Canova Tech

The Art of Silicon Sculpting

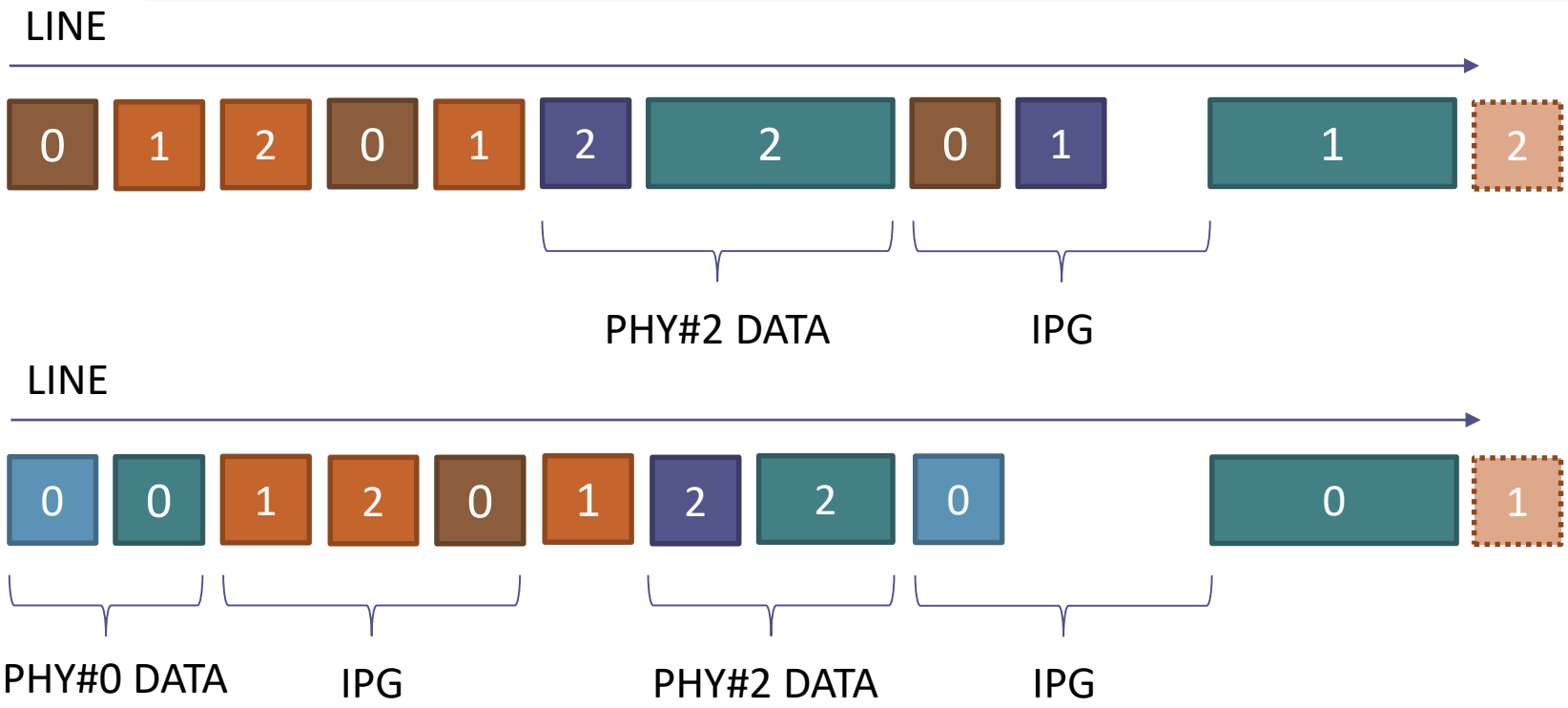
PIERGIORGIO BERUTO
ANTONIO ORZELLI

IEEE802.3cg TF

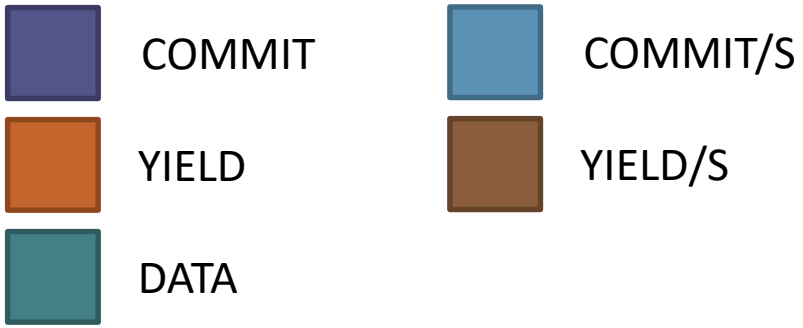
PHY-Level Collision Avoidance – Addendum #1

August 16th, 2017

Control signals



- **Control signals renamed since last pres.**
  - ACK → COMMIT
  - NACK → YIELD
- **Added new control signals**
  - COMMIT/S → same as COMMIT, sent by PHY #0 to allow synchronization
  - YIELD/S → same as YIELD, sent by PHY #0 to allow synchronization

