Performance Evaluation of an Optical Packet “Scheduling Switch”

Kyriakos Vlachos, Kyriaki Seklou, Emmanuel Varvarigos
Optical Packet Switches Architectures

Several innovative architectures including:

• Switches with recirculating loops
  *Startlite Architecture*, A. Huang IEEE GLOBECOM 1984

• Staggering Switch

• Switch with Large Optical Buffers (SLOB) architecture
  *D. Hunter et al IEEE/OSA J. Lightwave Technol. 1998*

• Wavelength Routing Switch – WRS
  *M. Renaud et al. IEEE Commun. Mag. 1997*

• Broadcast and Select Switch – BSS
  *M. Renaud et al. IEEE Commun. Mag. 1997*

However, work on new architectural concepts, node’s performance, and intelligent control have lagged behind progress in transmission speeds.
The “Scheduling Switch Architecture”

Concept:

Use a branch of delays to schedule packets in a T size frame and resolve contention.

Each delay branch consist of $2m-1$ delay blocks, where $m = \log T$.

The $i$th block consists of a three-state (or two 2x2) optical switch and three fiber delay paths, corresponding to delays equal to 0, $2^i$ and $2^{i+1}$ slots.

$T$ is assumed to be a power of 2 and corresponds to the maximum number of sequential packets from all incoming links that request the same output and can be served with no contention.
Traffic Assumptions

- We assume that the time axis on a link is divided into slots of equal length and every $T$ slots are virtually grouped to form a frame.

- A packet is an integer number of slots.

- A session is said to have the $(n,T)$ - burstiness property at a node if at most $n$ packets of the session arrive at that node during a frame of size $T$.

- The frame size $T$ can be viewed as a measure of the traffic burstiness allowed. The larger $T$ is, the less constrained (more bursty) is the incoming traffic allowed to be, and the larger is the flexibility -granularity– in assigning rates to sessions.

- Loss less operation of a scheduling switch network is obtained when

$$
\sum_{i=1}^{k} n_{i,j} \leq T
$$

for all $j \{1,2,..,k \}$ where $n_{ij}$ is the number of packets from input $i$ destined to output $j$.
Assuming that packets arrive independently at each incoming slot with probability \( p \), the probability of having \( i \) packets arrivals during the \( kT \) slots of the \( k \) incoming frames requesting the same output \( j \), \( j = 1, \ldots, k \), and assuming uniformly distributed destinations is:

\[
P [X = i] = \binom{kt}{i} \cdot \left( \frac{p}{k} \right)^i \cdot \left( 1 - \frac{p}{k} \right)^{kT-i}
\]

The packet loss ratio can then be easily calculated as:

\[
PLR = \frac{\sum_{i=T}^{kt} P[X = i] \cdot (i - T)}{p \cdot T}
\]

\[
= \frac{\sum_{i=T}^{kt} \left[ \binom{kt}{i} \cdot \left( \frac{p}{k} \right)^i \cdot \left( 1 - \frac{p}{k} \right)^{kT-i} \cdot (i - T) \right]}{p \cdot T}
\]

\[
= \frac{\sum_{i=T}^{kt} \left[ \frac{kT!}{i!(kT-i)!} \cdot \left( \frac{p}{k} \right)^i \cdot \left( 1 - \frac{p}{k} \right)^{kT-i} \cdot (i - T) \right]}{p \cdot T}
\]
For $T$ values higher than 32 and $p < 0.8$ the packet loss ratio is very low.

$T$ values of 32, 64 and 128 can be accomplished with all-optical technologies at low cost and with a low complexity.


Packet loss ratio for (a) $k=2$ and (b) $k=4$ input/output scheduler switch for binomial packet traffic and uniformly distributed destinations.
Packet loss ratio versus $T$ for (a) $k=2$ and (b) $k=4$ and for a utilization $\rho = \{0.1, 0.2, \ldots, 1\}$.

For $\rho=1$, packet loss ratio is $9 \cdot 10^{-3}$ and $11 \cdot 10^{-3}$, when $T=2^{10}$ for $k=2$ and $k=4$ respectively.
Performance Evaluation for Constrained \((n, T)\) Bursty Traffic

We assume that:

- Incoming traffic obeys the \((n, T_{\text{traffic}})\) smoothness property while the Scheduling Switch has been designed for \(T_{\text{switch}} \text{ with } T_{\text{traffic}} \geq T_{\text{switch}}\)

- \(T_{\text{traffic}}\), is an integer multiple of the corresponding \(T_{\text{switch}}\) parameter

- The ratio \(T_{\text{traffic}} / T_{\text{switch}}\) is viewed as an index of the traffic burstiness allowed in the network.

