Optical Metropolitan Multiservice Network : Access Protocols and Quality of Service

Tülin Atmaca and Viet Hung Nguyen

Networks and Services of Telecommunications Department

Institut National des Télécommunications

GET/INT, EVRY, FRANCE

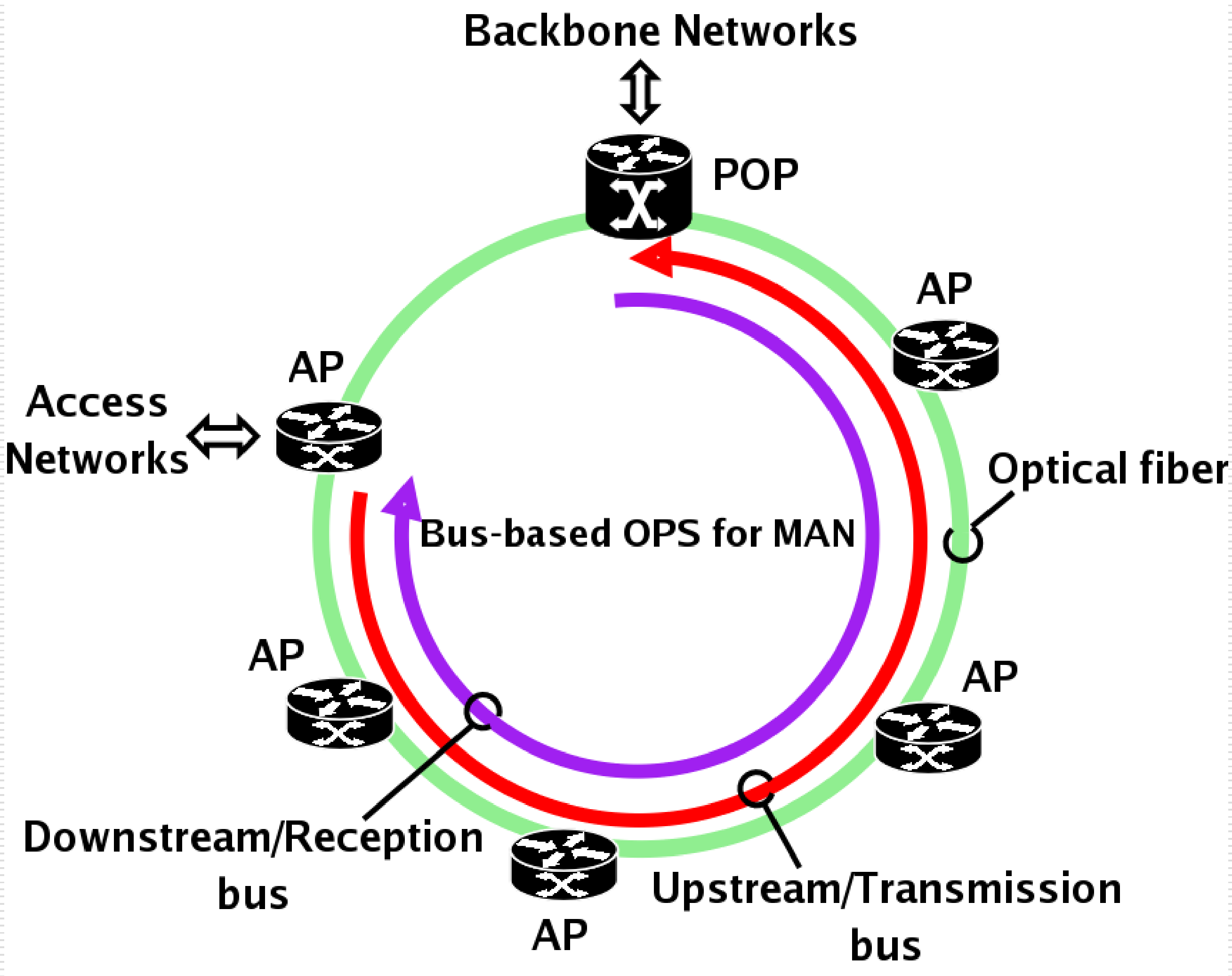
Outline

- Motivations
- Studied OPS Bus-based Network, OU-CSMA/CA Protocol and Related Problems
- Analytical Model for OU-CSMA/CA
- Advanced MAC Mechanisms:
  - Transmission Efficiency: Modified Packet Bursting
  - Fairness and Bandwidth segmentation : Di-MAC
- Transport Feasibility of TDM traffic over packet- switched network: Circuit Emulation Service
- Conclusions and Perspectives

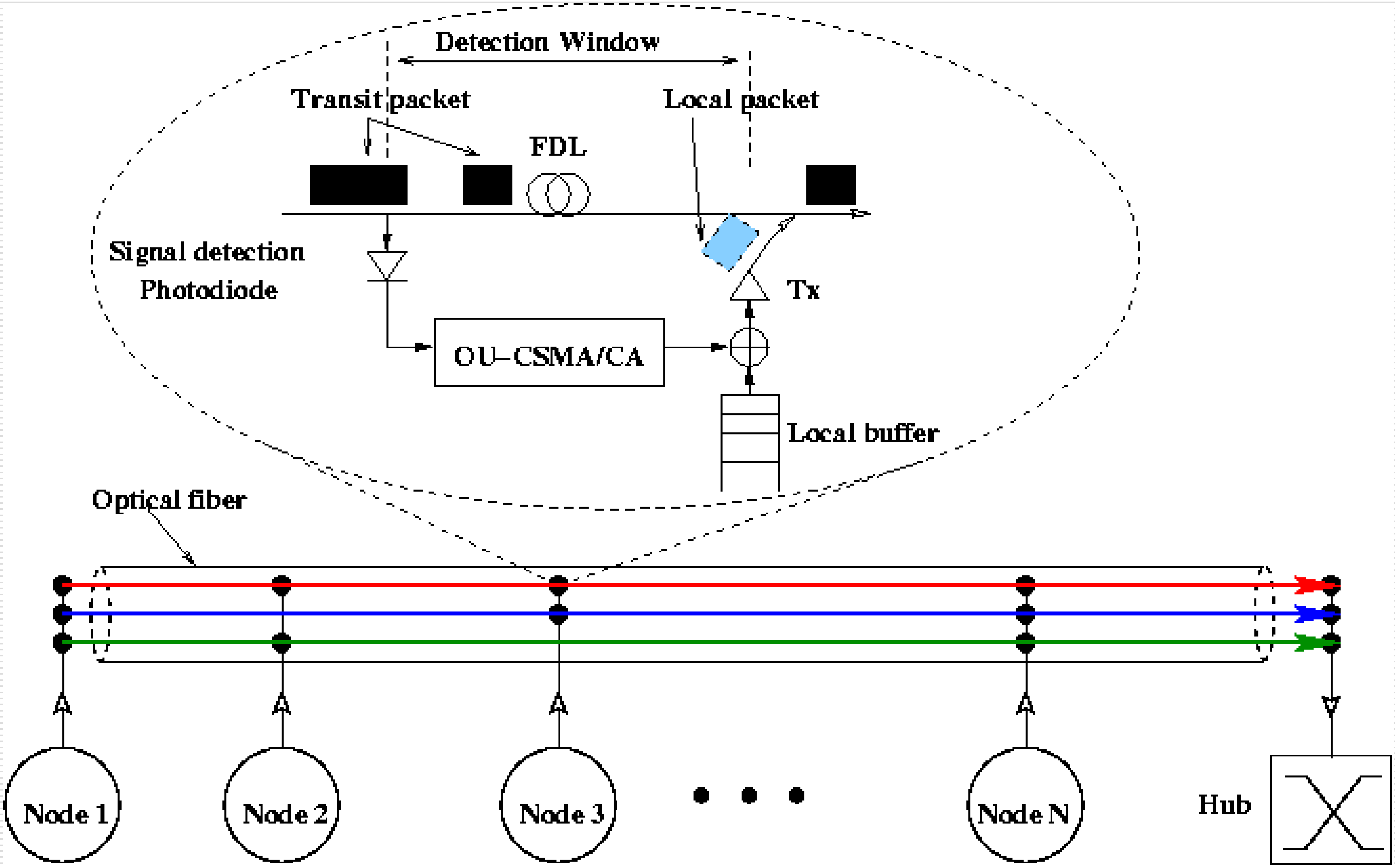
Motivations

- **Traffic Evolution** : data + video traffic volume >> voice.
- **More** bandwidh demand and QoS requirements → current MANs are not well designed for transport data + video traffic
- **Consequence**: MANs should be changed to maintain traditional circuit service while at the same time enabling new packet-based services
- **Study objective**:
  - Logical performance of a bus-based OPS MAN designed to support this new trend of QoS requirements

Studied Optical Metropolitan Network

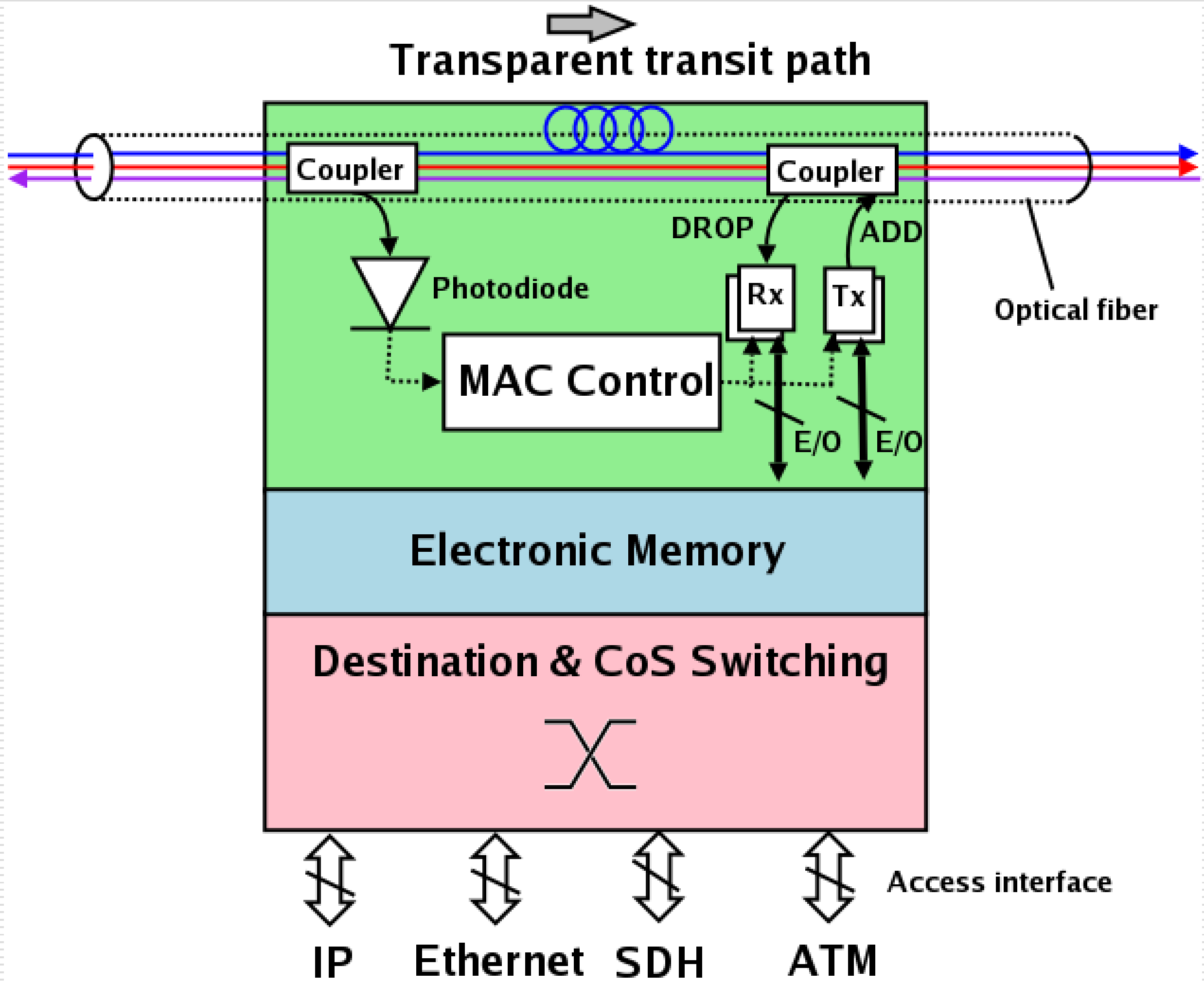


Focus: The Upstream Bus



Studied Optical Metro Network: Focus

Access Node Architecture



Studied Optical Metro Network: Node architecture

Optical Unslotted CSMA/CA Protocol

- OU-CSMA/CA protocol operation is based on the detection of idle periods (voids) on the transmission wavelength shared by several nodes
- Compared to CSMA/CD in Ethernet, the OU-CSMA/CA protocol provides more efficient resource utilization, thanks to avoidance of collision
- OU-CSMA/CA supports variable length optical packets thanks to its asynchronous nature

OU-CSMA/CA: Principle (1)

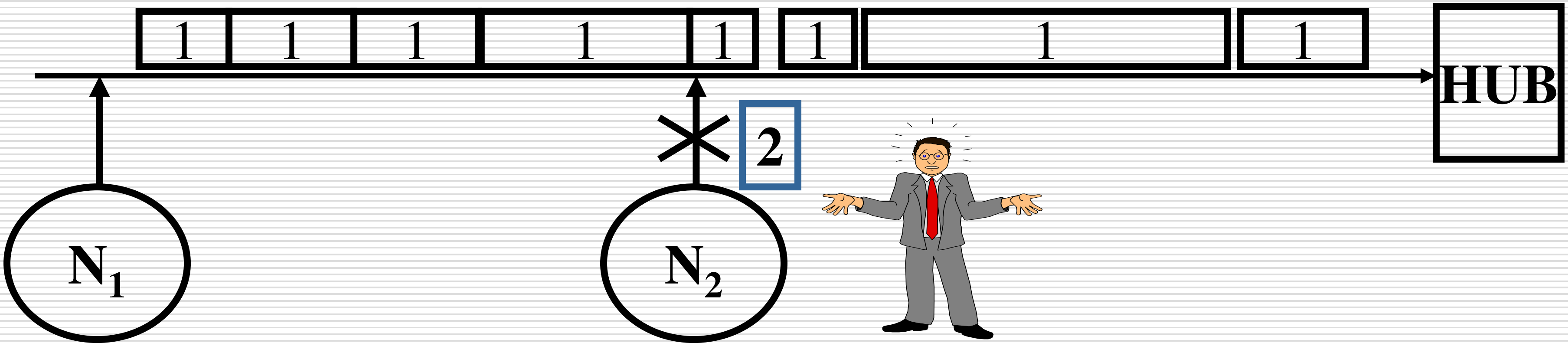
- To detect activity on a wavelength, a node uses low bit rate photodiodes (155MHz)
- Photodiode sends control information to the MAC logic controlling the operation of transmitters
- Avoidence of collision is performed by employing Fiber Delay Line (FDL) which creates on the transmission line a fixed delay between the control and Add/Drop functions

