ITC19/ Performance Challenges for Efficient Next Generation Networks 1907-1916
LIANG X.J. and XIN Z.H.(Editors)
V.B. IVERSEN and KUO G.S.(Editors)
Beijing University of Posts and Telecommunications Press

Analysis of Container Formation and Delays in the Centrally Scheduled Photonic Burst Switch

Amin R. Mazloom and Kalyan Basu
Center for Research in Wireless Mobility and Networking (CReWMaN)
University of Texas at Arlington, TEXAS, USA
{mazloom, basu}@cse.uta.edu

Abstract: *In this paper the concept of Container {fixed size macro data unit} is used to photonic burst switch network. This Container is tailored to integrate the existing multiple optical protocols stack as an intermediate layer between optical and link layers. The objective of Container based optical burst switching is to eliminate the O/E/O conversion in the core optical network to create a collision free stateless core. In addition the concept takes advantage of the wireless link's delay to minimize the necessity of long Fiber Delay Lines (FDL) in a mesh topology photonic network. This is made possible by a centrally scheduled edge switch that periodically assigns optical burst slots between the two nodes to carry the containers. The architecture is analyzed for its container formation delay and other performance dynamics like the container size.*
Keywords: Container, Photonic Network, Scheduler, Burst, SOA, G/D/1

INTRODUCTION

The optical transmission technology has brought a revolution in the transfer of information in the world, as a result the cost of the telecommunication services has dropped significantly over the last decade. In the same period, however the advances in switching technologies were marginal, as a result the cost of digital switches, ATM switches and Routers remained same or increased. The convergence of the switching the heterogeneous services likes SONET private line, voice circuits and data links is severally restricted by the architecture and design limits of the ATM switching. At the every switching point in the network, the optical transport is converted to electrical transport to rearrange the transport function. This O/E/O conversion causing serious power and space problem at the switching nodes. In order to reduce the cost and support the QoS of these everyday services in the backbone network, we need faster and more efficient switching hardware that can seamlessly integrate the optical transport and switching. In this paper we propose a new burst-switching concept that will be used at the edge of the core network. This architecture will create a unified transport mechanism for upper layer transport services for the national core network.

The justification for proposing this flat single hop core optical network is (1) There is no need for QoS function in the core network, as most of the core network delay is contributed by the propagation delay, (2) The lower cost of fiber optic transmission reduced the number of core switched significantly and that justifies a single hop fully mesh network, (3) The proposed higher modularity of the network connections in this mesh network will justify a static schedule that maintains the static link capacity for considerably longer time (say few hours), and (4) convergence of SONET, IP and Frame-relay traffic on a container removes the complexity of managing the individual traffic type in the core network.

We will follow the network architecture as proposed in [1], which is a Centralized Scheduled Optical Network (CSON) where the nodes are connected in single hop mesh topology. CSON guarantees a collision free switching environment for the swishes. Each switch will use the template to send their data units to the destination switch. This global schedule [1] will remain fixed for considerable longer time so we can consider it as a static routing plan. Global scheduler will periodically change the scheduler template of the switches to reflect the changes in long term traffic distribution.

In this work, we use the concept of container as an abstract data unit that converges the different heterogeneous data transport standards like SONET, IP, ATM and Frame-Relay. This will permit to create a transparent view of data through its optical transport network. We will detail the analysis of packet delay due to the process of containerization and container forwarding. The architecture, the definition of the container data unit and the scheduler algorithm that we propose will lead us to a fully optical collision free photonic network. The nodes of this network are state free and there is no need of O/E/O conversion for data units. The architecture eliminates the long FDLs by utilizing the optical links delay as the delay line for the scheduler.

In section 1 we introduce the details of our concept of container, section 2 will briefly describe the switch architecture. In section 3, a delay model for the container formation is developed. Throughout section 4, we discuss the container queue delay model based on the burst allocation of the scheduler. Section 5 will put the different delay subsystems together to come up with a cumulative switch delay. In section 6 we compare our analytical model with the simulation to validate the model and discuss our finding by some numerical result, we end with conclusion and future work.

