Impact of Assembly Algorithms on End-to-End Performance in OBS Networks

Maurizio Casoni

casoni.maurizio@unimore.it

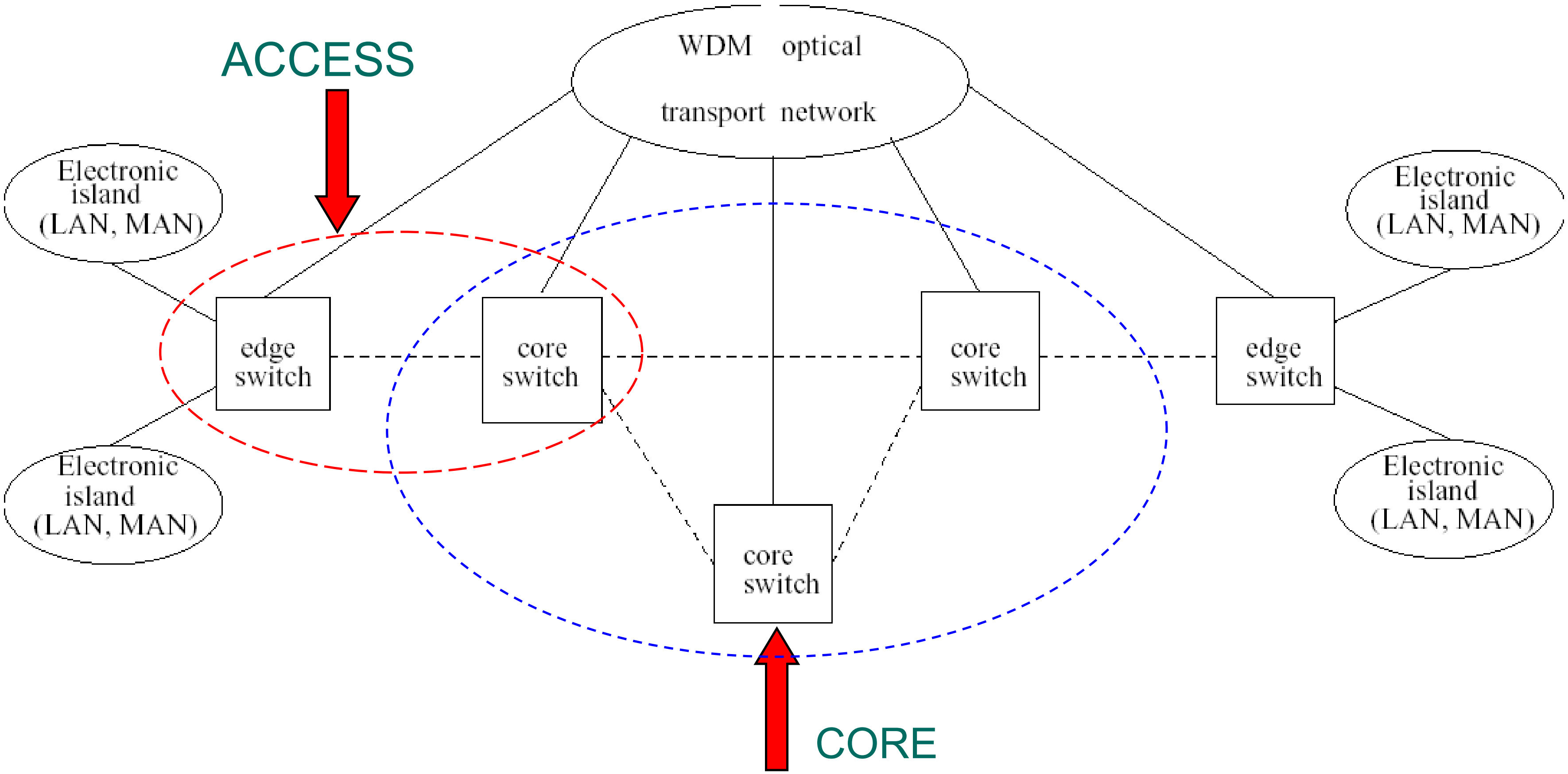
WONET - Wireless and Optical NEtworking Team

Department of Information Engineering
University of Modena and Reggio Emilia
Italy

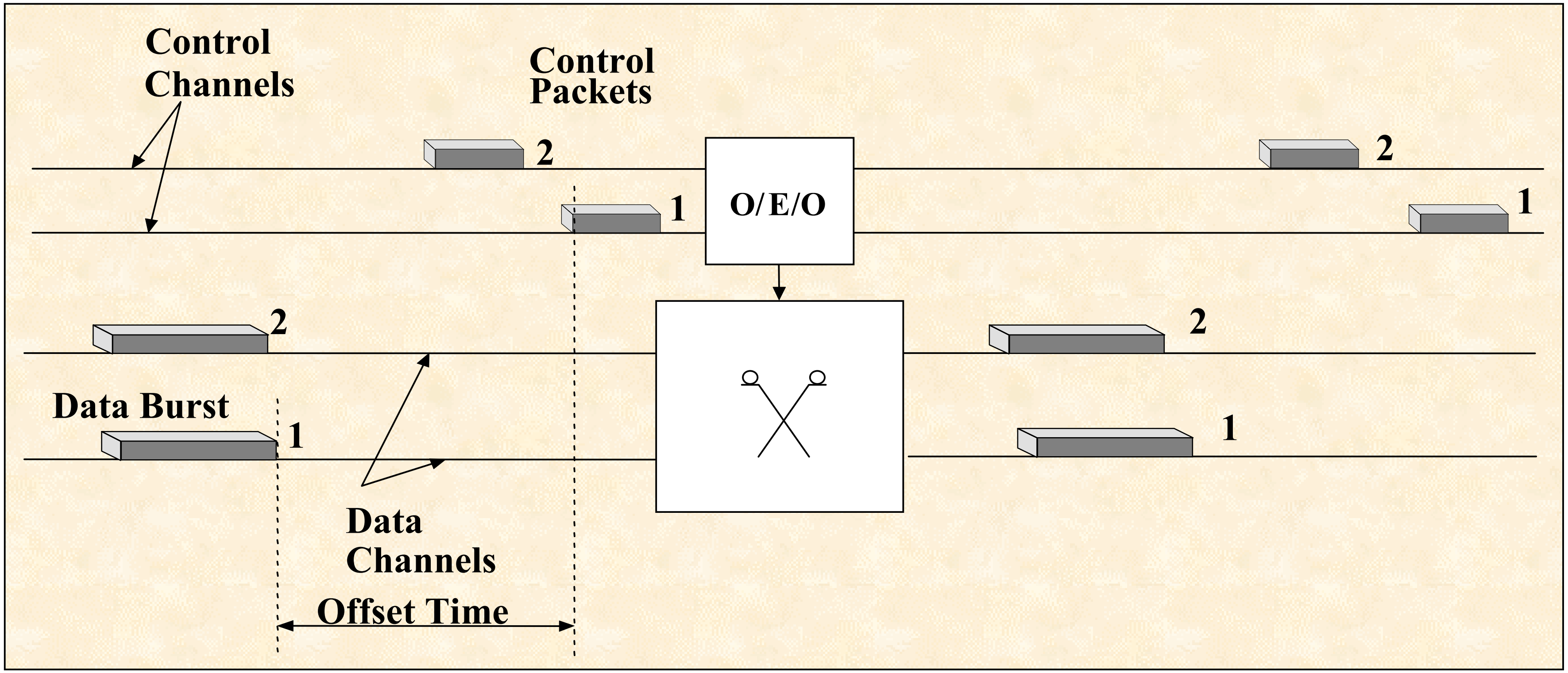
OUTLINE

- **Introduction: Scenario and Target**
- **Goal: effects of BA on e2e performance**
- **Some burst assembly algorithms**
- **MOBSim tool**
- **Case Study: Pan-European Network (Cost 266 simplified)**
- **Numerical results**
- **Conclusions**

