Routed Optical Networking: an alternative architecture for IP+Optical aggregation networks

Valerio Viscardi
Mass-Scale Infrastructure Group
Cisco
Milan, Italy
vviscard@cisco.com

Dirk Schroetter
Mass-Scale Infrastructure Group
Cisco
Richardson (TX), USA
dschroet@cisco.com

Moustafa Kattan
Mass-Scale Infrastructure Group
Cisco
Dubai, UAE
mkattan@cisco.com

***Abstract*—For many years the best strategy to optimize the Total Cost of Ownership (TCO) of an IP+Optical network has been to reduce as much as possible the utilization of IP routers' switch fabrics and interfaces. This can be achieved by means of optical by-pass using Reconfigurable Add/Drop Multiplexers (ROADMs). This strategy comes at the cost of a suboptimal wavelength utilization and longer (on average) optical links, running with a lower OSNR. In this paper we analyse alternative architectures which take advantage of the latest Network Processing Units (NPUs) in IP routers and pluggable 400G DWDM interfaces, which helps reducing the cost associated to packet processing.**

***Keywords*—IP, Optical, QDD, pluggable, modelling**

I. Introduction

ROADMs have been a standard building block in Optical Transport Networks. The major advantage of ROADMs is the possibility to bypass IP routers and connect source and destination with a single hop, saving in this way router fabric capacity and interfaces [1].

But the recent evolution of pluggable technology, integrating the coherent DSP on-board for coherent DWDM transmission, is completely changing the paradigm. As routers NPUs and interfaces increase their throughput, taking advantage from semiconductor manufacturing process node, resulting in a lower cost per bit, the saving of these resources is not necessarily translating in the lowest Network CapEx and OpEx. On the contrary, we found that a hop-by-hop approach can help reducing total network cost, simplifying the integration and the management of multi-vendor networks.

In this paper we present a model that allows calculating the required number of interfaces in open ring or bus network architectures, which are typically found in access and aggregation networks, and we compare strict hop-by-hop service provisioning against a full-bypass approach. We then apply this to a synthetic, statistical, model derived from realistic network deployments and we compute the total network cost, identifying areas of convenience, where hop-by-hop can be more efficient than bypass.

II. Aggregation Networks

A. Network Topology

The network topology we will be using in this paper is the linear bus topology (also "open ring"): not only can the traffic easily described thanks to the simplicity of the routing, but it is an important topology for access and aggregation networks where all the aggregation nodes need to be interconnected to the two hub locations (for redundancy).

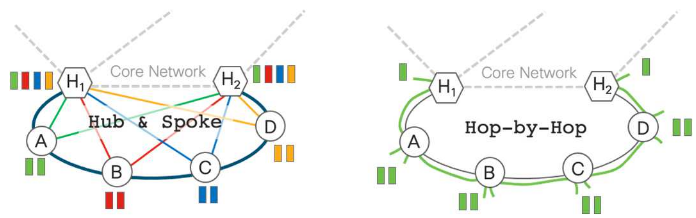

Figure 1: Aggregation network architectures – hub & spoke and hop-by-hop

The open ring is in fact the simplest topology which both minimize the amount of required fibre and still provides a redundant path to destination via dual homing the traffic.

Our goal is to compare two architectures:

1. *Hub & Spoke*: each aggregation node (the spoke) is connected to the two hubs by means of a dedicated wavelength to the two destination hubs (H1 and H2). The intermediate nodes propagate the wavelength to destination by means of Optical Add/Drop Multiplexers (OADMs).
2. *Hop-by-hop (daisy chain)*: each aggregation node delivers the traffic to its two adjacent next hops, where it is propagated via the L3 fabric.

The two architectures are shown in Figure 1. From the pictures we can clearly anticipate their pros and cons: while *hub & spoke* reduces the utilization of router fabrics in all the aggregation nodes to just the add/drop traffic (i.e. there is no transit traffic to be processed by the L3 fabric), it requires one dedicated interface per aggregation node in the two hubs. In addition, the long path to destination is affected by both propagation losses and concentrated losses in the OADMs, which may require a DWDM layer with optical amplification to overcome the overall optical path loss. Last but by no means the

978-3-903176-44-7 © 2022 IFIP

least, wavelength utilization is sub-optimal in case of low traffic requirements.