1. CONTAINER

Current optical transport network transports either SONET, or ATM or Ethernet protocols stacks as shown in figure 1 (left) and map their payload on the optical domain to carry the different types of traffic. For IP traffic, a variable data unit (DU) size is required and it also requires intermediate node DU processing. This requires state full intermediate nodes that have to work on electrical domain. This forces intermediate nodes to provide O/E/O conversion, which is very expensive for hardware, power consumption and adds additional delay to the data transmission task. We use the concept of container, like the bulk transport containers of the cargo ship, as the transport units of the core network that will contain all the traffic types inside it, as shown in figure 1 (right). As the delay in the core network is mostly propagation delay, all services are shipped independent of their priority (QoS criteria) between two points. This is similar to the method of the bulk transportation of cargo goods where the priority schemes are introduced at the edge to ensure the service QoS guarantee.

Container is a fixed size macro data unit that is made up by aggregation of the upper protocol data units. *Container* is converted to the optical domain to from a burst. The edge nodes of the photonic network will assemble and disassemble a container from upper data units based on their destination. The data units of the multiple sources destined to the same destination

3. DELAY MODEL FOR THE CONTAINER FORMATION

The total delay of figure 4 is composed of these delay components: (1) Network Delay, (2) Container Formation Delay, (3) Container Queue Delay and (4) Container Processing Time.

1) Networking Delay: This is the delay on an incoming packet that goes through the system before it reaches the output queue of the electrical switch fabric. The delay includes the ingress processing delay and the switching delay.

2) Container Formation Delay: Container formation is one of the major responsibilities of the edge router; also, it is the major source of the delay that might be applied to the packet to move it form source to destination. There are two major parameters in the container formation time and the interplay between these two factors will lead us to the optimized container delay. These parameters are: (a) Container size and (b) Container processing time of OBS

(a) Container Size: Since the container will need to be placed in to a frame time slot then we will have $1/\lambda \leq T_c \leq \Theta_t - \upsilon$ where T_c will be the frame slot time Θ_t, less than the training and guard times υ and inter arrival time $1/\lambda$ of the packets will be lower bound delay for container formation. If the line speeds were k bit per second then we get: $k/\lambda \leq kT_c \leq k(\Theta_t - \upsilon)$ and from there let $S_c = kT_c$ which will be the size of the container. Now we are to find the optimal size of the container since a large container will expose the onboard packets to excessive delay and a small container will waste the frame slot by using too much overhead time.

3) *Container Queue Delay*: This delay is the time from the container arrival to the queue until the processing time of that container, which depends on the traffic intensity and processing rate.

4) *Container Processing Time*: Container processing time will depend on the central scheduling algorithm and number of node in the mesh. The scheduler is described in [1] and assumed static for this analysis. Hence, we just assuming that it will be a constant.

3.1 Event in Container Formation Cycle

There are two events that will lead to indicate the end of single container formation process, and as a result will push the container to container queue for the further burst processing steps. Those two criteria are Container Formation Timeout (CFT) and Container Load Complete (CLC). These two events is demonstrated in figure 5, where S_C the container size w is the CFT value. In figure 5, events 1and 2 are the CLC events and events 3 is CFT event.

3.2 Packet Quantization

The IP packets are received in different sizes. For modeling purpose, to reduce the