INVESTIGATED SCENARIO



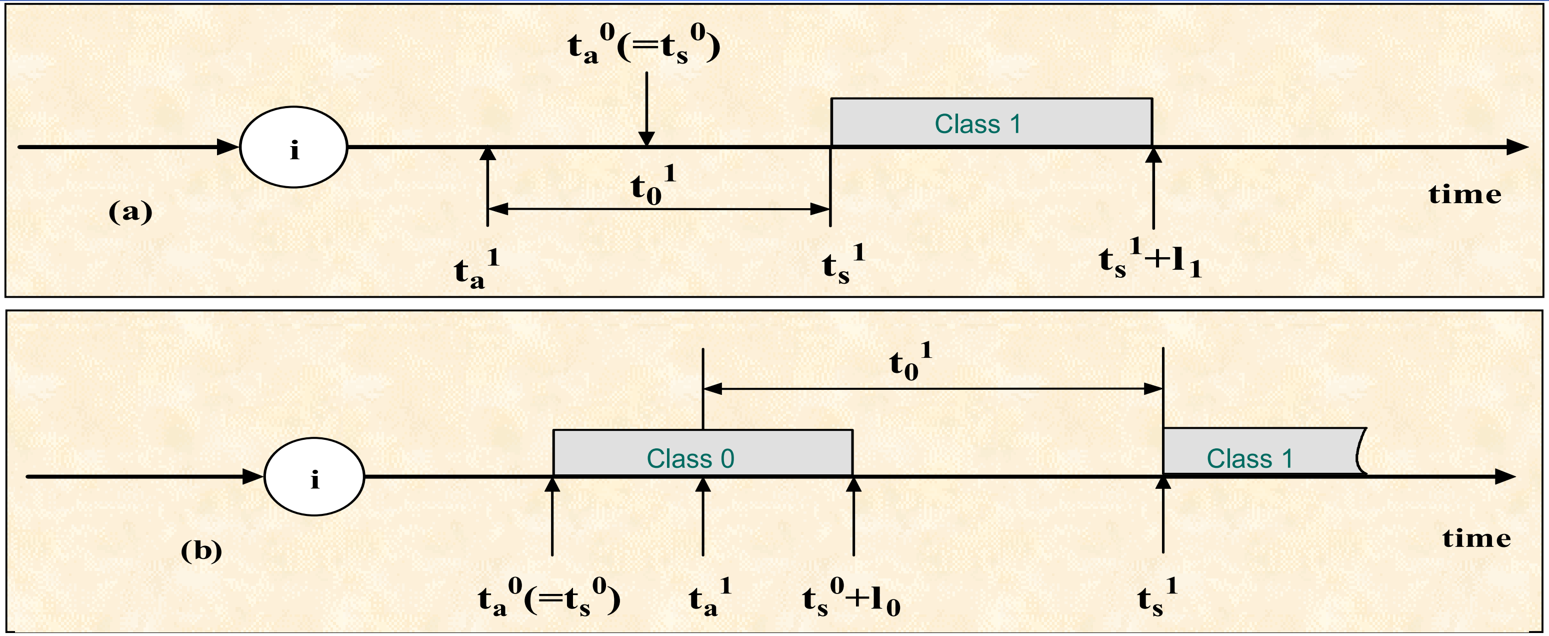
OPTICAL BURST SWITCHING (OBS)



- Dynamic setup of a wavelength path in presence of large data flows
- Data never leave the optical domain; control on separate channels
- Control precedes data by a basic offset time

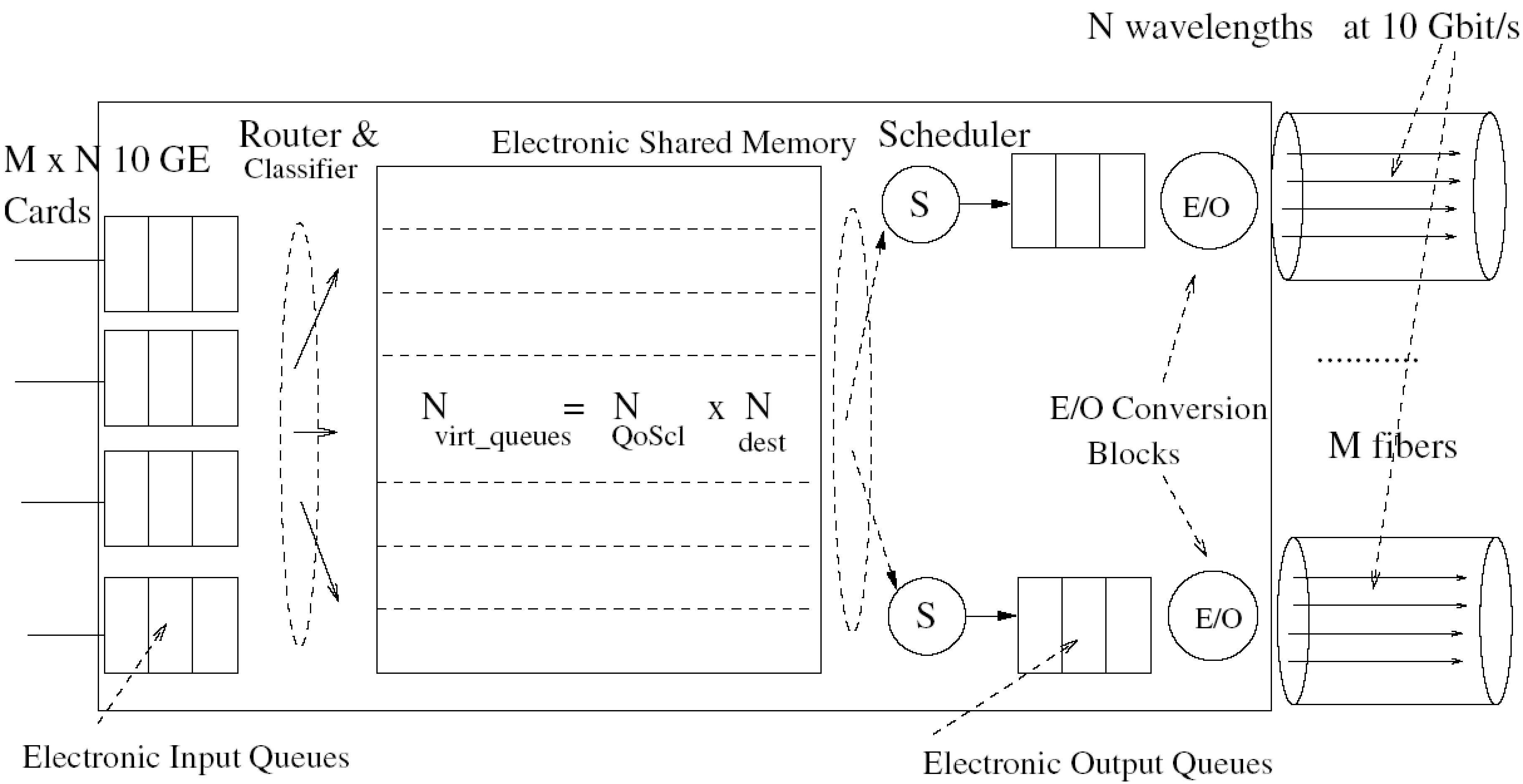
Good trade-off efficiency-feasibility → Mid-term solution

OBS with JUST ENOUGH TIME (JET)