OU-CSMA/CA: Principle (2)

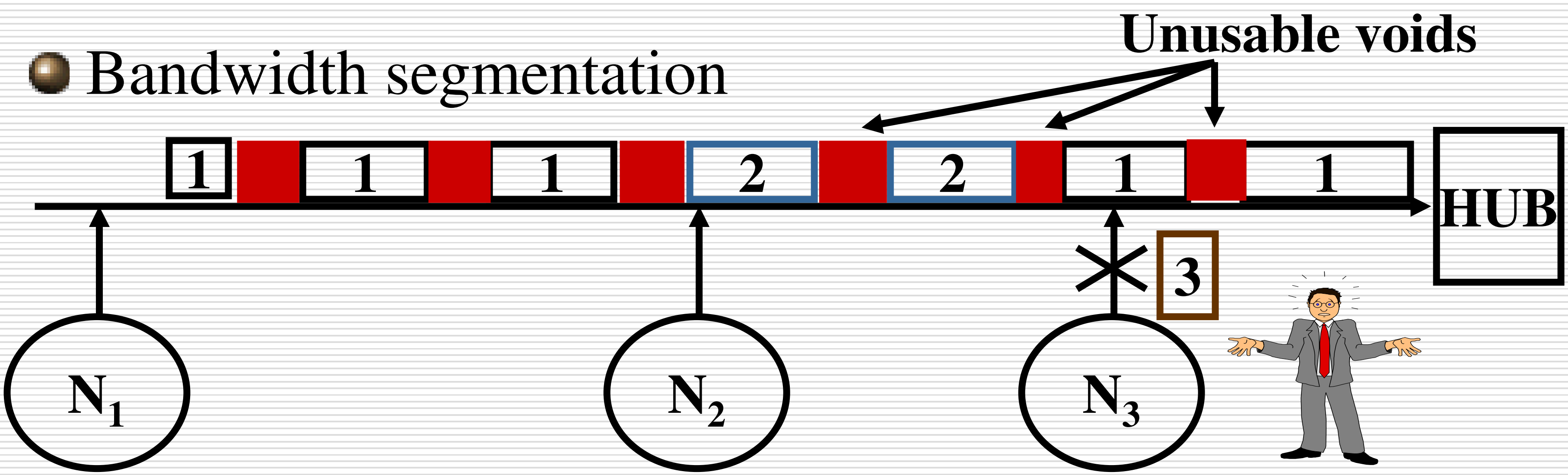
- FDL should be long enough to provide the MAC logic sufficient time to measure the medium occupency
- FDL storage capacity should be at least larger than the Maximum Transmission Unit (MTU) of the transport protocol used
- If the network is used to transport Ethernet Packets , then the FDL storage capacity is ≈1500 bytes

Problematics

- Unfairness (Positionnal priority)



- Bandwidth segmentation



Main Contributions

- Mathematical model for OU-CSMA/CA
  - Modelling (M/G/1 – PRI), numerical resolution, performance evaluation
- Modified Packet Bursting (MPB):
  - Improvement of transmission efficiency
- Dynamic Intelligent MAC (DI-MAC):
  - Solution to fairness and bandwidth segmentation problems
  - Proposal of a dynamic and distributed MAC
- Feasibility study: transport of TDM traffic (CES) over packet switched MAN

Mathematical Model for the OU-CSMA/CA Protocol

Difficulties and Existing studies

- **Difficulties:**
  - Interdependence among ring nodes renders difficult an exact performance analysis of OU-CSMA/CA
  - Variability of service times (packet lengths) and asynchronous transmission operations render complex the analysis
- **Existing studies** : approximate methods
  - M/G/1 Preemptive-Repeat-Identical (PRI) queue : measure mean response time (*Ex: [CHE 03, CCH 05, HGJ+ 05 and BBD+ 05]*)
  - Markov Chains (*Ex: [TOP 03]*)

Objectives and Proposed Method

- **Objectives**:
  - **Measure** the mean response time
  - **But also** measure the queue length distribution
- **Method**:
  - M/G/1 - PRI Queue
  - We analyze bus nodes one by one, we use a specific state description to reflect the interference from upstream nodes
  - Numerical solution

Modeling Principle (1)

- From the modeling perspective, the operation of OU-CSMA protocol is viewed as follows:
  - At the detection of a void, a node begins to transmit a local packet, but this transmission is interrupted whenever a transit packet arrives from an upstream node.
  - When transmission is interrupted, the packet returns to the head of the local buffer
  - Transmission of this packet is reattempted at the next void
  - This process is repeated until a large enough void is found and the packet is successfully transmitted.

Modeling Principle (2)

- Hence, a transmission of a packet may viewed as a number of interrupted transmission attempts followed by one successful transmission

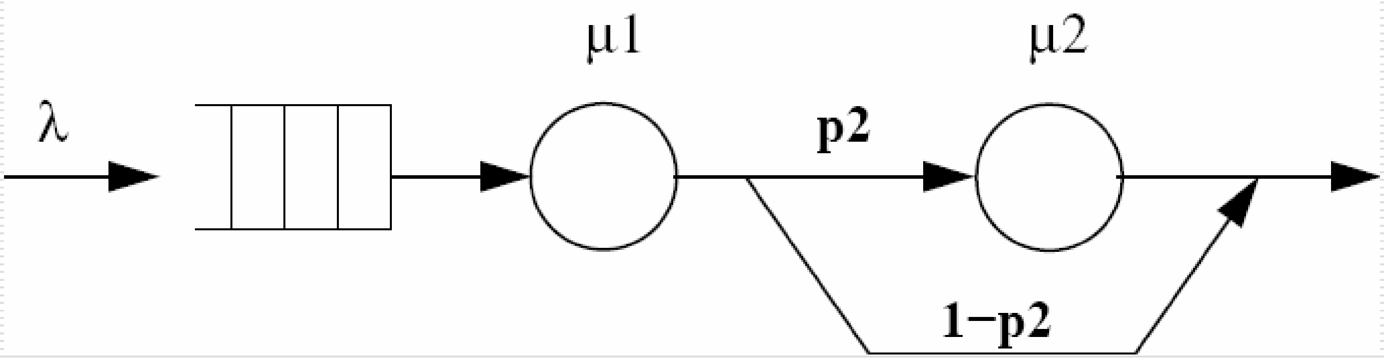
- This behavior can be modeled by a single server (sharing wavelength) with *N* priority queques (N ring nodes) with PRI (Priority Repeat Identical) service discipline

Model Description (1)

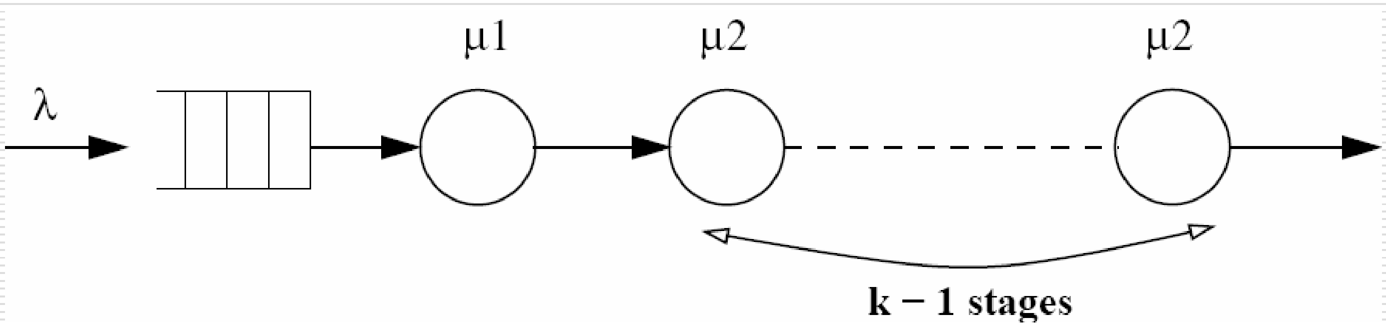
- **Hypothesis**:
  - N nodes, numerated from 1 (highest priority node) to N
  - Node i: infinite buffer, FIFO discipline, Poisson arrival process at node **i** with parameter λ_i, service time *i.i.d* with mean $\mathbf{m_i}$ and variance $\mathbf{Var_i}$

Model Description (2)

- **Modelling of original service time:** Model of Cox [COS 61]



E.g.: Cox system modelling distributions with Cv > 1

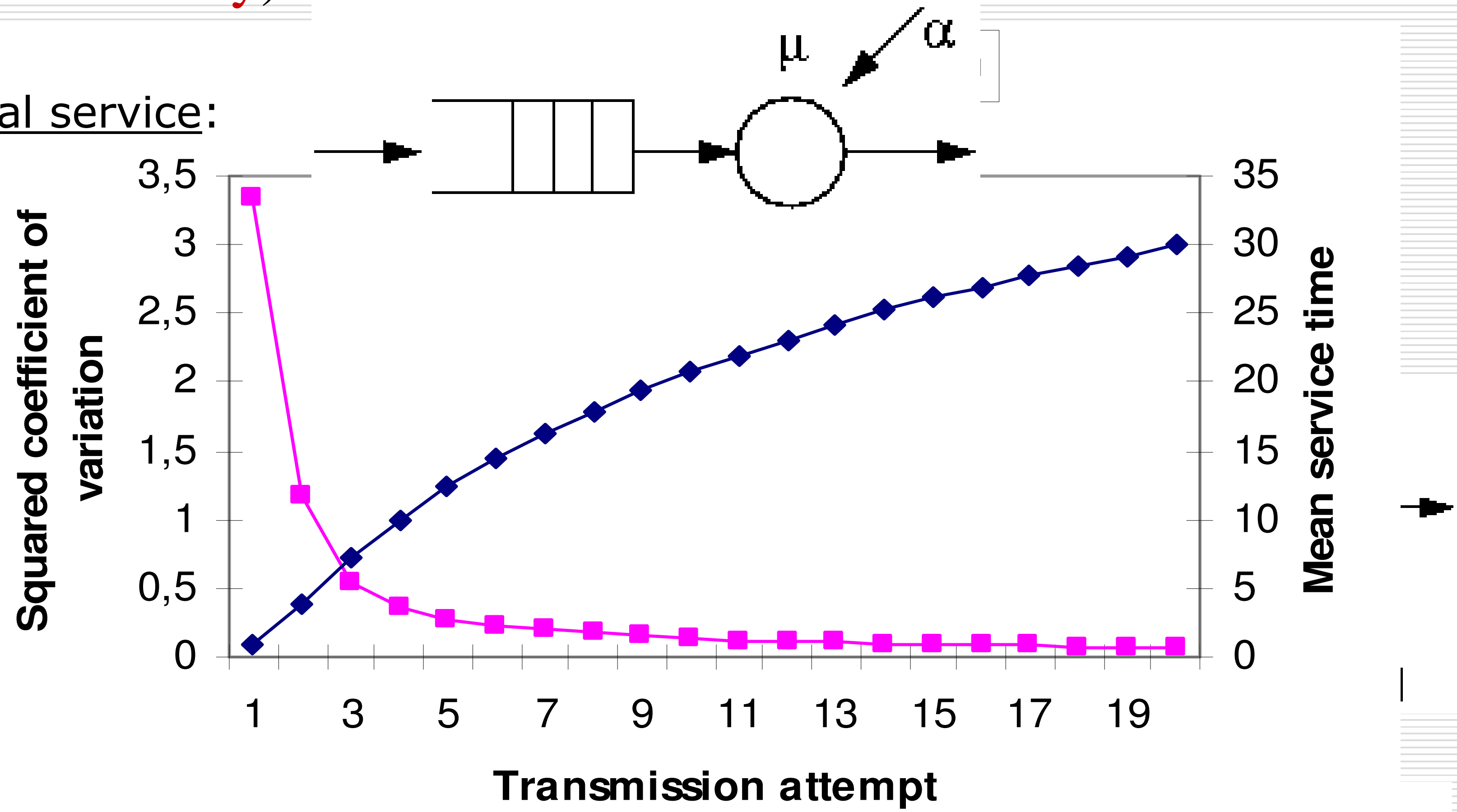


E.g.: Cox system modelling distributions with $1/\sqrt{k} \leq C_v < 1/\sqrt{(k-1)}$

Model Description (3)

Modelling of service time at each transmission attempt : (approximately)

Original service:



Se
att

q_1 ~ (

Model Description (4)

- **Solution for node 1:**
  - **Fixed Point Iteration**
  - Calculate: $\boldsymbol{u(n)}$ = service completion rate
  - $\boldsymbol{p(n)}$ = probability of having **n** packets, and β_2
- **Solution for node i >1:**
  - **Server disappears with rate α_i and reappears with rate β_i**
  - $\alpha_i = \sum_{j<i} \lambda_j$, and β_i computed approximately from the solution of node i-1
  - Calculate: $\boldsymbol{u(n)}$, $\boldsymbol{p(n)}$, β_{i+1}

Model Resolution (1)

- **Solution for node 1:**
  - State description: $\boldsymbol{p(n,l)}$, $\boldsymbol{n}$ = number of packets, $\boldsymbol{l}$ = stage number in Cox model
  - $\boldsymbol{u(n)}$ = service completion rate

$$u(n) = \sum_{l=1}^{k} p(l \mid n)\mu_l q_l$$

  - $\boldsymbol{p(n,l) = p(l|n)\, p(n)}$ and $p(n) = \frac{1}{G}\prod_{i=1}^{n} \lambda / u(i)$
  - Balance equations solved by fixed point iteration, calculate: $\boldsymbol{u(n_1)}$, $\boldsymbol{p(n_1)}$ ***et*** $\boldsymbol{\beta_2}$