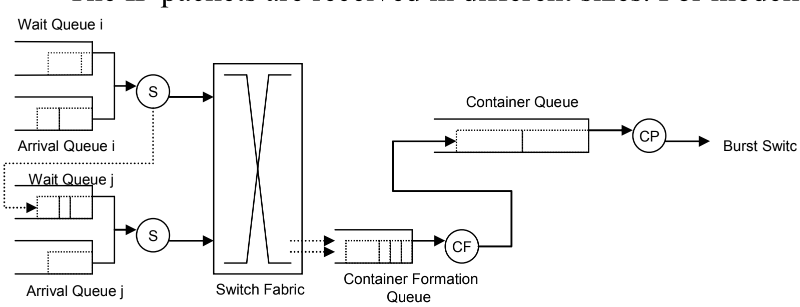

Figure-2 Simplified architecture of the switch

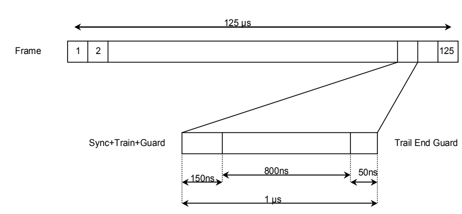

Figure-3 Structure of Frame

Table-1 IP Packet Quantization

ID	Packet Size Range	P(x)	Size (Byte)
1	[20,320]	0.676	320
2	[321,620]	0.294	620
3	[621,920]	0.003	920
4	[921,1220]	0.006	1220
5	[1221,1520]	0.021	1520

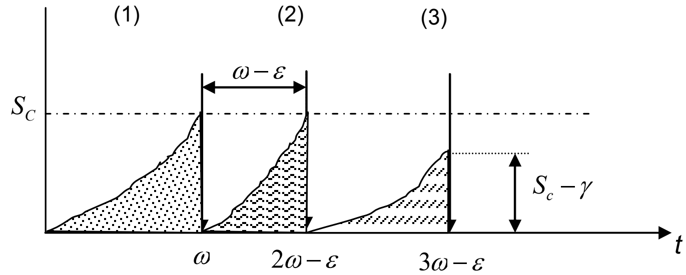

Figure-5 Events in the container formation

probability of occurrence in the IP packet distribution p(1) is available in this container then the with regarding Erlangian distribution at *h=6, overall* probability distribution contribution of this class of container in the overall container arrival time distribution is given by:

$$P[C(1,1,1,1,1,1)] = p(1)^6 E_r(t,6) \Rightarrow = p(1)^6 \frac{\lambda(\lambda t)^5 e^{-\lambda t}}{5!} \text{ or } P[C(4,2)] = p(4)p(2)E_r(t,2) \Rightarrow = p(4)p(2)\frac{\lambda(\lambda t)e^{-\lambda t}}{1!}$$

In order to find the PDF of the containerization delay we need to find the probability of every possible container combination, e.g. every path that begins at root and ends up with a leaf. Summation over all these single container random delays results containerization delay PDF. By further looking at the different container probabilities, we discover the following property:

$$P[C(1,1,1,1,2)] = P[C(1,2,1\,1,1)] = P[C(2,1,1\,1,1)] = \ldots$$

$$P[C(1,2,3)] = P[C(2,1,3)] = P[C(3,2,1)] \text{ or } P[C(1,1,3,1)] = P[C(3,1,1,1)] = P[C(1,1,1,3) = \ldots$$

So if π_i be the sets of container with Erlangian order of i then, for each non empty container tree *E-Set* will be set of all π_i sets of that container tree. $|E-Set|$ is the number of terms that will appear in the container delay PDF. In our example we have:

$$E-Set = \{\pi_2, \pi_3, \pi_4, \pi_5, \pi_6\} \qquad |E-Set| = 5$$

$$\pi_2 = \{C(4,2), C(2,4)\} \quad , \ \pi_3 = \{C(1,2,3), C(2,1,3), C(3,2,1)\} \text{ and } \ldots$$

If $n(PDF_C)$ be the maximum number of the terms in the PDF then

$$n(PDF_C) \le Tree\_Max\_Depth \quad \text{and} \quad n(PDF_C) = \left\lceil \frac{S_C}{\min\{packet\_size\}} \right\rceil$$

Hence, the highest power order of container PDF will be: $o(PDF_C) = \left\lceil \frac{S_C}{\min\{packet\_size\}} \right\rceil - 1$

