Power saving mechanisms in 1904.4

Glen Kramer, Broadcom

Problem statement

- In 1904.1, the power saving mechanism is based on sleep-wake cycles.
- Not widely used in deployed networks. The question is why?
  - Not enough improvement?
  - Detrimental impact on services?
  - Other reasons?

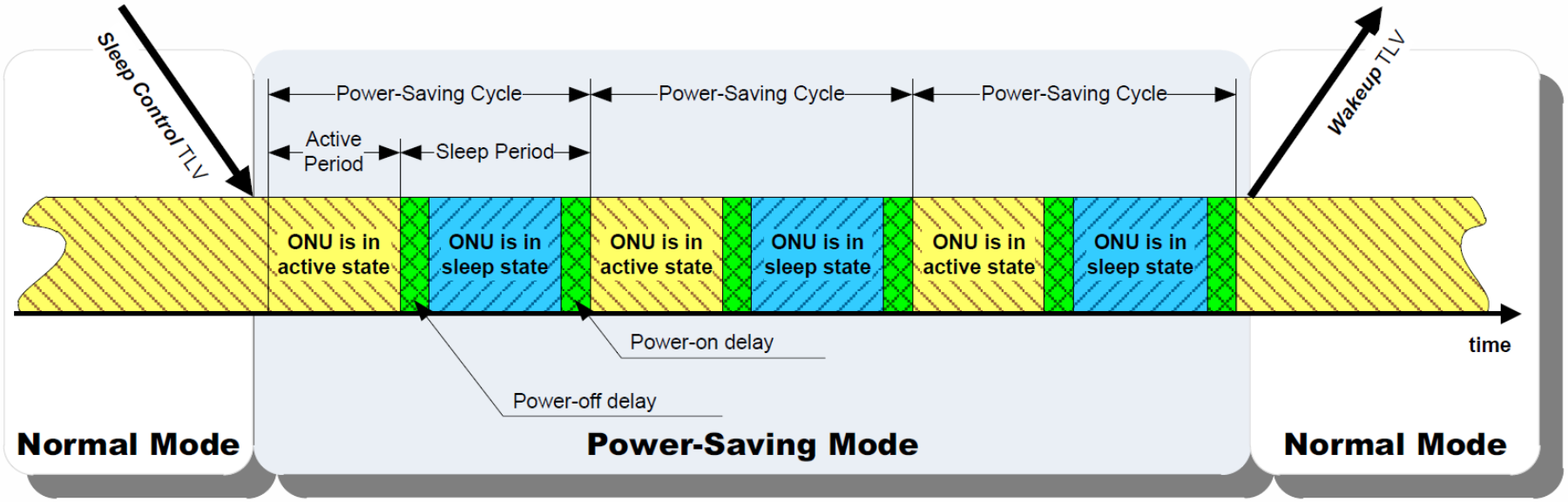


Figure 10-1—Timing diagram of the power-saving mode