Model Resolution (2)

- **Recurrent solution for node i > 1:**
  - **Server disappears with rate α_i and reappears with rate β_i**
  - $\alpha_i = \sum_{j<i} \lambda_j$, and β_i are calculated approximately
  - State description: ***p(n,j,l) with j*** = number of interruptions
  - So , like node 1, calculate: $u(n_i), p(n_i), \beta_{i+1}$

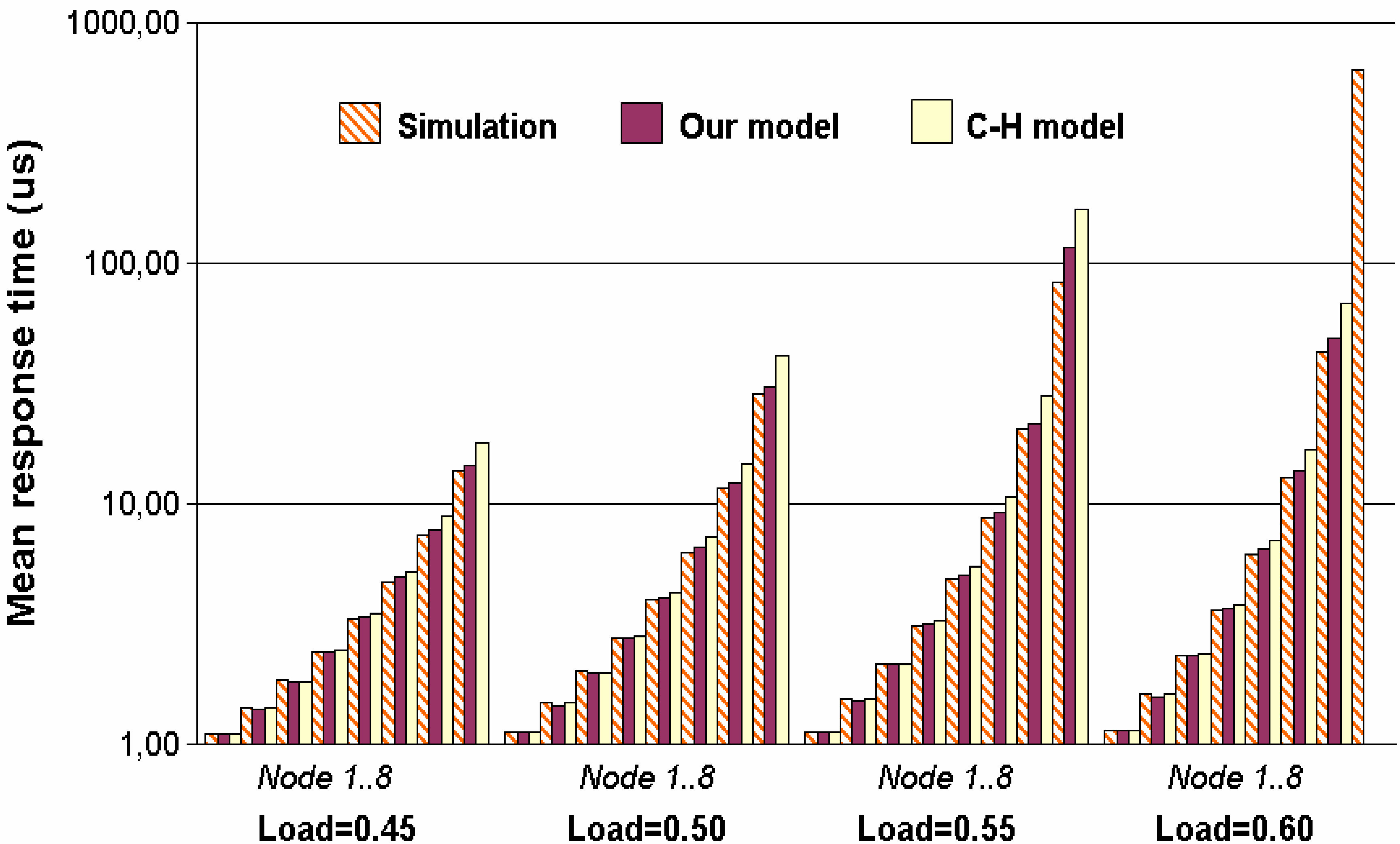
Precision of the Estimation of β and Impact of Higher Moments of Service Time

$\lambda_1 = \lambda_2$	Model	Mean packet number at node 2 (with distribution I)	Mean packet number at node 2 (with distribution II)
0.06733	Simul. with 2 Nodes	0.1073 ± 0.0002	0.1058 ± 0.0002
	Simul. with Node 2	0.1043 ± 0.0003	0.1023 ± 0.0003
0.13466	Simul. with 2 Nodes	0.41	: 0.0123
	Simul. with Node 2	0.4009 ± 0.0026	0.4038 ± 0.0038
0.20		2.5726 ± 0.3549	1.6053 ± 0.0750
		2.6417 ± 0.4280	1.4507 ± 0.1157

Impact of higher moments : almost 25% of difference

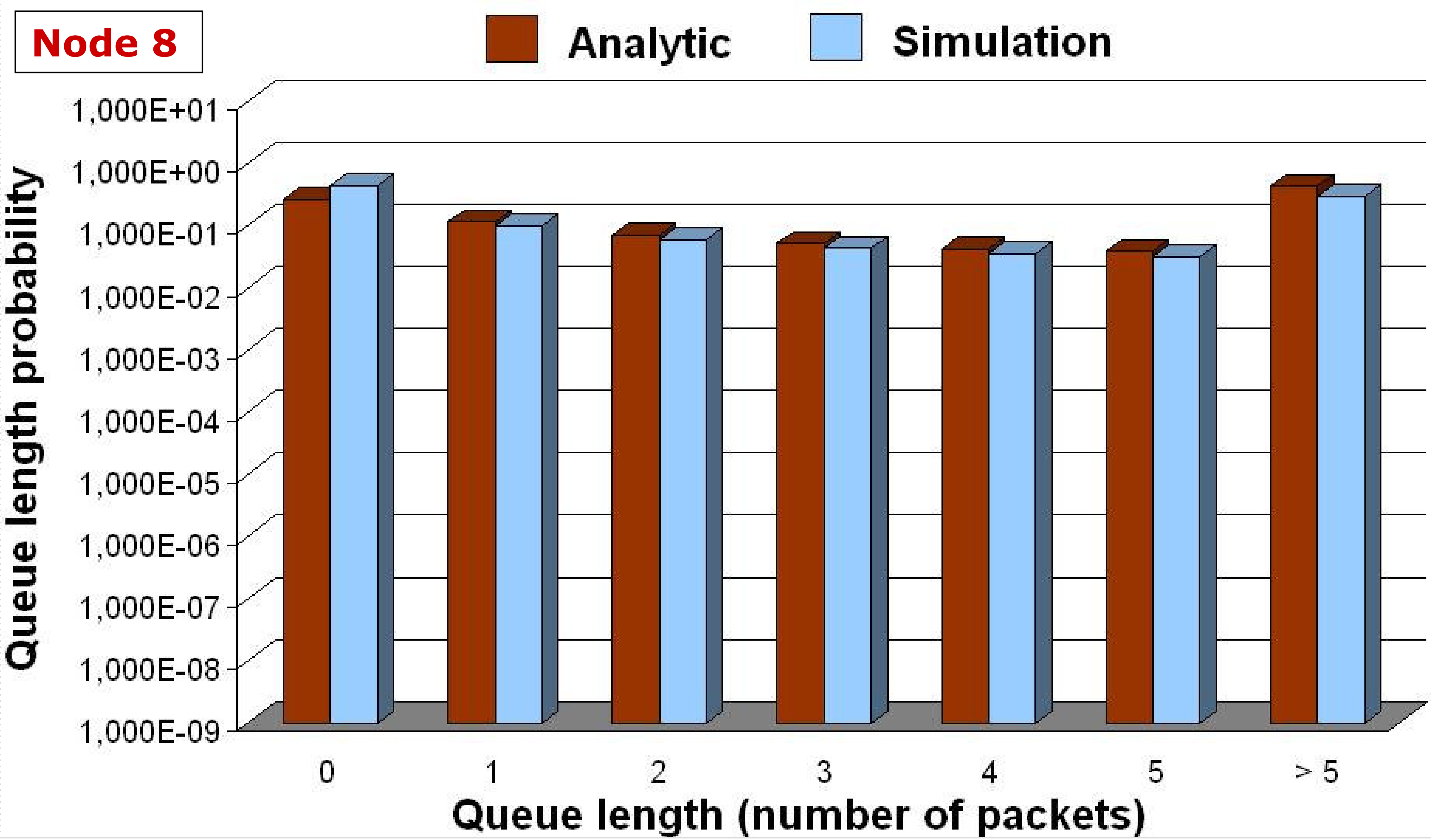
Precision of estimation of β: few % of difference

Mean Response Time vs. Offered Load



IP Packet size distributions [CAIDA]

Queue Length Distribution



IP Packet size distributions [CAIDA]

Conclusion on Analytical Model (1)

☺	😐
Better results vs. existing models: - Analytical results show unfairness among bus nodes due to positional priority -The mean response time increases rapidly as we move downstream on the bus	Approximate Model

Conclusion on Analytical Model (2)

- Our model can be generalized to dimension the buffer
- Network performance depends on the packet size distribution
- Worst case: uniform or uniformly decreasing traffic distribution

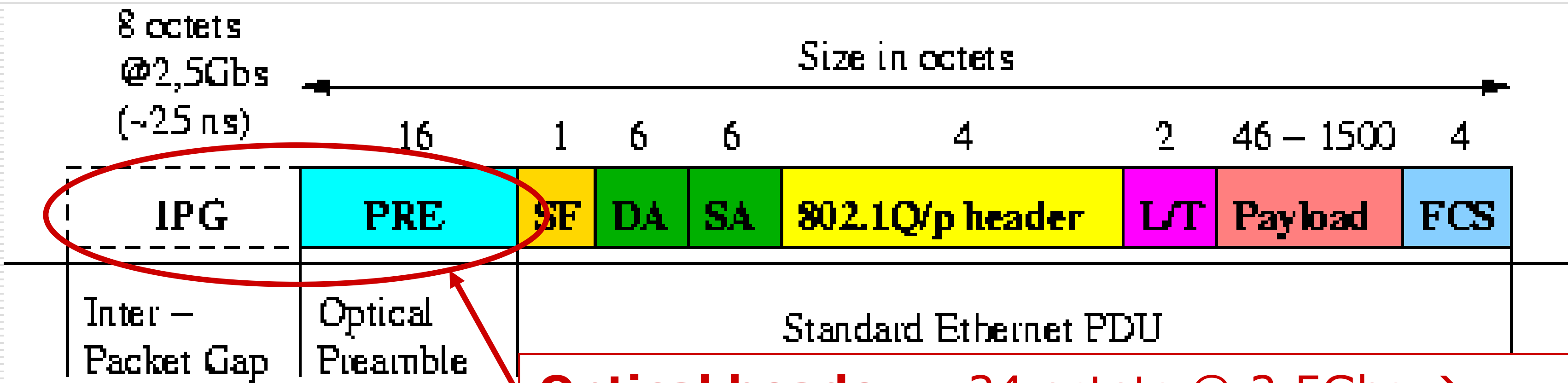
Enhanced MAC Protocols (1):

Packet Concatenation Mechanism - Modified Packet Bursting (MPB)

- Transmission efficiency
- Fairness

Encountered Problems

- OU-CSMA/CA: poor efficiency of transmission of small size packets
- E.g: optical packet format based on Ethernet



Optical header ≈ 24 octets @ 2.5Gbs → **transmission efficiency** of a packet of 50 bytes = **67%**

- Today, almost 60% of IP packets are of small sizes (≈ 50 bytes) [CAIDA]

MPB Principle

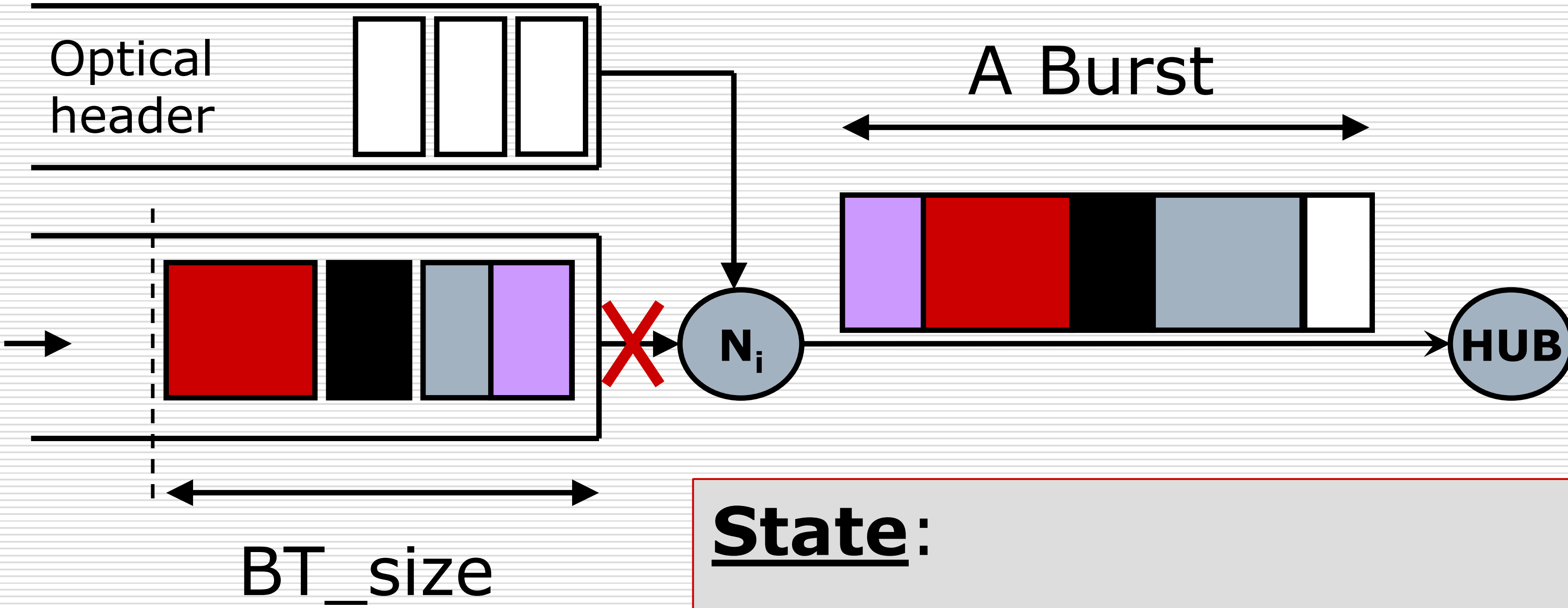
- MPB is Based on *Packet Bursting* in Gigabit Ethernet
- **Objective**: Reduce the volume optical overheads
- **How**:
  - Assembly / concatenate electronics packets into one optical « burst » **(≠ OBS)**
  - Burst length is only limited by: BW availability and # of available electronic packets

Bursting Timer

- **Note**: Transmission efficiency increases only if we have at least 2 electronic packets in buffer
- **Proposal of Bursting Timer**: which defines periods of assembly of electronic packets before building a burst

➔ Trade-off between additional delay and transmission efficiency !

