Deterministic Latency aspect of Modulo in Simulations

IEEE 802.3 September 2026 Interim Session

Lisbon, Portugal

Work by Steve Di Bella, General Motors
Presented by Yong Kim, General Motors

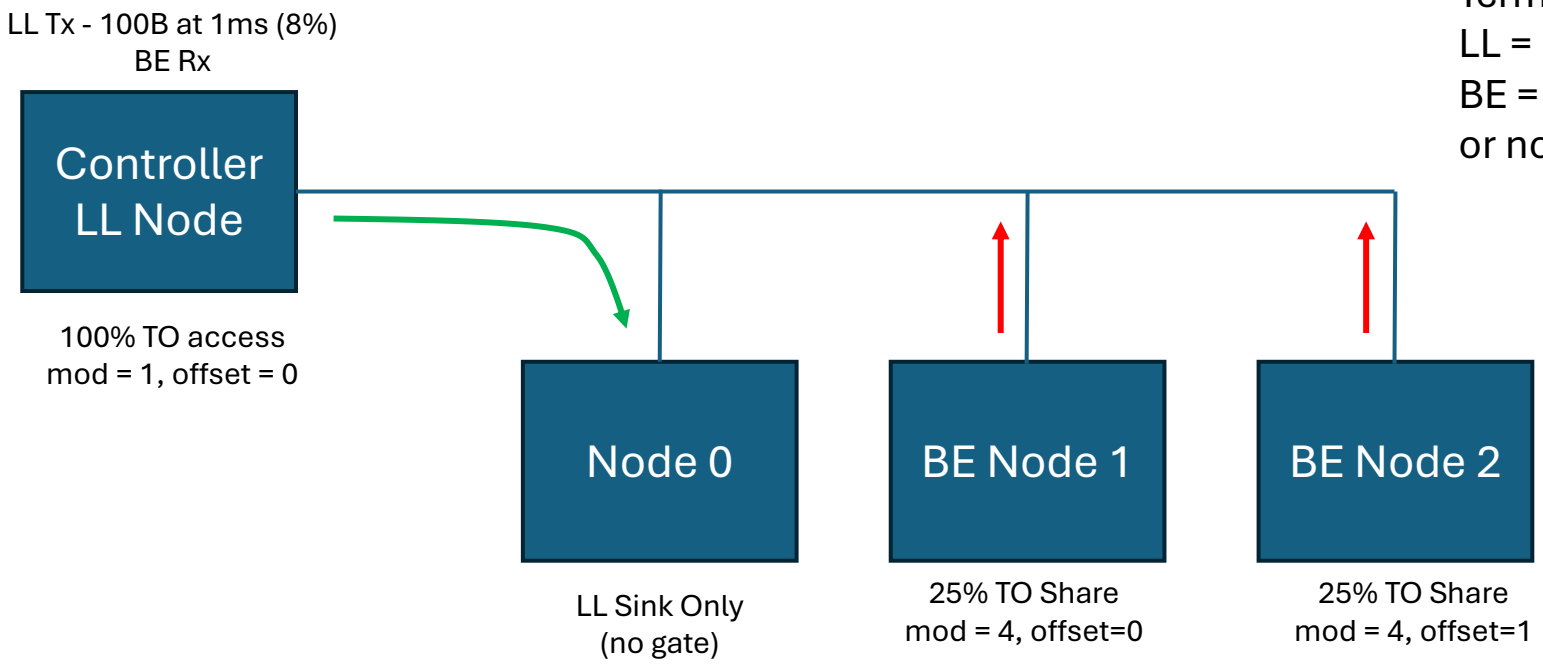
Background & Objectives

- Fair and equal access in PLCA is well understood – ”round robin”. This is the only scheme that provides deterministic latency to all nodes. Unfortunately the worst case latency is too long.
- Relative Priority Schemes are well-understood.
  e.g. multiple nodeID – may be used for “weighted round robin”
- Strict Priority Scheme is well-understood. It may be expressed using “precedence mode – terminate PLCA cycle as soon as one frame transmit occurs in a PLCA cycle”. Per CoS queue, or per message queue may be supported using multiple nodeIDs.

- Base standards of AVB/TSN strived for “deterministic” latency and latency jitter to help support requirements of time sensitive network applications.

- In this presentation, “modulo” scheme simulations are shown, to help examine whether deterministic and a more tightly bounded latency could be achieved. (see modulo technical feasibility summary here: https://www.ieee802.org/3/EPLCA/public/260827/jochim_eplca_01_260827.pdf)

A Simple Baseline - LL Protected from BE Load



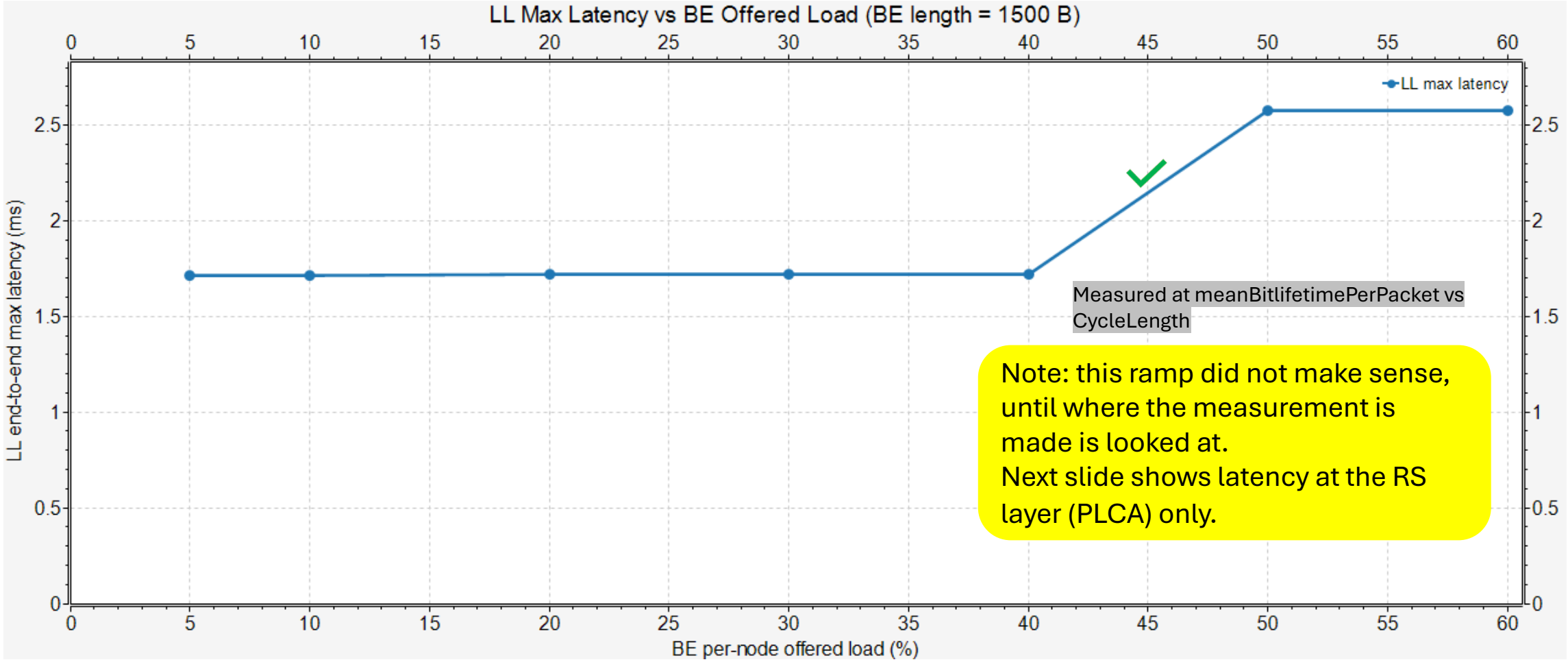
Terminology.
LL = Low Latency – time sensitive controls.
BE = Best Effort – less time sensitive controls, or non-time sensitive data.