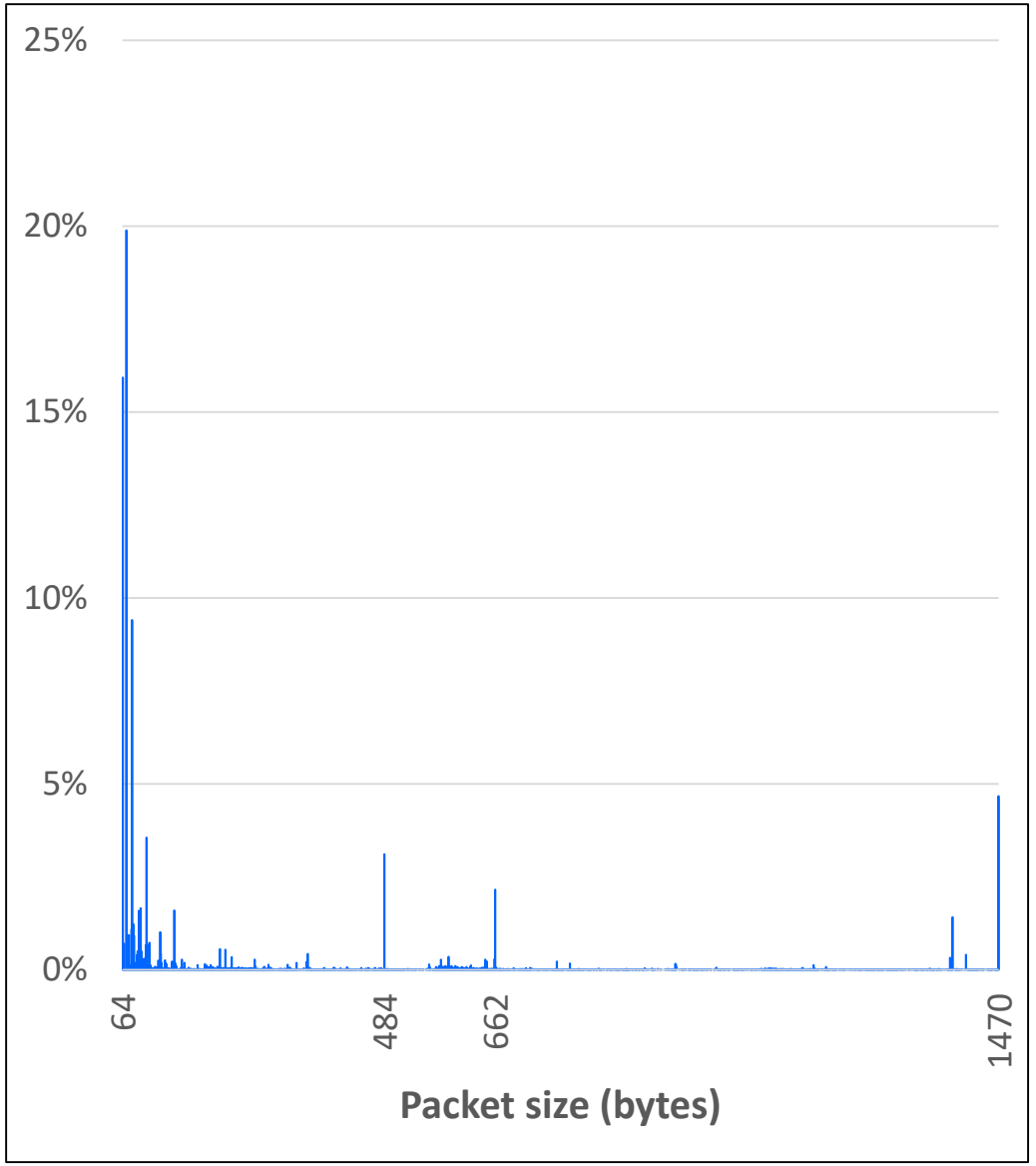
Traffic study

- Marek captured upstream traffic at local firewall
- Only traffic with external destination is captured
- Capture time from 9PM to 9AM
- Data captured on 1Gbps link

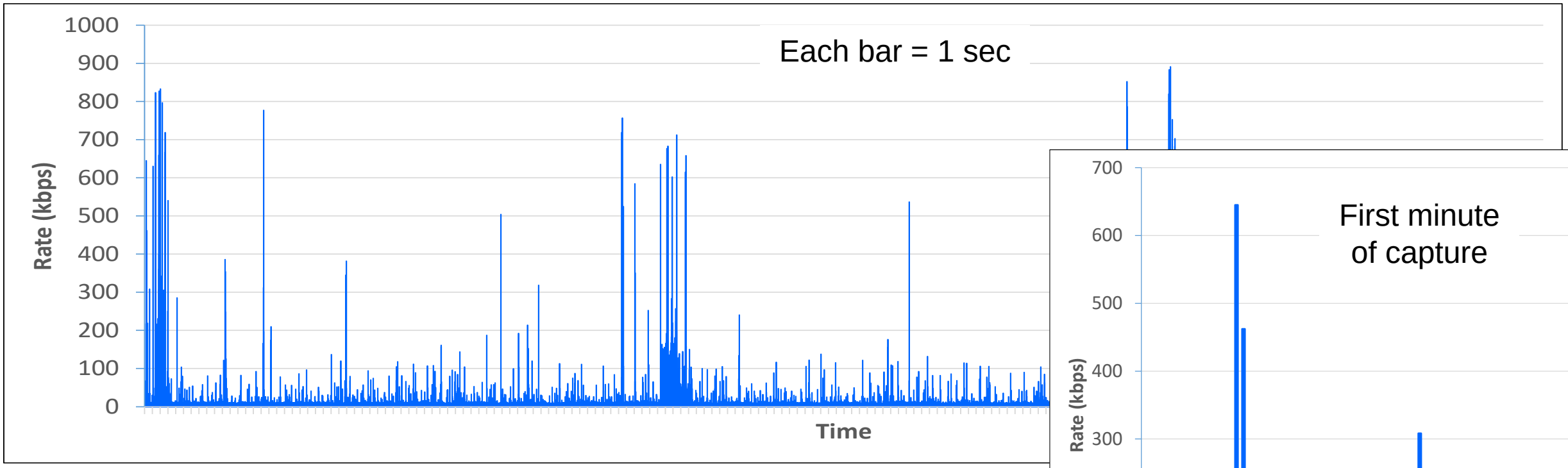
- For analysis, this presentation looked only at 3 hours of capture (1AM – 4AM)

