Modeling and Analysis of Latency Distribution in 40-100Gbps Dual-Mode Energy Efficient Ethernet

paper #33 (1570526636)

***Abstract*—Recently, The 40-100Gbps Energy Efficient Ethernet (EEE) is standardized by IEEE 802.3bj for energy saving. It involves not only the $DeepSleep$ state, which is just of 10% normal power consumption similar to the 1-10Gbps EEE, but also a new $FastWake$ state, whose power consumption can be 70% of the normal active state but the state transition time is much shorter. The default Dual-Mode strategy for the 40-100Gbps EEE should utilize these states suitably to make a good trade-off between the energy saving and the incurred latency of frames. Moreover, the interactive applications always place a soft or hard deadline on the transmission time of frames. Accordingly, the tail latency of frames, rather than the mean latency focused by existing work, should be emphasized in the 40-100Gbps EEE. Although numerous analytical works have been proposed for the strategies of EEE, most of them can only be applied to the condition of single power saving state, or just capture the average latency of frames. Therefore, we build an M/G/1 model for the latency distribution of frames in the 40-100Gbps Dual-Mode EEE in this work. The analytical results based on this model can guide the parameter configuration of the Dual- Mode strategy to fulfill the given tail latency requirement under different traffic load, as well as to make a good tradeoff between the energy saving and the incurred latency. Moreover, this model is powerful as it can generate the major results of several existing works under the corresponding simplified scenarios. Finally, excepting for the consistency with existing works, the analytical results are verified by simulations based on NS3 as a step further.**

***Index Terms*—Energy Efficient Ethernet, M/G/1, Latency Distribution, 40-100Gbps**

I. Introduction

Ethernet has been widely used and the link speed of Ethernet increases to 40-100Gbps recently [1]. With the increase of link capacity, the power consumption of each Ethernet interface also increases greatly. Consequently, the total energy consumption of Ethernet would also increase rapidly with the increase of link capacity and its popularity. However, most network interfaces keep on working with full power consumption even if there is no frame to be transmitted, leading to a significant waste of energy. According to [2], the utilization ratio of most network interfaces is only about 5% to 30%. Thus, it's urgent to create a more greening Ethernet environment.

In order to reduce the power consumption of Ethernet, IEEE 802.3az work group standardizes the Energy Efficient Ethernet (EEE) for 1-10Gbps links in 2010 [3] and IEEE 802.3bj work group standardizes the EEE for 40-100Gbps links in 2014 [8]. The basic idea of EEE is to turn off most components for power saving when there is no data transmission. For example, the Low Power Idle (LPI, renamed as $DeepSleep$ at 40-100Gbps) state defined for the 1-10Gbps links consumes only 10% of normal power consumption when there is no frame to be transmitted. However, time is needed for the transition from the LPI state to the normal $Active$ state. During the state transition, the interface works at full power and cannot transmit frames. What's worse, this state transition time increases from $4.48\mu s$ at 10Gbps [7] to $5.5\mu s$ at 40-100Gbps. The increase of the state transition time is very large, because the smaller latency is expected with the larger the link speed. Therefore, the $FastWake$ state, which needs less state transition time and consumes 70% of the normal power consumption, is suggested for the 40-100Gbps EEE. The Dual-Mode EEE depicted in [5] is the foremost strategy, determining when to enter into the $FastWake$ and $DeepSleep$ state for power saving in 40-100Gbps EEE.

Obviously, there is a tradeoff between the energy saving and the incurred latency of frames with the Dual-Mode strategy in 40-100Gbps EEE [5] [9]. Consequently, quantized analysis of the latency of frames and the energy saving is crucial, especially for guiding the parameter configuration. However, the two different power saving states of the 40-100Gbps Dual-Mode EEE complicate the quantized analysis, making it hard to adopt the analytical methods for the 1-10Gbps EEE [14]–[17] directly. In addition, the user-experience is highlighted nowadays and thus interactive applications put stricter requirements on the response time [4], [19], [20], which is generally denoted by the 99^{th} percentile tail latency of frames. Therefore, it is the tail latency rather than the mean latency of frames should be controlled within an acceptable range [18]. For example, the transmitted data will become useless if the whole data transmission cannot finish on time [19]. Amazon experienced a 1% drop in sales from a mere 100ms increase for web page loading [21]. And the estimation of Akamai shows that a 1s delay in web page response time can result in a 7% reduction in conversions, which adds up to $2.5 million in lost sales every year for an electronic commerce site making $100,000 per day [22]. But existing analysis works [10] for the 40-100Gbps Dual-Mode EEE focus on the mean latency instead of the tail latency of frames, as the mean latency can be obtained directly via the classic queuing theory. On the contrary, the tail latency is always greatly larger than the mean latency and it can only be obtained by deduction from the latency distribution of frames.