- Controller Node (LL) sends to Node 0 which only acts as a sink.
- 1500B frames are sent from each BE node at the rates shown in the table

BE Interval	per-node BE	combined BE	+ LL 8 %
24 ms	5%	10%	18%
12 ms	10%	20%	28%
6 ms	20%	40%	48%
4 ms	30%	60%	68%
3 ms	40%	80%	88%
2.4 ms	50%	100%	108%
2 ms	60%	120%	128%

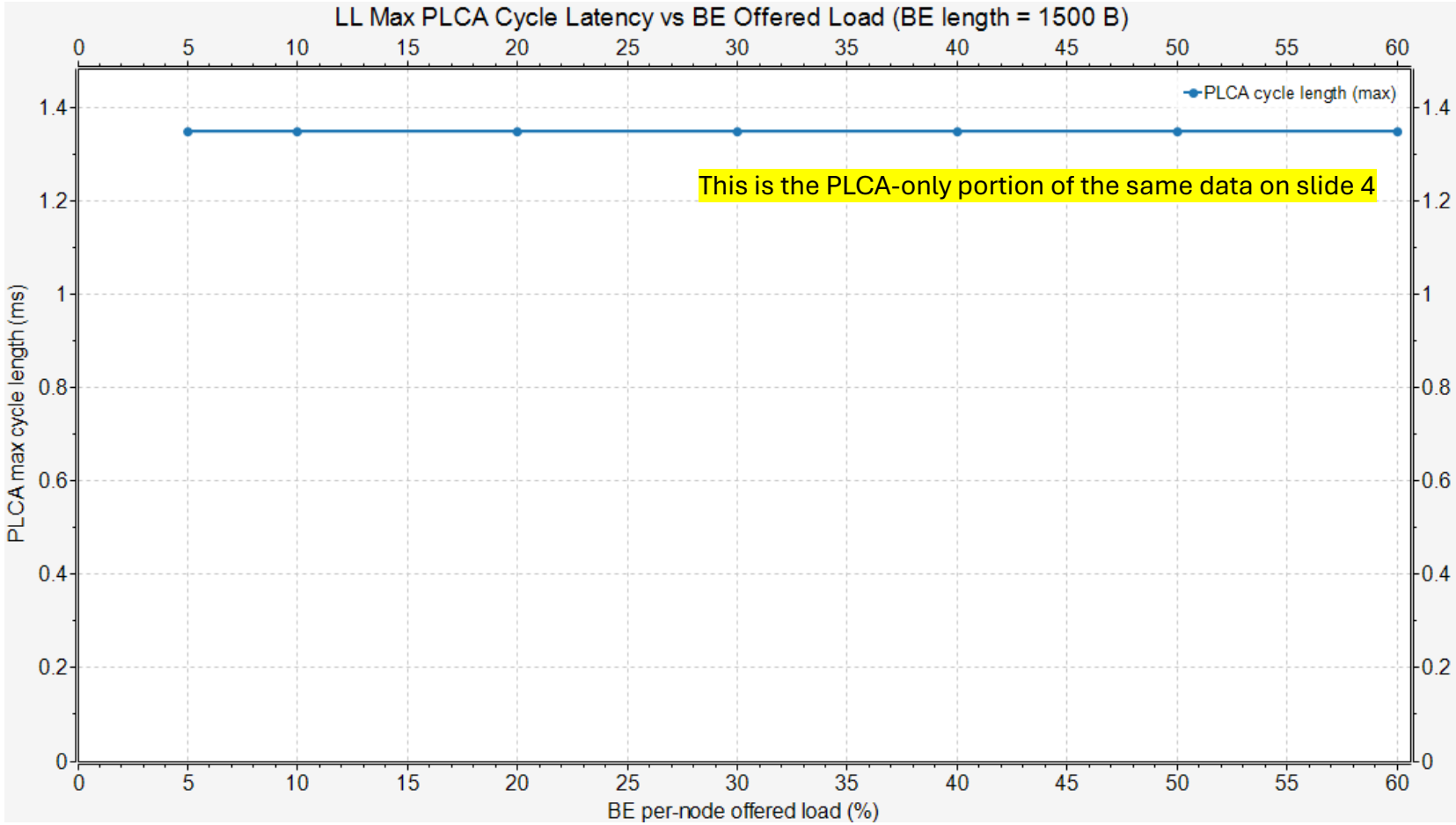
Boring Chart – 100 Byte LL Max Latency is Fixed.

