Grooming and Degrooming with Coordinated Universal Time (UTC)

M. Baldi†‡ and Y. Ofek†

†Synchrodyne Networks, Inc.
New York, New York, USA
{baldi,ofek}@synchrodyne.com

‡Turin Polytechnic, Computer Engineering Dept.
Turin, Italy
mario.baldi@polito.it

***Abstract* - Coordinated Universal Time (UTC) enables the efficient and cost effective implementation of a grooming and degrooming method. The number of times data packet headers are being decoded throughout a global network deploying a UTC is limited to the necessary minimum of one. The simple implementation of grooming and degrooming devices results in high scalability and low cost.**

1. INTRODUCTION AND MOTIVATIONS

The advances in optical transport lead to the realization of ever larger optical channels (40 Gb/s field trials are being conducted) that are filled up by many traffic sources. This results in the need for:

(i) *grooming* – to aggregate traffic of a large number of sources into one optical channel, and

(ii) *degrooming* – to separate traffic from one optical channel to many different destinations.

For example, considering a 10 Gb/s high-speed optical channel and a 1 Mb/s “high-speed” residential access, such as the one provided through xDSL services, the traffic from 10,000 sources is to be groomed in order to fill up a single optical channel. This creates the demand for a large amount of grooming and degrooming equipment with high interface density and low cost per interface.

Grooming of data packets from a large number of low speed sources to fill up a high capacity optical channel is a major challenge to network designers. Grooming and degrooming of optical channels cannot be performed in the all-optical domain, i.e., it should be done electronically with either a packet-based or a time-based approach for multiplexing and demultiplexing.

Packet-based grooming and degrooming has the following shortcomings:

1. Header processing in high-speed optical links is a complex function that impacts the cost, the scalability, and the physical size of grooming and degrooming equipment.
2. The service provided when asynchronously multiplexing flows of data packets does not have deterministic guarantees. Complex and not scalable packet queuing and scheduling algorithms that should be implemented in hardware are needed in order to ensure quality of service.
3. At the bandwidth mismatch point between high capacity optical channels and lower capacity edge/access channels large buffers are needed to absorb the inherent burstiness of asynchronous flows of data packets destined to the same low-speed channel. Such buffers represent a cost and a source of large delay, jitter, and possible data packet loss.

Time-based grooming and degrooming based on the SONET operation principles does not suffer from the above shortcomings affecting asynchronous packet-based grooming and degrooming. However, SONET byte multiplexing and demultiplexing is complex and, consequently, not scalable. High speed SONET equipment is expensive due to the demanding time resolution requirements and pointer-based frame alignment mechanisms.

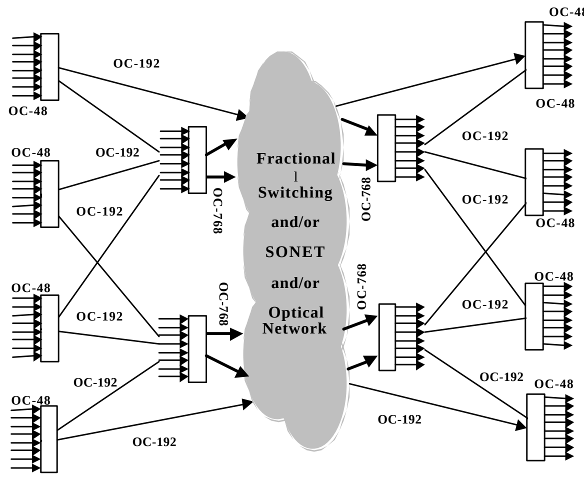

Fig. 1: Grooming and Degrooming Scenario.

This paper shows that the availability of Coordinated universal Time (UTC) enables the implementation of simple and scalable time-based grooming and degrooming equipment that will interface to various optical core networks, as shown in Fig. 1.

2. COORDINATED UNIVERSAL TIME (UTC) FOR GROOMING AND DEGROOMING

Coordinated Universal Time or *UTC* (a.k.a. *Greenwich Mean Time* or *GMT*) is a *time-of-day* international standard that is globally available at low cost, with accuracy of 1 μsec, through several distribution systems, such as, GPS (USA satellites system) [2], GLONASS (Russian Federation satellites system) [3], and in the future by Galileo (European Union and Japanese satellites system) [4]. The UTC second is partitioned into successive *time frames* (TFs), as shown in Fig. 2. Contiguous time frames are grouped into time cycles and contiguous time cycles are grouped into contiguous super cycles, wherein one super cycle is equal to and temporally aligned with one UTC second, as shown in Fig. 2.

