Sharing the bandwidth of a wavelength: Burst switching with and without reservation

Gregor v. Bochmann, University of Ottawa

This presentation is based on the paper
Retransmission schemes for optical burst switching over star networks** by

Anna Agustí-Torra*, Gregor v. Bochmann†, Cristina Cervelló-Pastor*

*Department of Telematics Engineering
Technical University of Catalonia (Barcelona, Spain)

†School of Information and Technology
University of Ottawa (Ottawa, Canada)

** published in Proc. 2nd IFIP Intern. Conf. on Wireless and Optical Communications Networks (WOCN), March 2005, Dubai, United Arab Emirates.
http://beethoven.site.uottawa.ca/dsrg/PublicDocuments/Publications/Agus05a.pdf

Abstract

- We propose optical burst switching with various retransmission schemes to recover burst losses due to contention. We consider a star topology network. Our simulation studies compare these different retransmission schemes and also provide a comparison with burst switching with advanced reservation which avoids contention losses completely. We will present comparative values for different performance parameters, such as average effective transmission delay and bandwidth utilization, for varying network parameters, such as number of edge nodes, traffic load, and propagation delays between the edge and core nodes.

Outline

- Introduction
- Retransmission schemes
  - Basic retransmission scheme
  - 1st proposal
    - Objective: Limit the number of retransmissions per burst
  - 2nd proposal
    - Objective: Maximize the network utilization
- Performance analysis
  - Performance of retransmission schemes
  - Comparison with a simple TDM approach
- Conclusions

Introduction

- Optical Burst Switching (OBS)
  - Data units are grouped generating bursts at the edge nodes
  - Control plane and data plane are separated
  - A control packet is sent and processed before the burst transmission
  - Header processing and switch control are performed at the electronic domain

- AAPN (Agile All-Photonic Networks) (an NSERC research network)
  - Network architecture of overlaid hierarchical stars
  - Time sharing of wavelengths: transmission by time slots (fixed size and synchronized at the core) (slot = burst)
  - Otherwise like burst switching:
    - No optical buffer capacity at intermediate nodes
    - No wavelength conversion

Burst switching with and without reservation

- Classical burst switching
  - No reservation
    - Contention at output links of a switch
    - Burst loss rate depending on traffic load
- TDM (sharing bandwidth in time domain)
  - Reservation for data flow before transmission
    - Traditional TDM: repeating frames, fixed bandwidth per flow
    - Variable bandwidth allocation per frame: frequent control messages for changing the allocation
  - Result:
    - No losses
    - End-to-end round-trip delay for signaling before data can be sent

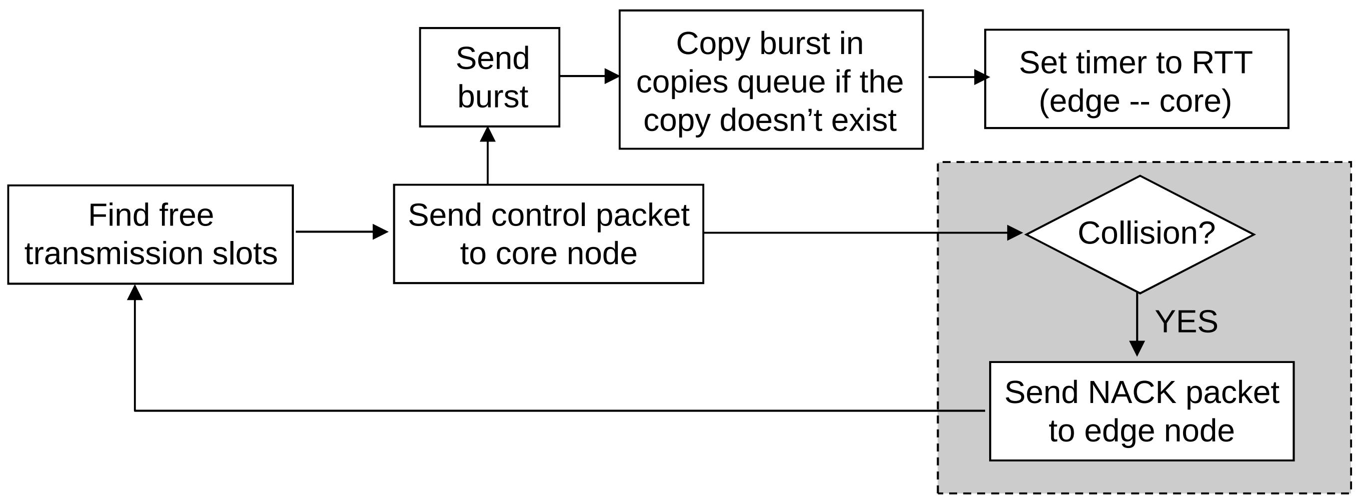
Comparison: reservation or not ?