Interfering BE Frame Length= 1500 Byte



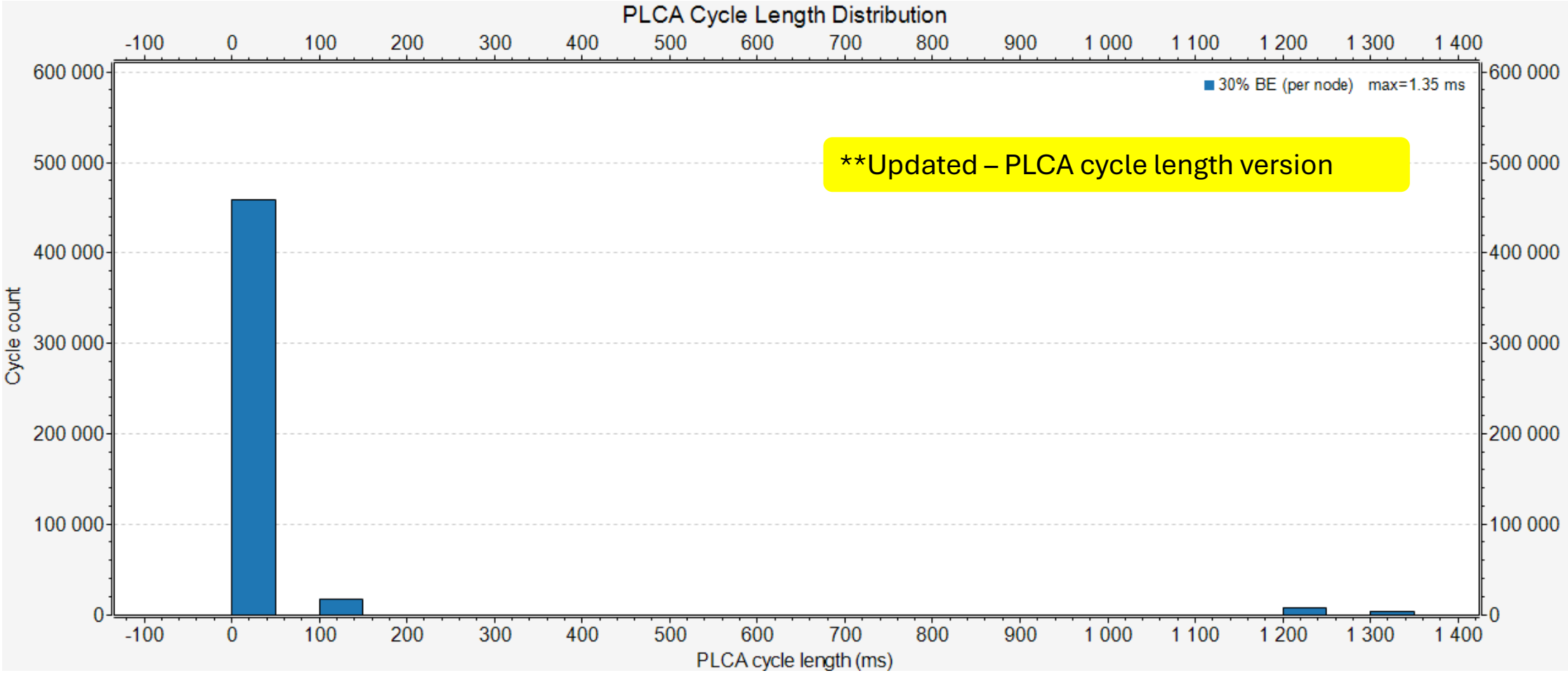
Boring Chart – 100 Byte LL Max Latency is Fixed.