Pipeline forwarding (PF) of time frames through a sequence of grooming devices, switches, and degrooming devices is realized by pre-scheduling the forwarding time of data packets contained in each time frame. Fractional lambda (λ) switching (FλS) [1] relies on PF to support the provision of dynamically switched fractions of optical channels (λs) across a global network. No packet header processing is necessary (for grooming, switching, and degrooming) once data packets have been associated to a time frame, i.e., to a fraction of λ (or fractional λ pipe, FλP) upon entering a FλS network.

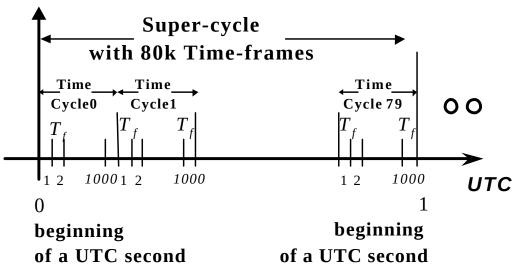

Fig. 2: UTC second divided in 12.5 μs time frames.

Data packets on the low capacity input channels of the grooming and degrooming devices depicted in Fig. 1 are organized in time frames. Due to the propagation delay on the links between (de)grooming devices, time frames are not aligned to the UTC at the receiving end of links. Thus, data packets received on each channel are first aligned to the UTC and then switched and multiplexed (or de-multiplexed and switched) onto the output channels. The result of the alignment operation is that the beginning and end of each time frame from each input channel occur at the same time as the beginning and end of the UTC time frames.

The time frame duration and its size (the number of bits that can be transmitted during one time frame) over the output channels can be different from the duration and size used over the input channels. In the most general case, time frame duration might not be the same on all the input channels. However, without loss of generality, in order to keep the following description simple it is assumed that the same time frame duration is used over all input channels and the same time frame duration is used over all output channels.

3. TIME AND FRACTIONAL l SWITCHING - FlS

This section studies the relationship between time measurements and scheduling in communications networks, in general, and FλS, in particular. The broad approach is taken in order to provide the rationale for FλS.

3.1 Why Time?

There are a number of answers; a simple one is that time minimizes latency through switches and grooming devices, as will be shown in Subsection 3.4. Minimizing latency translates in reducing the memory requirement, which is important for implementation in the optical domain since, as discussed in [1], the step from DRAM to optical memory corresponds to **an increase of more than 1,000,000 folds in the amount of silicon**.

3.2 Time Measurement

Measuring time between two events in the same location is performed locally by counting periodic rotations of various sorts. In ancient era the time was measured by counting the earth rotations, or, as some argued, the sun rotations around the earth. Since then, the measurement of time has improved dramatically.

According to the UTC *time-of-day* international standard time is measured by counting the oscillations of the cesium atom in multiple locations. In fact, 9,192,631,770 oscillations of the cesium atom define one UTC second. UTC is available everywhere around the globe from several distribution systems, such as, GPS (USA satellites system) [2], GLONASS (Russian Federation satellites system) [3], and in the future by Galileo (European Union and Japanese satellites system) [4], or can be distributed via communications satellites through the Two-Way Satellite Time and frequency Transfer (TWTFT) method [5].

3.3 Scheduling

Scheduling requires the ability to measure time. We consider scheduling with two time measurement methods:

1. Scheduling with local time based measurements. The delay between nodes cannot be measured, and therefore, the scheduling is based on local time. This method is used in circuit switching (e.g., SONET).
2. Scheduling with UTC-based measurements. The delay between nodes can be measured by using UTC and scheduling can be based on UTC. Scheduling with UTC implies no clock slips or drifts, and consequently, very simple implementation.