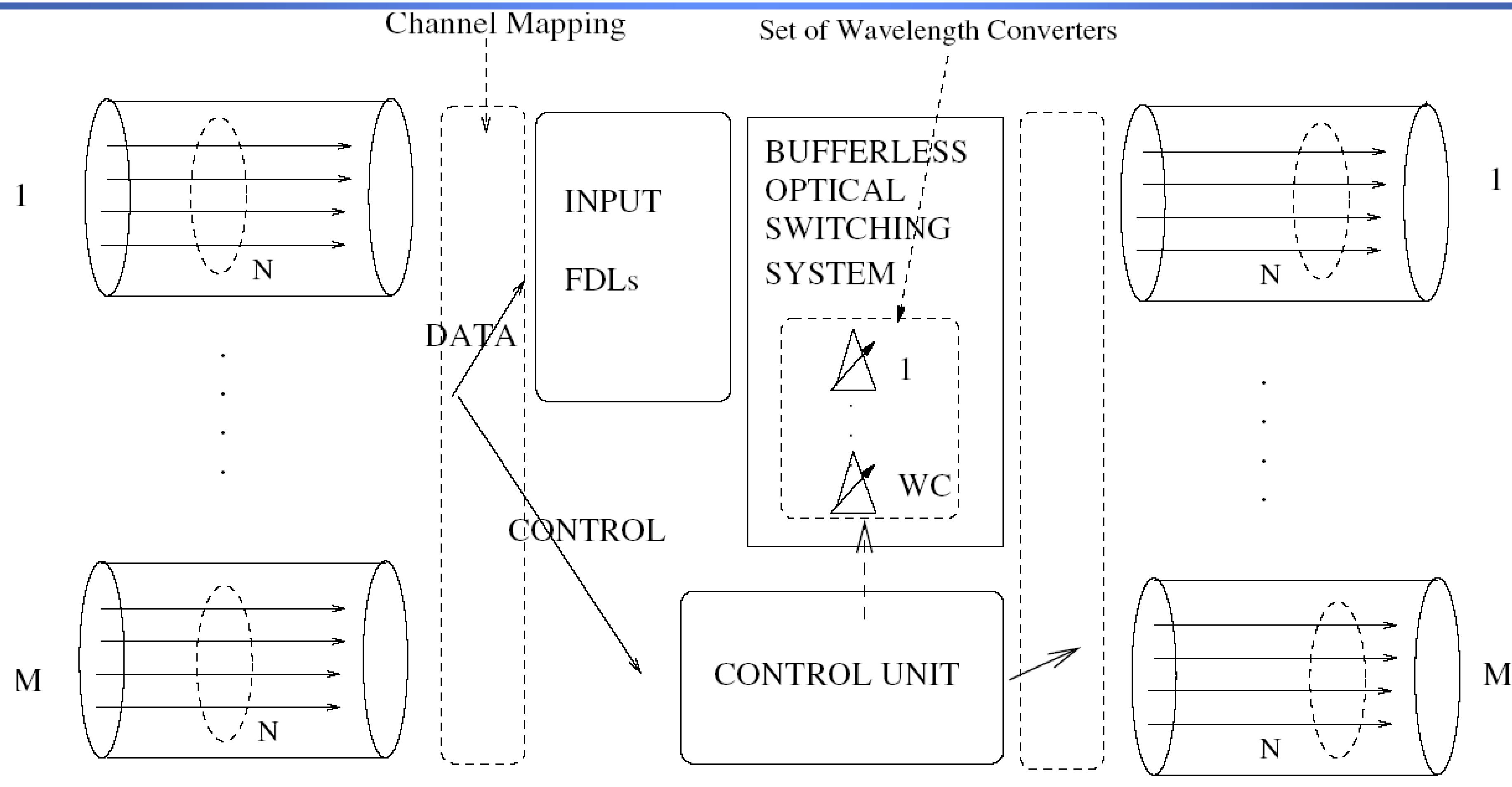
- OBS node reserves resources for the burst duration only
- Offset may include an optional extra-offset for QoS purposes
- Algorithms/protocols are required to properly manage optical resources
- IP&Optical control plane integration: MPLS paradigm
- MPλS maps LSPs into wavelengths
- LOBS: label carried by control packets releasing the wavelength resource

Edge node



N wavelengths at 10 Gbit/s
M x N 10 GE
Cards
Router &
Classifier
Electronic Shared Memory
Scheduler
S
E/O
N virt_queues = N QoScl x N dest
E/O Conversion
Blocks
M fibers
....
S
E/O
Electronic Input Queues
Electronic Output Queues

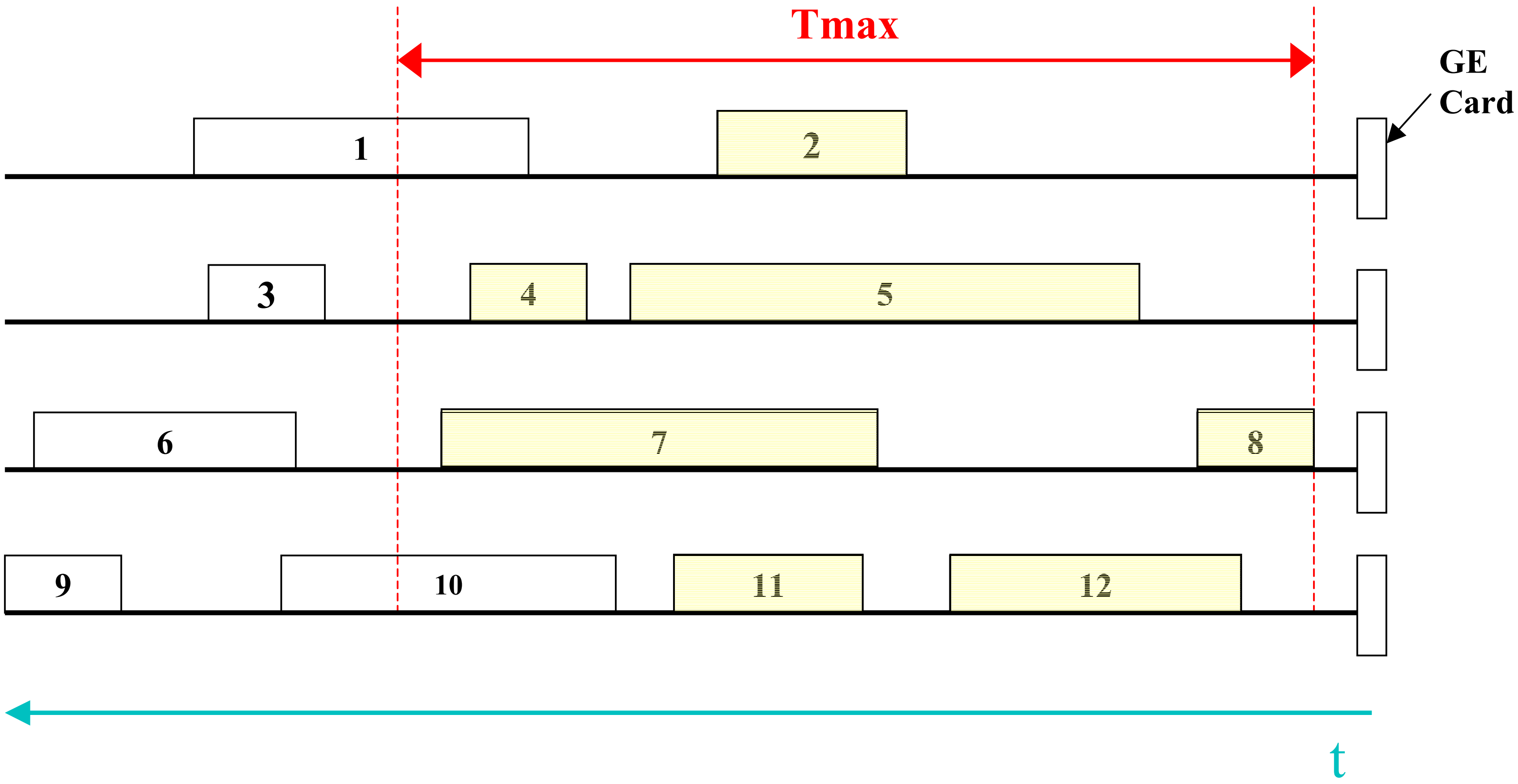
Core Router


Channel Mapping
Set of Wavelength Converters
1
N
M
N
DATA
CONTROL
INPUT
FDLs
BUFFERLESS
OPTICAL
SWITCHING
SYSTEM
1
WC
CONTROL UNIT
1
N
M
N

Edge Node: Investigated Algorithms

- ✓ Maximum Time
- ✓ Minimum Length - Maximum Time v.1
- ✓ Minimum Length - Maximum Time v.2