Interfering BE Frame Length= 1500 Byte. PLCA Latency Only



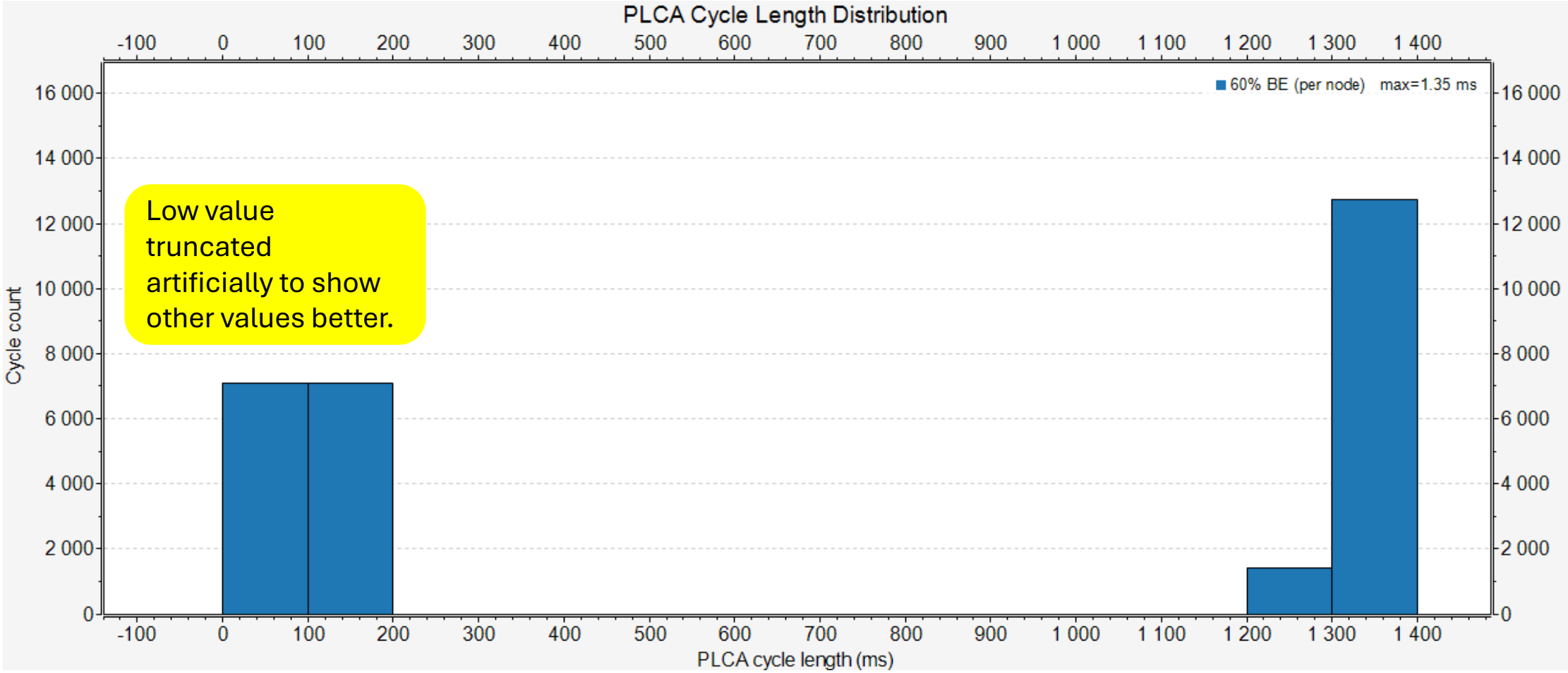
Worst Case Latency –

100 Byte LL Max Latency is Fixed. Interfering BE= 1500 Byte, fixed.



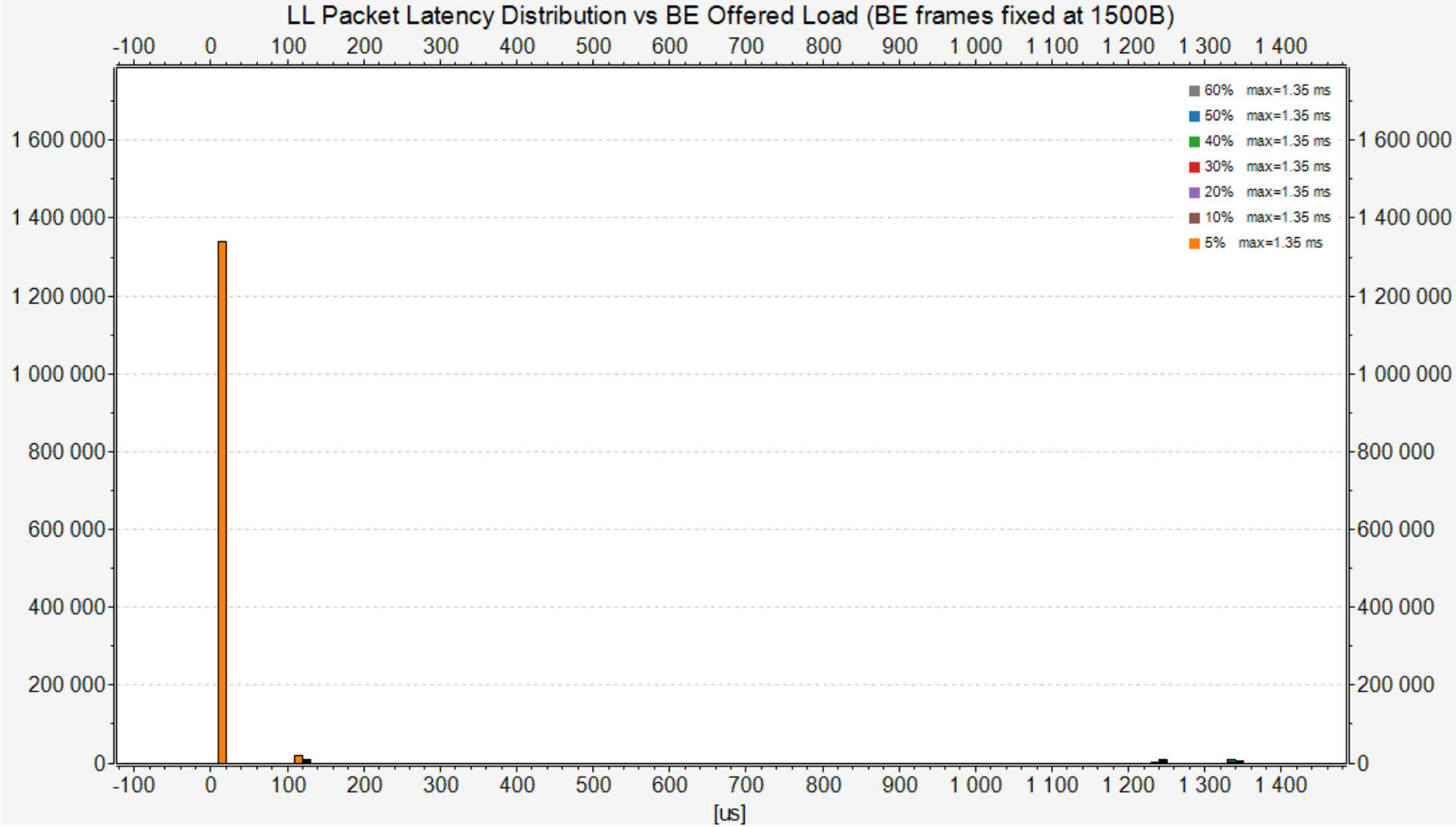
Example distribution for 30% - number of occurrences made plotting all on the same axis hard to read (see slide 8).

Example distribution for 60 % (saturation) – the same data set as slides 5 &6

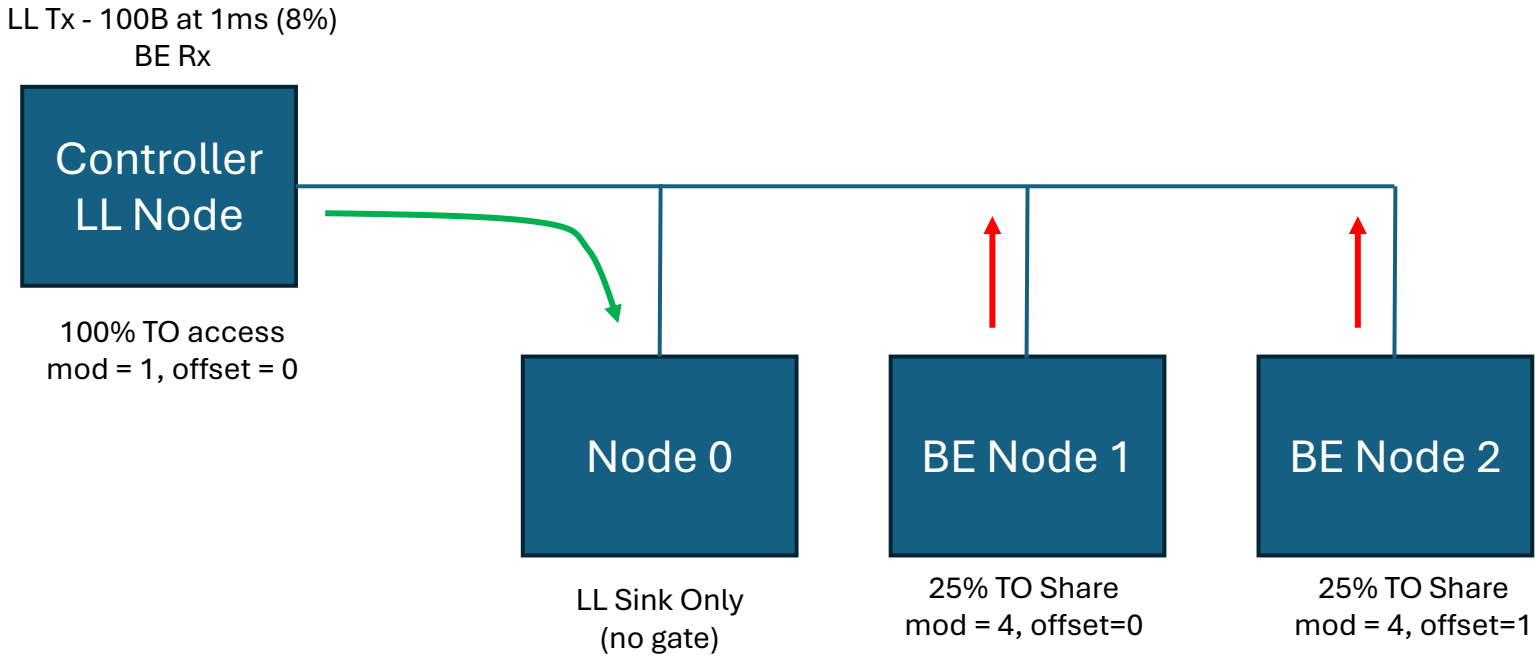


Example distribution for 60 % (saturation) – same data set as slides 5 & 6. number of occurrences made plotting all on the same axis hard to read (see slide 8).

