virtualization in the metro domain, facilitating the edge COs to be directly connected to DCs. We build an optimization model that facilitates mapping traffic to such a virtualized network, also comparing this network to a vanilla IP and MPLS network. Through simulations we show that DCI-Xponder is cost-wise better for OTT operators especially when the termination of the application is at a DC. This approach mitigates the need for duplication in packet switching and aggregation functions that are already provided as part of the CO/DC installations while requirements for packet-drop and for operational simplicity are met by this greatly simplified networking approach.

REFERENCES