Hop-by-hop can provide resource savings in the two hubs, as wavelengths are shared among the aggregation nodes. This approach can also simplify the optical design of the network as only single hops need to be engineered (there is no optical bypass of nodes). This architecture though makes more use of fabric capacity in the intermediate nodes and quickly scales up the interface count when the total traffic in the network exceeds the interface capacity.

We can therefore anticipate that *hop-by-hop* provides benefit at low traffic loads while *hub & spoke* is more beneficial at high traffic loads. We will now try to quantify this.

B. Dimensioning the fabric and interface requirements: uniform traffic

To calculate the traffic into the network we can start by simplifying the problem and assuming each aggregation node in the network is dual homed to the two hubs with an average capacity *C*. The traffic is uniform, which is a good starting point to compare the two architectures. Assuming we have *N* nodes in the network, the total number of aggregation nodes is *N-2* (2 are the hub nodes) and the total traffic in each section of the ring is:

$$(N-2)\cdot C \tag{1}$$

This is true both in the case of hop-by-hop and in the case of optical bypass (hub & spoke). The main difference between the two cases is in the total interface capacity that needs to be allocated in the routers.

In the case of *Hub & Spoke*, the hub nodes (*node 1* and *node N* in Figure 2) need to allocate enough interfaces to support the total traffic defined in (1). Each spoke node needs to support enough interfaces to terminate the local traffic, $2 \cdot C$. This is shown in (2) and (3):

$$TrunkI/F_{node_i}^{i=1,N} = (N-2)\cdot\left\lceil {}^{C}/_{C_{I/F}}\right\rceil \tag{2}$$

$$TrunkI/F_{node_i}^{i\neq 1,N} = 2\cdot\left\lceil {}^{C}/_{C_{I/F}}\right\rceil \tag{3}$$

Where $C_{I/F}$ is the throughput supported by the interface (e.g. 400 Gbps) and $\lceil\ \ \rceil$ is the ceiling function (e.g. round-up).

We can similarly calculate the number of required interfaces for the *hop-by-hop* architecture:

$$TrunkI/F_{node_i}^{i=1,N} = \left\lceil {}^{C}/_{C_{I/F}}\cdot(N-2)\right\rceil \tag{4}$$

$$TrunkI/F_{node_i}^{i\neq 1,N} = 2\cdot\left\lceil {}^{C}/_{C_{I/F}}\cdot(N-2)\right\rceil \tag{5}$$

Required fabric capacity can be similarly calculated for all the nodes in the two cases. We can now build a parametric model where we calculate the total required capacity and CapEx as a function of average node capacity *C* and total number of nodes in the network *N*.

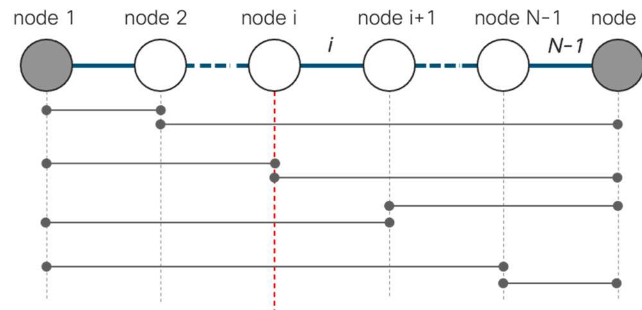

Figure 2: Aggregation Network Topology

III. CapEx analysis results

We can now build a parametric model to compare the two architectures when increasing the number of nodes *N* and the average node capacity *C*. The model has been built using relative pricing, similarly to what described in [1][2]. All the models built in this paper are based on 400G ZR+ modules, to allow propagation over longer distances. The relative price between the pluggable and the transponder is between 2.5 and 3, in-line with what used in [2].