LL Latency Distribution @ RS Layer (PLCA only)



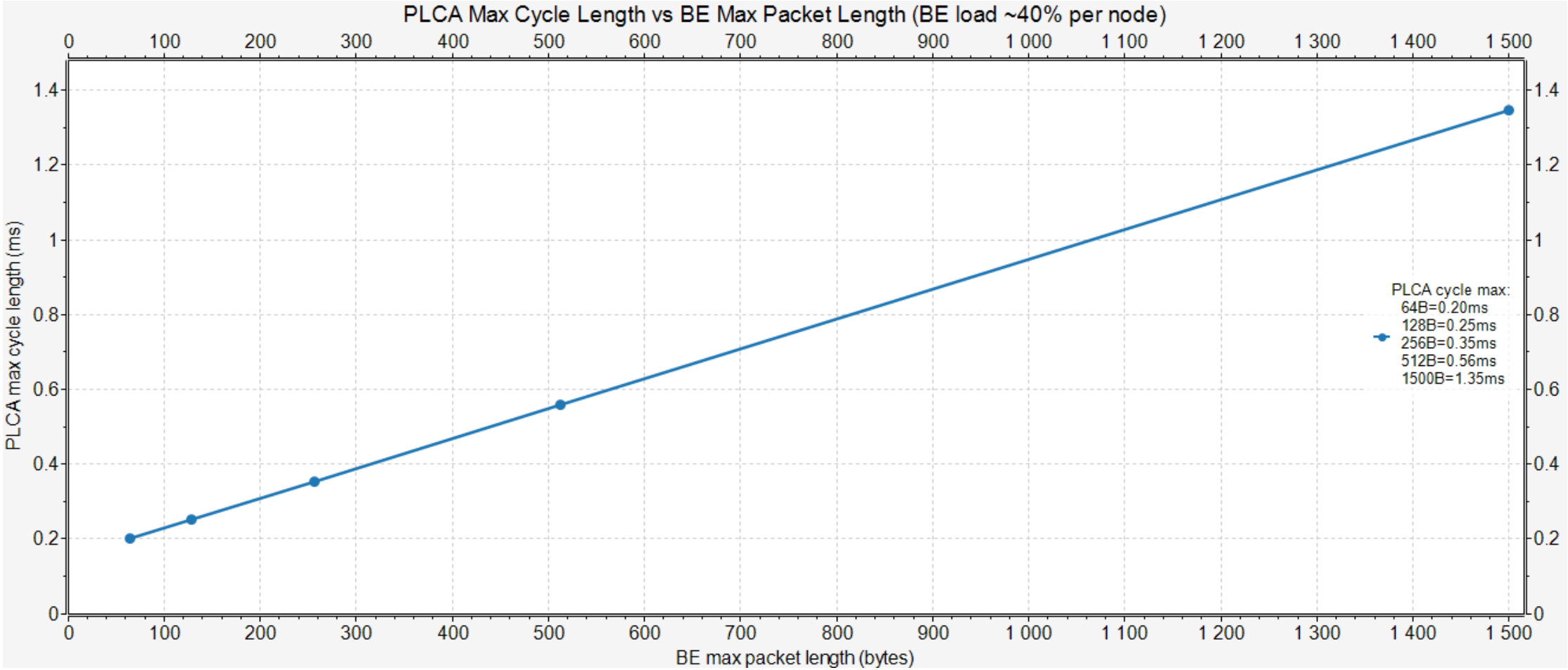
The Baseline - LL with random BE length interferences
(LL Depends Only on BE Length)



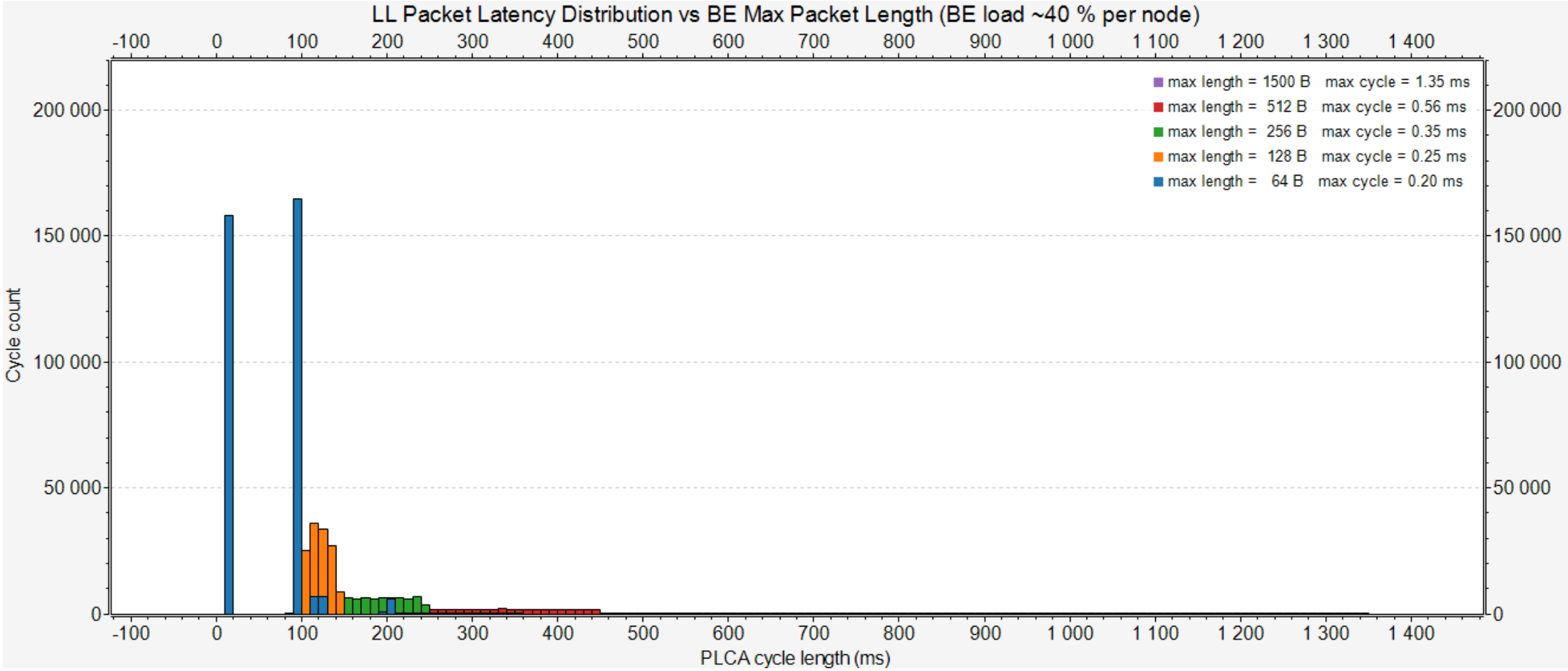
- Controller Node (LL) sends to Node 0 which only acts as a sink.
- Each BE node sends frames of random length intuniform(64 B, beMaxLen)
- BE periodic rate scales with mean length to keep BW ~40% per BE node

beMaxLen	mean len	BE Interval	per-node BE	combined BE	+ LL 8 %
64 B	64 B	128 µs	40%	80%	88%
128 B	96 B	192 µs	40%	80%	88%
256 B	160 B	320 µs	40%	80%	88%
512 B	288 B	576 µs	40%	80%	88%
1500 B	782 B	1564 µs	40%	80%	88%

