Cost-Effective Sub-wavelength Solution for Data Centre Location in Scaled Next-Generation Networks

Cs. Kiss Kalló*, J. Shields*, Conor O'Malley*, Raymond Carley*,

Víctor Lopez**, J. P. Fernández-Palacios**

* Intune Networks, Dublin, Ireland

** Telefónica I+D, Madrid, Spain

csaba.kisskallo|john.dunne|conor.omalley|raymond.carley@intunenetworks.com; vlopez|jpfpg@tid.es;

***Abstract*— This paper summarises the modelling results obtained with a sub-wavelength optical packet switching technology, called OPST (Optical Packet Switch and Transport), when compared against an IPoDWDM solution. In particular, the impact of data centre location in the network on the network cost is studied for two architectures based on different technologies. Data centre location, routing strategies, service mix, traffic growth and subscriber service take-up is modelled to obtain a broad view about the cost sensitivities in these networks.**

The main contribution of the paper is to demonstrate that resilience to data centre location and changing traffic patterns enable the sub-wavelength packet optical solution to achieve cost savings of 150% when compared to the IPoDWDM approach. Such flexibility of the packet optical solution also enables 100-300% of power consumption cost reduction and an average 150% savings on rack cabinets for typical configurations.

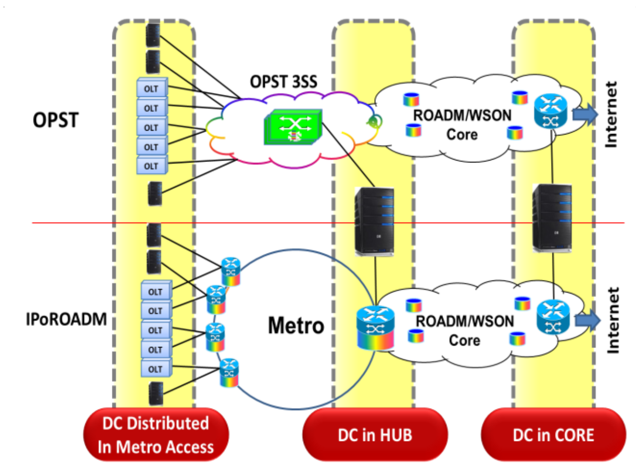

Fig. 1. OPST (top) and IPoWDM (bottom) architectures and data centre locations. Currently most of the content is sourced from data centres located in the Internet or the core. Interest for moving the data centres in the hub or even in the metro access is increasing, due to the performance benefits.

I. Introduction

There are more than 400 cities around the world today with more than 1 million inhabitants, while the population of 30 metropolises exceeds 10 million people. Such dense metropolitan areas are responsible for the majority of the traffic growth. According to recent traffic forecasts [1, 2], by 2015 the global data consumption will cross the Zettabyte (10^{21} bytes) threshold. The largest portion of this data is some form of video content. However, next generation cloud services are attracting a lot of attention and in the future they will progress toward commercial offerings. Combined with mobility, such cloud services can easily generate traffic flows capable of exceeding typical current networks capacity [3], even at a very low penetration rate. These trends lead the telecommunications industry to the realisation that *cost-effective scalability* is a crucial characteristic for future network architectures.

In addition to scalability, there is huge uncertainty around the level and pace of the growth as well as around the location for placement of the service platforms. New services, content-rich applications and mobility are only the top of the list of factors that can easily and instantly swing the traffic pattern in a network.

The location that content will be provided from, i.e. the location of the data centres (DCs, Fig.1) is also subject to conjecture. Network operators are heavily investing into data centres to maintain profitability in an environment where third-party content is increasing their cost, without impacting their revenues at the same extent. The ownership of the network infrastructure provides operators with a competitive advantage through control over performance that they are looking to monetise. This suggests that the source of content will migrate from the Internet - its current location - closer to the user, into the national backbone, metro head-end and/or even into the metro access sites. Given the huge amount of uncertainty pertaining to next generation services and traffic patterns, however, designing the a network that cost-effectively provides high performance for delivering these services to end-users is a major challenge to network operators.