- AAPN objective: Provide generic transport service
  - Varying bandwidth demands
  - Low delay
  - Negligeable losses
  - Single QoS, good for all applications
- Possible solutions
  - Burst switching **with retransmission** (to eliminate losses)
  - Flexible TDM with reservation
    - Implies delay for waiting for allocated time slot within frame
    - Involves end-to-end delay for changing reservations
- Comparison of solutions
  - For low traffic load, burst switching is better (no delays)
  - How do they compare for higher traffic load ?

Basic retransmission scheme

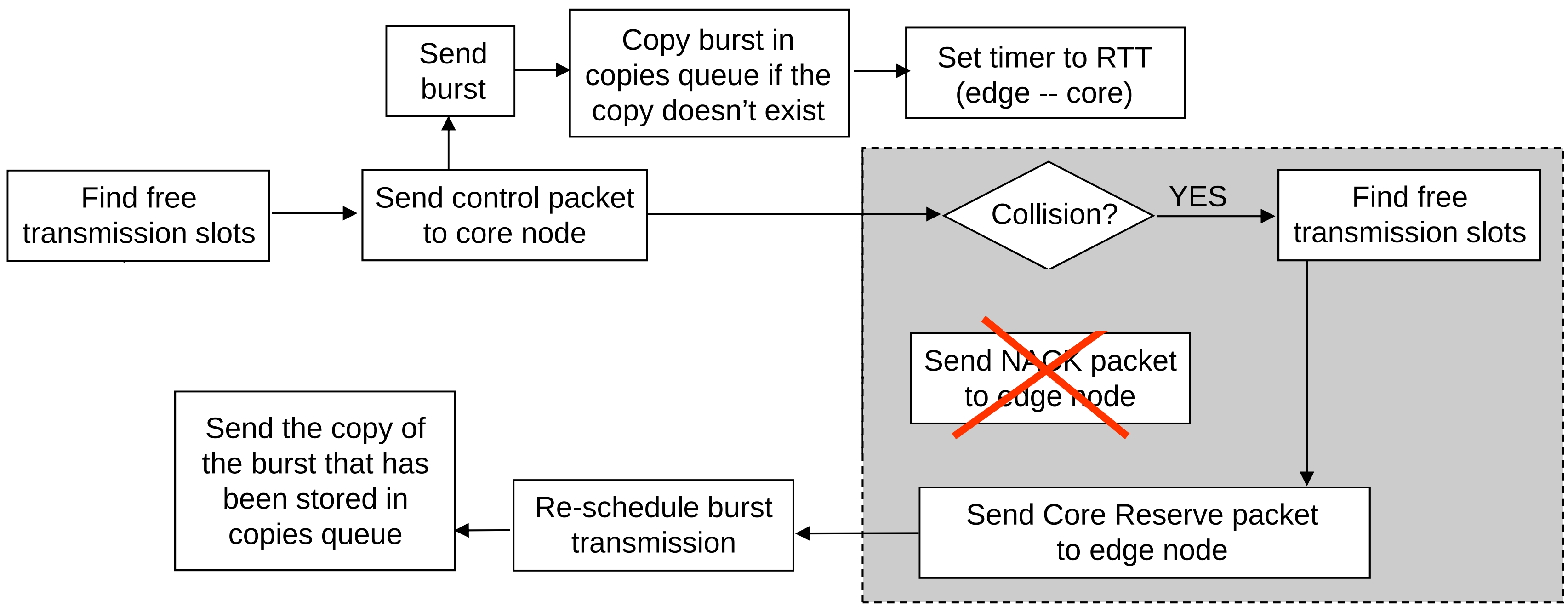
- Each edge node keeps a copy of every burst sent until its successful transmission
  - Copies of burst sent are stored in an additional queue at the edge node
  - A timer controls de lifetime of each copy. It is set to be larger than the RTT between edge and core node
- When the core node detects a collision, it will notify the concerned edge node by sending a NACK packet
  - The NACK packet contains the identifier of the discarded burst
- When the edge node receives a NACK packet, it will retransmit the discarded burst
  - The edge node still keeps the copy of the burst at the queue of copies
  - The timer is set again to its initial value