What the captured traffic looks like?

- 68110 frames
- Frame sizes from 64 to 1470 bytes
  - Not all frame sizes were observed
- Multi-modal frame size distribution
  - Major modes:
    - 64B – 16%
    - 70B – 20%
    - 79B – 9.4%
    - 102B – 3.6%
    - 484B – 3.1%
    - 662B – 2.2%
    - 1470B – 4.7%
  - These 7 sizes account for 59% of all packets.

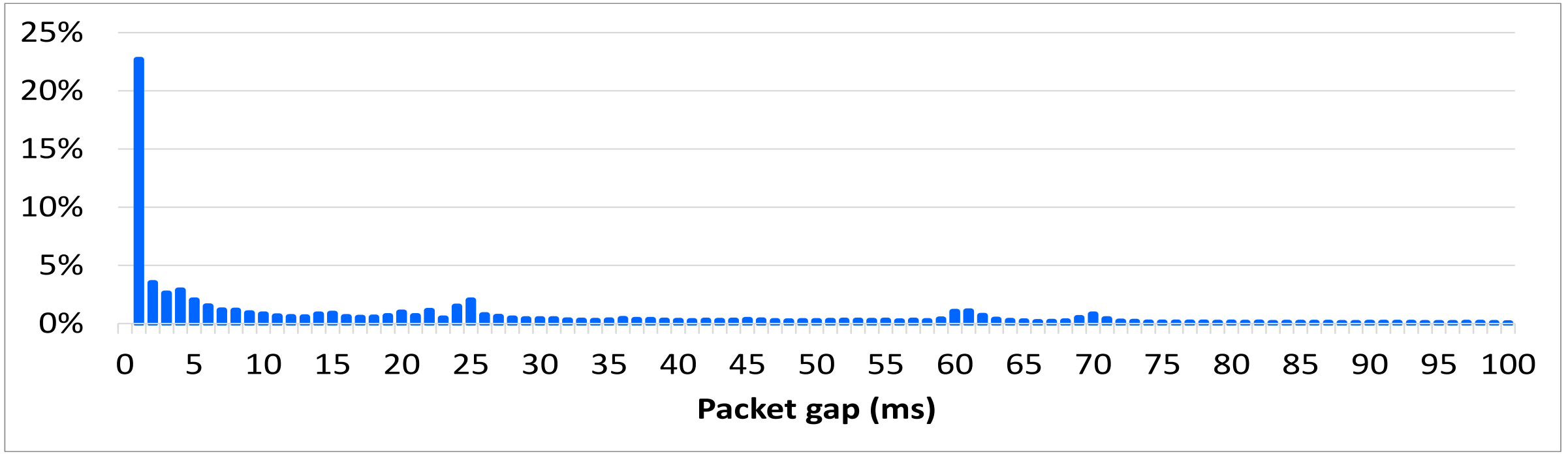


Temporal profile



- Fairly uniformly-spread packet arrivals with periods of higher activity and periods of no activity
- Average data rate = 15 kbps
- Maximum rate in 1 sec window = 890 Mbps