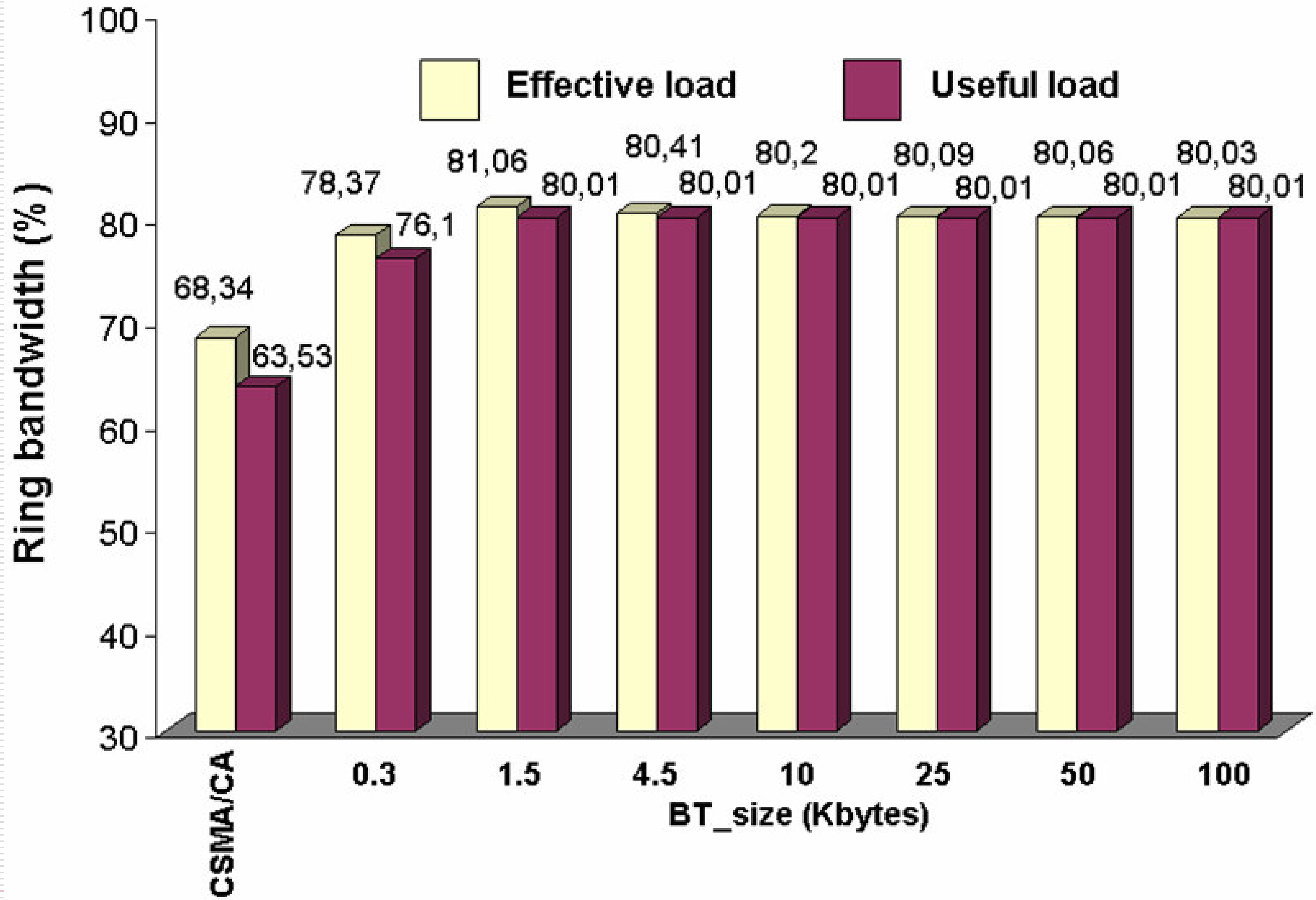
Illustration of MPB



State:

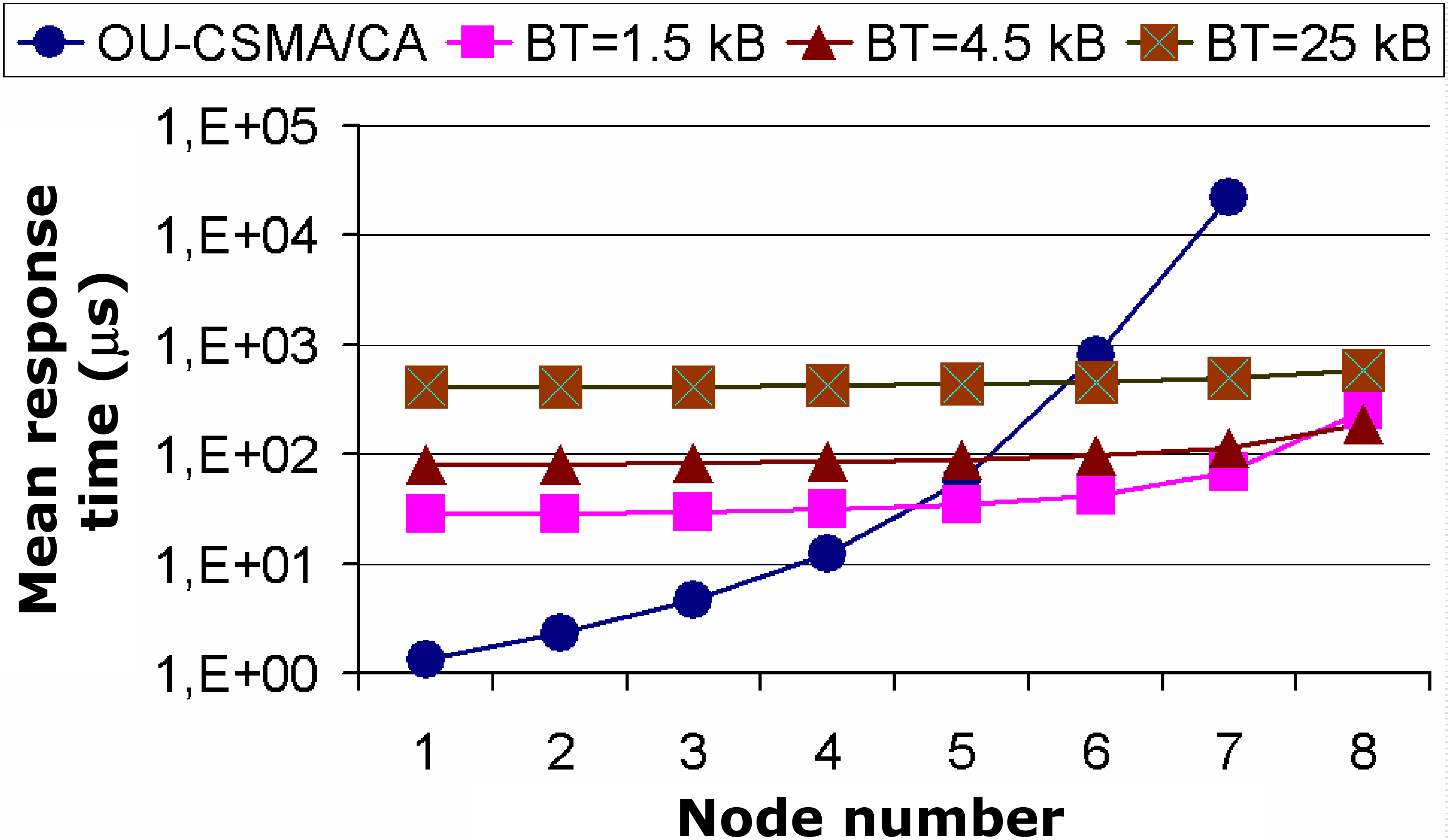
- A burst is in construction phase

Impact of BT_size (1)



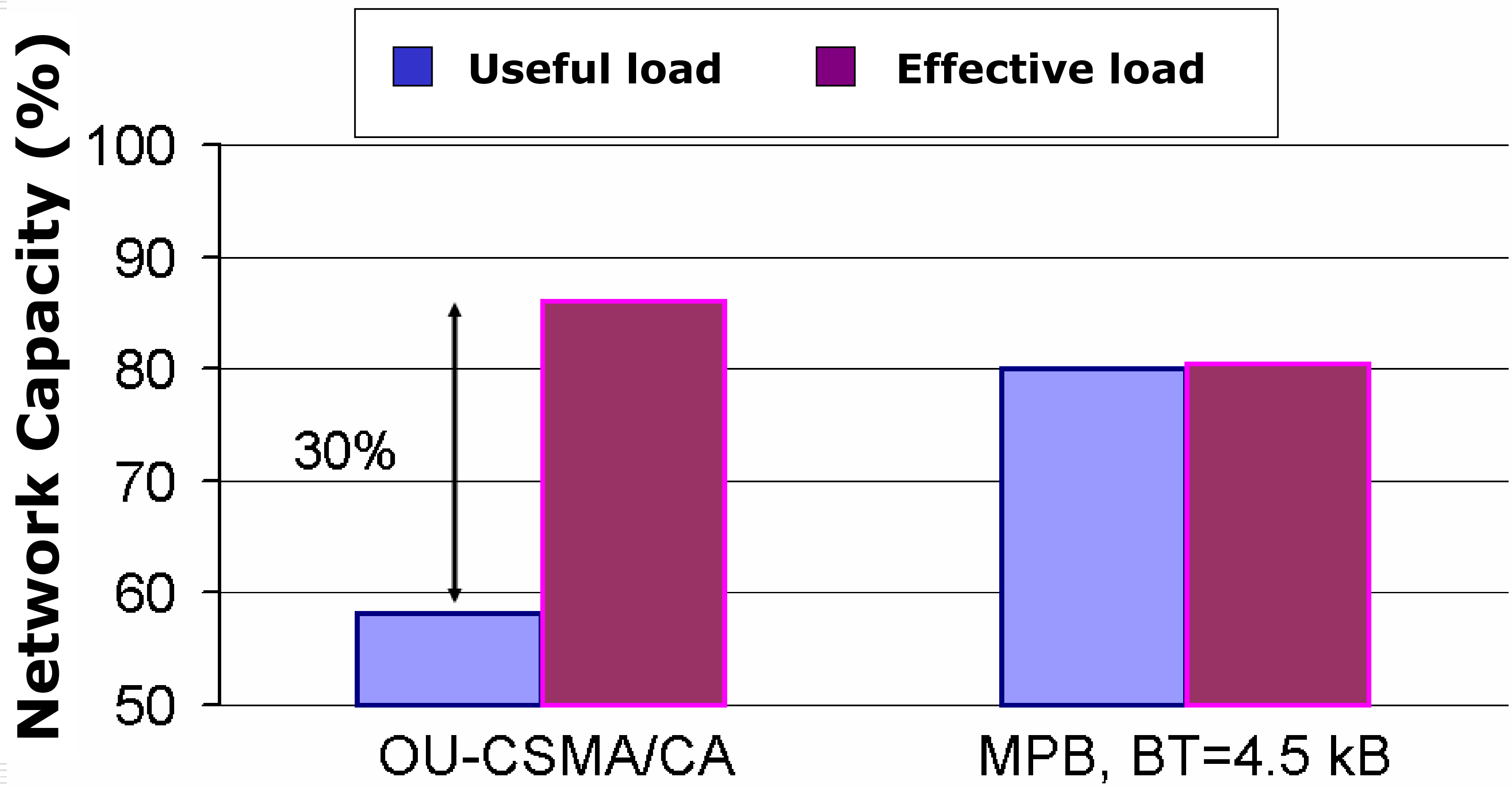
Params. Simul: Offered load=0.80, IP Packets

Impact of BT_size (2)