3.5 Containerization Delay PDF

If we sum over the normalized probability of individual *E-Set* element, times the corresponding Erlangian probability, the resultant expression will be PDF_C. Hence if we define K_m^j as the *mth* packet in the container of combination set *j* belonging to π_i and considering $\widetilde{N}(x_i)$ as the normalized value of x_i over set X and h = *Tree_Max_Depth* then we can write:

$$PDF_C(t) = \sum_{i \le h} \left[\widetilde{N}\left(\sum_{j \in \Pi_i} \left(\prod_{m=1}^{i} \Pr(\mathrm{K}_m^j) \right) \right) \frac{\lambda(\lambda t)^{i-1} e^{-\lambda t}}{(i-1)!} \right] \qquad (1)$$

In formula (5) ρ is the occupancy, C_A^2 is the coefficient of variation of the inter arrival time, C_S^2 is the coefficient of variation of the service time and E(S) is the mean service time. Also since our service rate is constant then $C_S^2 = 0$ and E(S) = D so (5), becomes: $\widetilde{W} = \frac{\rho}{1-\rho} \cdot \frac{C_A^2}{2} \cdot D$ (6)

now let's extend the formula to get that closer to the notation of delay based on the PDF of arrival time. From (5) if the expected inter arrival time denoted by E (A) then we have

$$\widetilde{W} = \frac{D/E(A)}{1-D/E(A)} \cdot \frac{E(A^2)-E(A)^2}{2E^2(A)} \cdot D \quad \Rightarrow \quad = \frac{E(A^2)-E(A)^2}{2E^2(A)(E(A)-D)} \cdot D^2 \tag{7}$$

From Laplace Transform and (7) we can write: $\widetilde{W} = \frac{L_{C(t)}^{(1)}(0)^2 - L_{C(t)}^{(2)}(0)}{2L_{C(t)}^{(1)}(0)^2 (L_{C(t)}^{(1)}(0)+D)} \cdot D^2$ (8)

Expression (8) could be used as general formula for container queue delay approximation. Also the above formula stays valid as long as $D < E(A)$, which is intuitively true, if we want to keep the system in equilibrium state. Eventually from (1) and (8) we can write:

$$\mathrm{E}_P = -L'_C(0) + \frac{L'_C(0)^2 - L''_C(0)}{2L'_C(0)^2 (L'_C(0)(0)+D)} \cdot D^2 \tag{9}$$

Formula (9) gives the mean packet delay in the container-based traffic and processing models. Notice that *subscript C* denotes the PDF of container arrival.

5. CUMULATIVE SWITCH DELAY

Any packet that enters the Container queue has either experienced the queuing delay and went through the switch fabric, or there was no queuing at all. The initial service time for a packet includes header check and switching delay, which could be considered as exponential service time τ' (service rate μ'). So with *M/M/1* queue model for pre-containerization delay:

$$\rho_i = \frac{\sum_i \lambda_i}{32 \times \mu'} \quad \text{and} \quad \widetilde{q}_i = \frac{\rho_i}{1-\rho_i} \quad \text{and} \quad \widetilde{w}_i = \frac{\rho_i}{1-\rho_i} \cdot \left(\frac{1}{32} \sum_i \lambda_i \right)^{-1} \Rightarrow \quad \widetilde{w}_i = \frac{32}{(32 - \sum_i \rho_i)} \tau' \tag{10}$$

Considering (10) and the Jackson model, we can easily generalize this delay for any non-uniform traffic model too. If we define γ_{ji} as the fraction of the packets the will arrive at port *j* but for any reason could be, load balancing [7], explicit destination or port ownership [7], will be forwarded to port *i* for further processing then we will have:

$$\Gamma_i = \sum_j \gamma_{ji} \lambda_j \quad , 0 \le \gamma_{ji} \le 1 \qquad \widetilde{w}_i = \frac{\rho_i}{\Gamma_i (1-\rho_i)} \tau' \tag{11}$$

If α is the switch delay then the mean pre-containerization of port *i* PCW_i will be:

$$PCW_i = \widetilde{w}_i + \alpha \tag{12}$$

Here we put every thing together so form (9), (11) and (12) we can derive Δ_i, the mean packet switching delay for port *i* of switch, as:

$$\Delta_i = \frac{\rho_i}{\Gamma_i (1-\rho_i)} \tau' + \alpha - L'_C(0) + \frac{L'_C(0)^2 - L''_C(0)}{2L'_C(0)^2 (L'_C(0)(0)+D)} \cdot D^2 \tag{13}$$