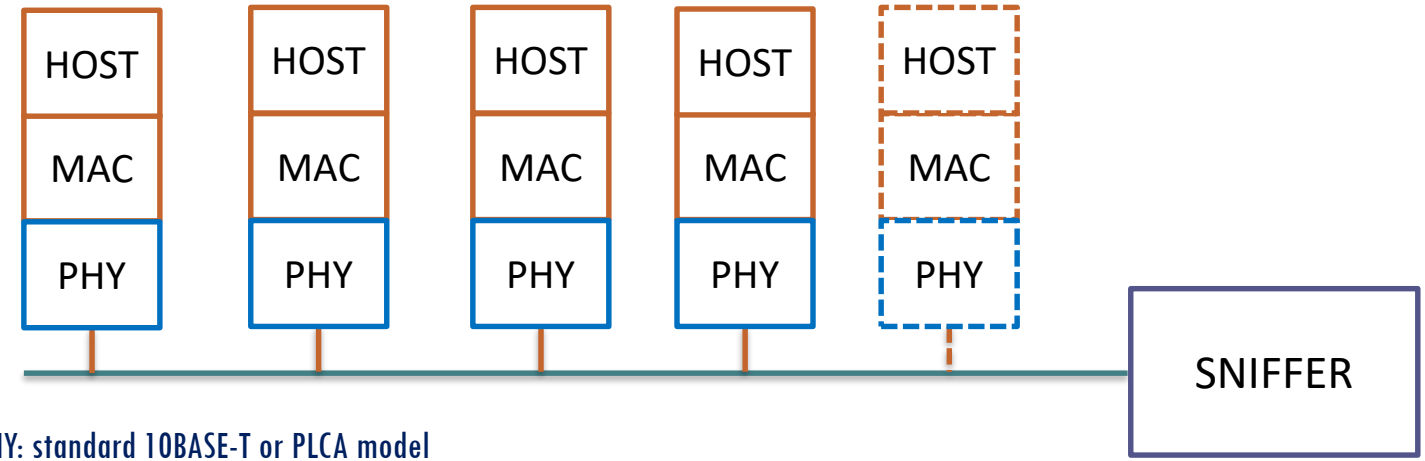


Determinism

- Packet Ordering
  - Node level: node "A" transmits packets #0, #1 and #2 at times T0 < T1 < T2. Packets shall be received in the same order exactly (#0, #1, #2) with unspecified latency.
    - This is **guaranteed**
      - MAC procedures are synchronous → once a frame is waiting to be transmitted, the MAC does not process other frames until the current one is either transmitted or discarded because of excessive attempts.
        - » In PLCA proposal packets are NEVER discarded in such way as the packet is always transmitted within at maximum two attempts.
      - PHY is normally supposed to deliver the frame as soon as it's fed be the MAC.
        - » In PLCA proposal, the PHY shall buffer a *very short* portion of the packet in order to reduce latency.
        - » Not enough to allow misordering (would need to buffer at least two whole packets)
  - Network level: node "A" transmits packets #0, #1 and #2 while node "B" transmits packets #3, #4, #5 at times T0 < T3 < T1 < T4 < T2 < T5. Packets shall be received in the following order: #0, #3, #1, #4, #2, #5.
    - Not how ethernet works
    - Out of scope of PLCA proposal
- Latency
  - Best case (min) occurs when the MAC initiate TX just before handshaking time (~ 0)
  - Worst case (max) occurs when the MAC initiate TX after handshaking time AND each other node commits to transmit a 1500 bytes packet (~ 7ms for a six node network)
  - Comparison with CSMA/CD (see next slides)

Simulations

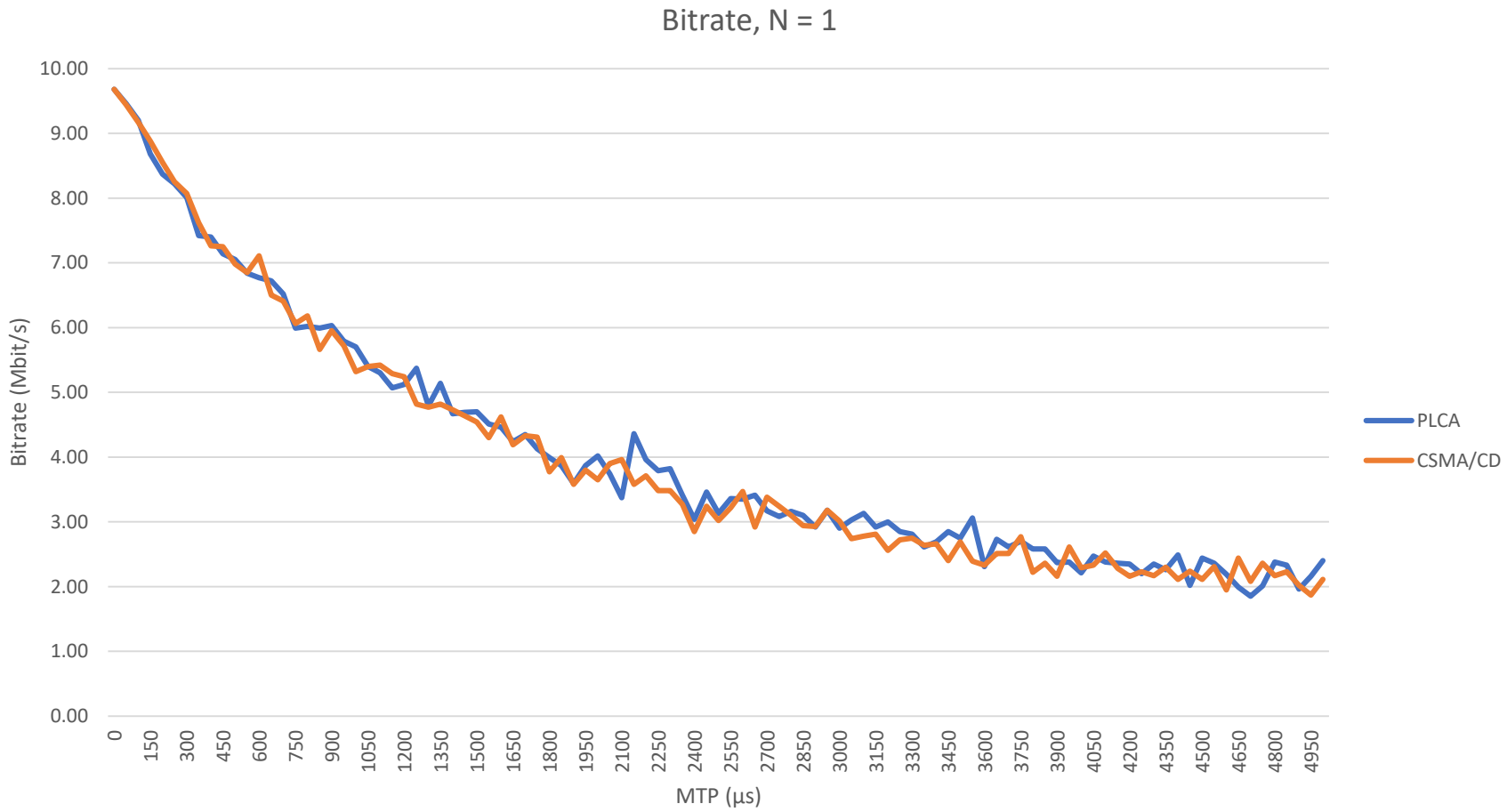
- Full digital simulation (verilog)
  - 4b/5b encoding + DME (25Mhz)
  - Use 5b codes S, R, T, H for COMMIT, YIELD, COMMIT/S, YIELD/S signaling