A. Single network

The first step is to focus on a single aggregation network and evaluate the CapEx savings (or additional costs) to build the same network using a hop-by-hop approach vs. a hub & spoke one. For now, we focus on calculating how many optical interfaces are required, either 400G ZR+ pluggables or transponders, assign a cost to them and calculate the network cost. We do not include for now the cost of DWDM (when required) and the cost of the linecards. These will be factored in when we build the synthetic network example (see section B).

To build a model we grow node capacity *C* from 10 Gbps up to 800 Gbps per node. Similarly, we increase the number of nodes from 3 to 12, which means the total number of aggregation (spoke) nodes goes from 1 to 10 (i.e. *N – 2*).

For each one of this (*N*, *C*) combinations, we calculate the total number of required interfaces and fabric capacity

We can calculate the total number of required interfaces and fabric capacity for all the combinations (*N*, *C*) and assign a total network cost for the two architectures. We can then plot a differential heatmap where each element of the matrix is equal to the difference between the *hop-by-hop* and *hub & spoke* total network costs

The results are shown in Figure 3: negative numbers indicate (*N, C*) combinations where a hop-by-hop architecture is cheaper. Positive numbers indicate (*N, C*) combinations where an optical bypass approach gives savings. We can identify an area, indicated by the red oval, that represent the ideal application for hop-by-hop routing. This simple analysis indicates that hop-by-hop can be an effective way to build aggregation network when the average capacity per node C is up to 120 – 140 Gbps and the number of aggregation nodes is up to 5 (7 nodes total including the two hubs).

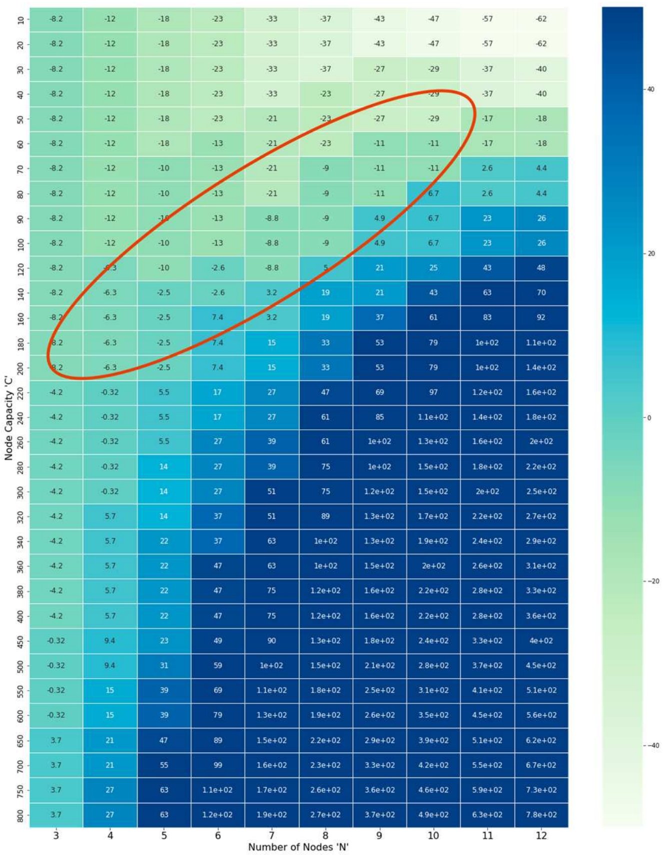

Figure 3: Hop-by-hop vs. Hub & Spoke differential cost heatmap

Increasing the average node capacity beyond this level would require the deployment of additional interfaces to the network. Every time the total network traffic exceeds the deployed capacity, we must deploy 2 · $(N - 1)$ additional interfaces to the network.

Increasing the number of nodes has a similar impact on total network cost: total network traffic is now spread among many nodes, resulting in a lower available average capacity. This quickly requires deploying additional capacity and, as explained above), the number of required interfaces increases linearly with the number of nodes N.