Params. Simul: Offered load =0.80, IP Packets

Impact of Volume of Small Packets



Offered Load = 0.80, 100% small packets

Conclusion on MPB

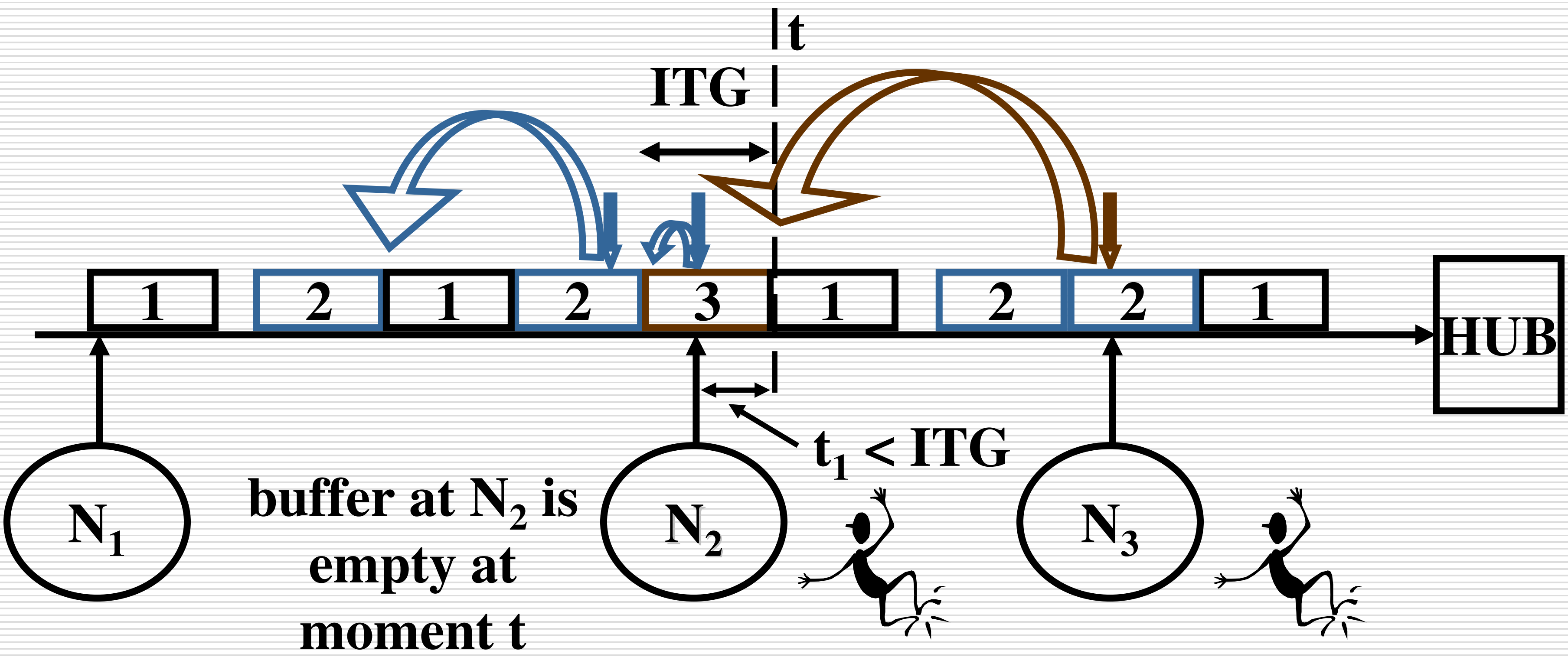
☺	😐
• Efficient: transmission efficiency ≈ 100% • ↗ maximum load ≈ 90% • Robust: network is more stable, insensitive to traffic changes	• Limited application fields ➡ Optical header does not contain any information (routing, control…) ➡ Tx/Rx in « burst mode »

Enhanced MAC Protocol (2): Dynamic Intelligent MAC (Di-MAC)

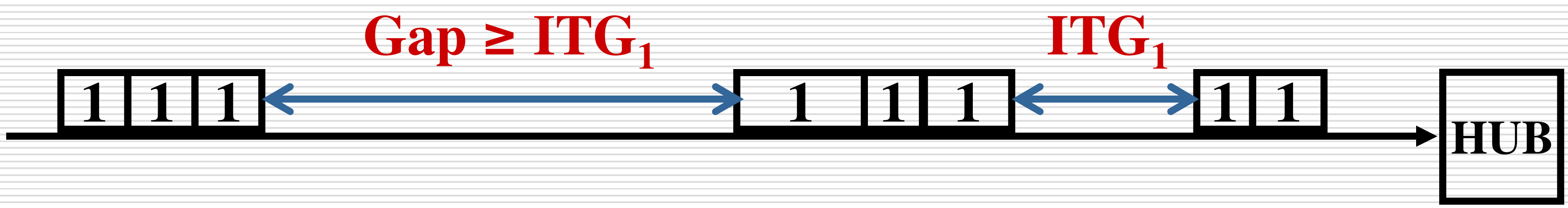
- **Objectives of the protocol:**
  - Overcome the problems of unfairness and of bandwidth segmentation
  - Distributed and Dynamic MAC

Principles of Di-MAC

ITG = Inter-Transmission Gap, should be ≥ MTU (Maximum Transmission Unit)



Importance of ITG



Bandwidth viewed by node 2

- Choice of ITG is very important
- Big ITG values at upstream nodes ➔ Very large bandwidth for downstream nodes ➔ Increase fairness and global performance !

Difficulties in the Choice of ITG

- How to find a compromise between **delay vs. ITG**
- **Distributed** operation : ITG adjustment must be done locally
- **Dynamic** adjustment: ITG adjustment must be dynamic in function of traffic/bandwidth occupation

Some Definitions

- **Threshold delay: Wmax** (chosen in function of SLAs)
- Node status: "**normal**", "**critical**" (excessive delay)
- **ITG blocking periods** (reservation periods)

Di-MAC Algorithm

- **Initially:** ITG = MTU (Maximum Transmission Unit)
- **ADD Process**:

« ITG ← ITG + MTU » **while** « Status = Normal »

- **SUB Process**:

« ITG ← MTU » **if** « Status = Critical »

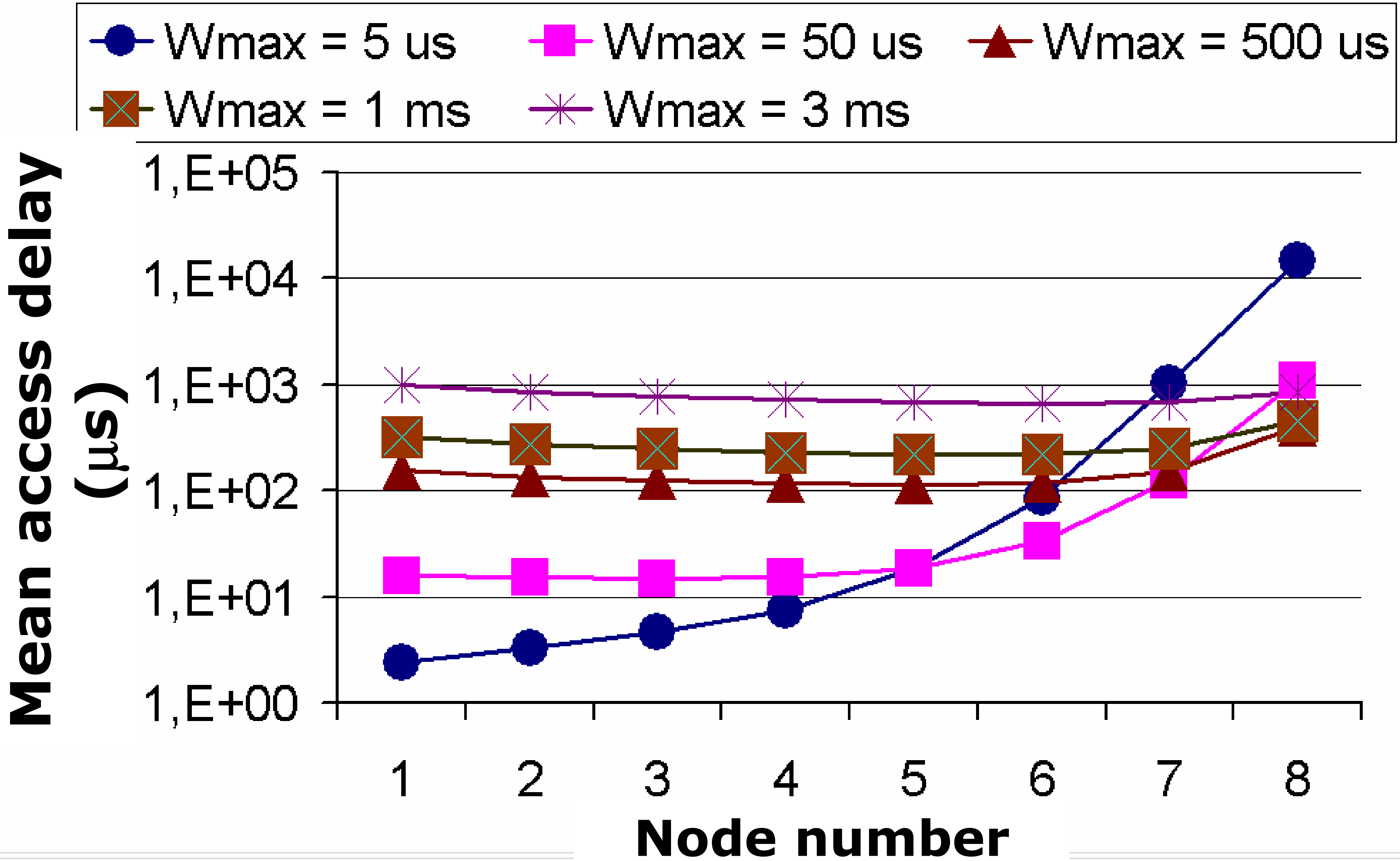
- **CONTROL Process**: Control of transmission based on the current value of ITG

DI-MAC vs. Multiservice

Multiservice version of DI-MAC

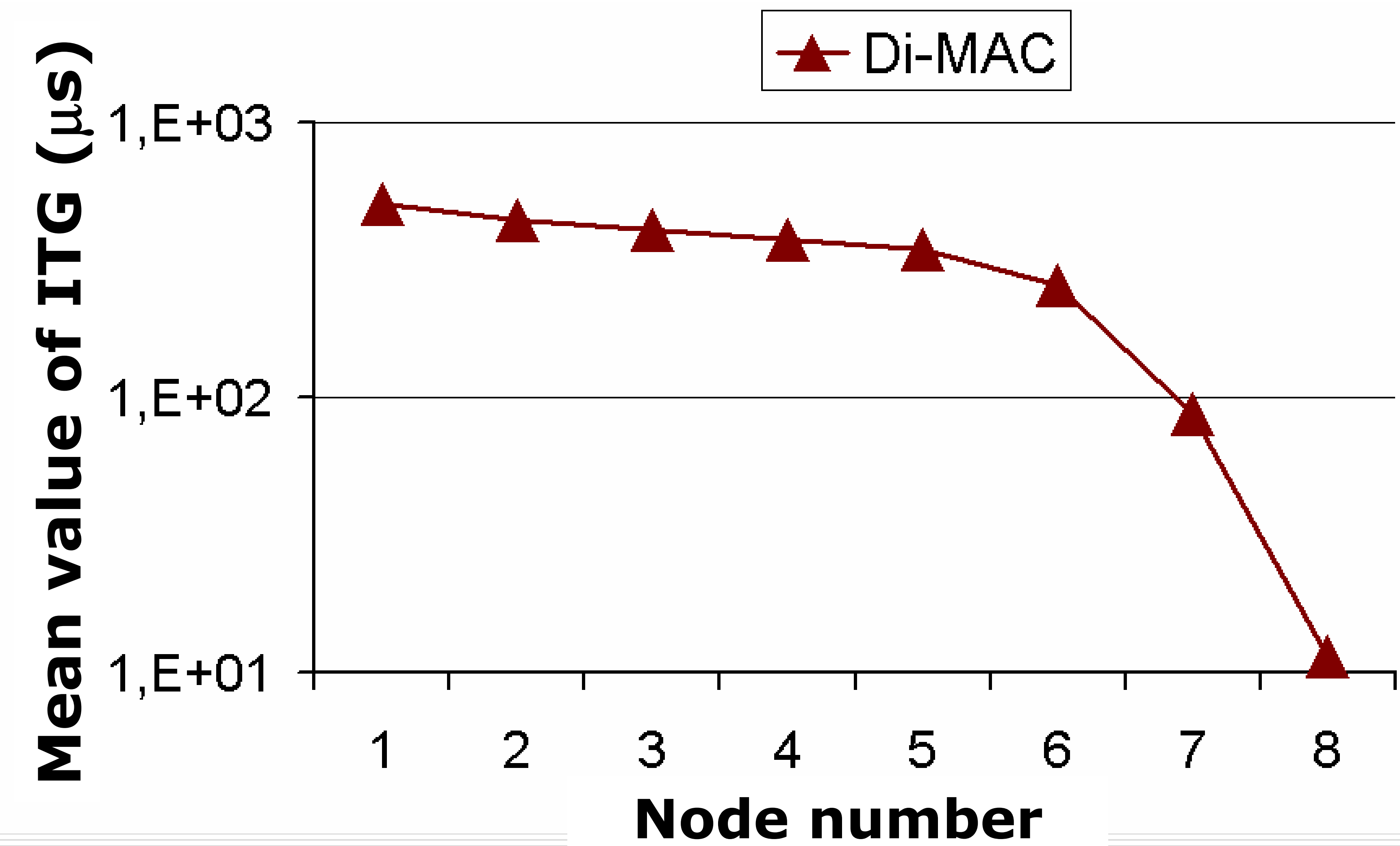
- The network supports a number of classes of service (CoS)
- Each service defines its own Wmax parameter
- High priority service must have small Wmax value
- Arrivals of high priority traffic end/interrupt ITG blocking periods of lower priority traffic
- Low priority traffic arrivals cannot influence the service of higher priority traffic