- PHY: standard 10BASE-T or PLCA model
- MAC: standard CSMA/CD capable MAC (802.3 clause 4)
  - host interface: DPRAM (one frame) + busy indication + size + trigger
  - PHY interface: MII (txd, txclk, txen, txer, rxd, rxclk, rxdv, rxer, col, crs)
- HOST: simple transmitter
  - wait for MAC BUSY = 0
  - wait random time between 0 and MTP (sim parameter, 0 = MAX speed)
  - write random data in DPRAM of size PKTSZ (sim. parameter 60 < PKTSZ < 1500) or random size
- SNIFFER: measuring throughput, latency
  - throughput: number of received bytes (excluding FCS, PREAMBLE) / total simulation time
  - latency: time between BUSY = 1 and BUSY = 0 for each node

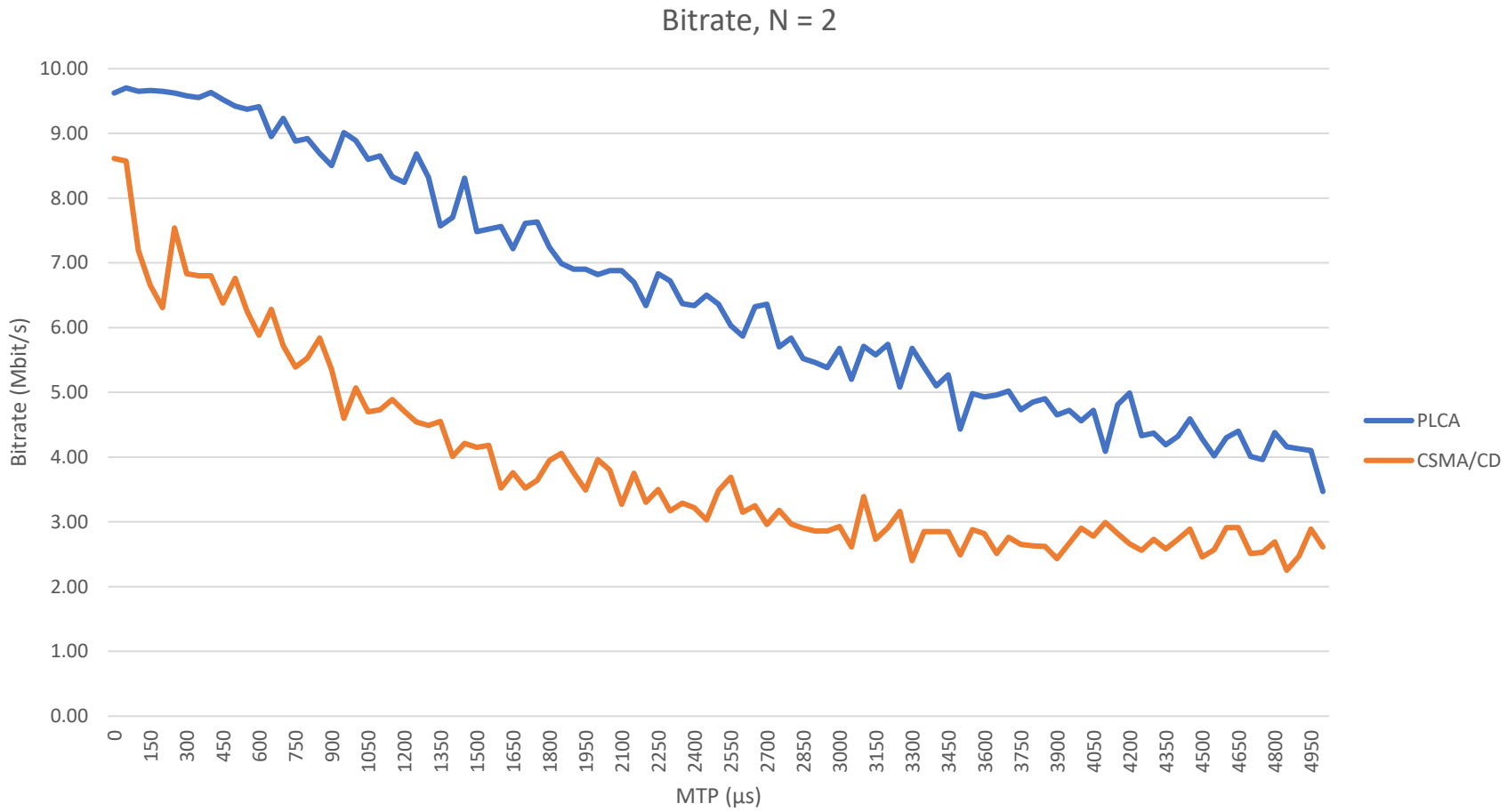
Throughput

- 100 pkts, random size (60, 1500), variable MTP and N (number of nodes)
- Comparison between 10base-T (simple CSMA/CD) and PLCA