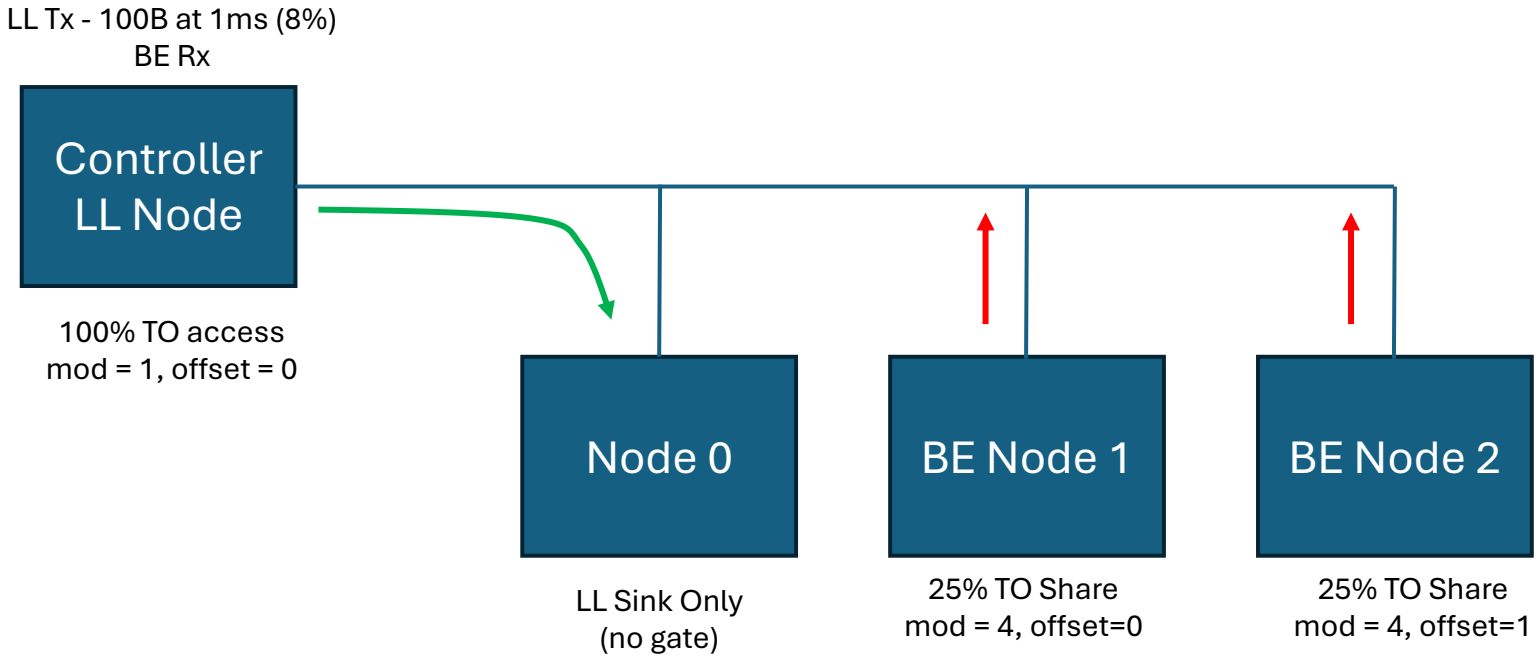
LL Max Latency, when BE length= RAND (64..length)



LL Latency Distribution (= PLCA Cycle Duration)



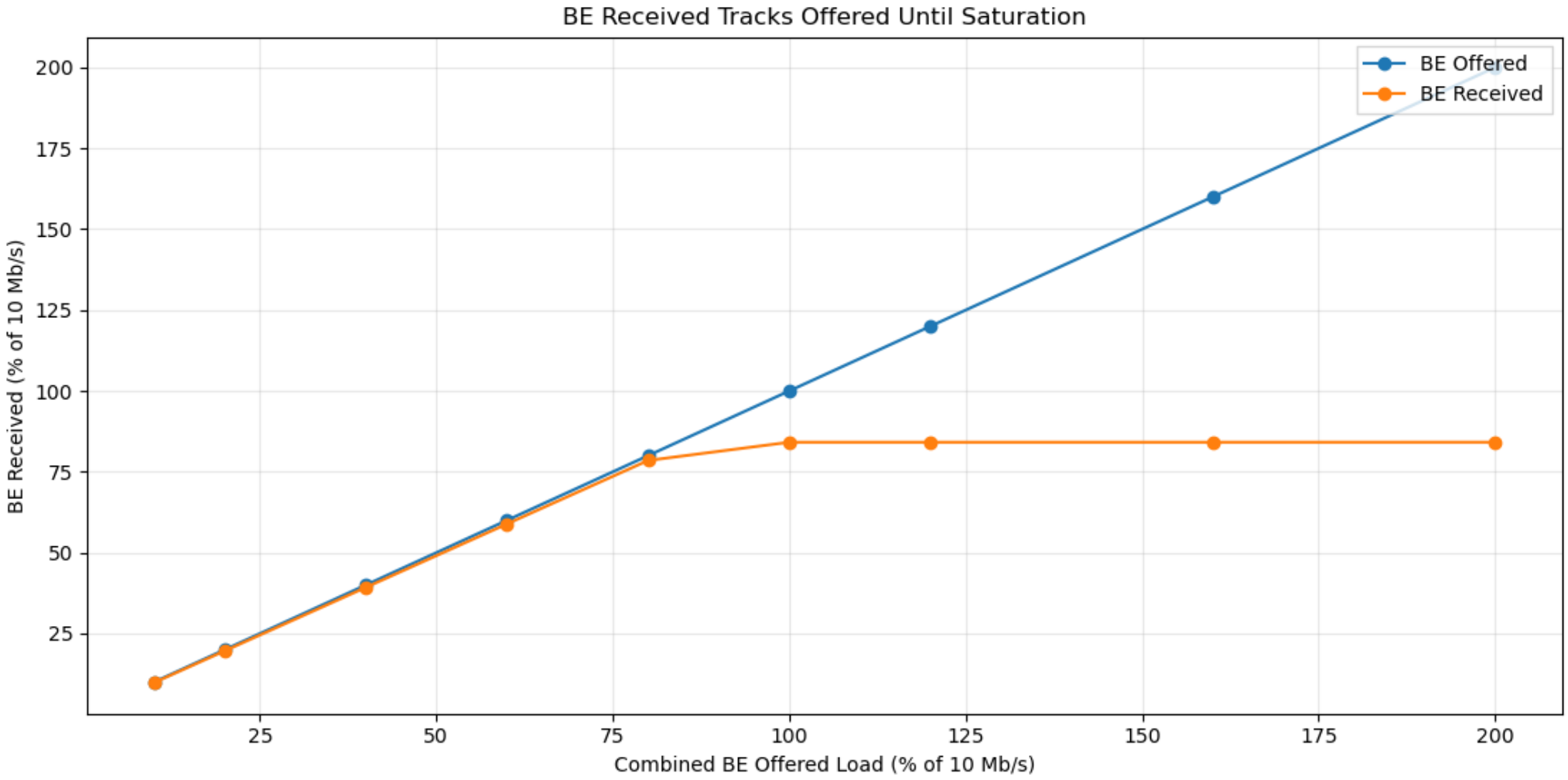
The Baseline – In previous case, BE latencies
(BE Not Starved by LL)