While this model gives a good indication about area of potential savings, this does not include the impact of router fabrics and DWDM equipment to overcome the insertion loss due to propagation in the fibre and discrete losses through OADMs. While router capacity can be easily calculated based on network traffic and routing architecture as explained above, to dimension an optical layer we need to make some assumption about the average distance between nodes. A good way to do this is to build a synthetic network based on average data from real networks, providing in this way a good representation of the problem.

B. Syntetic network analysis

Our goal is to build a synthetic model which is representative of the interconnection of all the aggregation networks to the core network. An example of this aggregation-to-core interconnection is shown in Figure 4. Each core node typically terminates multiple aggregation networks and has a dual function: hub for all the aggregation networks and transit core function. Our focus is to dimension the aggregation side.

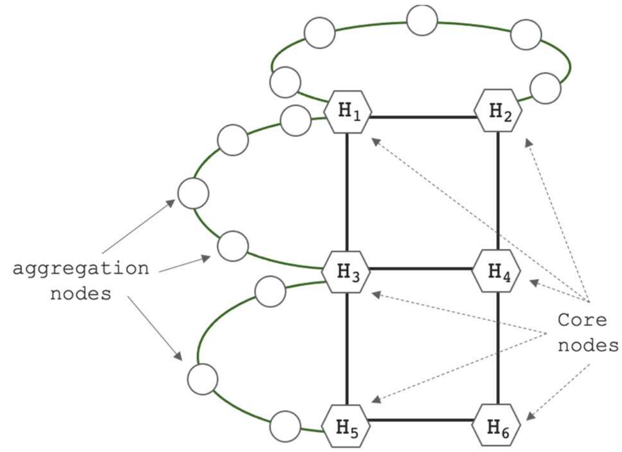

Figure 4: Core and aggregation networks

Aggregation networks can be classified based on their geographic location, to properly account for node distances.

The proposed model divides aggregation networks into six classes, based on their location: cities, urban areas, rural areas. Table 1 shows the six classes in which we have divided all the aggregation networks. For each network class the table shows the number of aggregation nodes (N – 2), the average distance between adjacent nodes, the maximum optical path length to either of the hub nodes and the relative weight of the class used in the model.

Network Class	**Network Characteristics**			
	***N* – 2**	***Avg. node distance***	***Avg. max. optical path***	***Relative weight***
Main city	4 – 5	< 10 km	< 30 km	20%
Medium city	3 – 4	< 20 km	< 60 km	22%
Small city	2 – 3	< 10 km	< 25 km	18%
Urban area	4 – 5	< 40 km	< 120 km	15%
Rural area	3 – 4	< 130 km	< 250 km	15%
Remote area	2 - 3	< 180 km	< 250 km	10%

Table 1: Aggregation network classes

The relative weight is used to quantify how many networks of this type are present, expressed as percentage of the total. Assuming 500 networks in total (roughly 1500 aggregation nodes, in total which is a reasonable number for many countries) the proposed model would consider 100 main city, 110 medium city, 90 small city, 75 urban area, 75 rural area and 50 remote area networks.

This model is parametric, and we can modify both the total number of networks in the model and their relative weight. Changing the weight allows adapting the model to countries which have a large extension of low density populated areas (e.g. Scandinavia). On the contrary, other countries might be mostly urban (e.g. Benelux) and require more weight moved to metro network classes. Density is somewhat indicated by the average distance between aggregation nodes: metropolitan networks have much shorter distance between adjacent nodes, which do not require optical amplification in the case of *hop-by-hop* routing, while *hub & spoke* might need amplification depending on the add/drop structure of the OADM nodes. Rural and remote networks are sparser and may have at least one span between two adjacent nodes exceeding the power budget of the interface, forcing the use of optical amplifiers even in the case of single wavelength applications (e.g. dark fibre). *Hop-by-hop* can therefore provide savings when building the DWDM optical infrastructure as in most of the network classes there is no need to install any optical amplifier, as L3 forwarding in the routers provides also signal regeneration.