Fig. 3 and Fig. 4 are examples of the above two scheduling methods[1]. In these examples, scheduling is periodic and time is divided into time frames (TF) of predefined duration T_f. For example, a time frame of

1 Without loss of generality, the propagation delay between *Switch i* and *Switch j* was ignored.

10 μseconds is obtained by dividing one UTC second by 100,000. For periodic scheduling time frames are grouped into *time cycles*; for example, 1,000 time frames of 10 μseconds create a 10 millisecond time cycle.

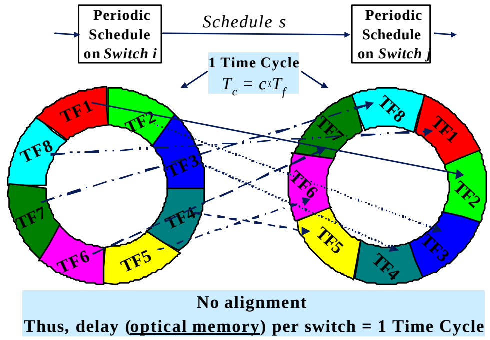

Fig. 3: Local time based-scheduling - **SONET**

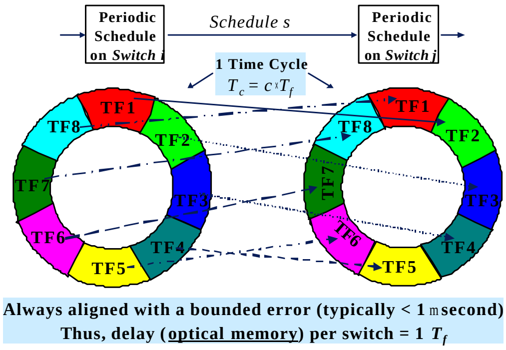

Fig. 4: UTC-based scheduling - **F l S**

3.4 Per Switch Delay and Memory Requirement

Let's assume that two neighboring switches, *Switch i* and *Switch j*, perform a given task — e.g., switching or transmitting data units — during predefined time frames according to a schedule, *Schedule s*. *Schedule s* repeats every time cycle, T_c, where $T_c = c \times T_f$. In the examples in Fig. 9 and Fig. 10, $T_c = 8 \times T_f$, and *Schedule s* on *Switch i* during time frame *k*, is scheduled on *Switch j* during time frame $(k+1)$ mod *8*.

When the scheduling on *Switch i* and *Switch j* is based on local time, the delay between *Schedule s* on *Switch i* and on *Switch j* is not known, and consequently, the delay between TF *k* and TF $(k+1)$ mod *c* is not known. Since the schedule repeats every time cycle, the maximum delay between a TF on *Switch i* and the corresponding TF on *Switch j* is one time cycle, T_c – where, $T_c = c \times T_f$.

When the scheduling on *Switch i* and *Switch j* is based on UTC, the delay between *Schedule s* on *Switch i* and on *Switch j*, is known, and consequently, the delay between TF *k* and TF $(k+1)$ mod *c* is known. Consequently, the maximum time between the execution of the aforementioned task in *Switch i* and in *Switch j* is only one time frame – T_f (which results from the actual data unit propagation delay between the two switches not being an integer number of time frames – a.k.a. quantization delay). Since data units need to be stored while waiting for the task execution in *Switch j*, the time between the two task executions determines the amount of memory required within the switches. Keeping memory requirements small is particularly important as all-optical networking becomes less far fetched.

3.5 SONET

SONET switches operate according to a reoccurring schedule that, as was mentioned before, is based on a local clock; consequently, data traversing a SONET switch are delayed up to a whole time cycle. Due to byte-by-byte channel multiplexing, the SONET time cycle is the time between the transmission of two successive bytes of the same channel. For example, the time cycle — hence the scheduling delay — of an STS-1 switch is 125/810 = 154 ns, independently of the line rate of its interfaces.

However, independently switching from input to output each incoming byte requires, for OC-192 line rates, switching and processing time well below 100 picoseconds. In order to overcome the picosecond accuracy requirements, a SONET switch might use a switching unit *x* times larger than the multiplexed one byte slot. However, this will imply a factor of *x* increase in the time cycle, and, since SONET scheduling uses local time measurements, in the per-switch delay and memory requirements. For example, STS-N frames can be byte-by-byte de-multiplexed into multiple STS-1 frames that are then switched as a whole. In this case, *x*=810 and the time cycle — hence the scheduling delay — of such an STS-1 switch is 125 μs.