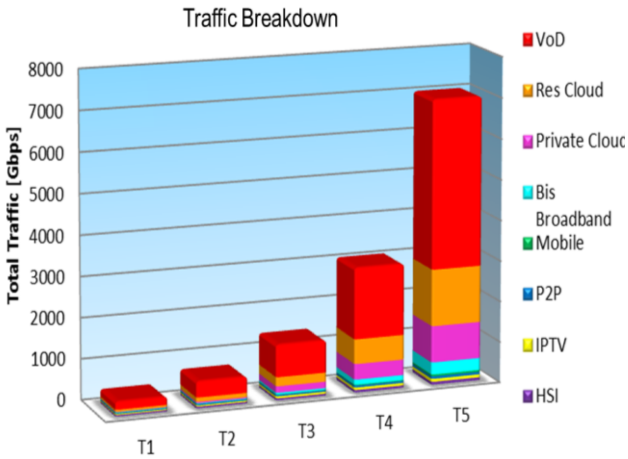

Fig. 2. Modelled traffic evolution per service.

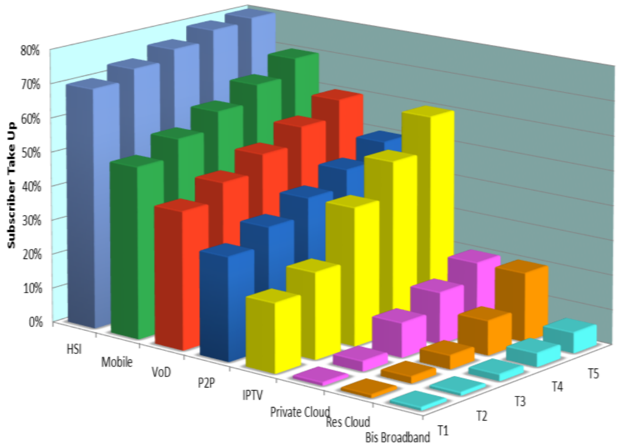

Fig. 3. Modelled subscriber take-up evolution per service.

This paper analyses the scalability and durability of next-generation networks with an emphasis on the economics. Incremental network deployment, data centre location, network architecture, service mix, traffic and subscriber take-up are all taken into consideration. Furthermore, next generation networking technologies like sub-wavelength optical packet switching and IPoDWDM/IPoROADM are compared to understand the cost implications.

In the remainder of this paper Section II describes the modelling approach and assumptions used. Section III presents the techno-economic comparison of the modelled architectures. Finally, Section IV concludes the paper.

II. Approach and Assumptions

A. *Generic Assumptions and methodology*

Two network architectures are modelled and compared to analyse their characteristics at scale. A subscriber base of 1 million is provided with a set of services. Each service is allocated bandwidth per subscriber. The models calculate the aggregation, switching, grooming and transport resources required in the network to deliver the services when the data centre is placed in the core, at the metro hub and distributed out in the metro. When the resources are derived, the models calculate, using typical unit pricing, the capital cost of the network. The models were run at five points in time (T1,..,T5) that represent growing uptake of services. The resulting cost output was used to compare an optical sub-wavelength packet forwarding technology called Optical Packet Switch and Transport (OPST) solution with an IPoDWDM one for each placement of the DC. OPST merges switching and transport into a single layer and thus provides a sub-wavelength optical packet forwarding platform that can aggregate, groom, switch and transport packets in a uniquely efficient manner [4, 5]. A next-generation cloud service over OPST case study is presented in [3], where a brief summary of the OPST technology and equipment is also available. On the other hand, IPoDWDM is a reasonable solution for carriers to reduce investment, while absorbing traffic increment [6]. Traffic breakdown and growth is based on internal Telefónica sources and aligned with the surveys in [1, 2] (see Fig. 2). Between T1 and T5 the traffic gradually grows from 300Gbps to 7Tbps to reflect a significant network scaling scenario. The time span of this scaling is variable for each network, depending heavily on local market dynamics, operator strategy and success.

B. *IPoDWDM Architecture*

In the IPoROADM architecture (Fig. 1, bottom) traffic is collected by IP access routers from Passive Optical Network (PON) head-ends, called Optical Line Terminals (OLTs). It is then aggregated by hub routers located at the metro head-end. A dynamic Wavelength Switched Optical Network (WSON) [7] built with ROADMs is used in the backbone to transport the traffic to core routers, which on turn are connected to Internet exchange points [8]. Routers with 2Tbps non-blocking switching capacity are used in the hub and core. Each router has 20x100G slots that are populated with 10x10G port cards. When the capacity of a router is exceeded, router matrices are built. Router matrices can be placed in the hub or in the core. All of the relevant combinations of router matrix and DC placement are analysed.