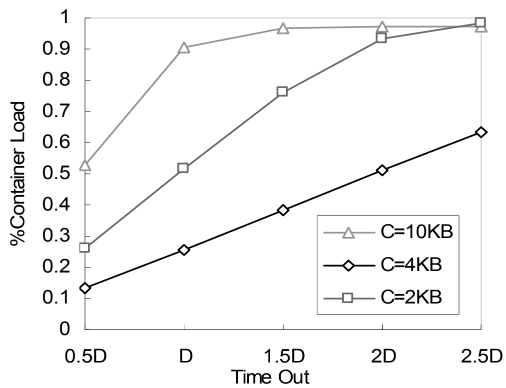

Figure-11 Container Load Efficiency vs. Time-Out

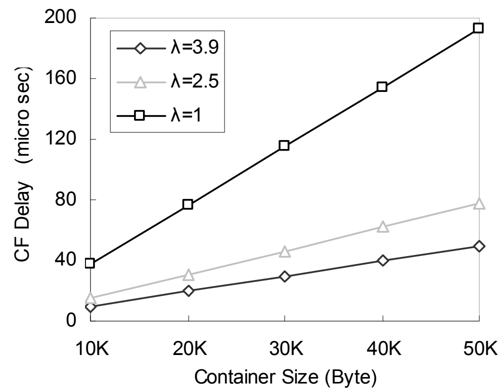

Figure-12 CF Delay vs. Container size

7. CONCLUSION AND FUTURE WORK

In this paper, we proposed the concept of container in optical communication that seamlessly integrates the electrical protocol stacks to optical burst switching. This integration eliminates further need for O/E/O conversion in the core and leads to a stateless core with respect to upper data units. The whole concept of container will significantly reduce the costs of processing; resource management and capital cost (power, space) that are required on current optical networks while maintaining the QoS requirements of the different services. We proposed Central Scheduler based photonic network that uses SOA switching matrix. Based on the defined functionality of the system, the different delay components are identified. We then study the behavior of container with respect to the key parameters of the architecture. Several other improvements that could be added to improve this model like packet shuffling, time out of the container in the model and other distributions of packet arrival time are part of future research.

REFERNCES

1. Preetam Ghosh; An Optical Burst Switched (OBS) Transport network to support long haul traffic of the core. MS Thesis, CREWMAN Lab; University of Texas at Arlington.
2. Sahri, N. Prieto, D. Silvestre, S. Keller, D. Pommerau, F. Renaud, M. Rofidal, O. Dupas, A Dorgeuille, F. Chiaroni, D.; A Highly integrated 32-SOA gates optoelectronic module suitable for IP multi-terabit optical packet router. : OPTO+, Alcatel Inc., Marcoussis, France, page(s): PD32-1- PD32-3 vol.4
3. Shlomo Ovadia, Christian Maciocco, and Mario Paniccia Intel CorporationGMPLS-Based Photonic Burst Switching (PBS) Architecture for Optical Networks.
4. Chuck Fraleigh, Sue Moon, Bryan Lyles, Chase Cotton, Mujahid Khan, Deb Moll, Rob Rockell, Ted Seely, Christophe Diot;" Packet-Level Traffic Measurements from the Sprint IP Backbone.
5. Leonard Kleinrock; Queuing Systems Volume I: Theory: John Wiley & Sons Publications 1974.
6. Internet. J. Prod. Res., 18 (1980), pp.761-773; On the approximations to the single-server queue.
7. Amin R. Mazloom;Container Formation and Architecture Design of the Centralized-Scheduled Time-Sliced Optical Burst Switch (CSTOBS); MS Thesis, CSE Department; University of Texas at Arlington; http://cseweb.uta.edu/~mazloom/msthesis.pdf.