Impact of W_{max}



Params. Simul: Offered load=0.80, Uniform Traffic, IP Packets

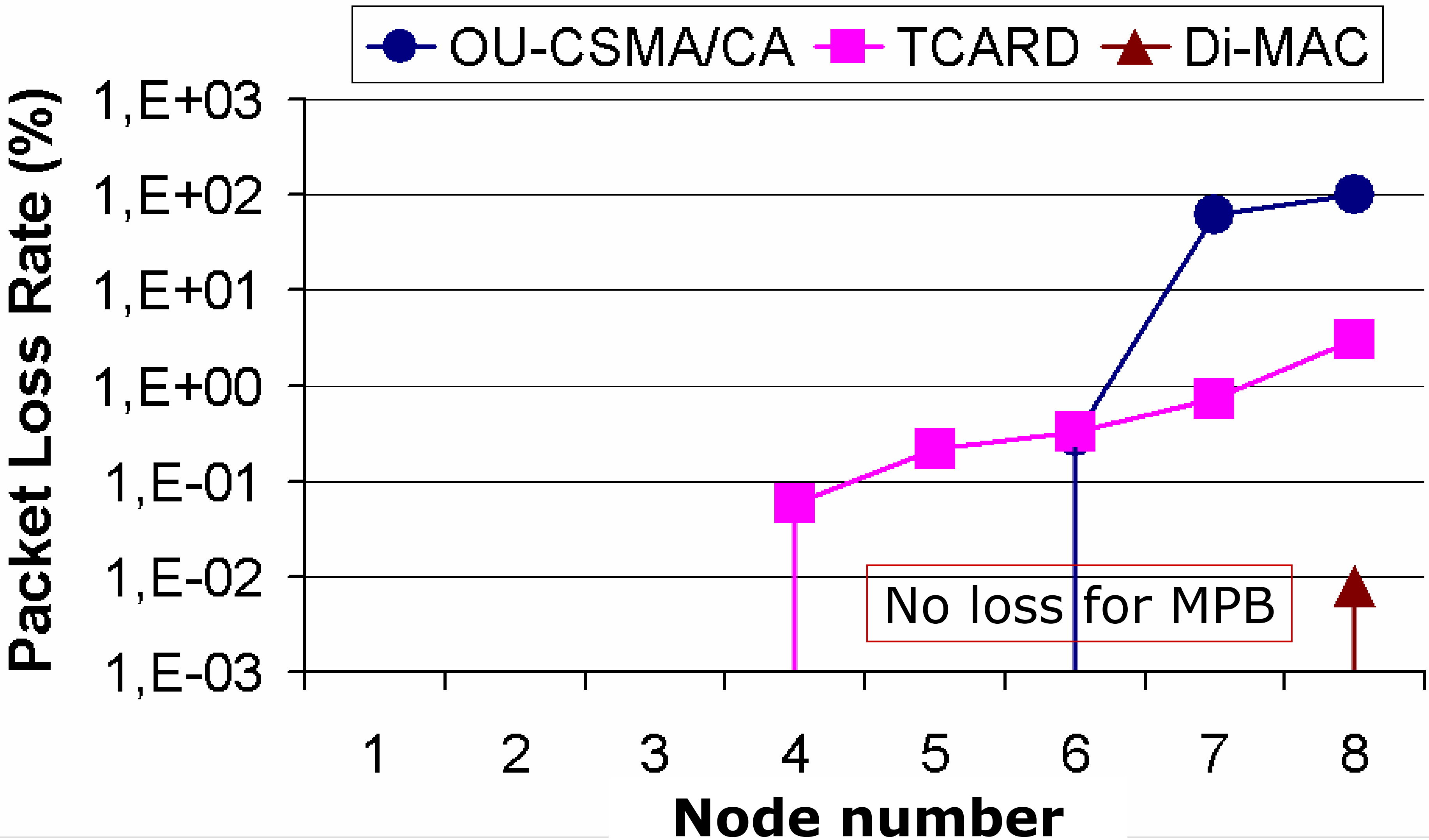
Mean values of ITG



Params. Simul: Offered load=0.80, Uniform traffic, IP Packets

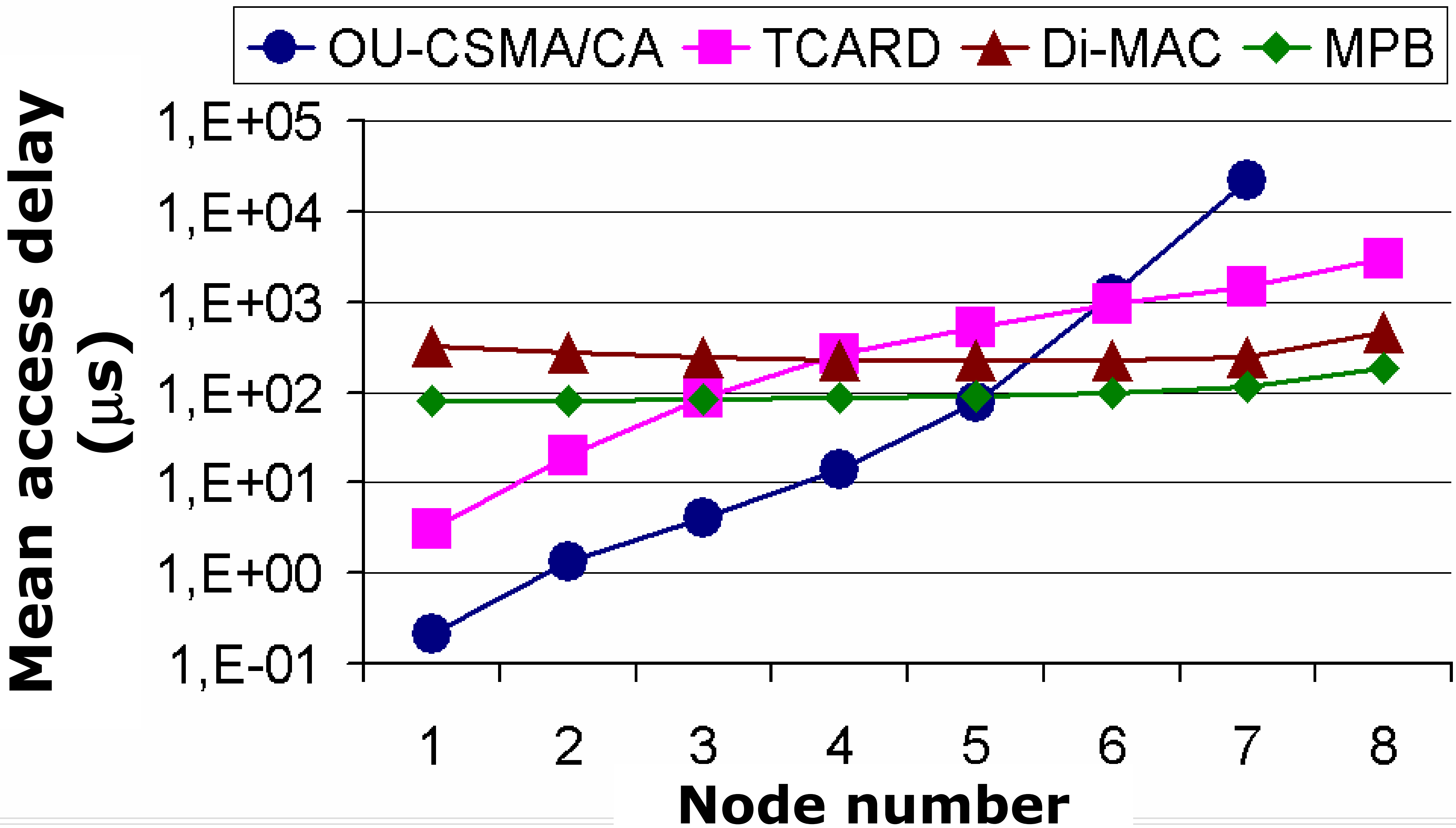
Wmax = 1 ms

DI-MAC vs. OU-CSMA/CA,TCARD, MPB



Params. Simul: Offered load=0.80, Uniform Traffic, IP Packets
Wmax = 1 ms; BT_size = 4.5 kB

DI-MAC vs. OU-CSMA/CA,TCARD, MPB

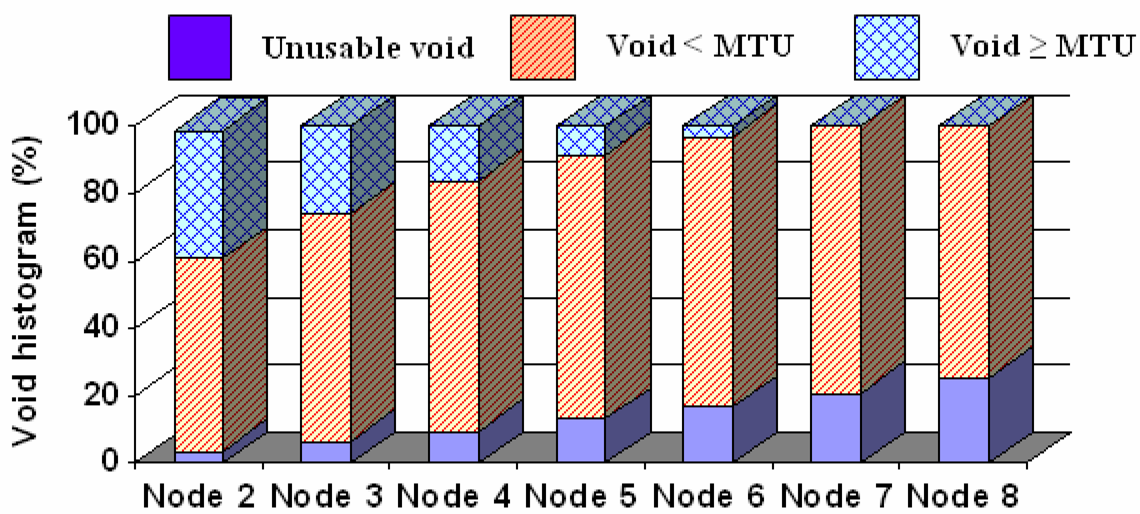


Params. Simul: Offered load=0.80, Uniform Traffic, IP Packets
Wmax = 1 ms; BT_size = 4.5 kB

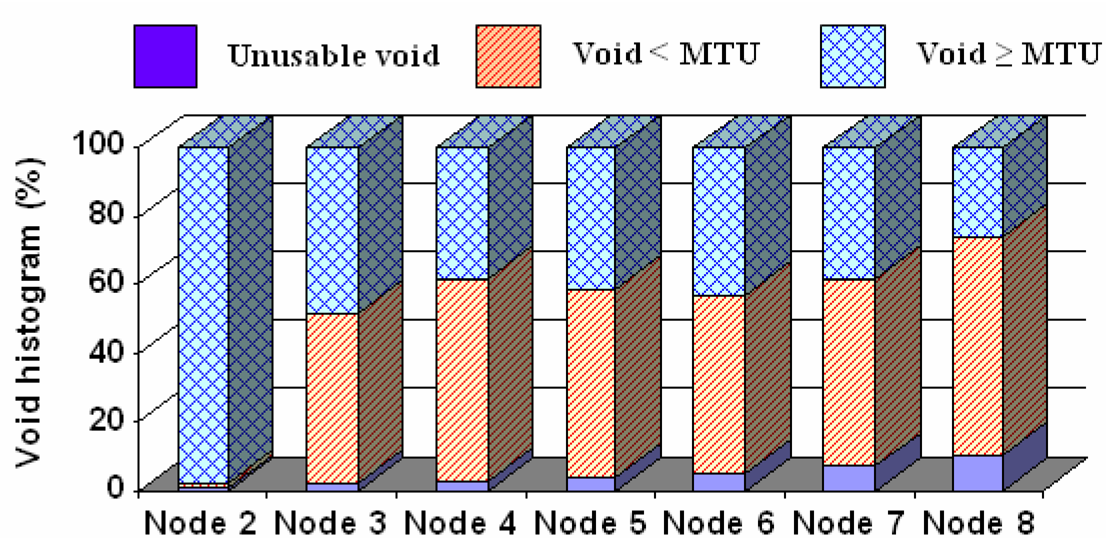
DI-MAC vs. OU-CSMA/CA: *Bandwidth Segmentation*

Histogram of voids

OU-CSMA/CA



DI-MAC



Conclusion on DI-MAC

☺	😐
DI-MAC = fair, dynamic and distributed Efficient: bandwith segmentation ↘; fairness ↗ ↗ maximum load ≈ 85% Robust: more stable network, insensitive to the trafic changes and to config. (ex:nodes add/drop)	Overbooking of bandwidth when the network is low loaded ITG versus very bursty traffic?

QoS:
TDM Transport Feasibility over Packet-Switched Network (Circuit Emulation Service)

Problems

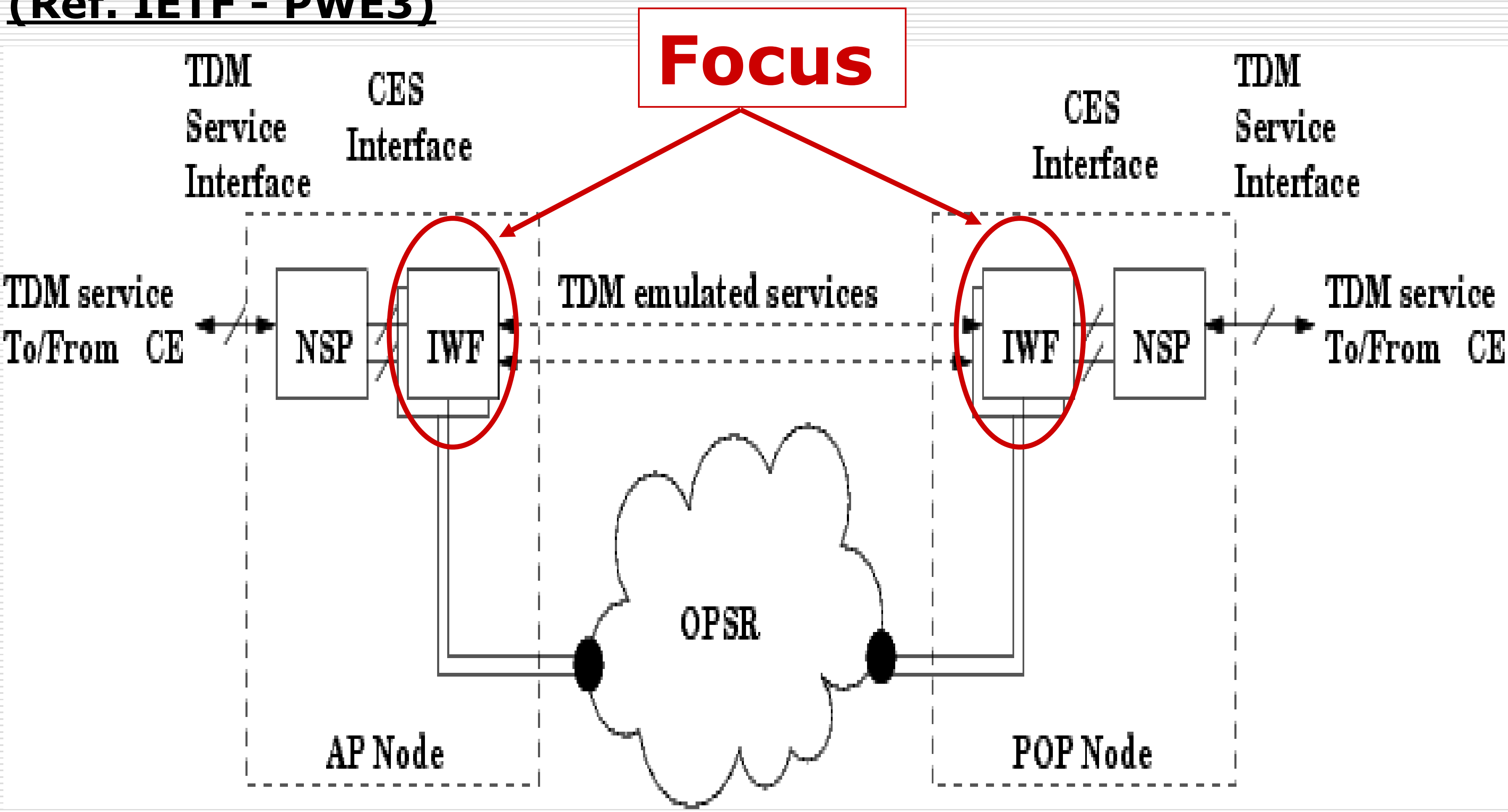
- **Tendency:** replacement of circuit-switched networks by Packets-switched networks
- **But:** circuit services constitute main revenues (incomes) of operators
- **So:** need of convergence of Circuit and Packet services over packet-switched networks
- **Question:** How to ensure QoS?