In the metro access sites the IP access routers have 8x10G slots for a total of 80Gbps non-blocking switch fabric capacity. Client-facing slots are filled with 4x10G port cards, displaying a 4:1 oversubscription. This is feasible, because it is assumed that the average utilisation of the OLT 10G uplinks is 25%. Therefore, 4 OLT uplinks can be served by 1x10G access router slot. The network-facing access router slots are filled with non-oversubscribed 1x10G port cards.

C. *OPST Architecture*

The comparative architecture (Fig. 1, top) is based on the *Optical Packet Switch and Transport (OPST)* sub-wavelength optical packet switching technology developed by Intune Networks [5]. OPST fully integrates the packet forwarding of layer 2 with the sub lambda transport of layer 1 and physical fibre of layer 0 into a single layer. The OPST switching fabric is interconnected

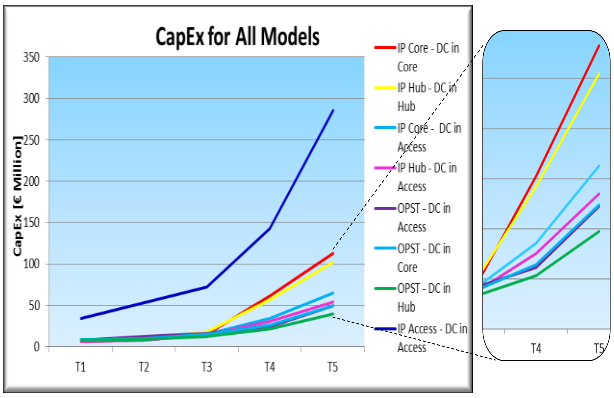

Fig. 4. CapEx for all variations of DC and switch capacity locations.

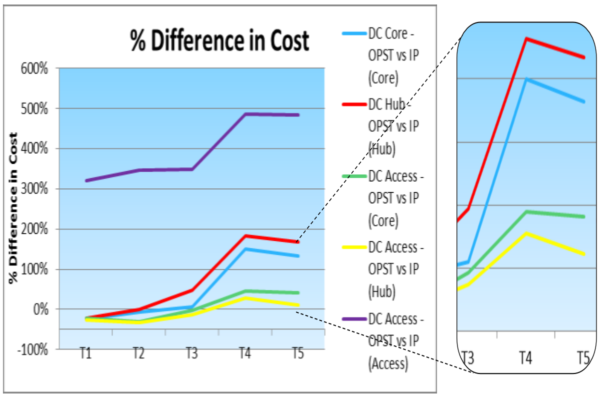

Fig. 5. Relative CapEx for all models

by a fibre pair into a DWDM ring of hundreds of kilometres in circumference where packets are wavelength-routed by destination. Each OPST port has a packet-responsive tuneable transmitter (50-100 nanosecond tuning time) that sends data packets on the wavelength assigned to the desired destination device when asked to do so by the local scheduler. Any port can transmit asynchronously on the same wavelength (i.e., to the same destination) on a time-division basis. Each port has a quasi-static tuneable receiver (built using a Wavelength Selective Switch from ROADM technology), which is assigned once at start of operation. Therefore any port can receive packets from any other port on the ring enabling a full mesh of optical connectivity.

Each OPST port is called a Distributed Switching Subsystem (DSS). Up to 8 x DSS units can be placed into one OPST node, called the *iVX8000*. Up to 80 x DSS units distributed in up to 16 iVX8000 nodes can be deployed on the same OPST ring. The DSS used in this paper has two 10G client ports. It also has 10Gbps transport capacity on each of the fibres, providing 10Gbps protected or 20Gbps unprotected capacity per DSS. Thus, a full OPST ring has a total of 800Gbps of protected switch and transport capacity.