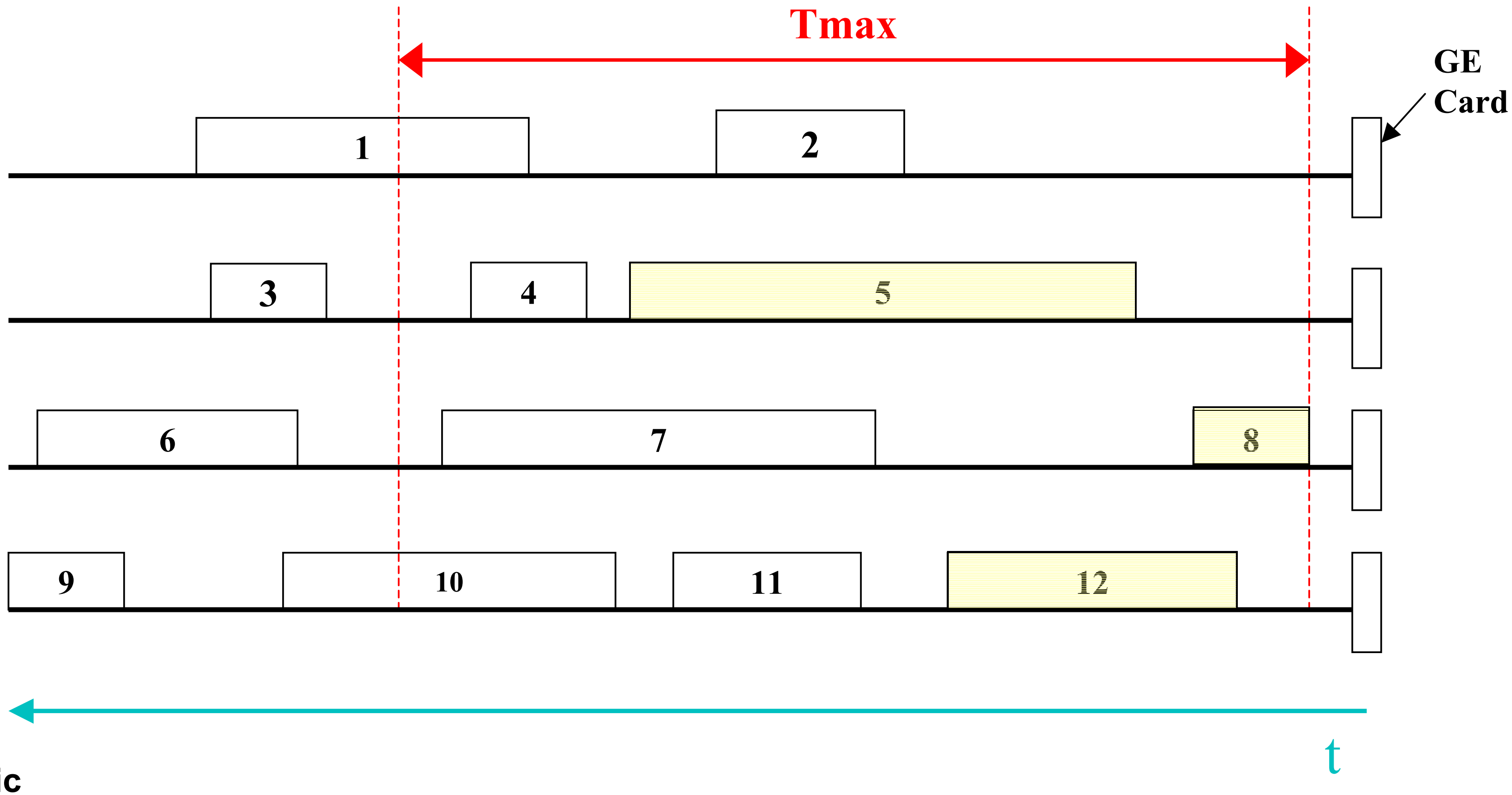
Maximum Time



IP traffic

- ✓ **towards the same destination**
- ✓ **belonging to the same QoS class**

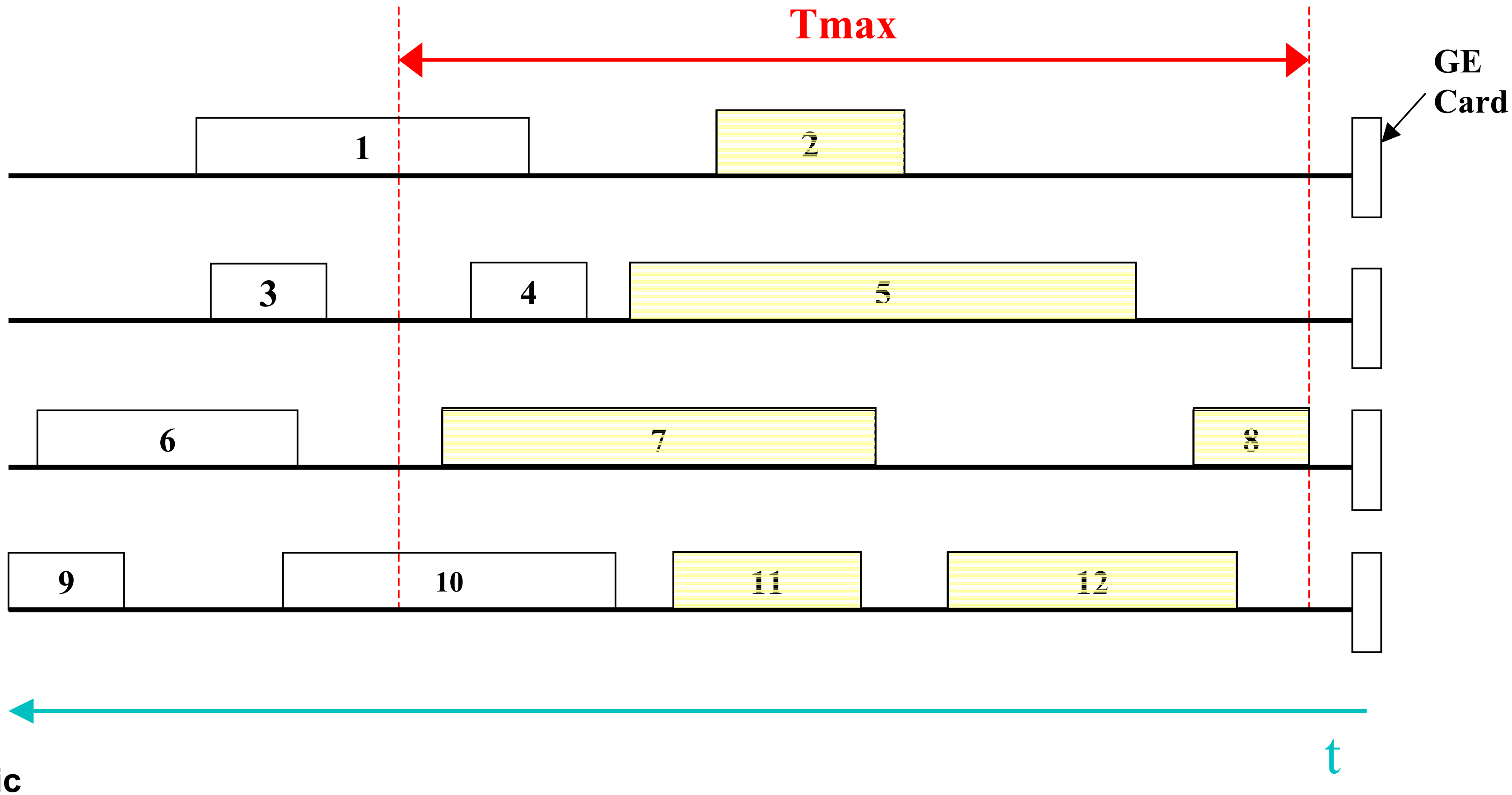
Minimum Length - Maximum Time v.1



IP traffic

- **towards the same destination**
- **belonging to the same QoS class**

Minimum Length - Maximum Time v.2