Circuit Emulation Service - CES

- **Objective:** provide a « TDM-like » service via a Packet-Switched Network, transparent for TDM classic customers.
- **Standard:** standardized by IETF, ITU-T, MEF, MFA Forum
- **NB:** CES ≠ VoIP!

Model of CES

(Ref. IETF - PWE3)



NSP = Native Service Processing; IWF = Inter-Working Function

TDM Frames Segmentation

- **Static Segmentation:** According to a **fixed threshold** (best value to be determined)
- TDM frames need CES overhead of 16 bytes (see IETF-PWE3)
- Performed by the **IWFs** blocs (**Inter-Working Functions**)

Globale Architecture of QoS

CoS	Priority	Network Performance		
		PLR	Delay	Jitter
CoS1 (Real Time, e.x.: TDM)	High	10^{e}-9	Strictly limited	Strictly limited
CoS2 (Loss sensitive)	Average	10^{e}-9	Limited	Limited
CoS3 (Best-effort)	Low	No guarantee	No guarantee	No guarantee

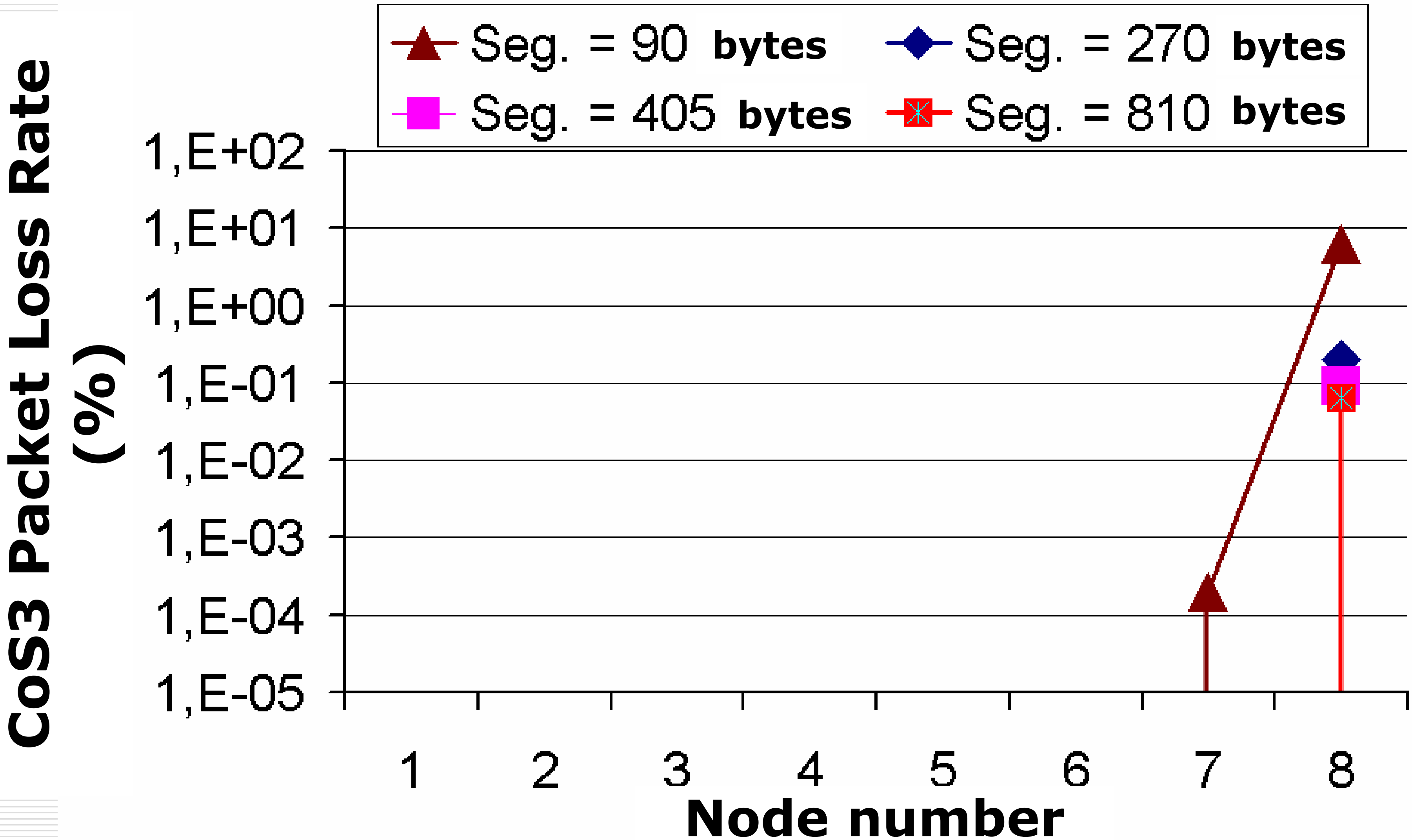
Services mapping

Client Applications	**Real-time Applications** Voice (Telephony, VoIP,…), Interactive Video on demand…	**Loss-sensitive Applications** Banking transaction, Critical data transfer….	**Best-effort Applications** File transfer, Web browsing…
Client Network Services	**PDH, SONET/SDH** (TDM), **ATM** (CBR, Real-time VBR), **IP** (IntServ), **Frame Relay** (RT-VFR), **Ethernet** QoS…	**ATM** (ABR, Non-RT VBR), **IP** (IntServ, DiffServ), **Frame Relay** (Non-RT-VFR) …	**Ethernet** (BE), **IP** (BE), **Frame Relay** (BE), **ATM** (UBR/+)…
MAN Services	**CoS1**	**CoS2**	**CoS3**

QoS Requirements of CES

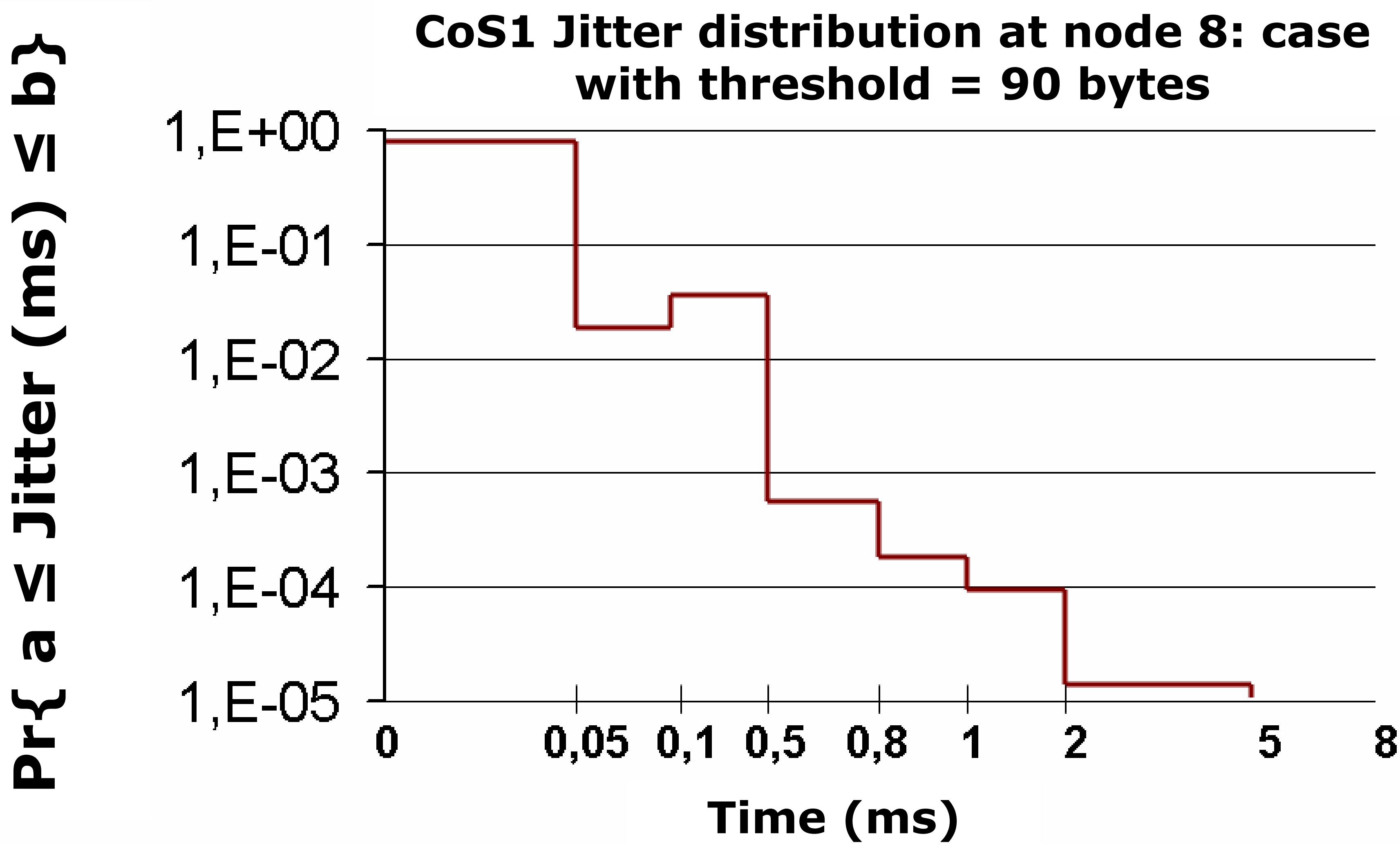
- According to Metro Ethernet Forum (MEF):
  - Packet Loss Rate and maximum end-to-end delay must be minimized
  - Maximum Jitter is limited to 10 ms.

Impact of segmentation threshold (1)

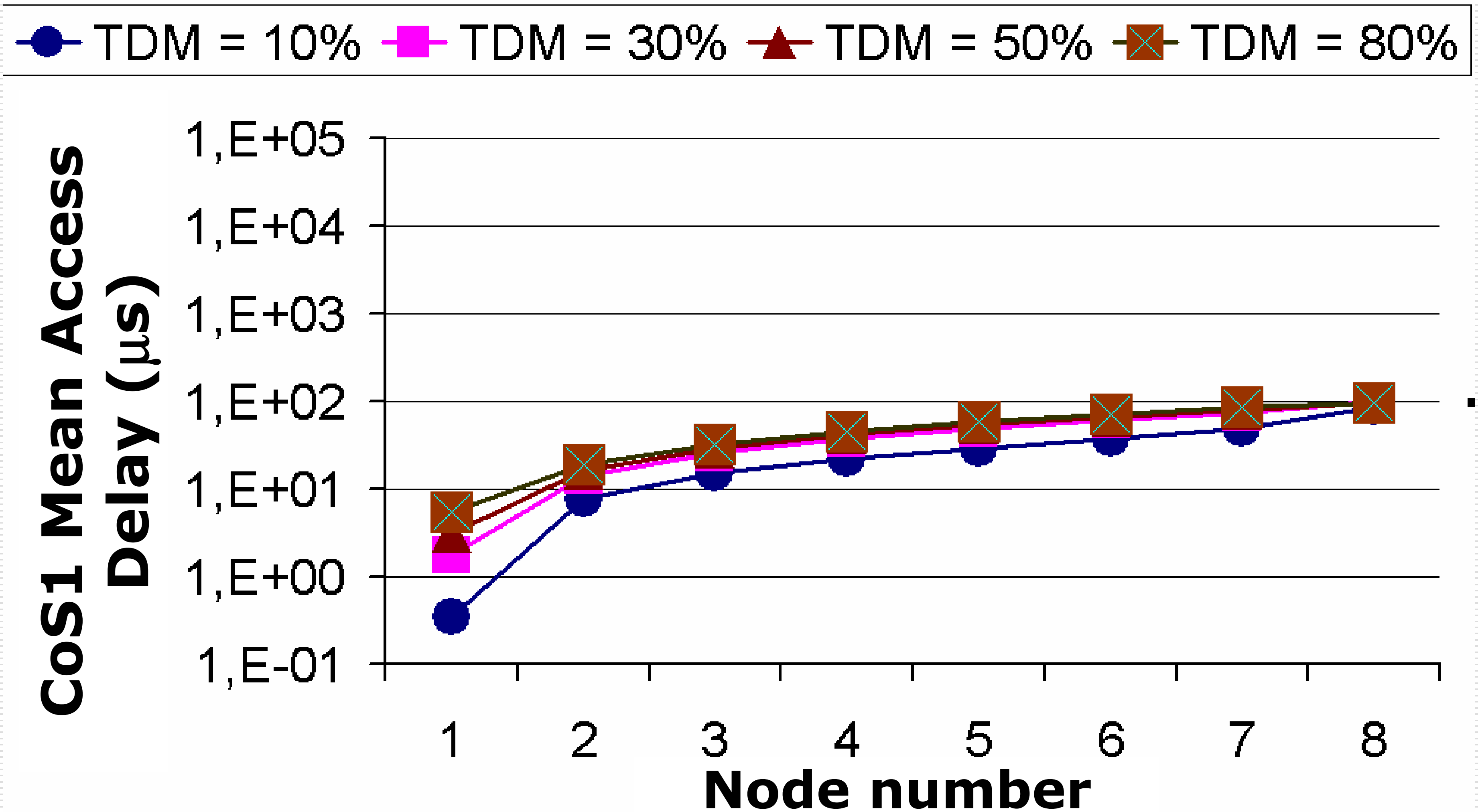


Load = 0.80, Uniform Traffic, with Di-MAC, CoS1 (15%) : CoS2 (20%) : CoS3 (65%), TDM = STS-3 (Frame = 2430 bytes)