Multiple OPST rings can be connected using MPLS-TP (Multiprotocol Label Switching – Transport Profile) or Intune's *Three-Stage Switch (3SS)* solution [5]. The 3SS consists of three sets of OPST rings. The first set of rings collects traffic from metro access sites. The third set provides connectivity to the backbone network. Finally, the second or middle set of rings provides interconnection for the first and third sets. Flows are managed through the 3SS using intersection controllers. There is one intersection controller for each middle stage ring that makes flow control through the entire 3SS automatic and transparent to the operator. From a service provisioning perspective, the 3SS appears as a single L2 switch, while it comprises multiple optical platforms interconnected on fibre rings from an operations and maintenance perspective. In the comparative OPST architecture (Fig. 1, top) a 3SS carries out the aggregation, grooming, switching and transport functions in the metro network. The 3SS substitutes all of the routers and ROADMs used in the IPoROADM solution for the metro.

D. Data Center Location

Data centre location can have a significant impact on the cost and performance of a network. In this study different configurations are analysed, where the data centre is placed in the core, in the hub or embedded/distributed in the metro access (Fig. 1). Currently most of the data centre traffic is coming from the Internet. However, with increasing service uptake and lower latency requirements for key cloud applications, the data centres in the metro hub and out into the metro can better respond to the performance and cost requirements. All of these DC location considerations are addressed in the models and discussed in detail in Section III of this paper, in conjunction with the modelling outputs.

E. Service Mix

Eight services types of residential, mobile and business type are modelled. These services are High Speed Internet (HSI, e.g. email, browsing), Internet television (IPTV, e.g. live television), Video on Demand (VoD, e.g, Content Delivery Network (CDN) or Youtube video streaming), Residential Cloud (e.g. Virtual PC, machine-to-machine and other next generation services), Peer-to-Peer (P2P, e.g. file sharing), Mobile Backhaul (e.g. LTE, UMTS, GPRS, GSM), Business Broadband (employee browsing) and Private Cloud (e.g. corporate cloud, includes private lines). For each service an average bandwidth per subscriber was assumed. The traffic is categorized by destination, meaning that traffic can arrive to the metro access sites from the Internet, from DCs located in different places of the network, local metropolitan area traffic and multicast traffic. Multicast traffic is assumed to be initiated from the DCs and form a separate category for the particular way it is handled in the network. Average traffic breakdown per service over time is shown in Fig. 2.

Each service is also associated a percentage of subscribers. This percentage represents the subscriber take-up, which is also an expression of the popularity of each individual service (see Fig. 3). The random values are generated according to a uniform distribution within each individual range.

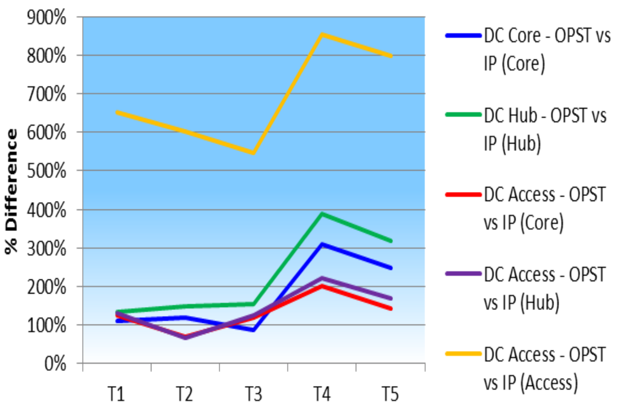

Fig. 6. Relative power consumption. Positive percentiles indicate that OPST enables power savings compared to the different IP0DWDM configurations.

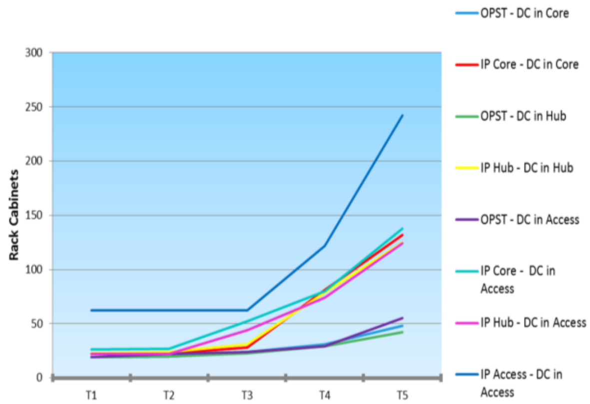

Fig. 7. Rack space requirements. The 3 OPST solutions are the 3 bottom curves, showing lower rack cabinet requirements independently from the scenario configuration