Therefore, in this paper, we build an M/G/1 model for the latency distribution in the 40-100Gbps Dual-Mode EEE. This model is powerful because the analytical results obtained based on it can degenerate to the major results of many existing analysis works [9], [13] and [18] under the corresponding simplified scenarios. Based on this model, we can derive

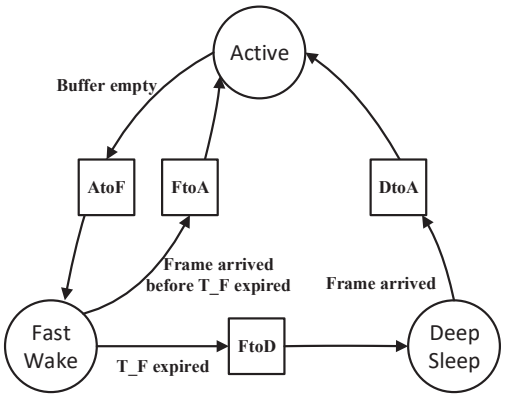

Fig. 1: State transitions of Dual-Mode EEE

both the mean latency and tail latency of frames, as well as the desired parameters with the given requirements on both the mean and the tail latency. As a step further, based on the M/G/1 model, we deduce the latency distribution of frames to guide the parameter configuration of the Dual-Mode EEE. During the deduction, the service time is assumed to be an exponential distribution, because the computation process for general distribution is too complex to be involved in the limited space. Based on the latency distribution of frames, the numerical analysis shows that the timer of the $FastWake$ state should be set large enough to avoid the large wake-up time for the $DeepSleep$ mode and achieve good mean and tail latency. What's more, different values are specified for the timer under different traffic load based on the model. Excepting the consistency of our analytical results with existing works, we also verify the accuracy of our results by simulations based on NS3 under different kinds of traffic generators.

The rest of this paper is organized as follows. Section II outlines the background and related work. Subsequently, section III builds the M/G/1 model for the latency distribution and section IV presents the analysis of the latency distribution. Next, verifications of the analytical results are provided in Section V, and numerical analysis and simulations are conducted in Section VI. Finally, Section VII concludes this paper.

II. Background and Related Work

A. Background

For the complement of this paper, we would give a brief introduction about the 40-100Gbps Dual-Mode EEE at first. More details can be found in [5].

As shown in Fig. 1, the 40-100Gbps Dual-Mode EEE interfaces have three different states: $Active$, $FastWake$ and $DeepSleep$. In the $Active$ state, the interface works with full power and frames are transmitted normally. In the $FastWake$ state, the energy consumption is only 70% of the normal $Active$ state [6]. And in the $DeepSleep$ state, most components of the 40-100Gbps EEE interfaces are shut off and its energy consumption is just 10% of the normal $Active$ state, similar to the LPI state of the 1-10Gbps EEE interfaces [7].

TABLE I: State transition time of 40-100Gbps EEE

Time	T_{AtoF}	T_{FtoD}	T_{DtoA}	T_{FtoA}
Value	$0.18\mu s$	$0.72\mu s$	$5.5\mu s$	$0.34\mu s$

Moreover, when the state transitions happen for energy saving, time is needed for turning on or shutting off components of the EEE interfaces. The specified times needed for each state transition are listed in Table I. Besides, during the process of state transitions, energy consumption is the same as the normal $Active$ state. Obviously, similar to the 1-10Gbps EEE interface, the strategy of determining when the 40-100Gbps EEE Interfaces enter or leave the $FastWake$ state or the $DeepSleep$ state is of great importance for energy saving. What's more, the $FastWake$ state enables 40-100Gbps EEE interface to be woken up quickly.

The default Dual-Mode strategy of the 40-100Gbps EEE Interfaces works as shown in Fig. 1. When there is no frame to transmit, the 40-100Gbps EEE Interfaces will transform from the $Active$ state into the $FastWake$ state in a fixed time T_{AtoF}. If there are frames arrive during the state transition process, the 40-100Gbps EEE Interfaces will be woken up immediately and return back to the $Active$ state after a fixed time T_{FtoA}. Otherwise, the Interfaces will enter to the $FastWake$ state and set up a timer T_F. If there are no frames arrive before the timer T_F expires, the 40-100Gbps EEE Interfaces will transform into $DeepSleep$ state after a fixed state transition time T_{FtoD}. Similar to the process of $AtoF$, the Interfaces will also be woken up when the state transition finish, if there are frames arrives during the process of $FtoD$. Otherwise, the Interfaces will enter into the $DeepSleep$ state and be woken up immediately upon the arrival of a frame. The state transition time from the $DeepSleep$ state to the $Active$ state is denoted by T_{DtoA}.