- Assuming that the link utilization is \(p\) then the number of packets \(n\) that may arrive during a frame \(T_{\text{traffic}}\) and request the same outgoing switch port is:

\[
\sum_{i=1}^{k} n_{i,j} = pT_{\text{traffic}} \quad \text{for all outputs } j.
\]

The \(pT_{\text{traffic}}\) packets that arrive per incoming frame and request output \(j\) are evenly distributed within the frame of size \(T_{\text{traffic}}\).
Performance Evaluation for Constrained \((n,T)\) Bursty Traffic

The \(pT_{traffic}\) can arrive in any of the \(\binom{KT_{traffic}}{PT_{traffic}}\) possible combinations.

Thus, the probability of having \(i\) packets within the \(T_{switch}\) first slots \(P_i\) is:

\[
P_i = \binom{KT_{switch}}{i} \cdot \binom{KT_{traffic} - KT_{switch}}{PT_{traffic} - i}
\]

And the corresponding Packet Loss Ratio:

\[
PLR = \sum_{i=1}^{pT_{traffic}} \left( \binom{KT_{switch}}{i} \cdot \binom{KT_{traffic} - KT_{switch}}{PT_{traffic} - i} \right) \cdot (i - T_{switch})
\]

Equation is valid only for \(pT_{traffic} > T_{switch}\), while for \(pT_{traffic} = T_{switch}\) or \(T_{traffic} = T_{switch}\), the packet loss ratio is zero for any utilization factor \(p\).
Performance Evaluation for Constrained \((n,T)\) Bursty Traffic

Packet loss ratio for (a) \(T_{\text{switch}} = 2\) and (b) \(T_{\text{switch}} = 16\), versus the \(T_{\text{traffic}} / T_{\text{switch}}\) ratio for a \(k=2\) and \(k=4\) scheduling switch and a utilization \(p = \{0.25, 0.5, 0.75, 1\}\).

\(T_{\text{traffic}}\) is varied from \(2 \cdot T_{\text{switch}}\) to \(2^{10}\).

Packet loss ratio decreases when \(T_{\text{traffic}} / T_{\text{switch}}\) increases (beyond 2). This is primarily due to the burstiness averaging as a result of the numerous possible packet distributions within a \(T_{\text{traffic}}\) frame.
Performance Evaluation for Pareto traffic

- Packets arrive in bursts (ON periods), which are separated by idle periods (OFF periods).
- ON periods is burst – train of packets with a Pareto distribution. The min. burst size is 1, corresponding to a single packet arrival.
- OFF periods with a min. size of \(b_{\text{off}}\)

Formula we used:

\[
X_{\text{PARETO}} = \frac{b}{x^{1/a}}
\]

where:
- \(x\) is a uniformly distributed value in the range \((0, 1]\),
- \(b\) is the minimum non-zero value of \(X_{\text{PARETO}}\), denoted by \(b_{\text{on}}\) and \(b_{\text{off}}\) for the packet train and idle period respectively and
- \(a\) the tail index or shape parameter of the Pareto distribution.

Especially for computer simulation the \(b_{\text{off}}\) must be defined due to the finite range of \(x\).
Performance Evaluation for Pareto traffic

Starting from: 
\[ p = \frac{\text{ON}_{\text{period}}}{\text{ON}_{\text{period}} + \text{OFF}_{\text{period}}} \]

and:
\[ X_{\text{Pareto}}^{\text{max}} = \frac{b}{x_{\text{min}}^{1/a}} \]

We calculate:
\[ E(x) = \int_{b}^{x_{\text{Pareto}}^{\text{max}}} xf(x)dx = \int_{b}^{x_{\text{Pareto}}^{\text{max}}} x \frac{ab^a}{x^{a+1}}dx = \frac{ab}{a-1} \left[ 1 - x_{\text{min}}^{a-1} \right] \]

and thus:
\[ b_{\text{off}} = \frac{a_{\text{off}} - 1}{a_{\text{off}}} \cdot \frac{1 - x_{\text{min}}}{a_{\text{on}} - 1} \cdot \frac{a_{\text{on}} - 1}{a_{\text{off}}} \cdot \left( \frac{1}{p} - 1 \right) \]
Packet loss ratio for (a) k=2 and (b) k=4 versus link utilization for $T \in [2 \ldots 64]$ and $T = 1024$. $a_{ON} = 1.7$, $a_{OFF}= 1.2$. 
Simulated setup:

4 edge routers, generating Pareto traffic with load $p$.
Within ER VQO is implemented.

Scheduling Algorithm: Round Robin for selecting an ER.
FIFO within each ER.

The FIFO property within each ER is relaxed only when equation $\sum_{i=1}^{k} n_{i,j} \leq T$ is violated

*The algorithm is designed to minimized holding times and maximize link load (all slots of an outgoing frame are filled)*
We have simulated four ERs, each with an input load $p \in [0 \ldots 1]$ and $T \in [T \ldots 1024]$.

Simulations have been carried out for a workload per source value of 1.

**Conclusions:**

The induced delay is relatively small and that the incoming–outgoing packet process enters its steady state within a few thousand outgoing frames with a worst-case finite holding time.
Thank you!!

Work supported by EU FP6 via the Network of Excellence e-Photon/ONe project