III. RESULTS

The most important results of this study are related to scalability of network solutions in function of data centre (DC) location. Fig. 4 presents the CapEx (i.e., capital expenditure) of all meaningful solutions for DC and switching capacity locations. The most frequently used solution today is represented by the "*IP Core – DC in Core*" curve, whereby most of the traffic is backhauled to core routers to which the DCs are also connected. Most often the DCs are located in a remote network in the Internet. From a router port requirement perspective, however, this is similar to having the DCs located in the national core network. For this reason these subcases are addressed together.

In the "*IP Core – DC in Core*" solution huge routing and grooming capacities are built in the core network. Traffic collected from client interfaces on the client edge of the metro network is aggregated by further switches and/or routers and then groomed in the core to ports dedicated to different services operated on DC servers. This is the most expensive solution (apart from "*IP Access – DC in Access*", presented later) due to the large number of expensive core router ports required for routing and grooming the traffic. Such high costs fall on the network operators, while at the same time the revenues obtained from content services (e.g., VoD, social media) go to the content providers (i.e., not to the network operators). This asymmetry of the cash flow has a negative impact on the profitability of the network operators.

To address the cash flow problem, network operators work on bringing the service platforms into their networks, either by building their own service platforms or by providing premium services for content providers through creating local distribution points for third-party content. Also, by moving the content closer to the end-user, a significant amount of transit traffic is eliminated from the core network, while the service performance (e.g., delay, jitter) improves. This can be achieved by placing DCs in the hub and switching the bulk of the traffic in hub routers (see "*IP Hub – DC in Hub*" in Fig. 4). However, router port requirements remain relatively high. Therefore, although the performance and cash flow symmetry improves, the solution CapEx is still high.

Further improvement on the CapEx and/or performance can be achieved by moving the DCs even closer to the end-user through distributing them out to metro access sites. Interconnection of the sites can be achieved by performing the routing function a) in the core ("*IP Core – DC in Access*") or hub ("*IP Hub – DC in Access*"); or b) by building out direct optical connections (i.e., optical mesh) between the metro access sites ("*IP Access – DC in Access*"). As shown in Fig. 4, these different options have significantly diverging CapEx implications. Although IP core and hub solutions achieve important CapEx reductions, all traffic is switched in the core or hub, respectively, which implies multihop connections and, hence, higher latency and jitter. On the other hand, the "*IP Access – DC in Access*" full optical mesh solution provides single-hop optical connectivity between the metro access sites, which implies the lowest possible delay and jitter. However, such optical connectivity requires a very expensive optical layer as well as a high number of L2/L3 ports. This provides by far the most expensive solution (see Fig. 4). Therefore an important decision needs to be made in the IPoROADM solutions between providing high performance or low cost as high performance at low cost is not possible in these networks.

The collapsed switching and optical layers of OPST provides a solution to this cost/performance problem. Fig 3 shows that OPST provides the lowest CapEx network independently of the location of the DC. The "*OPST – DC in Core*", "*OPST – DC in Hub*" and "*OPST – DC in Access*" curves provide significant CapEx savings for all these scenarios, which makes OPST a durable solution that scales cost-effectively with the increasing traffic demand. The reason for this is that OPST is a distributed switch based on a full and dynamic optical mesh of connectivity. This uses shared optical paths between ports allocating resources in real time to where and when they are needed, thereby optimizing both switching and transport port costs simultaneously. Furthermore, the OPST switch's utilization of this full mesh of direct optical connectivity ensures single-hop forwarding between ports thereby inherently minimizing latency. Thus OPST supports cost-effectively services that require stringent performance guarantees.

CapEx advantages are at the order of 150% on the favour of OPST when the DC is located in the core or hub (see Fig. 5). When the DC is distributed in the metro access, IPoROADM solutions can close the CapEx gap with OPST only by compromising performance ("*IP Core – DC in Access*" and "*IP Hub – DC in Access*" in Fig. 5). However, as the importance of delay and jitter guarantees increase, the CapEx moves higher up toward the "*IP Access – DC in Access*" curve of Fig. 5, being bound by a configuration that uses a full mesh of fixed optical connections between the metro sites. The CapEx of this latter solution is almost 500% higher than that of the OPST one.