Basic retransmission scheme



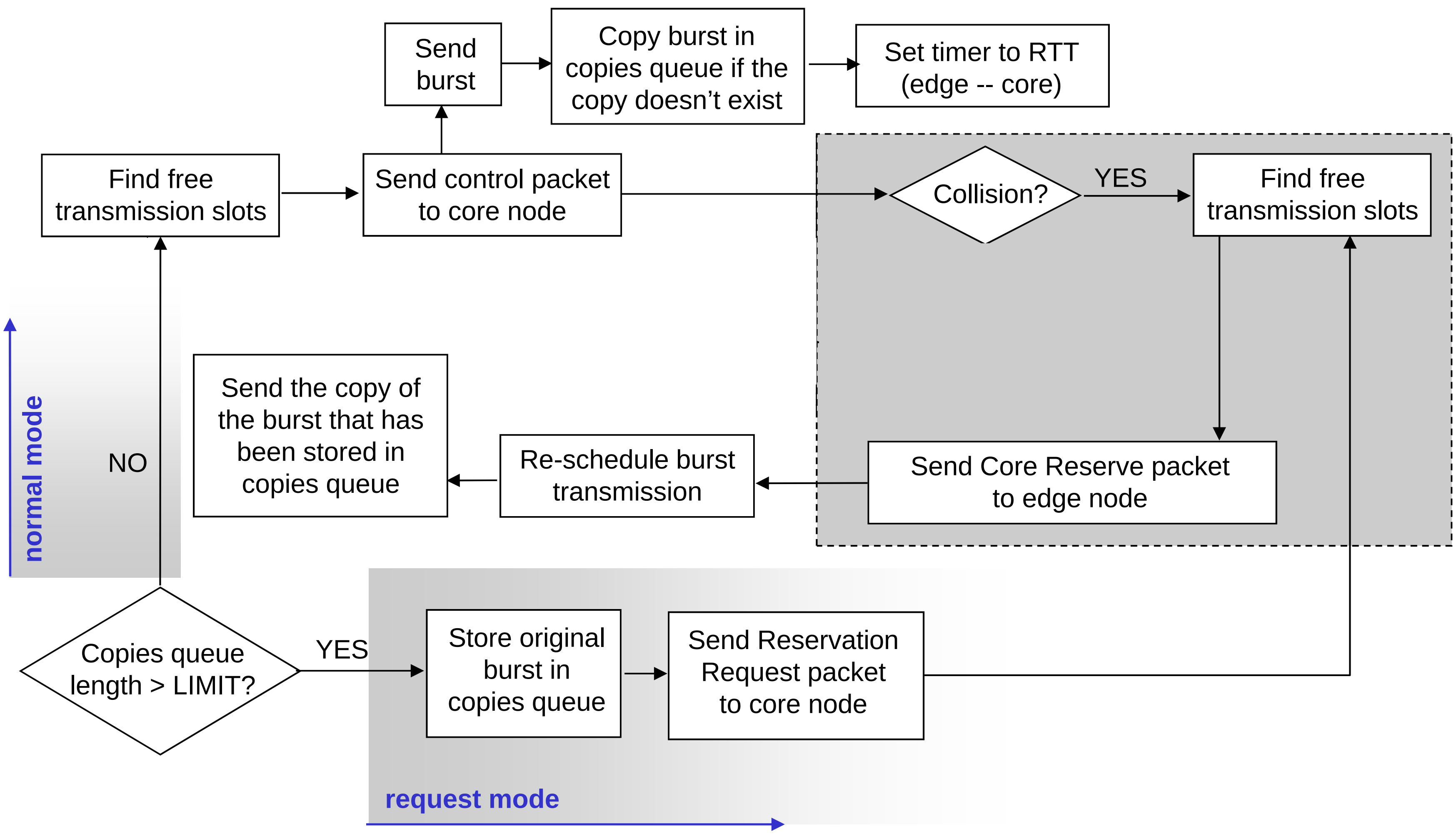
- Advantages
  - No losses. Retransmission of dropped bursts without intervention of higher layer protocols
  - Simplicity
- Drawbacks
  - Number of retransmission per burst unknown. High collision probability $\Rightarrow$ effective throughput decreases
  - Larger storage capacity at the edge nodes as compared with pure OBS

Retransmission scheme (version 2)



- Operation: get reserved time slot when burst was lost once
- Advantages
  - Maximum number of retransmission per burst = 1
- Drawbacks
  - High offered load $\Rightarrow$ throughput less than 100%