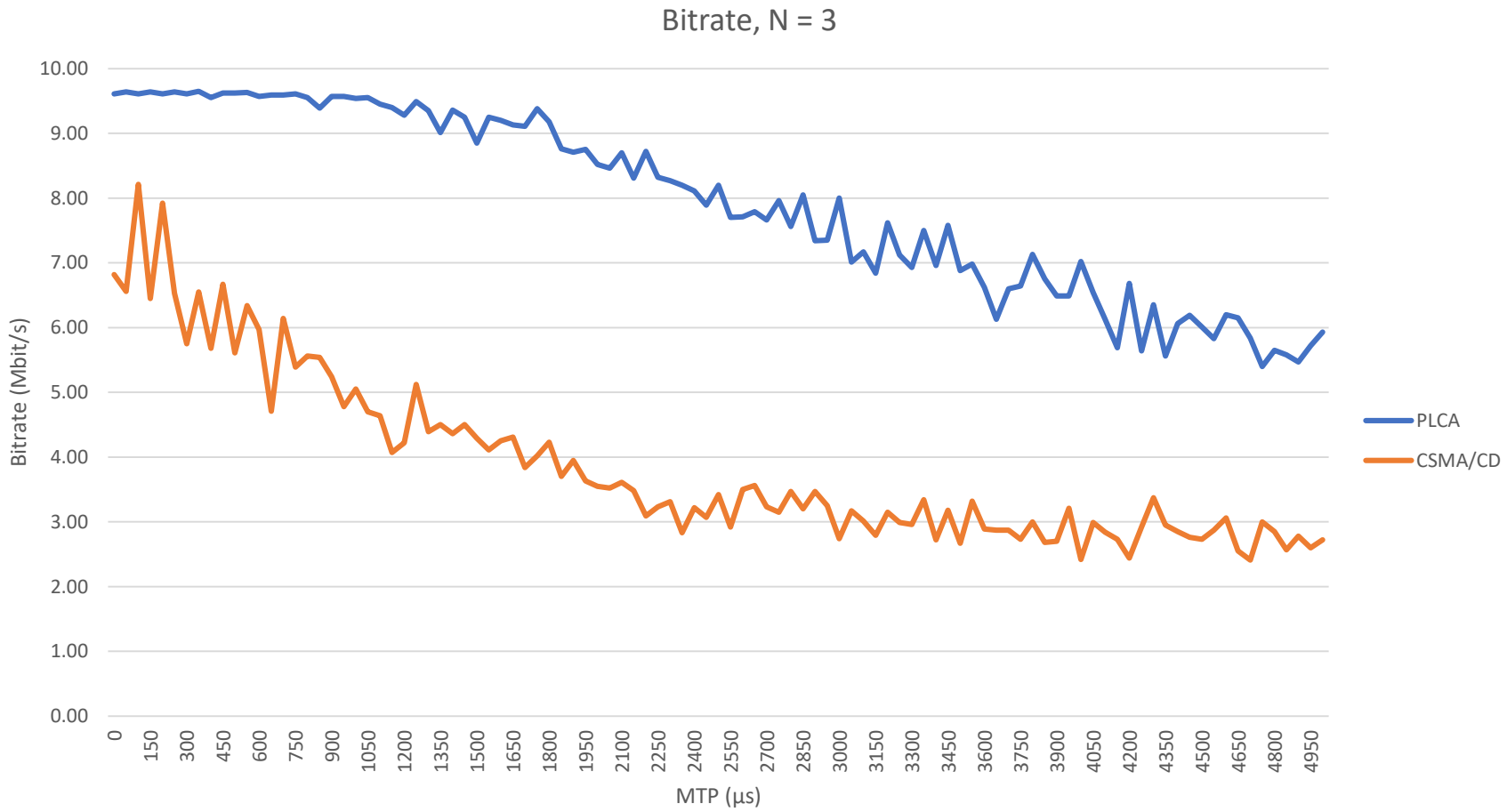
Throughput

- 100 pkts, random size (60, 1500), variable MTP and N (number of nodes)
- Comparison between 10base-T (simple CSMA/CD) and PLCA



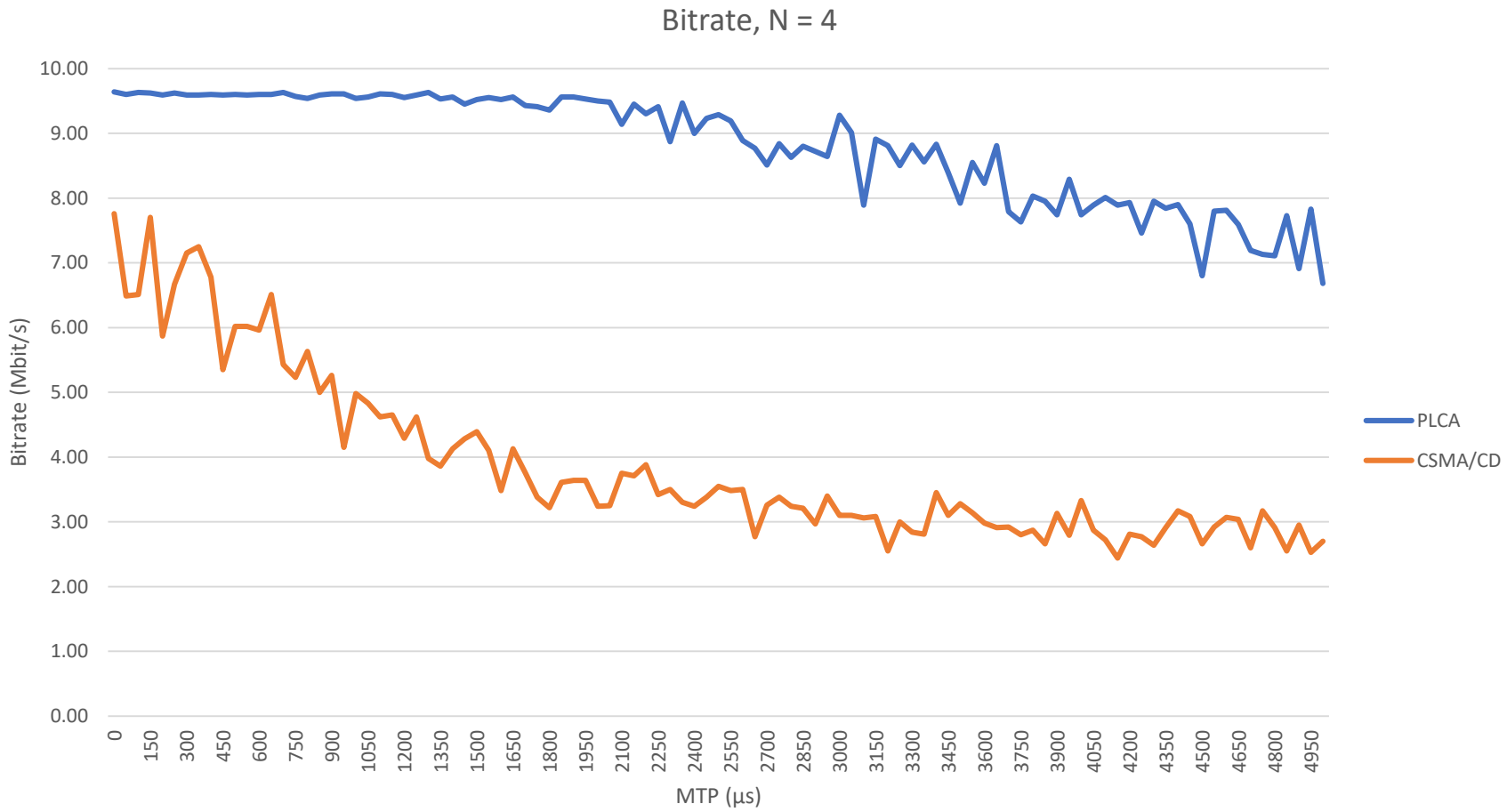
Throughput

- 100 pkts, random size (60, 1500), variable MTP and N (number of nodes)
- Comparison between 10base-T (simple CSMA/CD) and PLCA