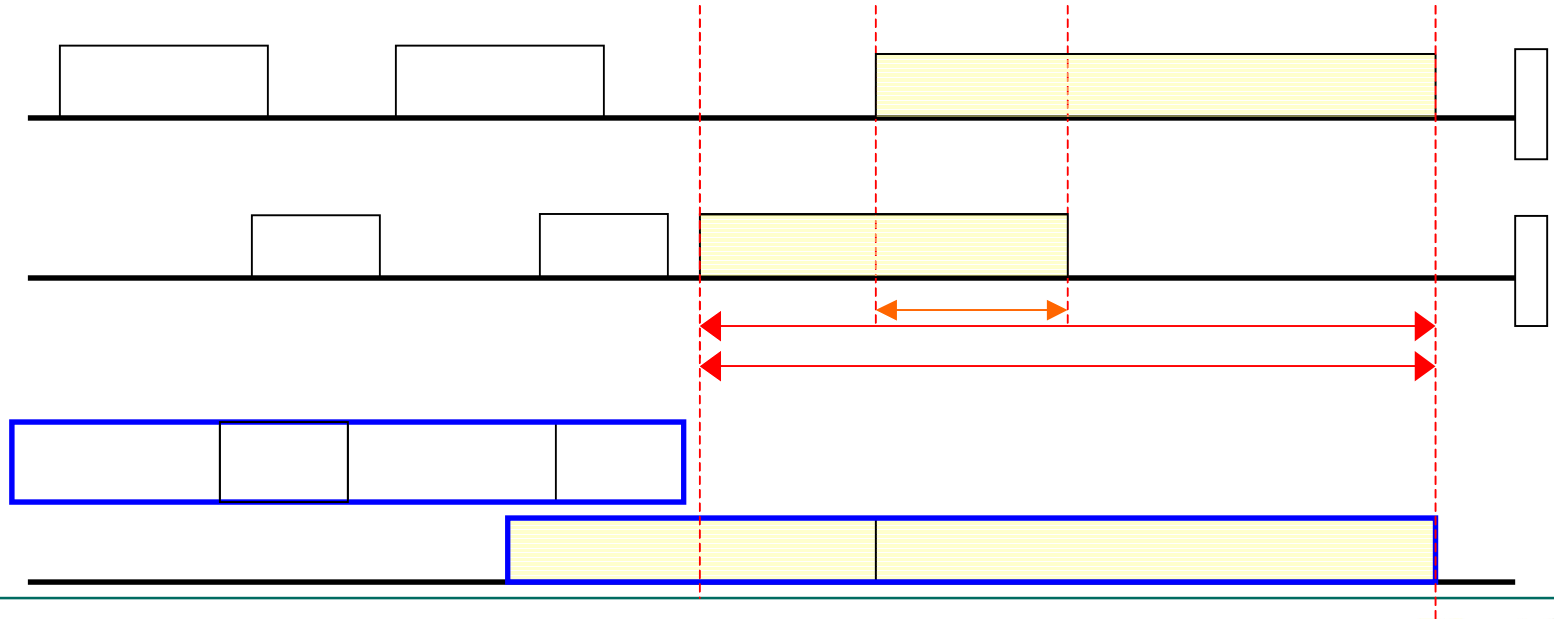


IP traffic

- **towards the same destination**
- **belonging to the same QoS class**

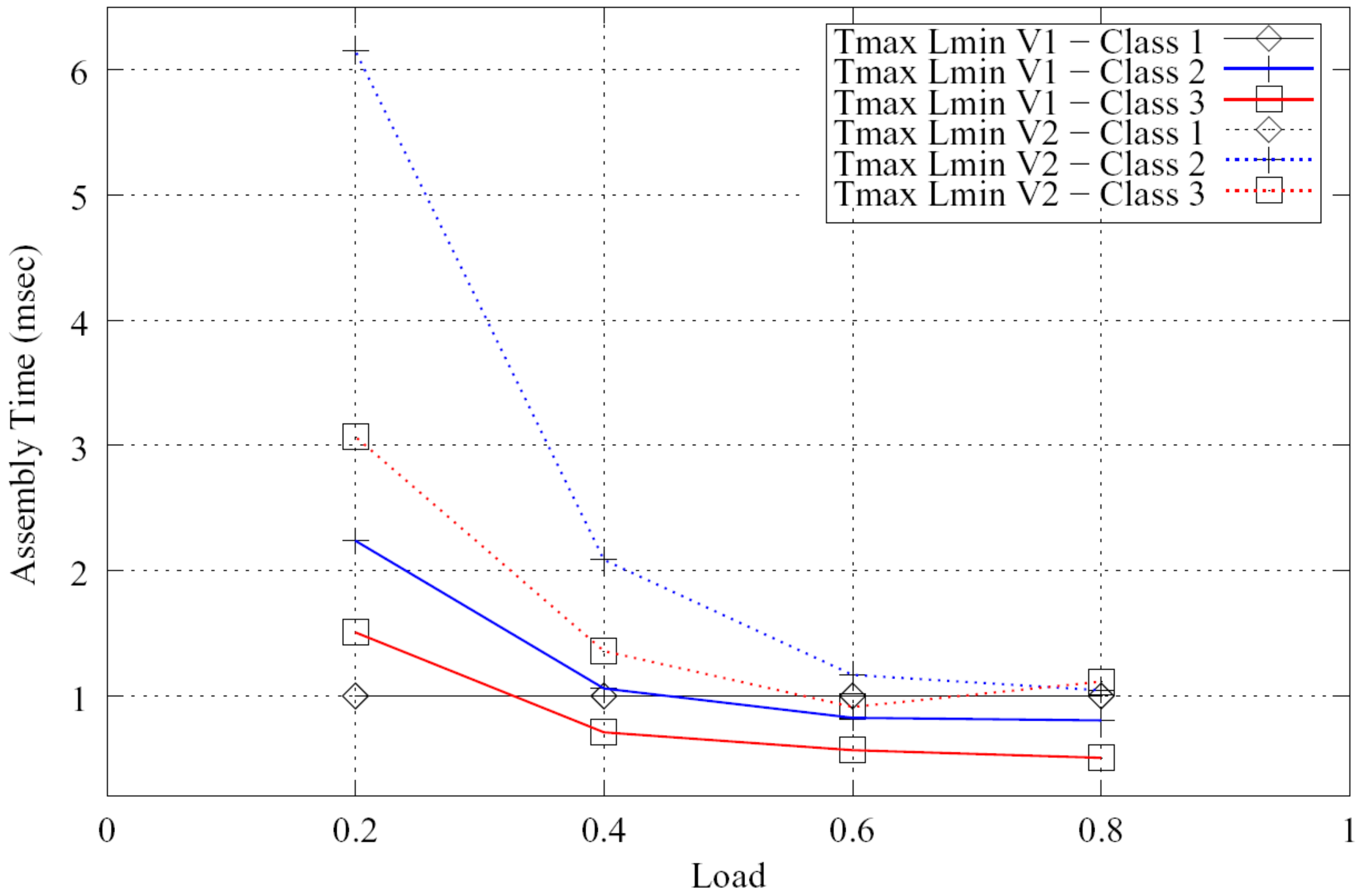
Main Investigated Parameters

- Input queue delay
- Assembly Time
- Assembly Length
- Overlapping bursts during transmission
- Output queue delay