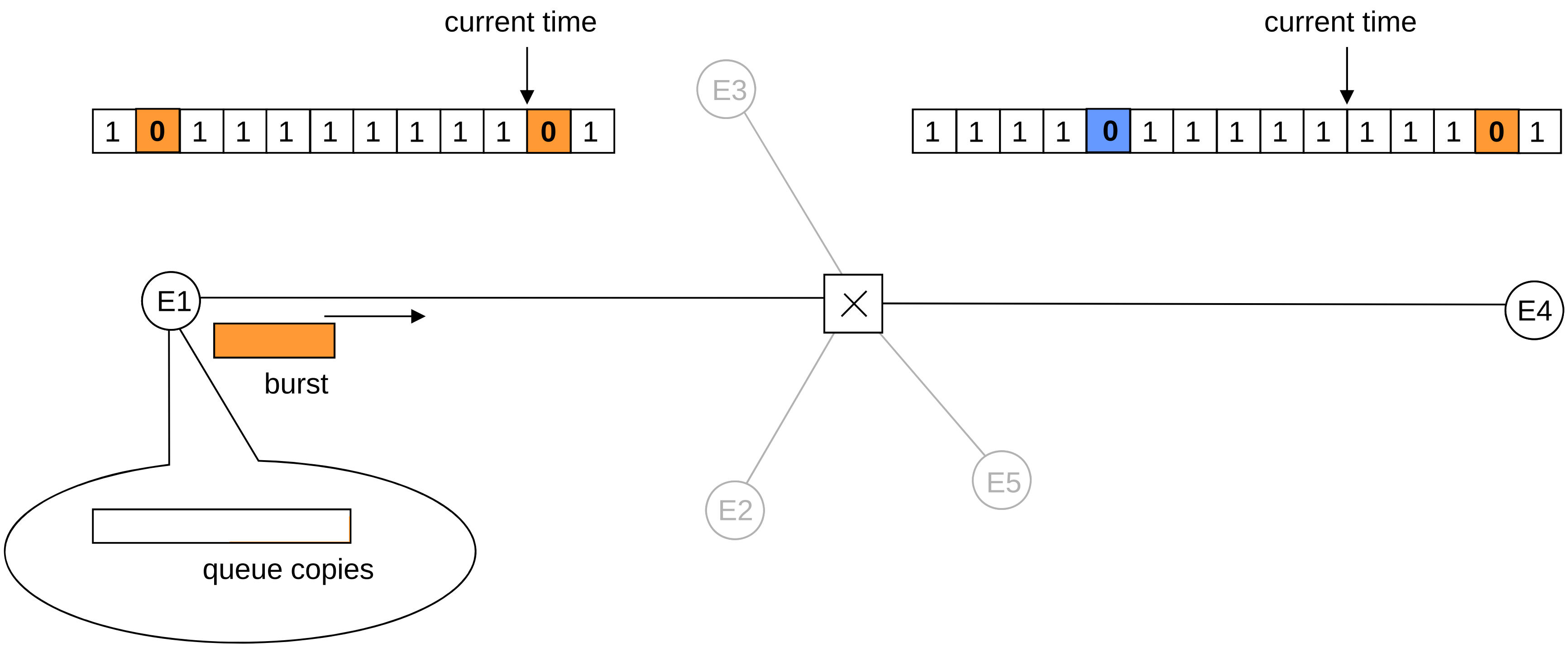
Retransmission scheme (version 3)



- Operation: if average load is below threshold, use version 2; otherwise get reserved time slot before transmitting the burst the first time (full reservation mode)