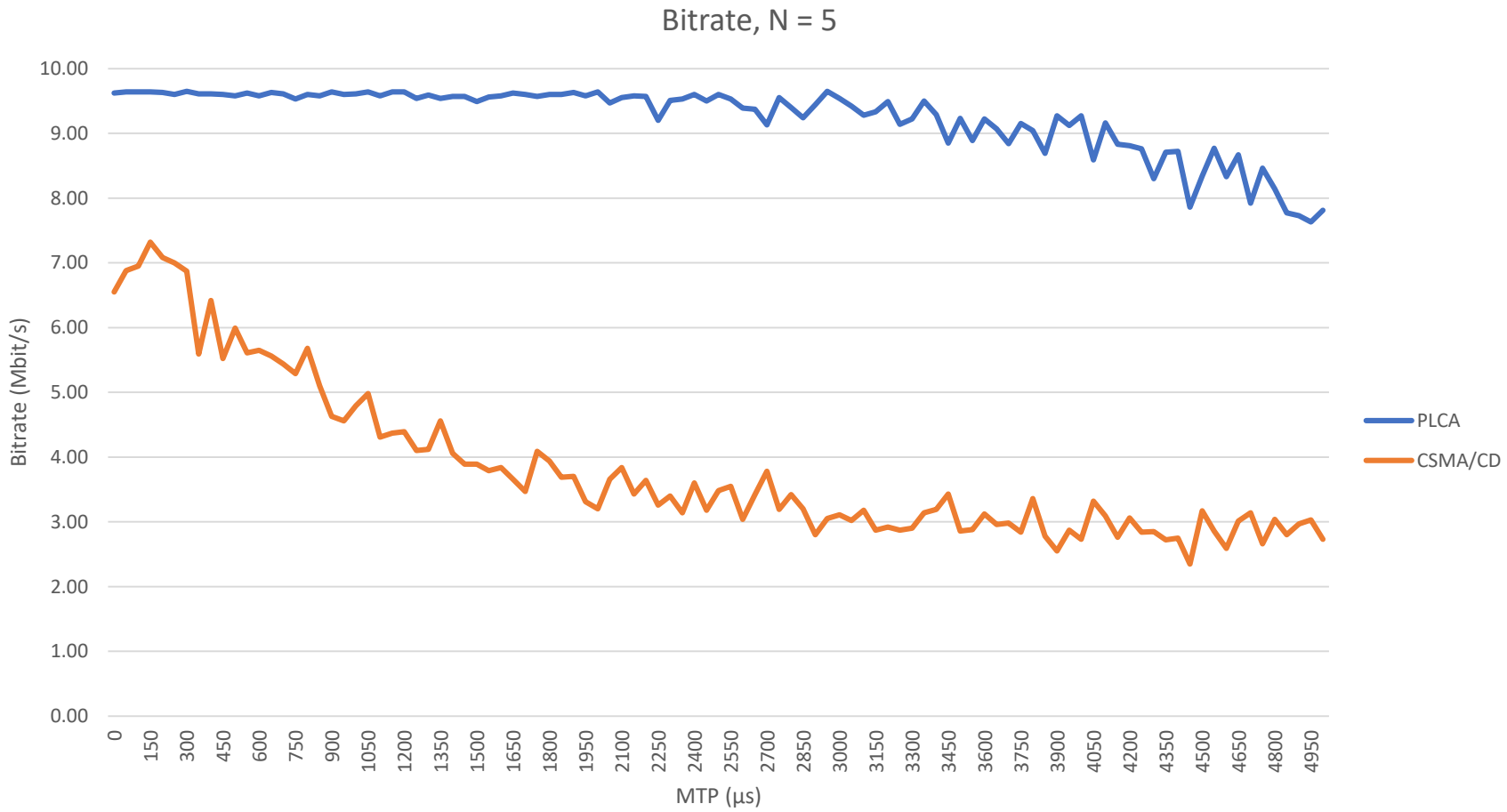
Throughput

- 100 pkts, random size (60, 1500), variable MTP and N (number of nodes)
- Comparison between 10base-T (simple CSMA/CD) and PLCA