Note that SONET also needs to process overhead information, such as, pointers to Synchronous Payload Environments. These pointers are needed since local time measurements on different switches are continuously drifting from one another.

4. Grooming and Degrooming Methods

Grooming is a process of multiplexing, where the number of inputs is (much) larger than the number of outputs—possibly one. On the other side, degrooming is a process of demultiplexing, where the number of inputs is (much) smaller—possibly one—than the number of outputs.

4.1 Intra-time frame grooming.

The time frame duration on the output channel is the same as on the input channel, however, the time frame size on the output channel is larger than on the input channel. For example, if the output channel capacity is *N* times larger than the input channel capacity, time frames on the output channel are *N* times larger than those on the input channel. Multiplexing is achieved by combining on the output channels *N* time frames worth of data packets received from the various input channels during the same time frame, as shown in Fig. 5(a). Data packets received from different input channels are multiplexed in the same time frame within separate *sub-time frames*. In the example depicted in Fig. 5(a) the time frame size on the output channel is four times the size of the time frames on the input channels. The time frames on the output channels are divided in four sub-time frames, each one with the same size as the input channel time frames.

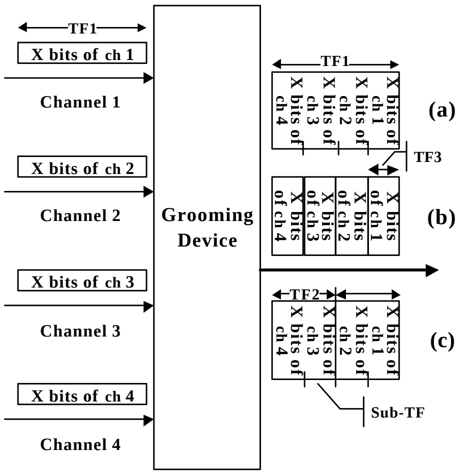

Fig. 5: (a) Intra-Time Frame Grooming, (b) Inter-Time Frame Grooming, and (c) Combined Approach.

4.2 Inter-time frame grooming.

The time frame size on the output channel is the same as on the input channels, however, due to the higher capacity of the output channel, the time frame duration is proportionally shorter. For example, if the output channel capacity is *N* times larger than the input channel capacity, time frames on the output channel are *N* times shorter than those on the input channels are. Multiplexing is achieved by interleaving on the output channels time frames worth of data packets received from the various input channels, as shown in Fig. 5(b). In the example depicted in Fig. 5 the output channel capacity is four times the capacity of the input channels (e.g., the output channel is an OC-192 channel, while each input channels is an OC-48 channel).

4.3 Combined Approach

Grooming that combines the above two methods, as shown in Fig. 5(c), where grooming is obtained by interleaving time frames containing data packets from different input channels (inter-time frame grooming), each time frame being composed of two sub-time frames (intra-time frame grooming).

4.4 Degrooming

Fig. 6 shows how degrooming is performed. Data packets are received on a high capacity input channel during each time frame. When *intra-time frame degrooming* is performed, as shown in Fig. 6(a), after data packets received from the input channel during one time frame have been aligned, they are transmitted during the same time frame on multiple output channels, data packets of each ingress sub-time frame over each channel. When *inter-time frame degrooming* is performed, as shown in Fig. 6(b), data packets received in *multiple consecutive time frames* from a high capacity input channel are transmitted during *one time frame* over *multiple lower capacity output channels*. Finally, as shown in Fig. 6(c), a combination of the two degrooming methods can be deployed.

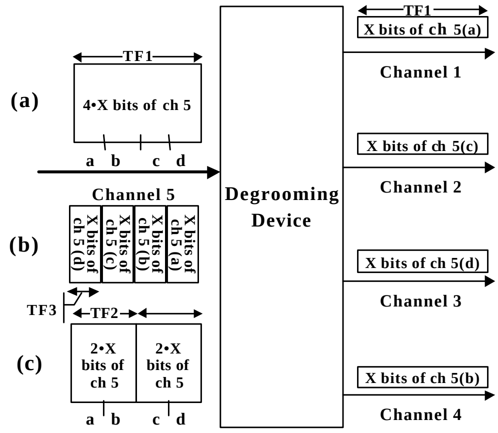

Fig. 6: (a) Intra-Time Frame Degrooming, (b) Inter-Time Frame Degrooming, and (c) Combined approach.