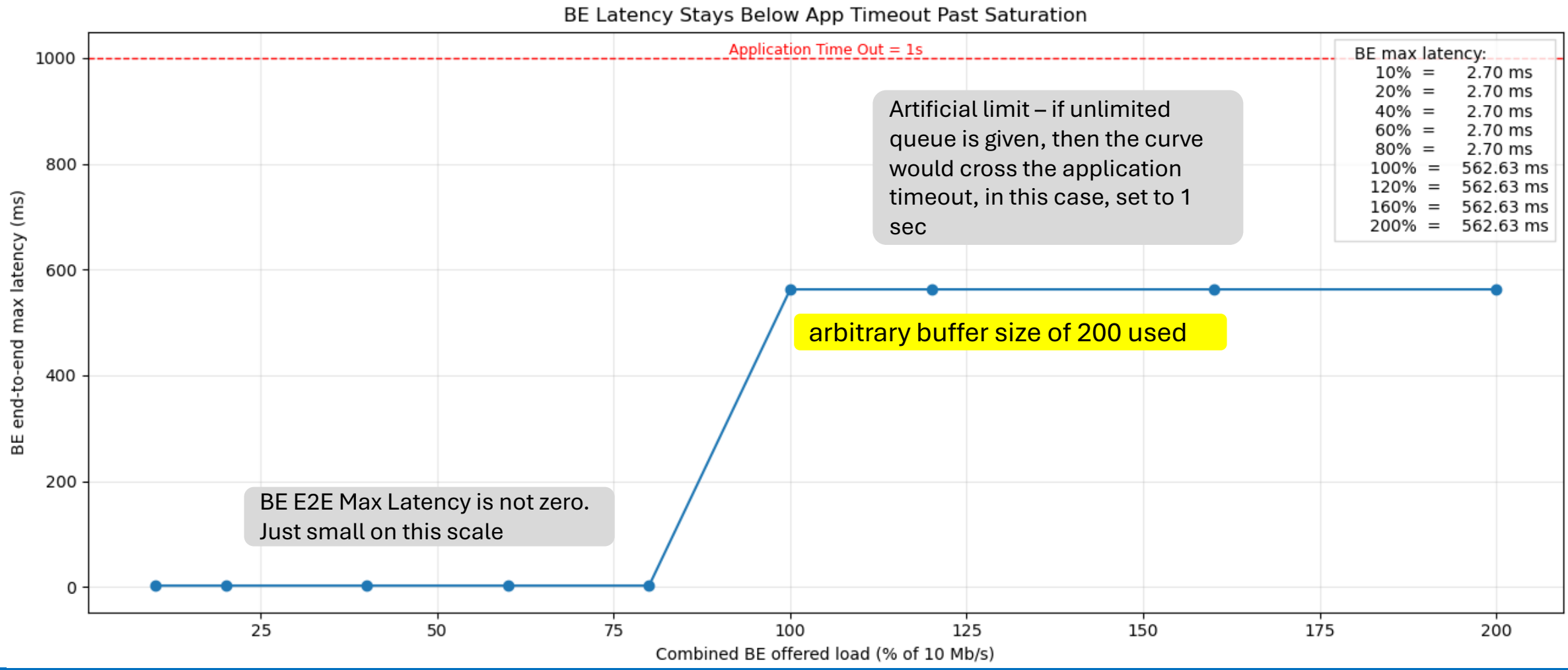
BE Interval	per-node BE	combined BE	+ LL 8 %
24 ms	5%	10%	18%
12 ms	10%	20%	28%
6 ms	20%	40%	48%
4 ms	30%	60%	68%
3 ms	40%	80%	88%
2.4 ms	**50%**	**100%**	**108%**
2 ms	**60%**	**120%**	**128%**
1.5 ms	**80%**	**160%**	**168%**
1.2 ms	**100%**	**200%**	**208%**