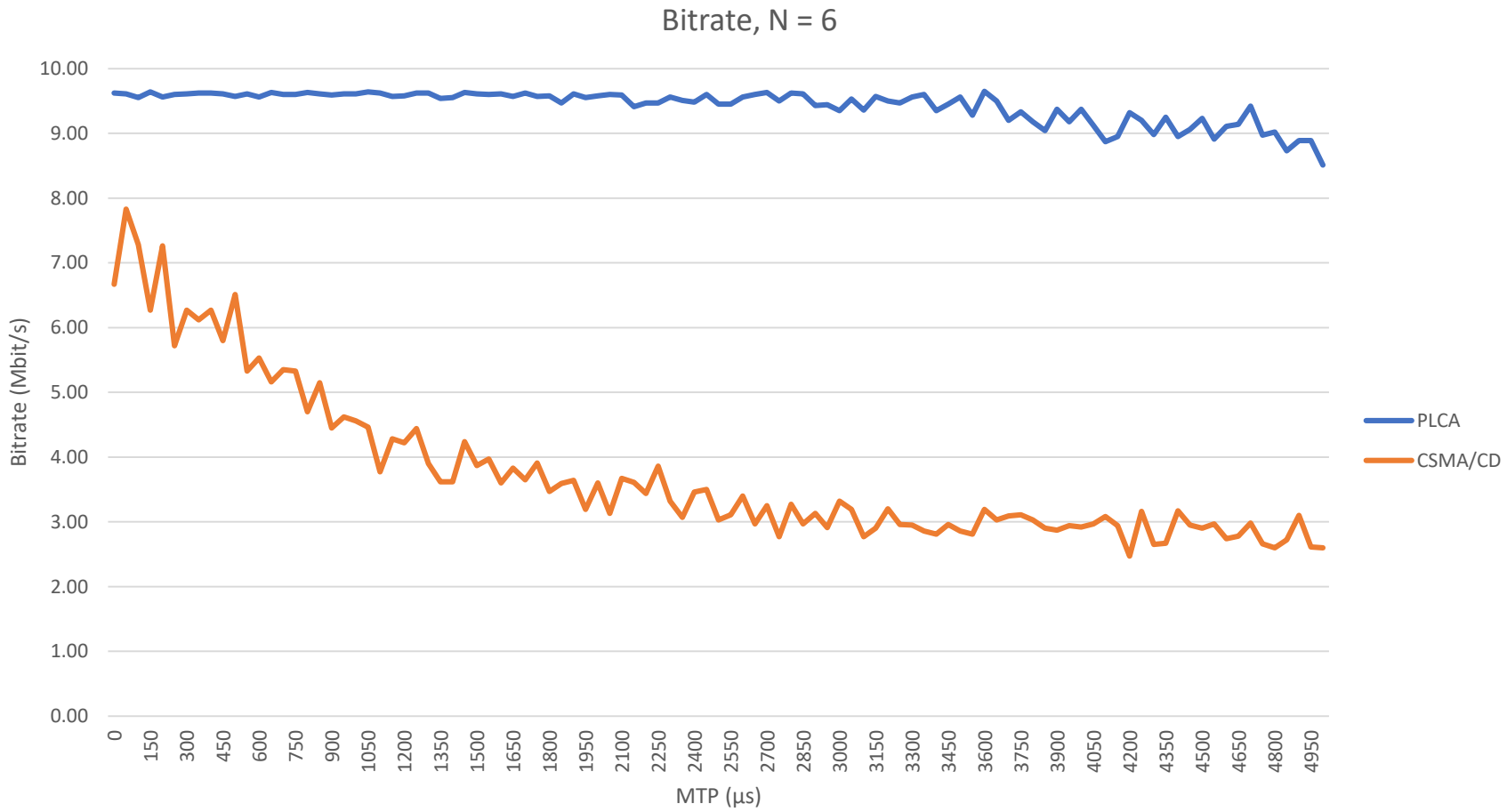
Throughput

- 100 pkts, random size (60, 1500), variable MTP and N (number of nodes)
- Comparison between 10base-T (simple CSMA/CD) and PLCA



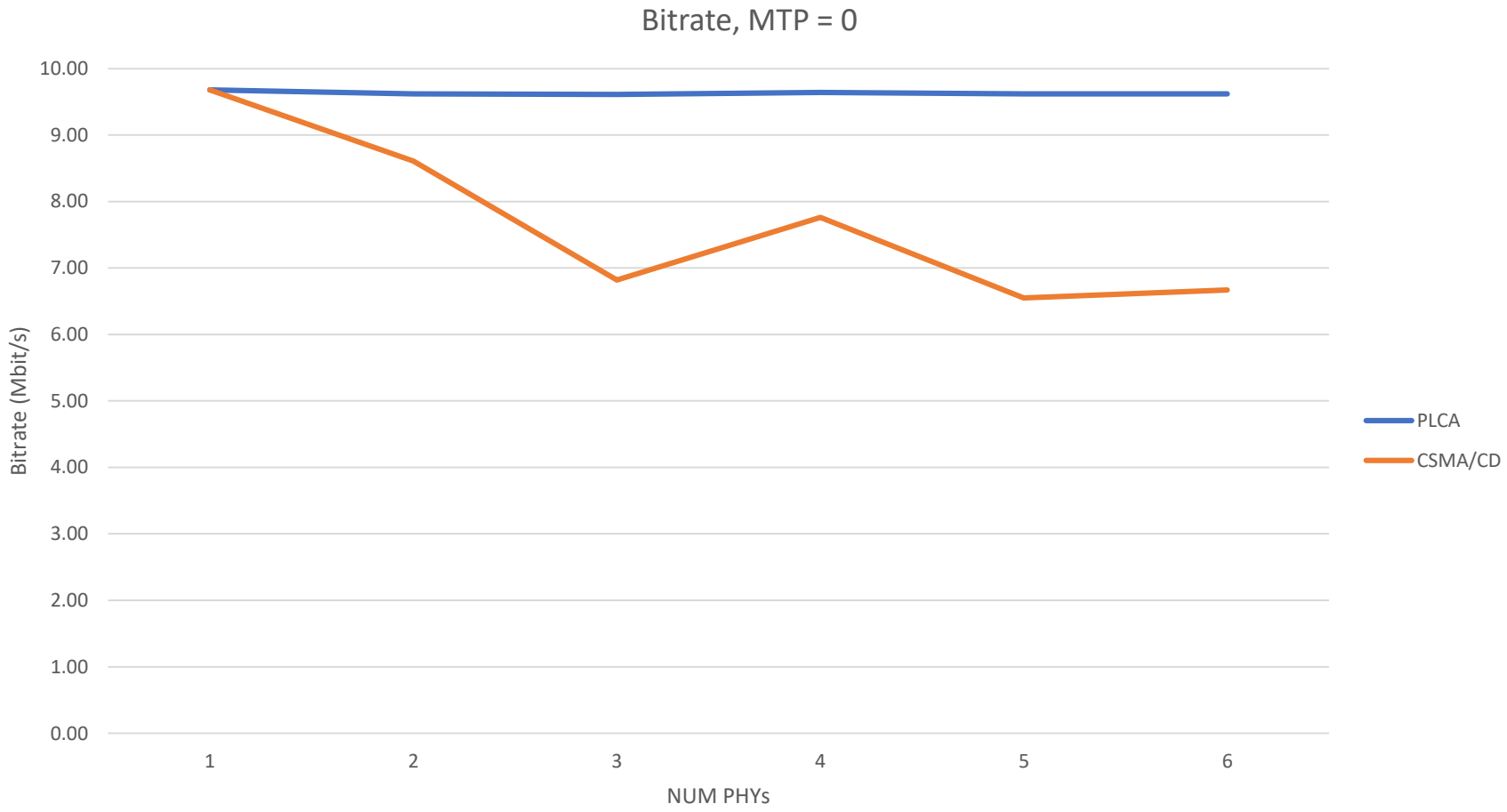
Throughput

- 100 pkts, random size (60, 1500), variable MTP and N (number of nodes)
- Comparison between 10base-T (simple CSMA/CD) and PLCA