Impact of segmentation threshold (2)

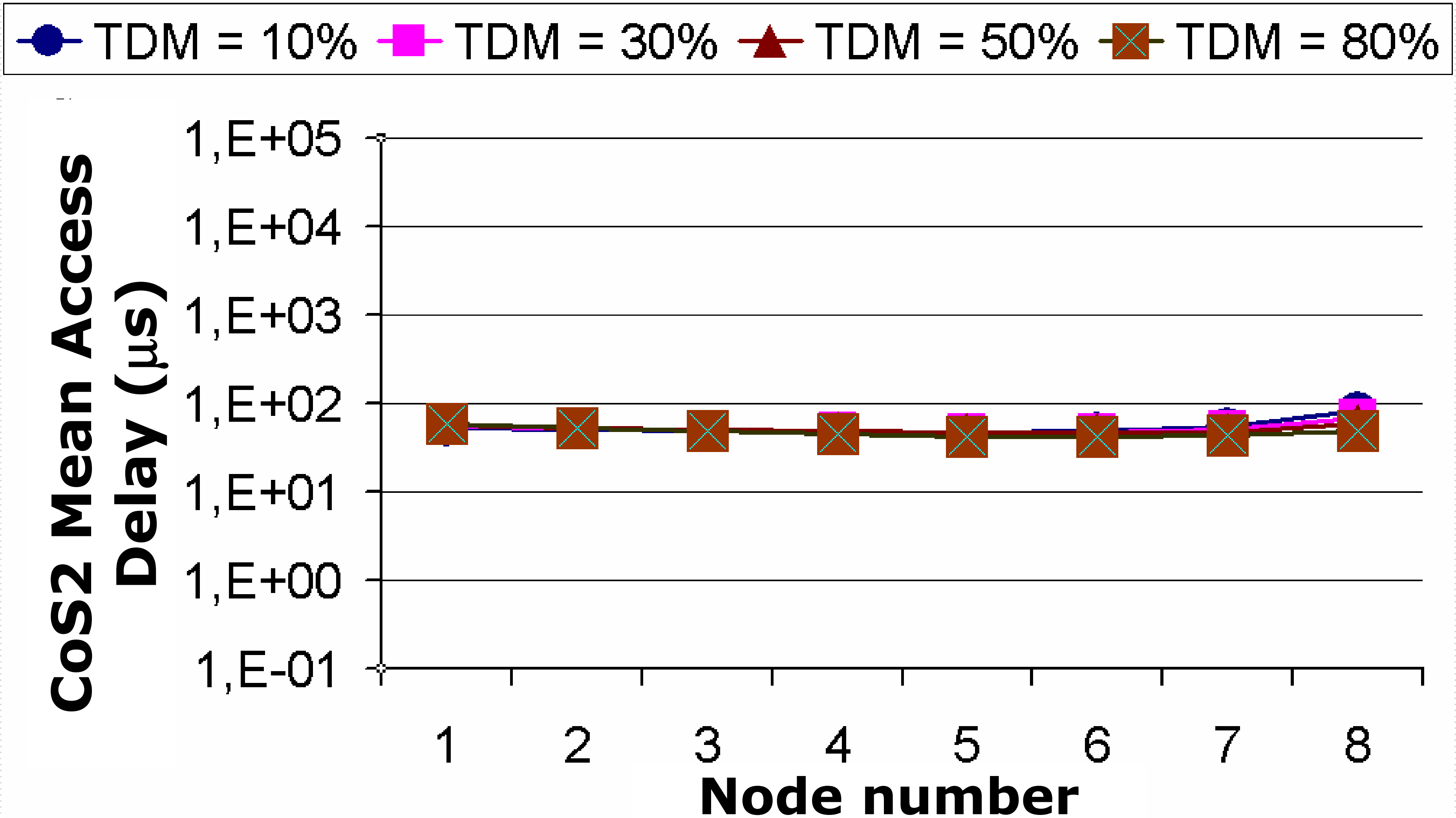


Limite of TDM Volume



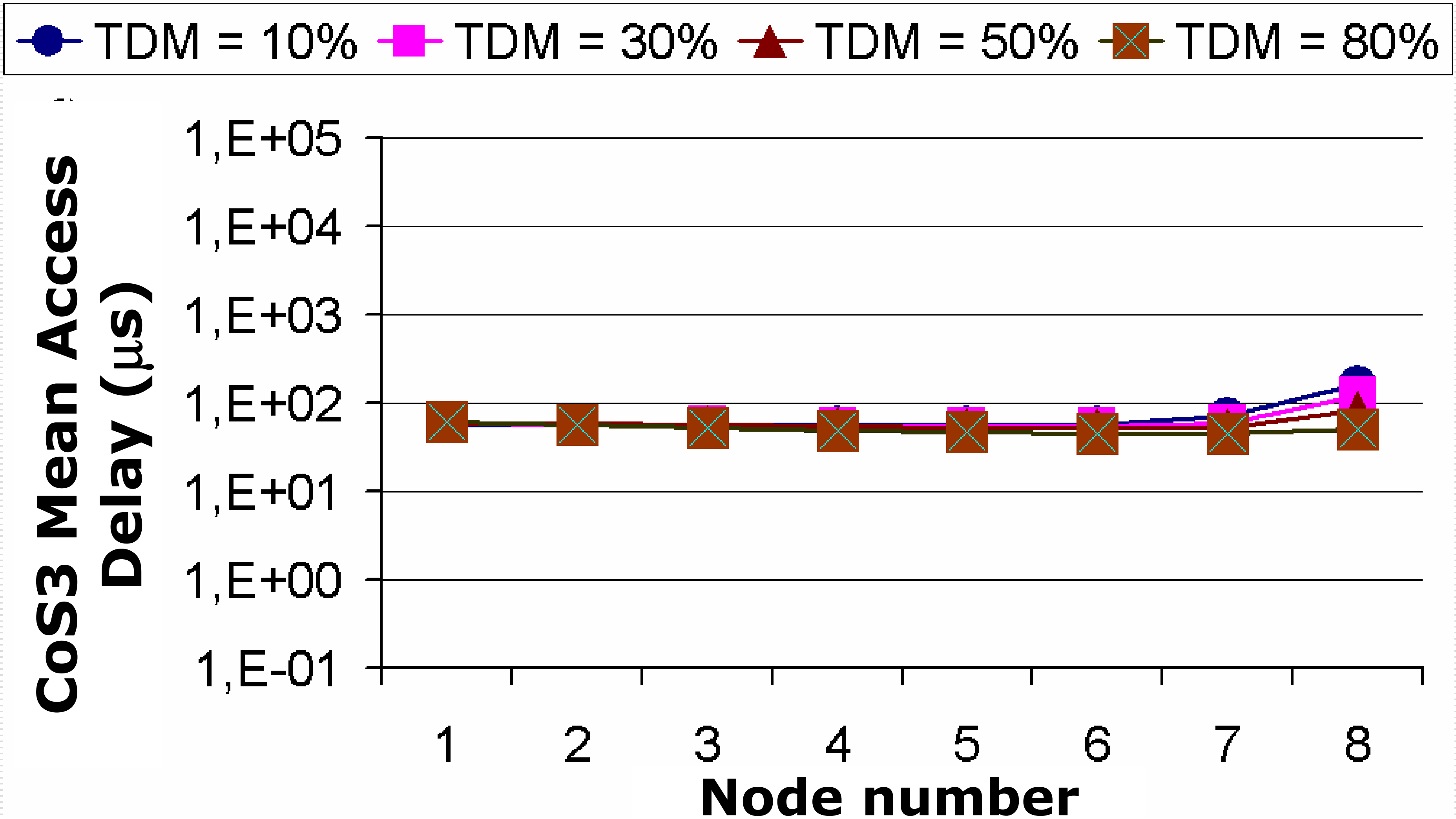
Load = 0.80, Uniform Traffic, with Di-MAC, TDM=810 bytes, CoS2=250 bytes, CoS3: IP packets

Limite du Volume de TDM (2)



Load = 0.80, Uniform Traffic, with Di-MAC, TDM=810 bytes, CoS2=250 bytes, CoS3: IP packets

Limite du Volume de TDM (3)



Load = 0.80, Uniform Traffic, with Di-MAC, TDM=810 bytes, CoS2=250 bytes, CoS3: IP packets

Conclusion on CES

- CES is feasible over the studied network
- Threshold segmentation: 810 bytes for SONET/SDH
- With OU-CSMA/CA:

  TDM is guaranteed BUT **data** is lost.

➔ The Network should use Di-MAC or MPB

Global Conclusions

- Study of optical passive bus-based MAN in « double bus »
- Proposal of Analytical model for OU-CSMA/CA + New solution of M/G/1 Queue with PRI
- Proposal of two advanced access mechanisms : MPB and Di-MAC (Transmission Efficiency, Fairness, Bandwidth Segmentation)
- Feasibility study of Circuit Emulation Service (CES) → convergence of Circuit and Packet

Perspectives (1)

- On MPB:
  - Find an algorithm allowing to configure the MPB parameters for a given network
  - Due to the changeable nature of traffic, choosing fixed values for MPB parameters might not be convenient due to high variability of the offered traffic ➔ An « Adaptive MPB » solution would be needed
  - Application of MPB in WDM environment :
    - Interaction between MPB parameters and load balancing mechanisms will need to be deeply studied

Perspectives (2)

- What will be the performance of MPB when it is generalized to be applied to others topologies (e.g. mesh, tree,…)?
- Will BT mechanism help MPB to improve the network performance with a very big BT_size?

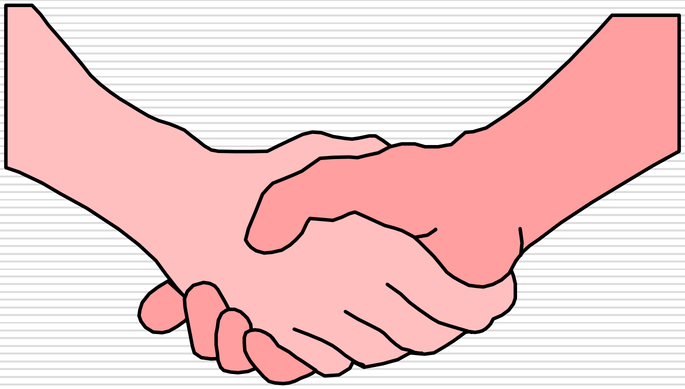
Perspectives (3)

- On Di-MAC:
  - Bursty self-similar traffic may cause the ITG parameters to be adjusted more frequently ➔ oscillation?
  - Impact of WDM dimension on the performance of DI-MAC: WDM increases the transmission capacity of the network, it may influence the parameter adjusting process of DI-MAC
  - What will be the performance of Di-MAC when it is generalized to be applied to others topologies (e.g. mesh, tree,…)?

References

- [BBD+ 05] N. Bouabdallah, A.L. Beylot, E. Dotaro, and G. Pujolle, Resolving the Fairness Issue in Bus-Based Optical Access Networks, *IEEE Journal on Selected Areas in Communications*, Vol. 23, No. 8. August 2005
- [BDC+ 04] N. Bouabdallah, E. Dotaro, L. Ciavaglia, et *al.*, Resolving the fairness issue in bus-based optical access networks, *IEEE Communications Magazine*, Volume 42, Issue 11, Page(s):S12 - S18, Nov. 2004.
- [CAIDA 99] KC. Claffy and S. McCreary, Internet Measurement and Data Analysis: Passive and Active Measurement, CAIDA - Cooperative Association for Internet Data Analysis, 1999, Available at: http://www.caida.org/outreach/papers/1999/Nae4hansen/.
- [CCH 05] H. Castel, M. Chaitou, and G. Hébuterne, Preemptive Priority Queues for the Performance Evaluation of an Optical MAN Ring, *Proceedings of Performance Modeling and Evaluation of Heterogeneous Networks (Het-Net'05)*, 2005.
- [CHE 03] H. Castel and G. Hébuterne, Performance Analysis of an optical MAN ring for variable length packet traffic. In *Proceedings of Photonic in Switching Conference 2003*.
- [COS 61] D.R. Cox and W.L. Smith, *Queues*. John Wiley, New York, 1961.
- [HGJ+ 05] G. Hu, C.M. Gauger, and S. Junghans, Performance Analysis of the CSMA/CA MAC Protocol in the DBORN Optical MAN Network Architecture, *Proceedings of the 19th International Teletraffic Congress (ITC 19)*, 2005.
- [NBA+ 04] V.H. Nguyen, M. Ben Mamoun, T. Atmaca and al., "Performance evaluation of Circuit Emulation Service in a metropolitan optical ring architecture", *Telecommunications and Networking – ICT 2004*, LNCS 3124 Vol., pp. 1173-1182, Fortaleza – Brazil, August 2004.

- [NGA 05] V.H. Nguyen, T. Atmaca, "DI-MAC – A Fair Access Protocol for Optical Metro Networks", Presented at *IFIP WG 7.3 International Symposium on Computer Performance, Modelling, Measurements and Evaluation - Performance 2005*, Juan-les-Pins, France, October 3-7th, 2005.
- [NGA+ 05] V.H. Nguyen, T. Atmaca, "I-MAC: A fair access protocol for passive asynchronous optical ring networks", *3e Performance Modelling and Evaluation Of Heterogeneous Networks (Het-Nets'05),* Technical Proceeding, p06/1-p06/10, July 2005.
- [NGA++ 05] V.H. Nguyen, T. Atmaca, "Performance analysis of the Modified Packet Bursting mechanism applied to a metropolitan WDM ring architecture", *IFIP Open Conference on Metropolitan Area Networks*: Architecture, protocols, control and management, Technical proceeding, pp.199-215, HCMC Viet Nam, 11-13 April 2005.
- [NGA 06] V.H. Nguyen, T. Atmaca, "Dynamic Intelligent MAC Protocol for Metropolitan Optical Packet Switching Ring Networks", *IEEE International Conference on Communications – ICC 2006*, Istanbul, Turkey, 11-15 June, 2006.
- [TOP 03] J. Tomasik and D. Popa, "Stochastic Automata Network Model of an Edge Node in All-Optical Asynchronous Double Bus Ring Network", *Photonic in Switching 2003*, Versailles, France, 2003.



Thank you!

Questions ?