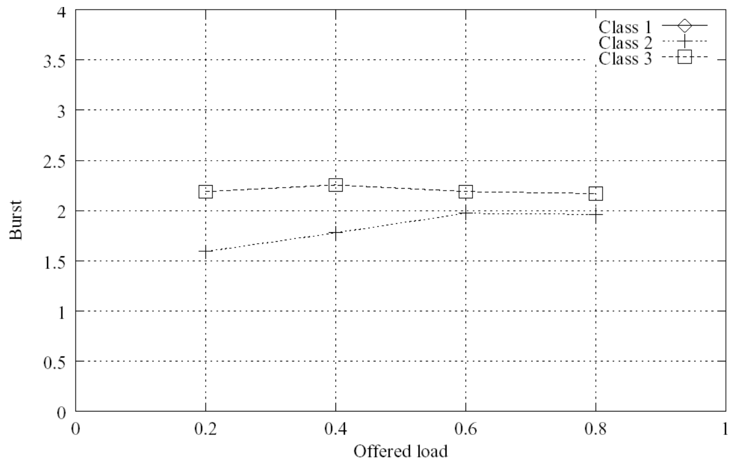
Assembly time

- T_{max1}= 1 ms
- $T_{max2/3}$ = 20 ms
- L_{min} = 640 kbytes



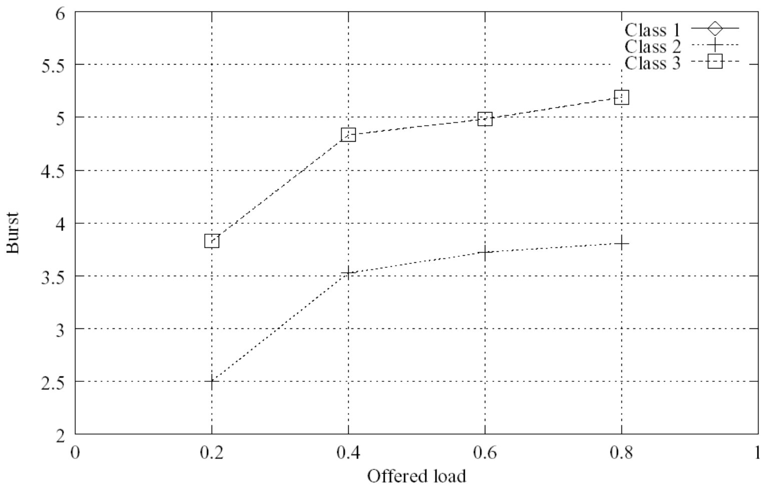
Overlapping Bursts

Tmax

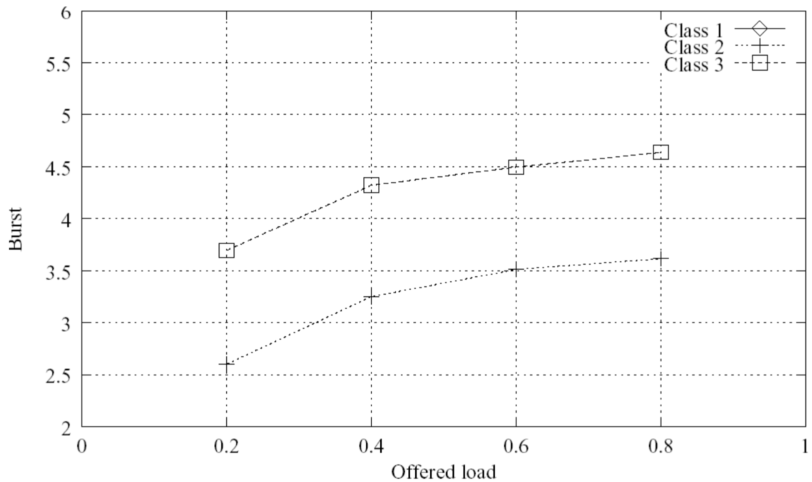


- ✓ Tmax has less but very longer overlapping bursts
- ✓ v.2 gives less than v.1

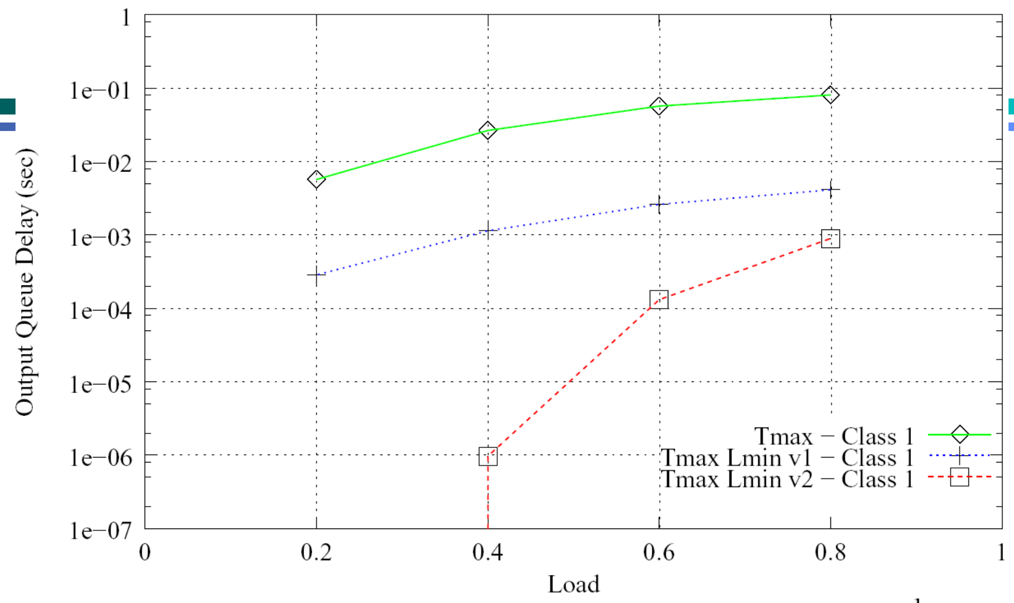
Tmax Lmin v.1



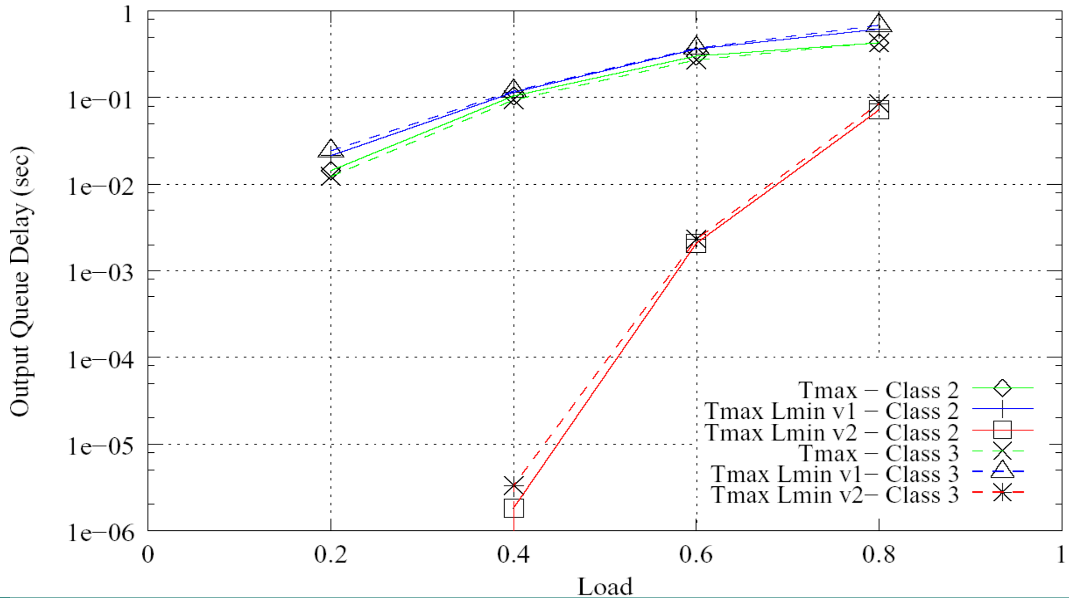
Tmax Lmin v.2



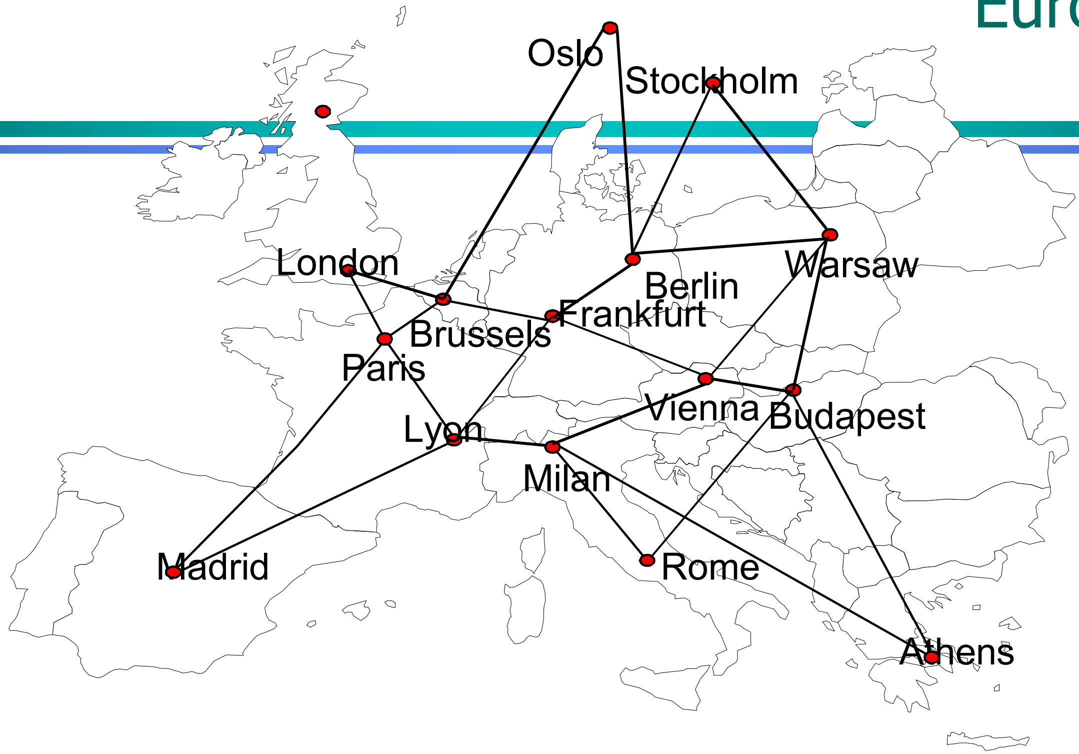
Output queue delay



v.2 seems to provide the best trade-off between burst length and number of overlapping bursts