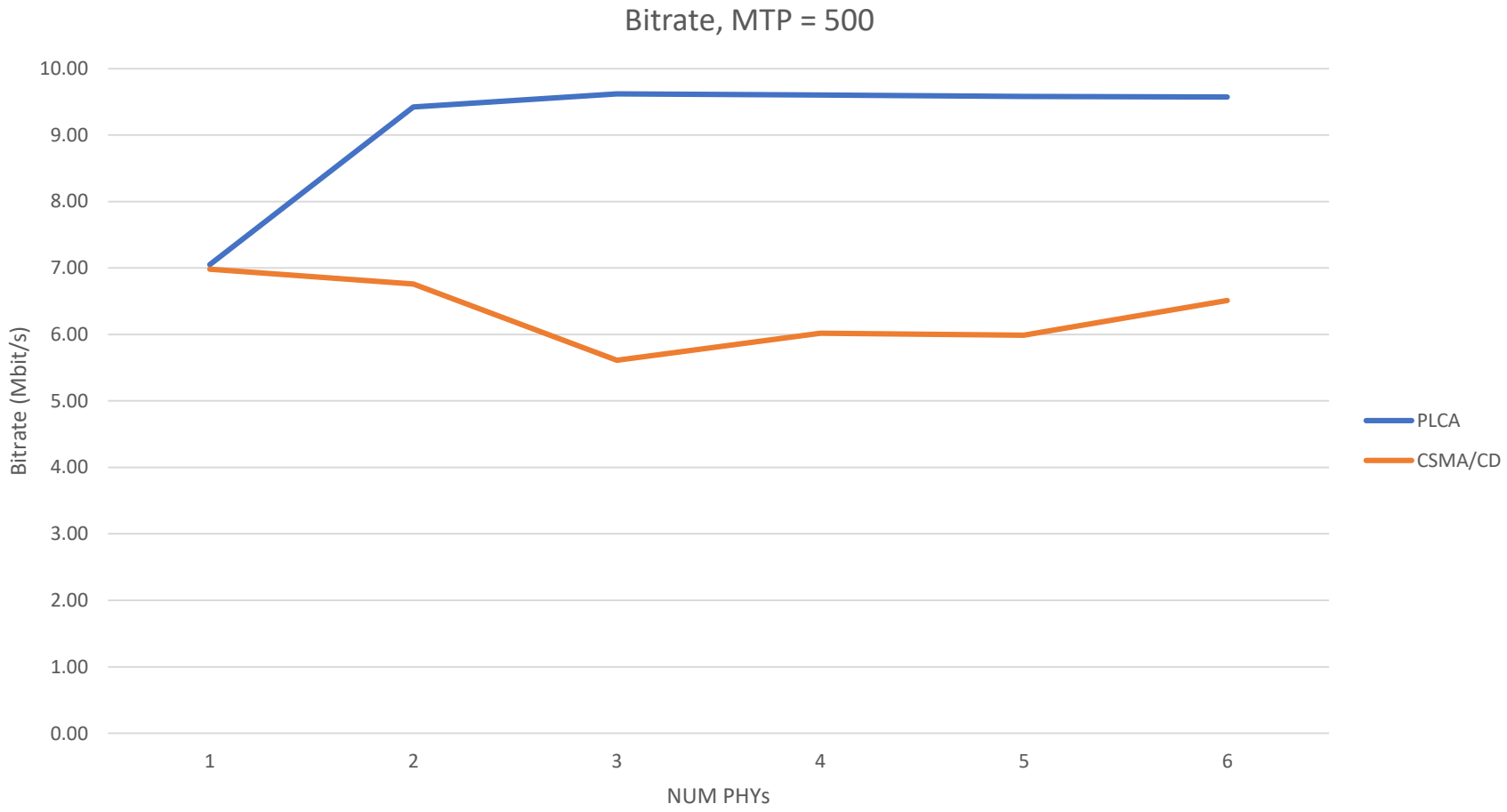
Throughput

- 100 pkts, random size (60, 1500), variable MTP and N (number of nodes)
- Comparison between 10base-T (simple CSMA/CD) and PLCA



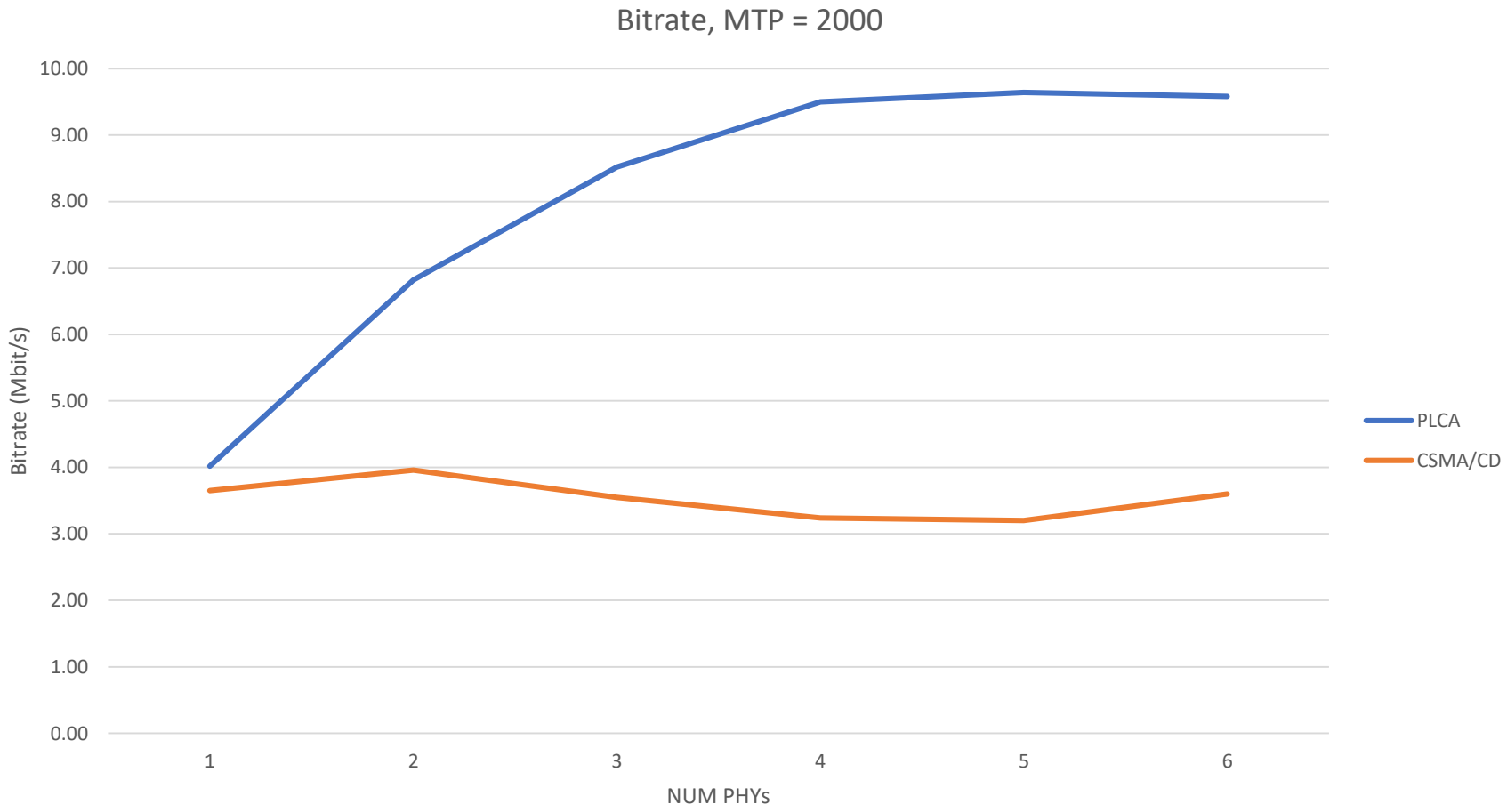
Throughput

- 100 pkts, random size (60, 1500), variable MTP and N (number of nodes)
- Comparison between 10base-T (simple CSMA/CD) and PLCA