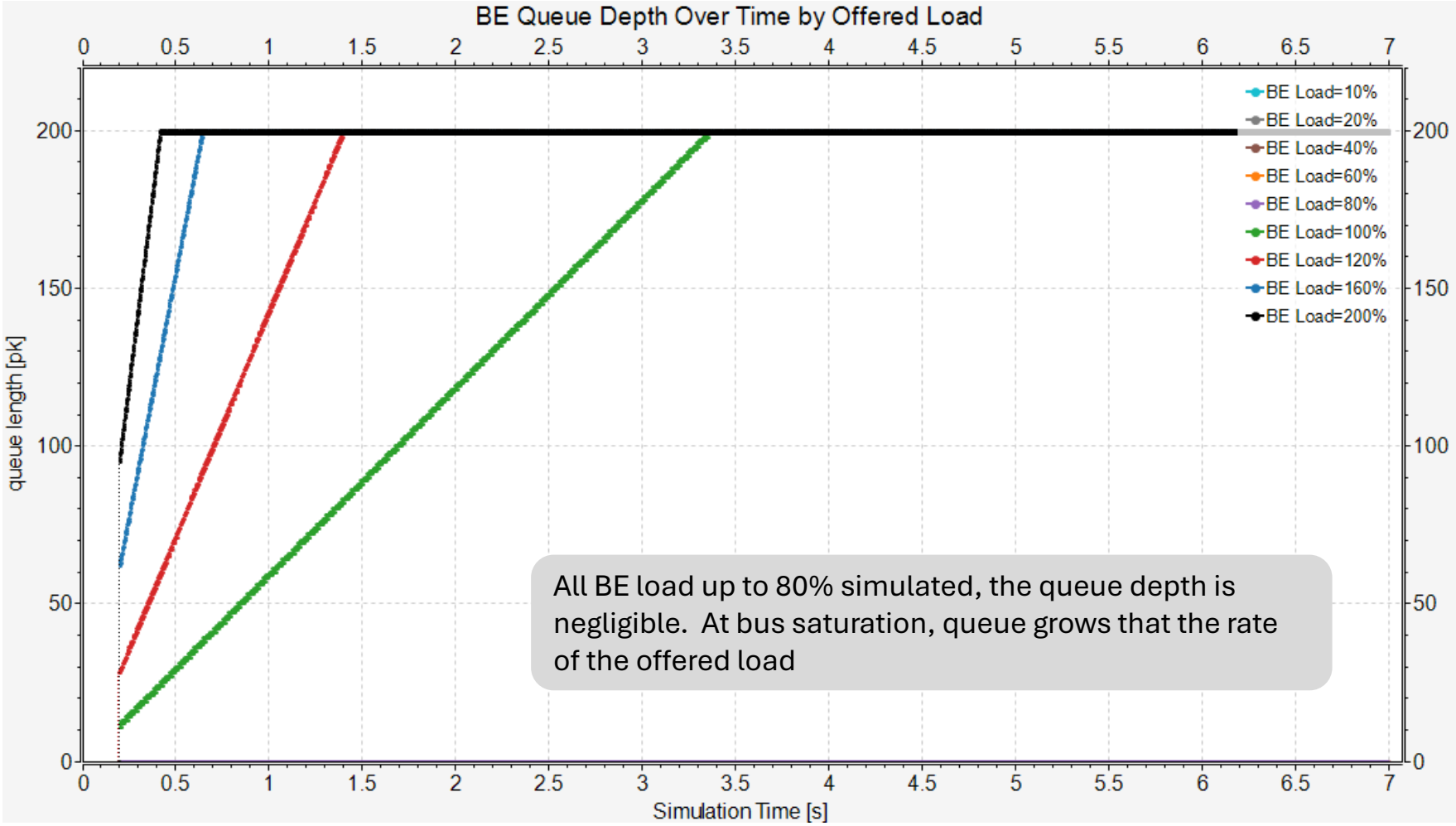
- Controller Node (LL) sends to Node 0 which only acts as a sink.
- 1500B frames are sent from each BE node at the rates shown in the table
- The sweep extends **past saturation** so that BE received plateaus at the schedule's capacity.

BE Offered Load – no starvation >> 75% BW

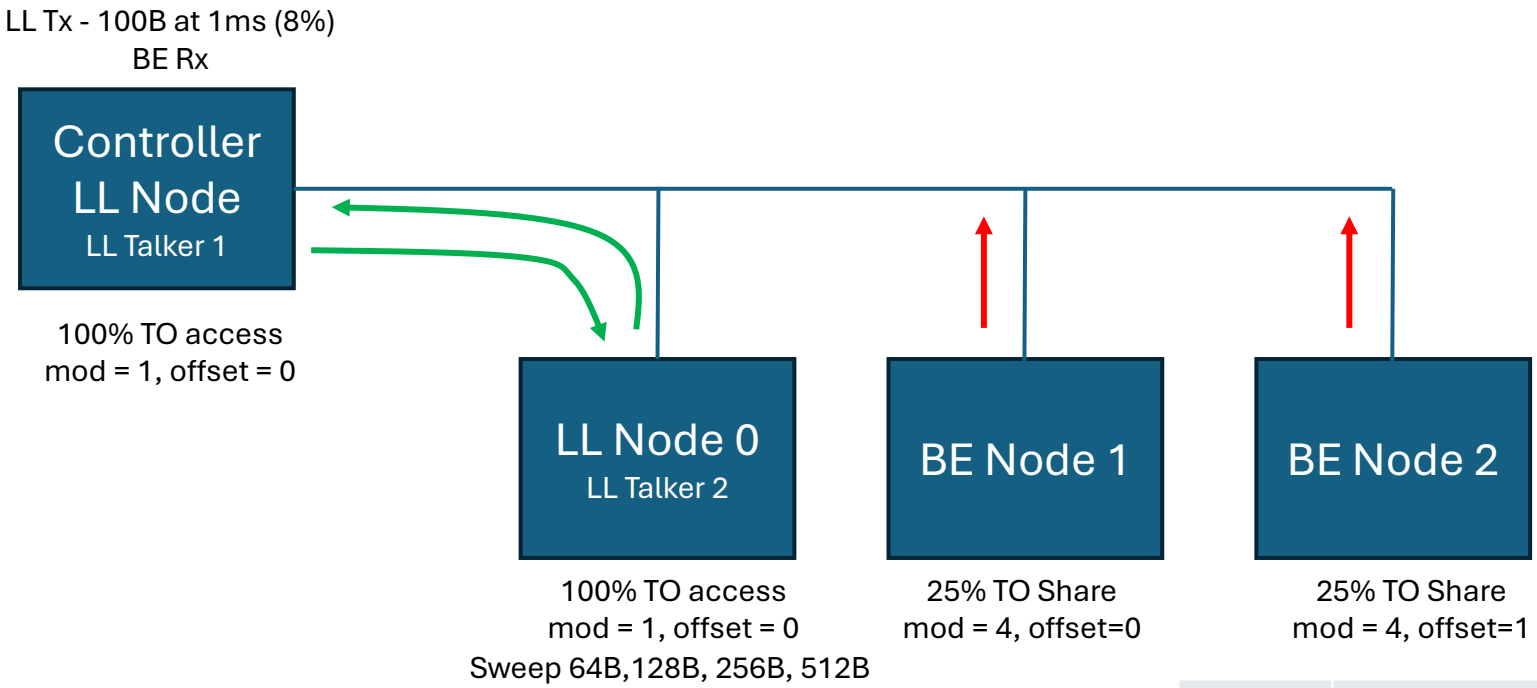


BE E2E Max Latency




BE Queue Depth Over Time by Offered Load
queue length [pk]
Simulation Time [s]
BE Load=10%
BE Load=20%
BE Load=40%
BE Load=60%
BE Load=80%
BE Load=100%
BE Load=120%
BE Load=160%
BE Load=200%
All BE load up to 80% simulated, the queue depth is negligible. At bus saturation, queue grows that the rate of the offered load

Two LL Talkers, Each with respective latency contract.



- Controller Node (LL 1) sends to Node 0
- Node 0 (LL2) sends to Node 1
- BE Traffic held constant at 40% - rand(64, 1500)

LL2 Size	LL2 Load	+ LL1 (8 %)	+ 2× BE (80 %)	Total offered
off (silent)	0%	8%	88%	88%
100 B	8%	16%	96%	96%
250 B	20%	28%	108%	108%
500 B	40%	48%	128%	128%
1000 B	80%	88%	168%	168%
1500 B	120%	128%	208%	208%

LL1 Latency influenced by LL2 Length only
LL2 Latency scales with frame size.