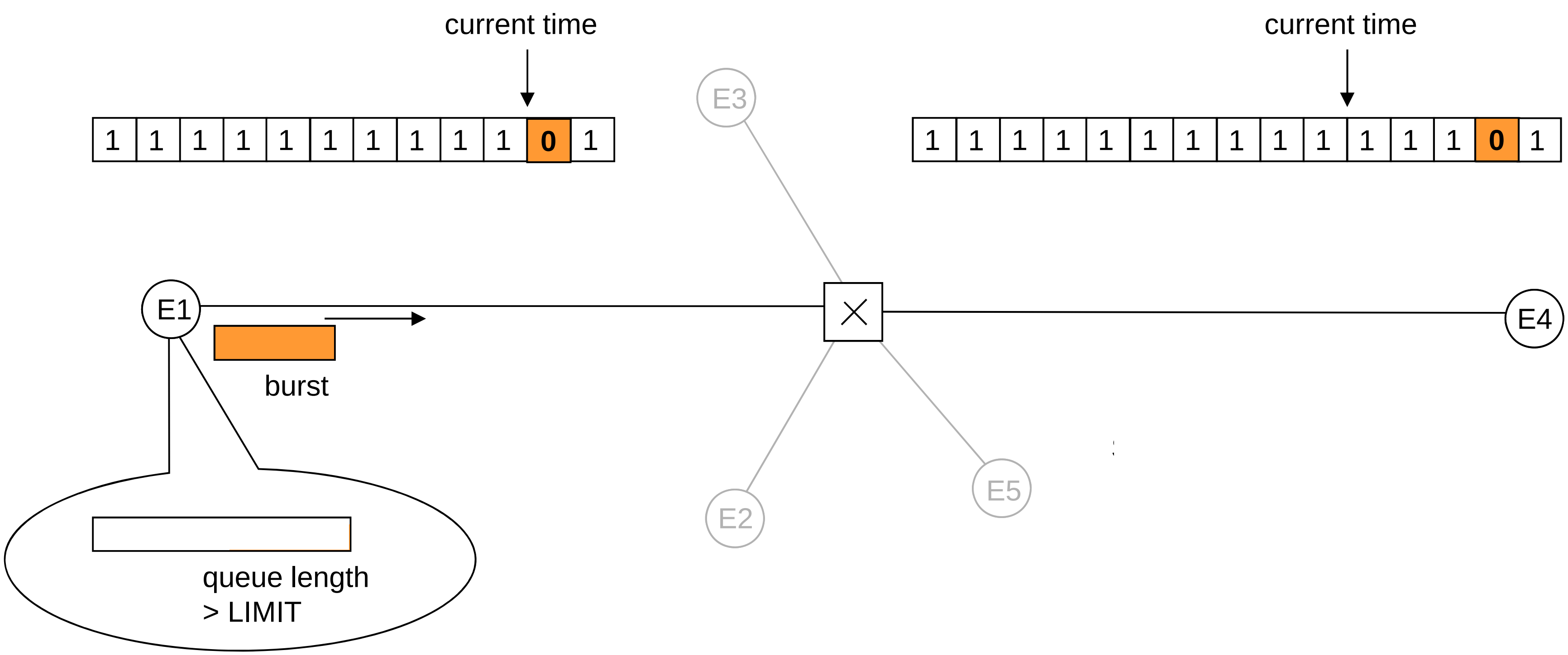
Retransmission scheme (version 3)

Normal mode



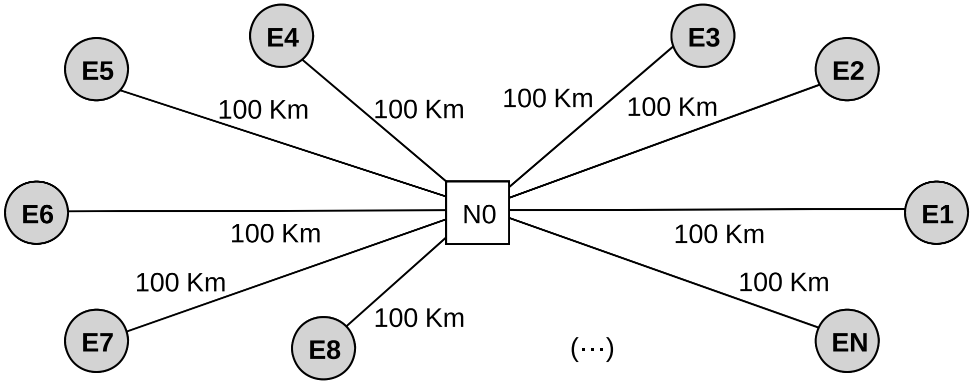
Retransmission scheme (version 3)