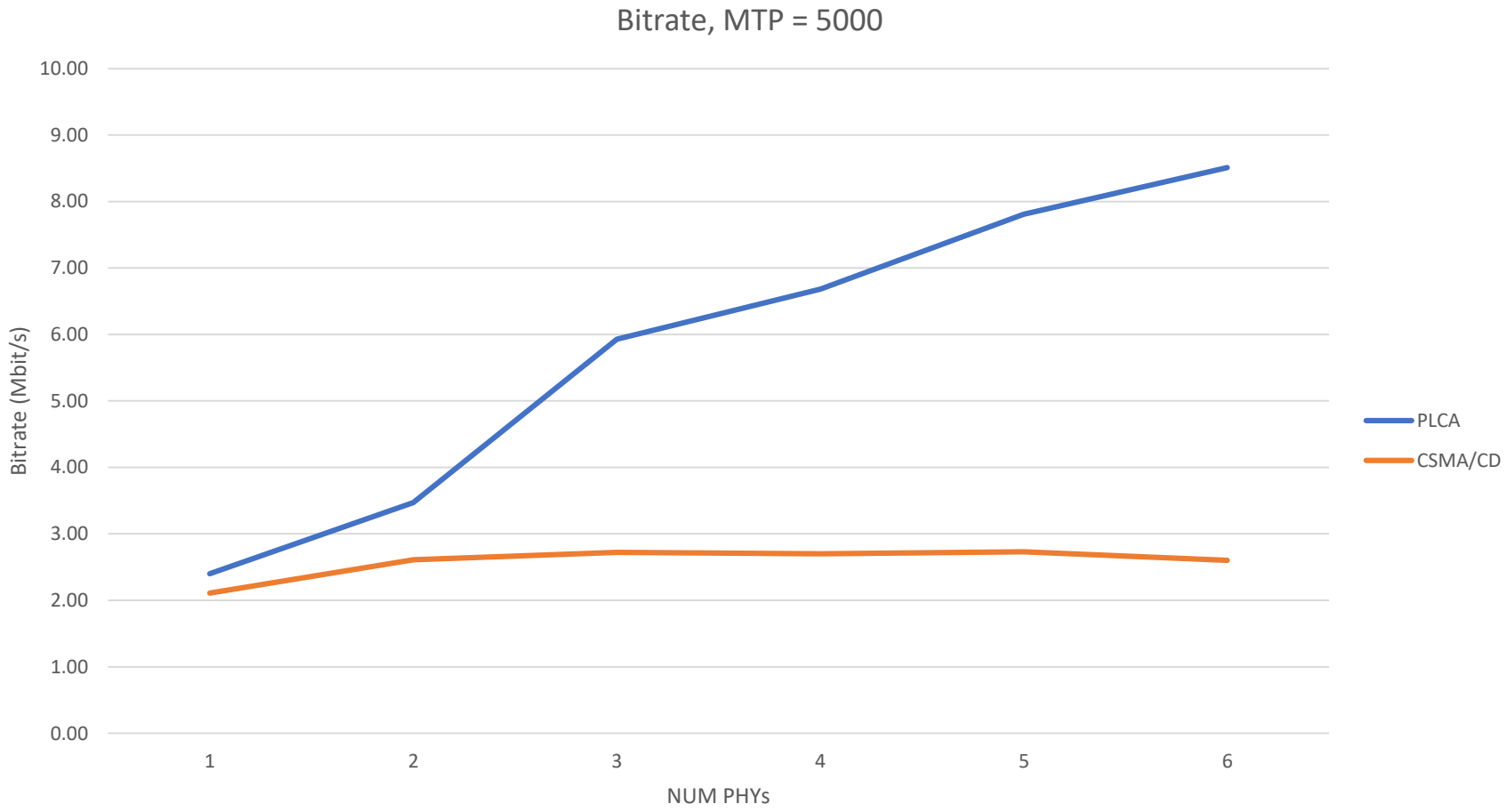
Throughput

- 100 pkts, random size (60, 1500), variable MTP and N (number of nodes)
- Comparison between 10base-T (simple CSMA/CD) and PLCA



Throughput

- 100 pkts, random size (60, 1500), variable MTP and N (number of nodes)
- Comparison between 10base-T (simple CSMA/CD) and PLCA



Latency

- 500 pkts, size = 60B, variable MTP, 6 nodes. Latencies in µs.
- Comparison between 10base-T (simple CSMA/CD) and PLCA

MTP	MAX_LAT	AVG_LAT	STDEV
0	57595.6	553.3	4826.0
500	59692.8	1034.2	4637.4
2000	29387.5	618.9	2298.2
5000	19645.4	264.0	1035.7

CSMA/CD

MTP	MAX_LAT	AVG_LAT	STDEV
0	443.4	441.1	26.2
500	546.4	186.4	90.7
2000	269.2	74.8	31.6
5000	223.7	64.0	17.8

PLCA

- Average latencies not so different at high loads
- Huge difference on MAX latency and STD deviation (due to collisions / backoff)
- PLCA is much more “deterministic”

Join/ Leave

- Still static configuration of the IDs
- How to handle "disabled" nodes?
  - COMMIT/YIELD is expected to be sent within a determined time (HS timer).
  - On timeout, next node will COMMIT/YIELD as appropriate
  - HS timer shall be long enough to accommodate for PHY latency but as short as possible not to degrade performance
  - Simulations performed with HS_TIMER = 20µs
    - Allows for ~16 µs of HS latency
    - Simulated with random missing nodes
      - Negligible impact on performance
- PHY with ID = 0 (master) use COMMIT/S and YIELD/S for handshaking.
  - Slave nodes synchronize on xxx/S signals before joining the HS mechanism
    - link_status shall not be signaled until first sync
  - Nodes receiving bad COMMIT/YIELD shall not transmit and re-synchronize
    - if this happens to the master, this one shall wait until there is no data on the line for a certain amount of time, then re-start transmitting HS.

Next Steps

- Simulations
  - More use-cases?
- Robustness
  - Simulate errored ACK/NACK handling
  - Corner cases study
- Optional
  - Auto-negotiation of IDs (instead of static config)?
    - not actively working on that
    - possible IDEA: keep a "virtual slot" at the end of the HS.
      - PHYs can send specific code + ID + CRC to get enumerated at that moment
      - monitor for collisions and backoff
  - On-the-fly election of new master for failover?

Thank You !