European Network

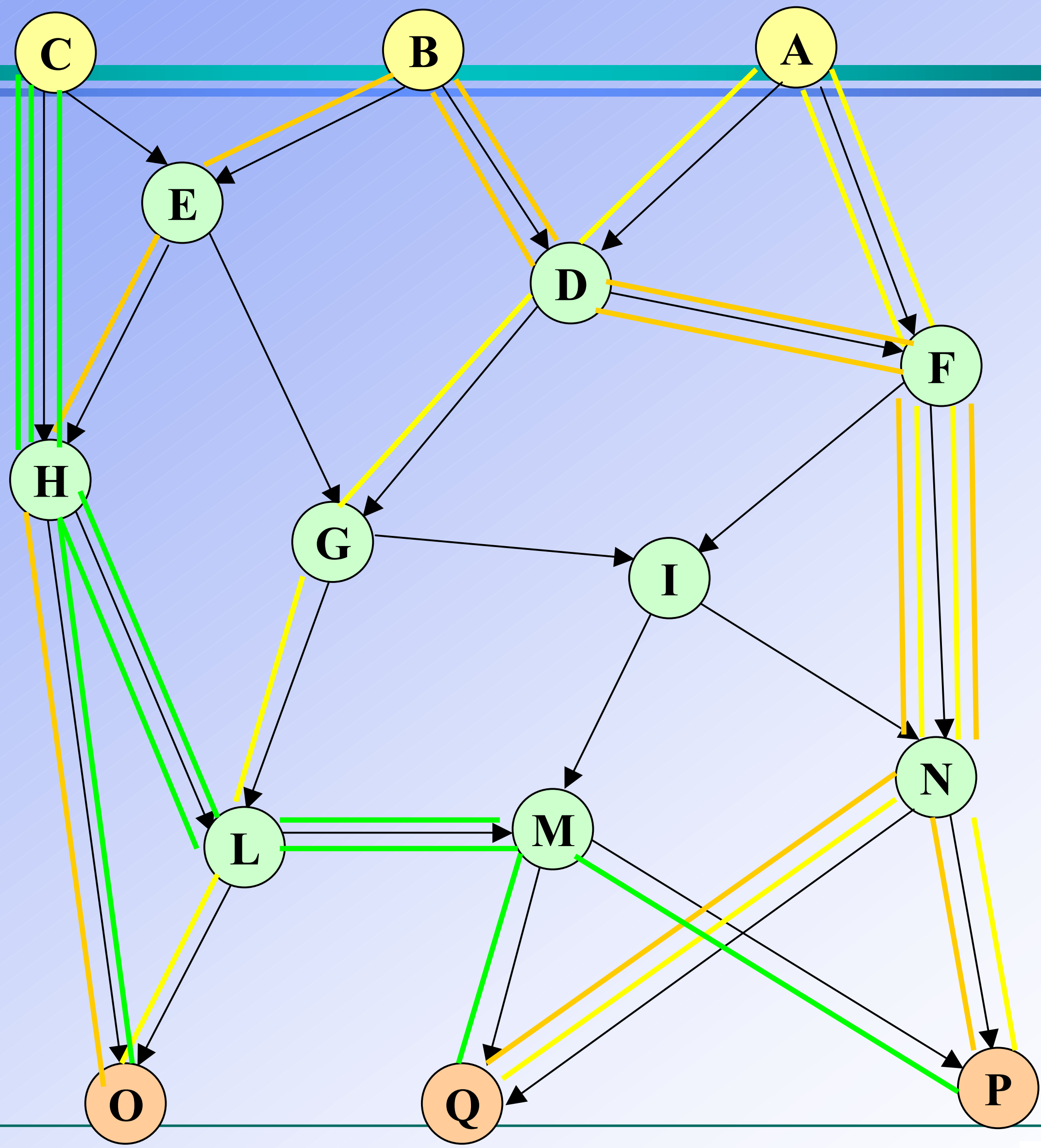


16 λs per fiber

QoS class for burst	Average Pareto ON period for incoming datagrams	Percentage	JET extra offset (8 μs) use	N. of shared wavelength converters	Deflection routing
Class 1	218 bytes	50%	Y	32	N
Class 2	10 Kbytes	20%	N	32	Y
Class 3	10 Kbytes	30%	N	0	Y

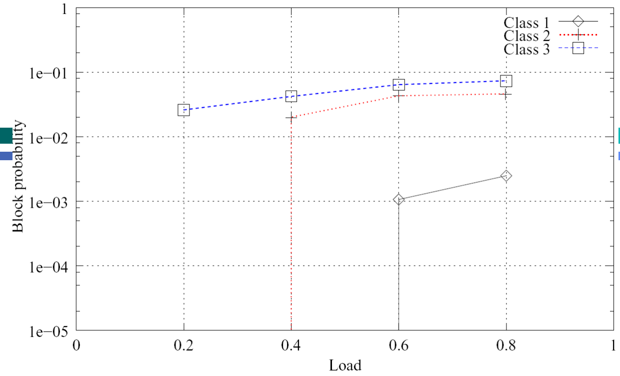
European Network

A – Stockolm
B – Oslo
C – London
D – Berlin
E – Bruxelles
F – Warsaw
G – Frankfurt
H – Paris
I – Vienna
L – Lyon
M – Milan
N – Budapest
P – Athens
Q – Rome
O - Madrid