Inter-packet gaps



- 483 gaps = 96ns (12B), i.e., back-to-back packets
- 22.7% of IPGs are under 1 ms
- 39% are under 10 ms
- 60% are under 50 ms
- 71% are under 100 ms

- Average IPG = 158 ms
- Maximum observed gap = 4.15 sec.
- 3 gaps > 4 sec
- 39 gaps > 3 sec
- 420 gaps > 2 sec
- 2484 gaps > 1 sec

Uplink model

- During the 3-hour observation window, the channel was busy for a total of 170 ms.
- The packets were spread in time, such that in 25G-EPON, a burst would most often contain only a single packet
- Each burst would incur extra overhead
  - PHY overhead: LaserOn, LaserOff, SyncTime
  - FEC overhead: Parity = 257b x 10 ≈ 100 ns.
  - REPORT MPCPDU (optional)

“Fixed Grant Interval” simulation setup

- Model a fixed grant interval.
- Each burst contains a REPORT and maybe one of few frames of data.
- Analyze overall utilization (and potential for shutting down the channel between transmissions)
- Larger cycle time is better for power saving, but increases the frame latency. Where is the balance?

- **Parameters**
  - LaserOn = 64 ns
  - LaserOff = 64 ns
  - SyncTime = 128 ns
  - Grant interval (cycle) = 20 ms

Observations

- 92% of bursts are 386 ns long and only contain one REPORT each
- 99.5% of bursts are under 1 us.
- Of the bursts that contained any data packets, 76.1% contained only one data packet and 13.7% contained two packets.
- Maximum observed burst was 23.9 us and it contained 84 70-byte packets

- In the upstream, ONU adds a lot of extra overhead (burst mode, FEC, REPORTs)
- The total channel busy time has increased from 170 ms at 1 Gb/s ONU UNI to 215 ms at 25G ONU's PON port
  - 32x increase in the number of transmitted bits (incl. burst OH)