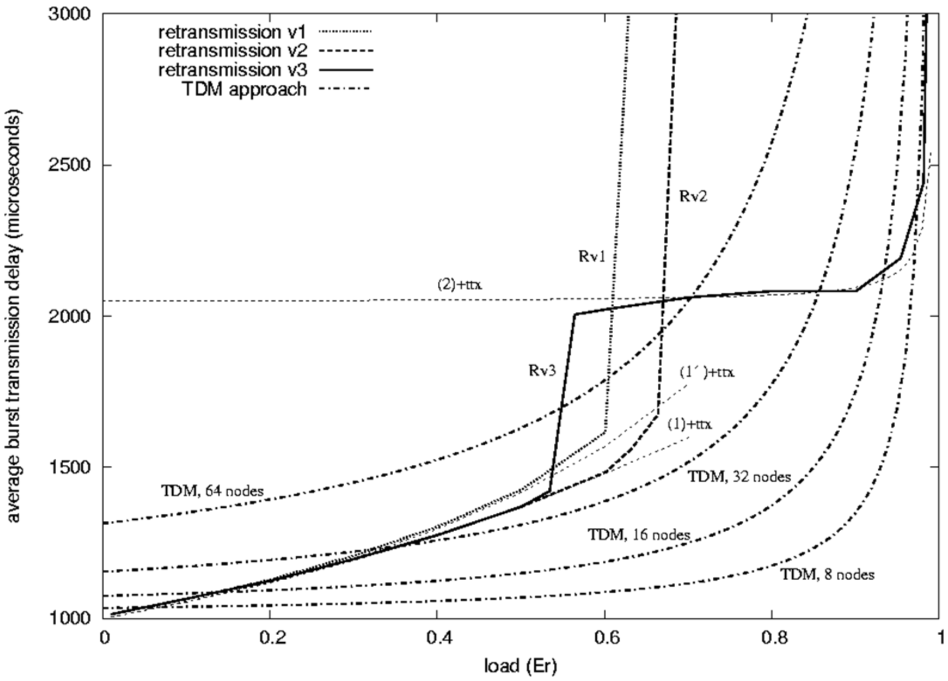
Request mode



Performance analysis

wavelengths = 1
speed = 10 Gbps
1 TAU = 1 τ = 10 μs
ON-OFF sources (expo.)
burst length = 1τ
all edge nodes are sources and destinations
(destination chose with same probability)





additional delay ≅ delay introduced by queuing and retransmission in addition to propagation delay and the burst transmission time.

$d = \Delta_1 + z * RTT$ (1´) Rv1

$d_1 = \Delta_1 + z * (RTT + \Delta_2)$ (1) Rv2 & Rv3 normal mode

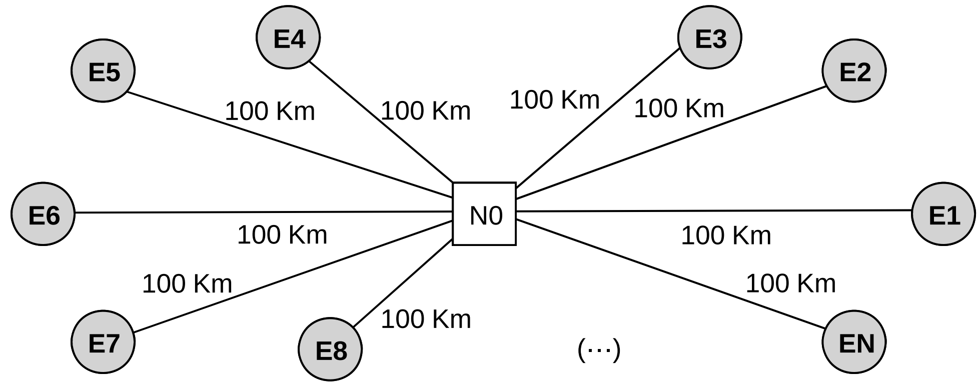
$d_2 = \Delta_1 + RTT + \Delta_2$ (2) Rv3 request mode

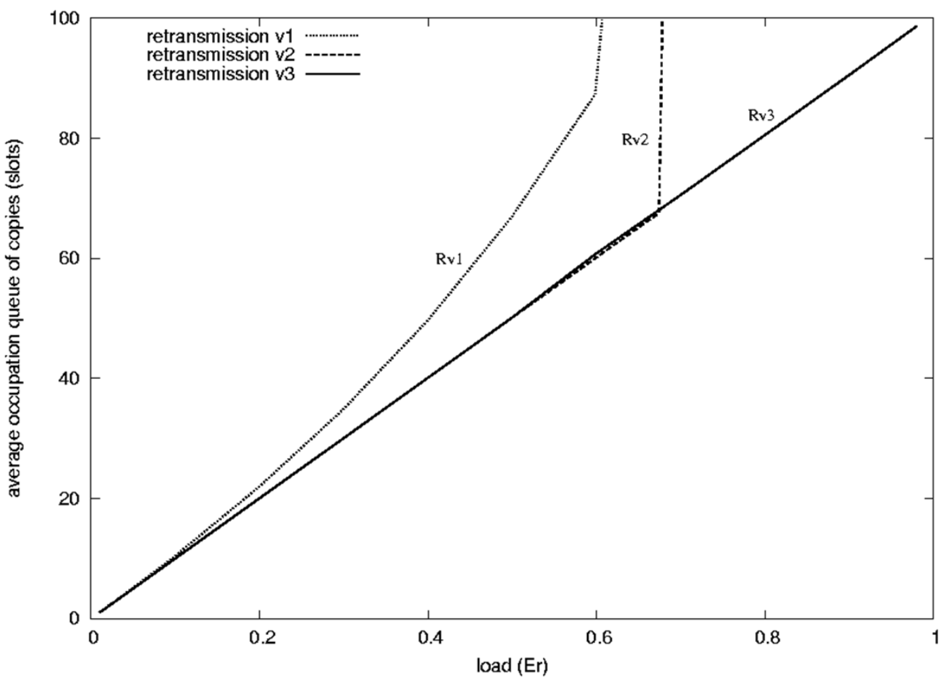
$d_{TDM} = 0.5 * (N-1) * slot\_length / (1-\rho)$

z = number of retransmissions per burst
RTT = round trip time edge – core
ρ = offered load
$\Delta_1 = (\rho' * slot\_length) / (2 * (1-\rho'))$
Δ_2 = (maximum offset +1) * slot_length
N = number of edge nodes