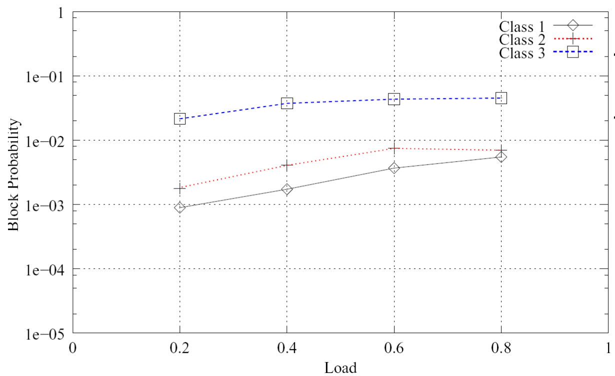


Blocking Probability from London to Rome

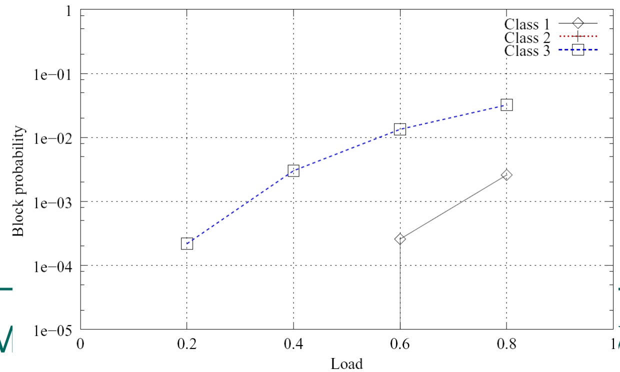
Tmax

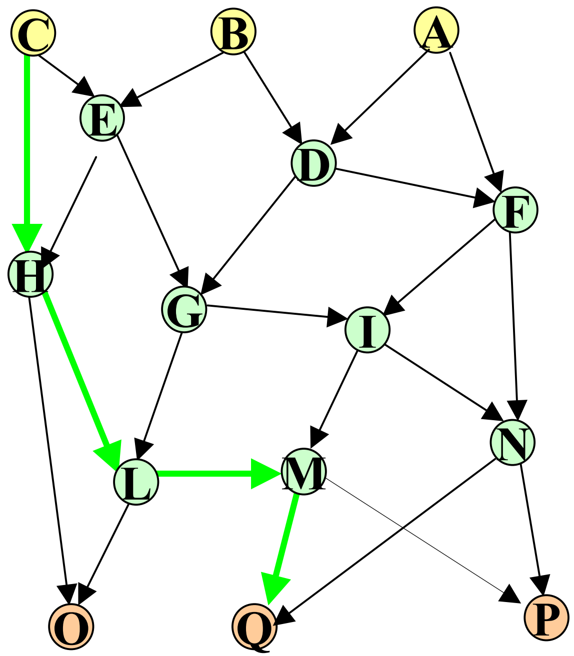


Tmax Lmin v.1



Tmax Lmin v.2

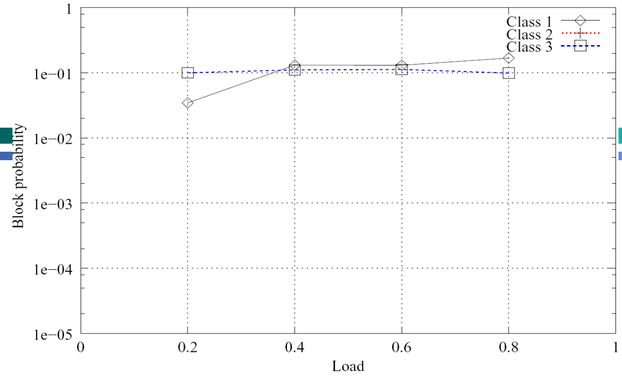




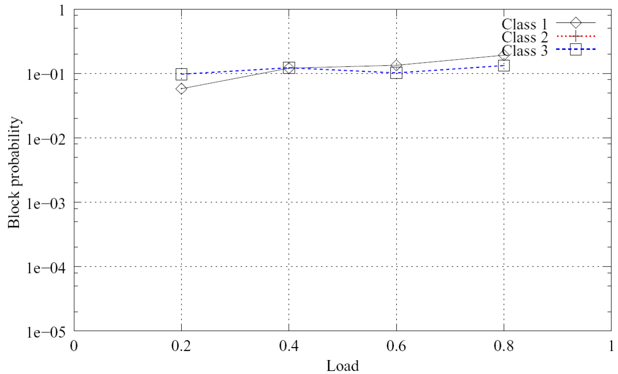
- Tmax Lmin v.2 seems to be the best

Blocking Probability from Stockolm to Rome

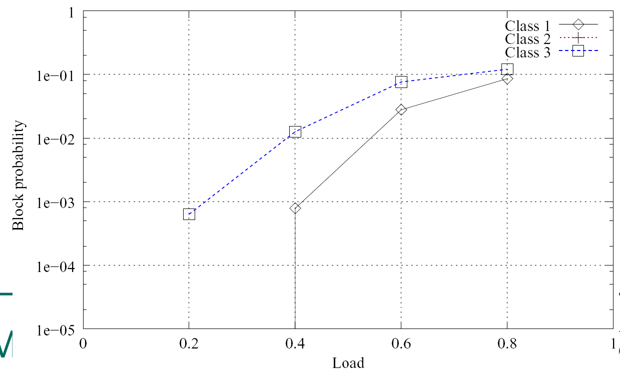
Tmax

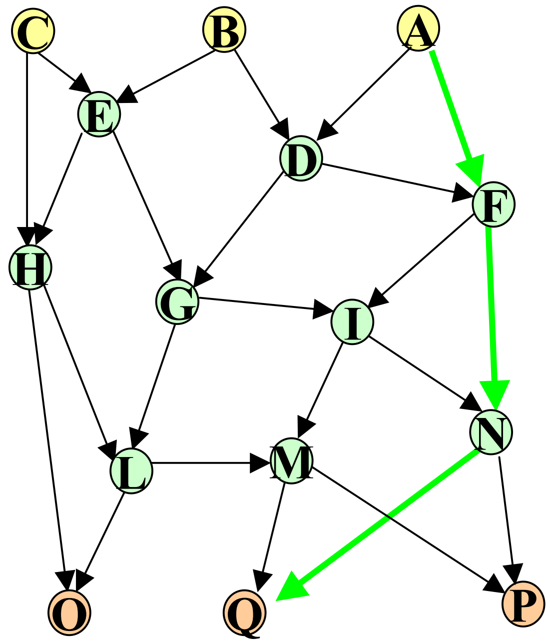


Tmax Lmin v.1



Tmax Lmin v.2





Performance depend on the path: in case of severe congestion (node F) degradation can be remarkable

Delay in Edge Nodes

- Incoming IP datagrams undergo a cascade of delays by input, shared and output buffers
- Let us assume a 0.6 load

Tmax Lmin v.1:

class 3: total delay = 10 + 0.5 + 250 ≅ 260 ms

class 2: total delay = 8 + 1 + 250 = 259 ms

class 1 : total delay = 0.1 + 1 + 2 = 3 ms

Tmax Lmin v.2:

class 3: total delay = 2 + 1 + 2 = 5ms

class 2: total delay = 1 + 1.2 + 2 ≅ 4 ms

class 1 : total delay = 0.1 + 1 + 1 ≅ 2 ms

Tmax:

class 1: output delay=80 ms !!!

End to End Performance

TCP Reno:

$$Thr = \frac{MSS}{RTT\sqrt{\frac{2bp}{3}} + T_o \min\left(1,3\sqrt{\frac{3bp}{8}}\right) p\left(1+32p^2\right)}$$

- Edge-to-edge one way delay:

$$T_{e2e1way} = T_{assembly} + N_{hops} \times T_{hop} + T_{disass}$$

- RTT $\approx$ 2 x $T_{e2e1way}$
- Assuming 800 km average link length ($T_{hop} \approx$ 4 ms) and N_{hops} in [3..5] then N_{hops} x $T_{hop} \leq$ 20 ms
- Regarding p, assuming "slow" TCP sources implies that at most one segment per connection is in a generated burst: p $\approx$ burst blocking prob.

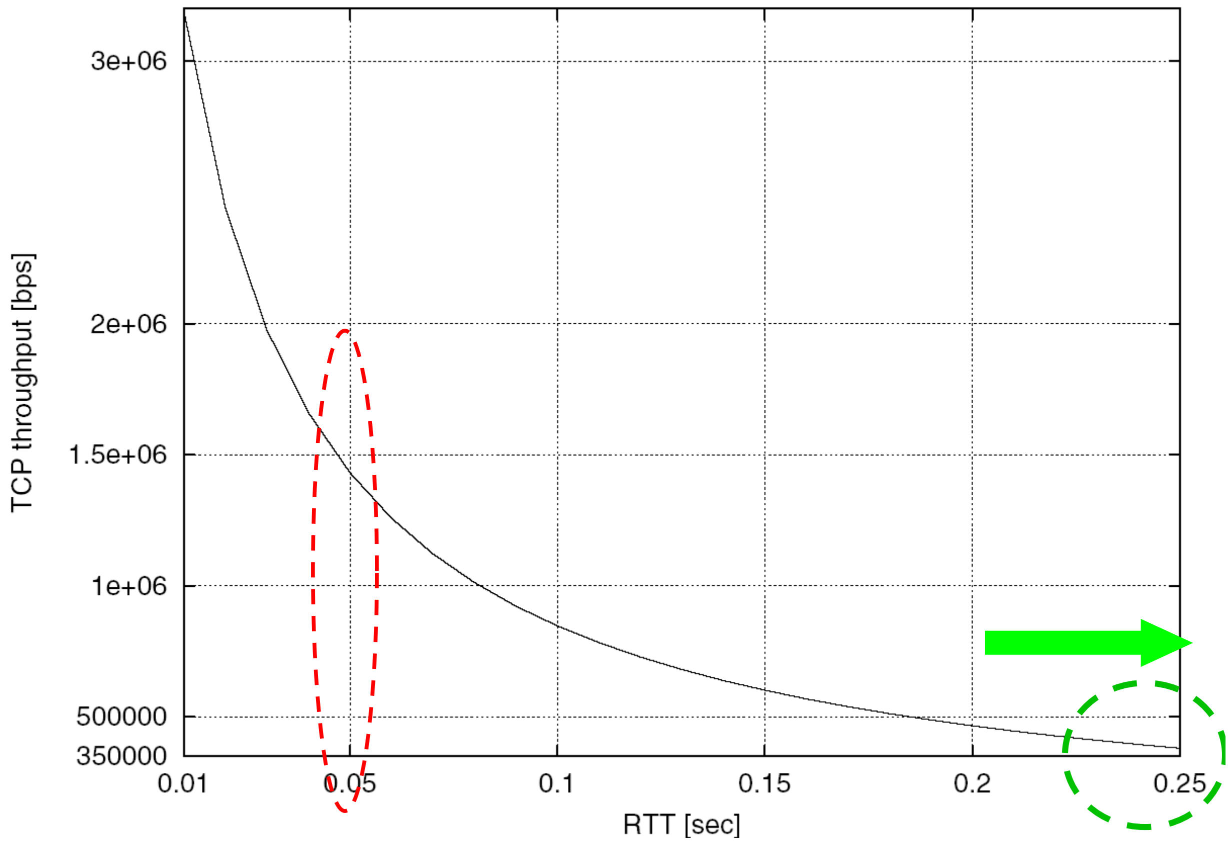
TCP Throughput

- T_o = 1 s; b = 2; MSS = 1500 bytes;
- p = 1%
- ***Tmax Lmin v.1***:

  Class 2/3: RTT > 300 ms
  (assembly driven)

- ***Tmax Lmin v.2***:

  Class 2/3: RTT ≈ 50 ms
  (propagation driven)



CONCLUSIONS

- Investigation of the effects of BA on end-to-end performance
  - JET resource reservation mechanism
- QoS differentiation through
  - Different extra-offset settings
  - Different employment of a limited set of wavelength converters
  - Different routing (deflection)
- "Realistic" traffic patterns
- Study of a Pan European Network
  - burst blocking probabilities
  - edge-to-edge delays

- Important/critical parameter for network performance/design:

$N_{virt\text{-}queues}$ X n. of overlapping bursts ~ n. of λs to avoid output contention

- Extreme attention must be paid in the burst assembly phase in order to be efficient, not to penalize loss sensitive data and to avoid to put time sensitive applications out of order

- A simple approach for service differentiation (extra-offset+converters management+class based routing) seems to be effective to provide insights for traffic and network engineering

Future works/collaborations

- ✓ Determine the end-to-end performance provided by TCP over OBS networks: TCP enhancements
- ✓ All nodes are both edge and core nodes
- ✓ Investigate OBS performance with protection/restoration mechanisms (in general in presence of failures)
- ✓ MOBSim Multi-thread: benefits? (today reliable simulations last one week on most recent Intel platforms)

THANK YOU FOR YOUR ATTENTION

casoni.maurizio@unimore.it

http://www.dii.unimo.it/casoni

… suggestions are very very welcome