"Silence Suppression" method

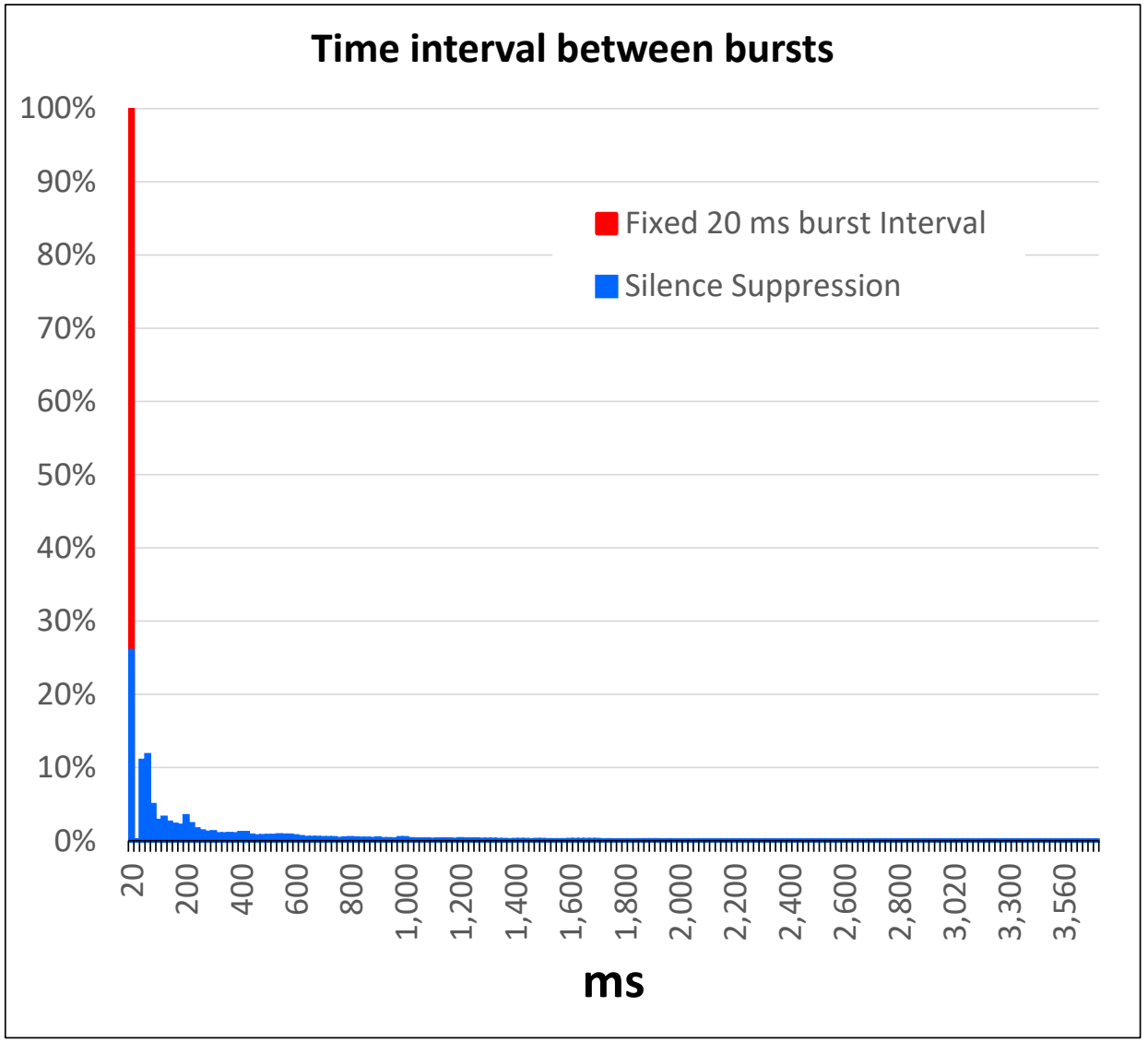
- In previous model: "92% of bursts are 386 ns long and only contain one REPORT each"
- Issue unsolicited grants at fixed intervals
  - GrantSize = CycleTime/Number of ONUs
- The *ForceReport* flag is not set
- If ONU has no data, it does not turn on the transmitter (eliminates 92% of bursts)

- **Parameters**
  - LaserOn = 60 ns
  - LaserOff = 60 ns
  - SyncTime = 129 ns
  - Grant interval (cycle) = 20 ms
  - Number of ONUs = 64
  - GrantSize = 106292 EQs (850336 bytes)

	Parameter	Units	Value	Notes
A	Bandwidth allocation cycle	ms	20	
B	ONUs		64	
C	Burst time per ONU	ns	312500	= ***A***×10⁶ / ***B***
D	Burst size per ONU	257b	31348	= ⌊***C***×25.78125 / 257⌋
E	LaserOn	257b	6	≈ 60 ns
F	LaserOff	257b	6	≈ 60 ns
G	SyncTime	257b	13	≈ 130 ns
H	FEC-protected length	257b	31323	= ***D*** – (***E***+***F***+***G***)
I	FEC codewords (incl. partial)		475	= ⌈***H*** / 66 ⌉
J	Payload area per burst	257b	26573	= ***H*** – ***I*** × 10
K	Payload area per burst	bytes	850336	= ***J*** × 256 / 8
L	Fixed grant size	EQs	106292	= ***K*** / 8

Observations (No-Report method)