Performance analysis

wavelengths = 1
speed = 10 Gbps
1 TAU = 1 τ = 10 μs
ON-OFF sources (expo.)
burst length = 1τ
all edge nodes are sources and destinations
(destination chose with same probability)





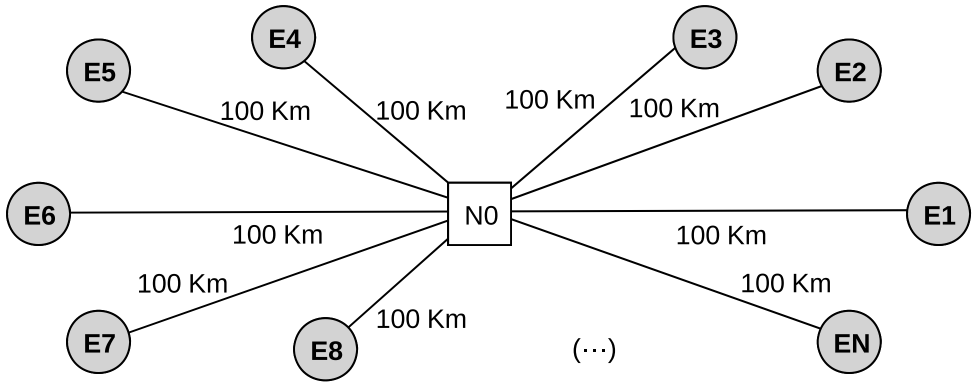
Queue length (total storage capacity needed to store the copies of bursts in one edge to all destinations):

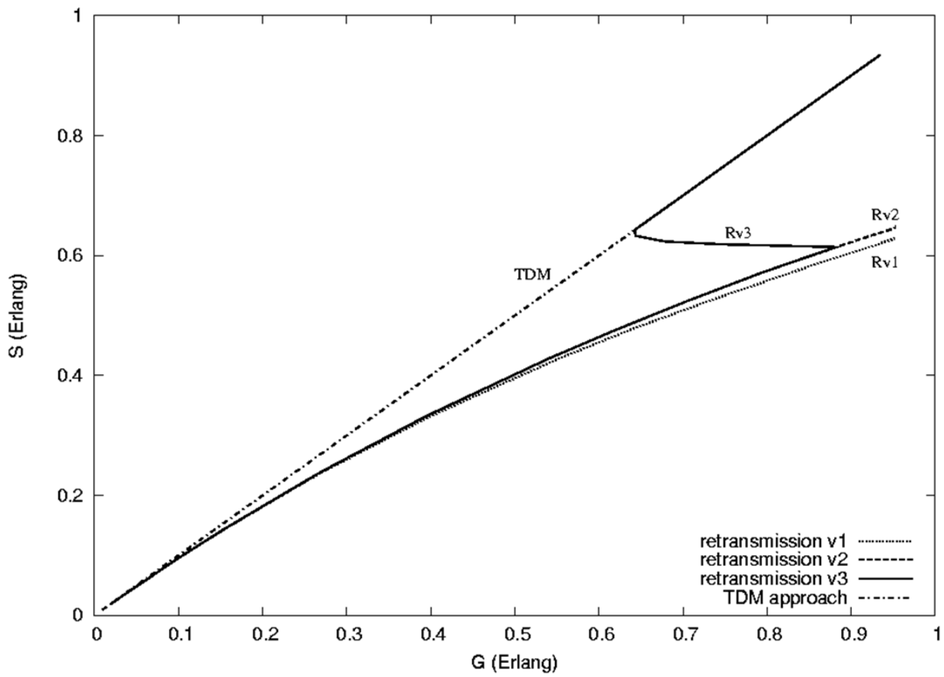
$q = \rho' / (\text{slot\_length}) * RTT$ Rv2 & Rv3