As an ancillary output of the modelling performed, the power consumption (Fig. 6) and rack space (Fig. 7) requirements were also quantified. These two important operational expenditure (i.e., OpEx) drivers display similar trends to that of the capital costs. At scaled traffic requirements the power consumption advantages provided by OPST are at the order of 100%-300% for the most of the scenarios. Also in this case, if OPST performance is emulated over an IPoROADM solution this latter consumes up to 800% more power.

The rack space requirement gap between the two compared solutions is also quite significant at scale (i.e, at T4 -T5). With the DC in the core and hub and routing capacity focussed either to the core or hub, the gap gets up to 80 rack cabinets at a total of 130. This translates into 150% more cabinets used for the IPoROADM solution. In the case of an optical mesh in the metro network that connects a distributed DC (to emulate OPST performance) up to 500% more rack cabinets are needed for the IPoROADM solution. Thus, beyond the CapEx savings, significant OpEx reductions can also be achieved with a sub-wavelength optical packet switching technology.

Finally, to ensure that specific traffic, service and take-up forecasts as well as network design trends do not invalidate the results in this paper, an extensive analysis was conducted based on wide input parameter ranges. This was then fed into a Monte Carlo simulation [9] to gain visibility into impact of variable network conditions on the economic performance. 1 million uniformly spread configurations were analysed and the short, medium and long-term trends observed. The results and details of this analysis are reported in another paper, currently under review.

IV. Conclusion

This paper demonstrates that the dynamic nature of an optical sub-wavelength packet forwarding technology called OPST enables CapEx savings of 150% in typical network scenarios compared to IPoROADM solutions. Furthermore, when very low delay and jitter is required by the services of a network operator, OPST performance can be matched by a significant IPoROADM overbuild that costs 500% more than the OPST solution. Such results underpin that optical sub-wavelength packet forwarding technologies scale cost-effectively under increasing traffic requirements and provide a future-proof solution for next-generation services and networks independently from the future location of data centres in the network. Finally, services with stringent performance requirements are supported much more cost effectively (in terms of CapEx, power consumption and rack space) and without compromise when compared to IPoROADM technologies.

Acknowledgment

This work has been supported by the EC's Seventh Framework Programme through the IST STREP project MAINS (INFSO-ICT-247706).

References

[1] Cisco White Paper, *Cisco Visual Networking Index: Forecast and Methodology, 2010–2015*, June 2011.

[2] Cisco White Paper, "Entering the Zettabyte Era", June 2011.

[3] Cs. Kiss Kalló, M. Basham, J. Dunne, J. P. Fernández-Palacios, *Cost Reduction of 80% in Next-Generation Virtual PC Service Economics Using an OPST Sub-Wavelength Metro Network*, In Proceedings of NOC 2011.

[4] Intune Networks, "*Verisma iVX8000 Product Data Sheet*", 2010. http://www.intunenetworks.com/home/shape-up/verisma_ivx8000

[5] Intune Networks, "*Optical Packet Switch and Transport. A Technical Introduction*", 2009. http://www.intunenetworks.com/home/shape-up/core_innovation/opst_technical_introduction/

[6] J.E. Gabeiras, V. López, J. Aracil, J.P. Fernández Palacios, C. García Argos, Ó. González de Dios, F.J. Jiménez Chico and J.A. Hernández: *Is Multi-layer Networking Feasible?*, in Elsevier Journal on Optical Switching and Networking, April 2009, Vol. 6, Issue 2, Pages 129-140.

[7] Y. Lee et. al., "*Framework for GMPLS and PCE Control of Wavelength Switched Optical Networks (WSON)*", IETF Draft, Feb. 2011.

[8] J. P. Fernández-Palacios, L. Perez, J. Rodriguez, J. Dunne, M. Basham, *IP Offloading over Multi-granular Photonic Switching Technologies* In proc. of ECOC 2010.

[9] N. Metropolis, S. Ulam, "*The Monte Carlo Method*", September 1949