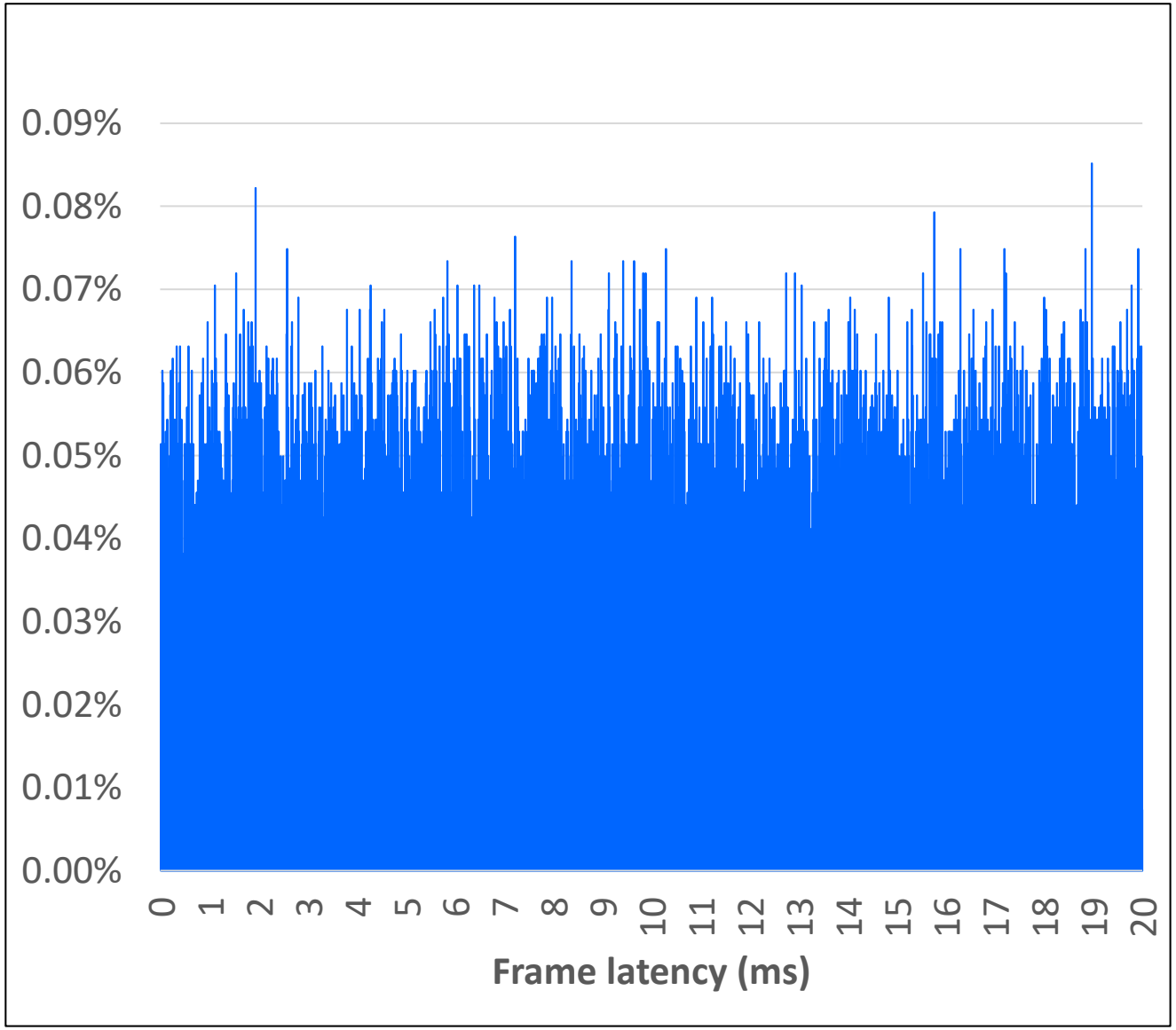
- The total number of bursts has decreased from 539993 (every 20 ms) to 42886 (only when data present)
- 95.55% of bursts are under 1 us.
- Total transmission time reduced from 215 ms to 22.4 ms.
- Maximum observed burst was 23.84 μs and it contained 84 70-byte packets



Packet Latency

- Both Fixed Cycle and Silence Suppression show identical packet latency distribution
- Packet latency is uniformly* distributed between 1 us and 20 ms (one grant period)

*theoretically uniform, but the number of samples is too small – 68K frames per 2000 bins



Conclusion and further questions

- During the observed time interval, most of the time, the upstream channel was not being used
  - Fixed Grant Interval: 99.998% idle
  - Silence Suppression: 99.9997% idle
- Sleep-wake cycles could provide measurable benefit
- Shutting down the Tx path in the optical module is relatively easy, but provides only limited power savings
- Shutting down Tx path in a SoC I harder and may impact traffic
  - UNI can receive data at any time. What SoC blocks can be shut down?
  - How long the recovery takes when the path is activated again?

Thank you