5. TIME FRAME HIERARCHY FOR GROOMING AND DEGROOMING

The time frame duration to be deployed on each communications channel is to be chosen wisely according to the channel capacity. In fact, a small time frame size is not desirable because the grooming and degrooming system operation is simpler and more efficient if any data packet transmitted on the channels fits within a single time frame, and even more if a few data packets are contained in each time frame. On the other hand, too large time frames are undesirable as well since in order to align received data packets with the UTC, all the data packets received in a time frame need to be buffered. Consequently, the larger the time frames, the larger the amount of memory needed to perform UTC alignment.

For what time frame duration is concerned, short time frames result in the need for a high speed control within each grooming and degrooming device, while long time frames result in longer alignment times and consequently longer delay introduced by grooming and degrooming equipement.

Table 1 shows a possible set of choices for the duration of time frames to be deployed over channels with capacities as defined by the SONET transmission hierarchy; choices for Gigabit Ethernet (GE) and 10 Gigabit Ethernet (10GE) channels are given as well. The last column of the table provides the ratio *s* between the time frame size deployed on a channel and the one deployed on the lower capacity channel, according to the presented time frame duration hierarchy. For example, when a time frame duration of 62.5 μs is used on an OC-12 channel and a time frame duration of 31.25 μs is used on an OC-48 channel, the time frame on the latter channel contains twice the amount of data packets as the time frame on the former channel. In other words, a intra-time frame multiplexing of two lower level time frames can be done on the OC-48 channel.

TABLE 1: TIME FRAME HIERARCHY BASED ON THE SONET CHANNEL HIERARCHY.

Channel Capacity **[Mb/s]**		***TF Duration*** **[ms]**	***TF Size*** **[Kbyte]**	***s***
51.84	OC- 1	250	1620	
155.52	OC- 3	250	4860	3
622.08	OC- 12	62.5	4860	1
2488.32	OC- 48	31.25	9720	2
9953.28	OC- 192	7.813	9720	1
1000	GE	50	6250	
10000	10GE	10	12500	2

6. DISCUSSION

This work has discussed how the deployment of UTC (Coordinated Universal Time) enables an efficient network architecture for grooming traffic from low speed edges to high capacity backbones, and vice versa (degrooming). The proposed grooming and degrooming method does not require processing packet headers, which is consequently confined to the low speed network edges. The paper discusses the advantages of using UTC in communications networks as a rationale for the presented UTC-based grooming and degrooming network architecture. Various grooming and degrooming methods are presented, trade-offs in the choice of their parameters are addressed, and a possible configuration is proposed.

The simple operation of the presented grooming and degrooming devices results in:

1. High scalability — a desirable property given the capacity increase brought by optical technologies in today's backbones.
2. Low cost — particularly important in a time in which telecom and service provider aim at minimizing their expenditures.

Moreover, the deterministic performance offered by the presented grooming and degrooming architecture — especially when coupled with backbones deploying Fractional λ Switching (FλS) [1] — enables a plethora of new high revenue services, such as high quality videoconferencing, video-on-demand, distributed games.

REFERENCES

[1] M. Baldi and Y. Ofek, "Fractional Lambda Switching," ICC 2002, Symposium on Optical Networking, Apr. 2002.

[2] National Institute of Standards and Technology (NIST), "GPS Data Archive," USA, http://www.boulder.nist.gov/timefreq/service/gpstrace.htm

[3] Russian Federation Ministry of Defense - Coordination Scientific Information Center, "Global Navigation Satellite System – GLONASS," Russian Federation, http://www.rssi.ru/SFCSIC/english.html

[4] European Union, "Transport-Satellite Navigation," Union Policies, Brussels, Belgium, June 2001, http://europa.eu.int/scadplus/leg/en/lvb/l24205.htm

[5] National Physical Laboratory, "Two-Way Satellite Time and frequency Transfer (TWTFT)," Teddington, Middlesex, UK, http://www.npl.co.uk/npl/ctm/twstft.html