RTT = round trip time edge – core
ρ' = offered load per edge node

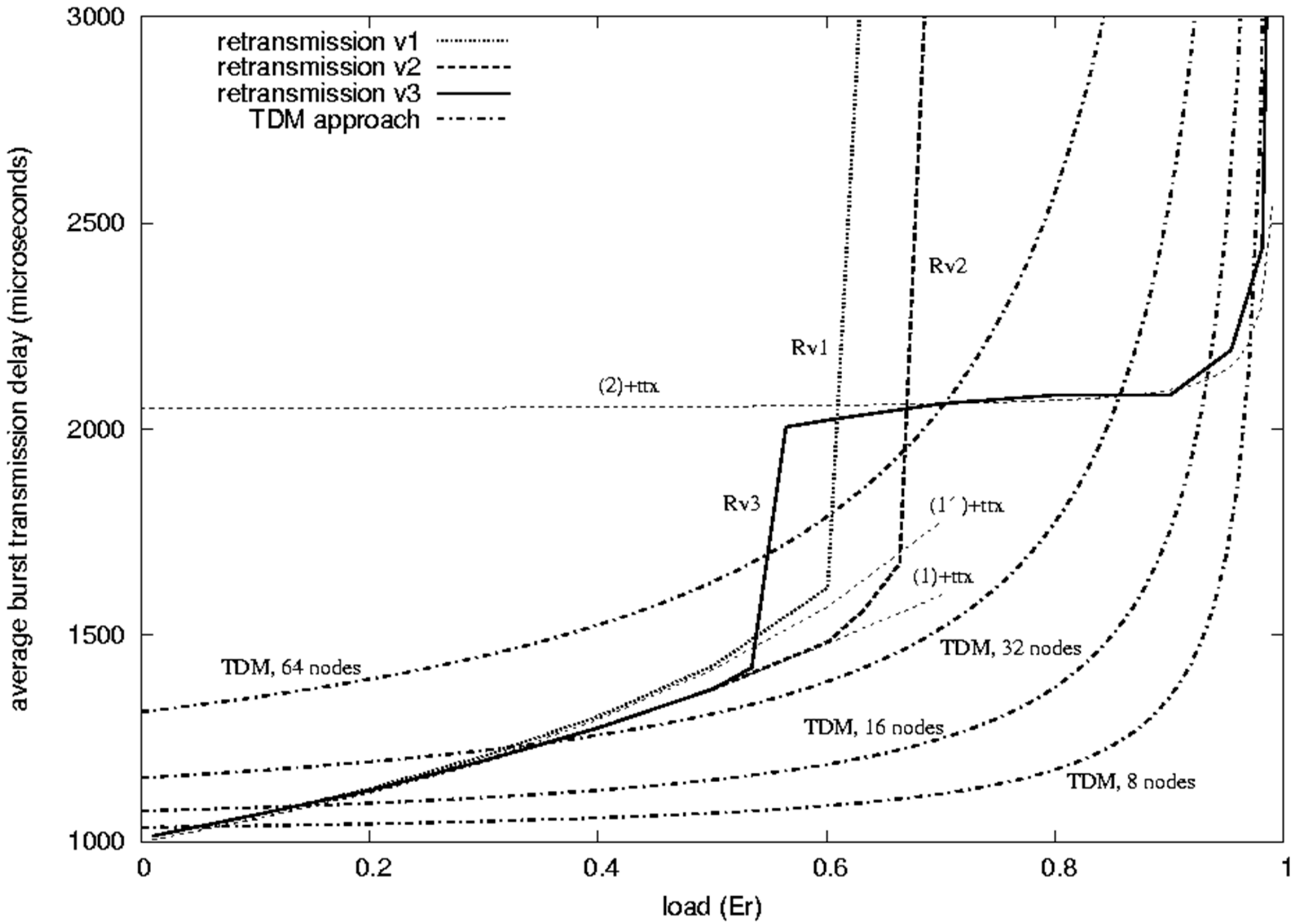
Performance analysis

wavelengths = 1
speed = 10 Gbps
1 TAU = 1 τ = 10 μs
ON-OFF sources (expo.)
burst length = 1τ
all edge nodes are sources and destinations





Overview of performance predictions



Concluding Remarks

- Proposed modifications of a basic retransmission scheme
  - Version 2: limits the maximum number of retransmission per burst to 1
  - Version 3: allows the effective use of the full transmission capacity
  - Disadvantage: higher complexity of the network control
- Performance analysis
  - Burst switching
    - Average burst transmission delay is independent of the number of nodes but highly dependent on the round trip delay between the edge node to the core
    - Size of retransmission buffer at source edge nodes also dependents on the same round trip delay
  - TDM approach
    - Average burst transmission delay depends on number N of nodes that share the output link (N x slot time)
    - No buffering for retransmission, only buffering to equalize the flow of incoming packets
- Open questions
  - Which slot assignment algorithm should be used in reservation mode?
  - What is the best threshold to switch between normal and reservation mode?
- Work in progress
  - Extend the comparison to other TDM approaches
  - Evaluate the behavior under:
    - different types of traffic (uniform, non-uniform, self-similar)
    - different link length
  - Analytical model of the retransmission probability