Network Class	**Node's L0 Composition**		
	Hop-by-Hop	***Hub & Spoke***	***L0 Savings***
Main city	Passive Mux/Dmx Passive OADM	Passive Mux/Dmx Passive OADM Optical Amplifiers	80%
Medium city			80%
Small city			80%
Urban area			80%
Rural area	Passive Mux/Dmx Passive OADM Optical Amplifiers		40%
Remote area			10%

Table 2: Optical infrastructure (L0) savings

Table 2 compares each network class and indicates the required optical infrastructure components. As indicated, *hop-by-hop* regenerate the signal at each hop and does not require any optical amplifier in 4 network classes out of 6. Amplifiers are always required in rural and remote areas, even though not on every span, but the savings are reduce compared to metropolitan and urban areas. *Hop-by-hop* optical CapEx savings (i.e. L0) are indicated in the table.

We now have all the data to build a network level model. The results are shown in Figure 5: total network costs (CapEx) for both *hop-by-hop* and *hub & spoke* architectures are plot as a function of average node capacity *C*. The network CapEx is reported as arbitrary unit (a.u.), being built using relative prices, as indicated in section III.

The two networks cost line cross when the average node capacity is in the range 200 – 300 Gbps. If average node capacity is lower than this value, a *hop-by-hop* architecture is overall cheaper. When node capacity exceeds this threshold, *hub & spoke* is overall cheaper. This is in line with what seen in section A, but the additional savings in the optical infrastructure provides additional advantages to *hop-by-hop*.

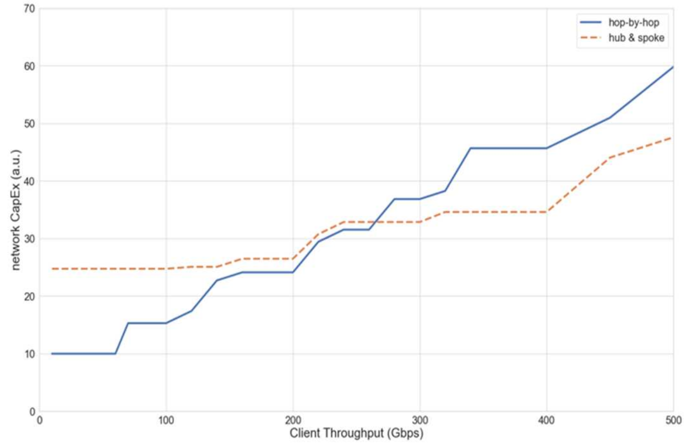

Figure 5: Network cost comparison

IV. Conclusions

We have shown that both the *hub & spoke* model as well as the *hop-to-hop* model may yield a cost-optimal design for a given network. The critical parameters that determine which of the two models yields the optimal design are:

1. The traffic matrix of the network.
2. The topology of the network.
3. The relative price points of the components making up the network.

We have demonstrated that focussing on the relative price points of a single components of the overall network architecture and taking this comparison as "pars pro toto" may lead to non cost-optimal designs.

We demonstrated that the crossover point between the two architectures is at an average node capacity of around 200 – 300 Gbps. This crossover point is partially dependent upon the achievable interface speed. With 800 Gigabit pluggable DWDM optics becoming available in the foreseeable future, that crossover point is likely to move up to higher average node capacities.

Acknowledgments

The authors would like to thank Alberto Donzelli, Cisco, for his invaluable technical support on IP architectures.

References

[1] J. M. Simmons and A. A. M. Saleh, "The value of optical bypass in reducing router size in gigabit networks," 1999 IEEE International Conference on Communications (Cat. No. 99CH36311), 1999, pp. 591-596 vol.1, doi: 10.1109/ICC.1999.768007.

[2] P. Wright, R. Davey, A. Lord, "Cost Model Comparison of ZR/ZR+ Modules Against Traditional WDM Transponders for 400G IP/WDM Core Networks", ECOC 2020.