B. Related Work

With the increase of link speed, the time of state transition between the $DeepSleep$ state and the $Active$ state becomes relatively large. Accordingly, the amount of energy wasted during the process of state transition cannot be ignored and prevents Ethernet from being energy efficient under certain traffic pattern. To alleviate this problem, the $FastWake$ state is suggested for the 40-100Gbps EEE. The effect of the $FastWake$ state has been confirmed by the simulation study on the Dual-Mode EEE [5]. As a step further, work [9] builds an M/G/1 model for the energy consumption in Dual-Mode EEE, and then conduct an analysis of the average power consumption under different traffic loads. However, this model cannot capture the latency of frames. Therefore, work [10] builds an M/G/1 model for the mean waiting time of frames and verify it by real trace-driven simulations. Unfortunately, this model cannot reflect the tail latency of frames, which directly associated with the user experience. To address this problem, we model and analyze the latency distribution of frames and accordingly can obtain not only the mean waiting time but also the tail latency of frames. With these results, we

can guide the parameter configuration of the Dual-Mode EEE, so as to make a good trade-off between the energy saving and the incurred latency.

On the other side, although there are many analytical works about the latency of frames in the 1-10Gbps EEE up to now [14]–[17], these works cannot be directly applied to the 40-100Gbps Dual-Mode EEE, which adds the $FastWake$ state. Moreover, a fewer of them focus on the tail latency of frames [18], which is crucial to interactive applications.

Instead of modeling and analyzing the average power consumption or the latency of frames in the Dual-Mode EEE as [9] and [10], existing works for 40-100Gbps EEE [11] and [12] all focus on improving the Dual-Mode strategy for better energy saving. In other words, the Dual-Mode strategy is the basis to the 40-100Gbps Ethernet. Therefore, it's meaningful to model and analyze the latency distribution of frames in the 40-100Gbps Dual-Mode EEE.

III. The M/G/1 Latency Distribution Model

In this section, we will build the M/G/1 model for the latency distribution of frames in Dual-Mode EEE based on the assumption that frames arrive following a Poisson process.

A. Basic Idea

Defining λ as the average arrival rate of the Poisson process. For any frame, let random variable T represent the waiting time, random variable T_Q denote the queueing time, and random variable D represent the service time of this frame. Then, there is $T = T_Q + D$. Obviously, T_Q and D are independent of each other. Therefore, the Laplace transform of the above equation can be expressed as

$$\widetilde{T}(s) = \widetilde{T}_Q(s) * \widetilde{D}(s) \tag{1}$$

On the other hand, $\widetilde{T}(s)$ can be calculated as follows. Defining the random variable N as the number of remaining frames in the Dual-Mode EEE when one frame departs, and A_T as the number of frames arriving within the waiting time T of this departure frame. Then, N=A_T. Moreover, because the arrival of frames follows the Poisson process and the minimal number of frames that arrive during waiting time T is 0, there is $\widehat{A}_T(z) = \widetilde{T}(\lambda - \lambda z)$ according to [14], wherein $\widehat{A}_T(z)$ denotes the Z-transform of random variable A_T. Accordingly, we can get the following expression

$$\widehat{N}(z) = \widehat{A}_T(z) = \widetilde{T}(\lambda - \lambda z) \tag{2}$$

This expression indicates that $\widetilde{T}(s)$ can be derived from $\widehat{N}(z)$. Moreover, $\widehat{N}(z)$ can be obtained by constructing a Markov chain as shown in the next subsection. After getting $\widehat{N}(z)$, we can deduce $\widetilde{T}(s)$, and obtain the probability density function of the latency of frames by computing the inverse Laplace transform of $\widetilde{T}_Q(s)$. Finally, the latency distribution of frames can be obtained by conducting integration over the probability density function.

B. Markov Chain for $\widehat{N}(z)$

Firstly, according to the definition of Z-transform,

$$\widehat{N}(z) = \sum_{j=0}^{\infty} \pi_j z^j \tag{3}$$

where π_j represents the probability of j frames remaining in the Dual-Mode EEE when a frame leaves. In order to compute the expression of π_j, let's consider the condition that two consecutive frames X and $X+1$ leave. For any integers i and j, let π_i and π_j denote the probabilities that there are respectively i and j frames remaining in the Dual-Mode EEE when these two frames leave. Because the arrival of frames follows the Poisson process, the variation of the number of remaining frames is a Markov chain. Subsequently, we will discuss the transformation probability from any state i to state j, i.e., the transformation probability matrix of the Markov chain.

When $i \neq 0$, there are frames in Dual-Mode EEE when X leaves. At this moment, the Dual-Mode EEE port keeps staying at the $Active$ state and starts to transmit the frame $X+1$. Because the number of frames becomes j after the service time D of frame $X+1$, there must be $j-i+1$ frames which arrive at this Dual-Mode EEE port during the service time D. Obviously, the maximum value of i is $j+1$. When $i = j+1$, no frame arrives during the service time D.

When $i = 0$, there are no other frames in Dual-Mode EEE after the frame X leaves. Hence, the newly incoming frame will be $X+1$ and j frames arrive within the waiting time of the frame $X+1$. Obviously, the frame $X+1$ can arrive at the $FastWake$ state, or the $DeepSleep$ state, or in the process of $AtoF$ or $FtoD$. In total, these four different conditions can be illustrated in Fig. 2, wherein symbol I represents the time interval from the departure of the frame X to the arrival of the frame $X+1$. According to the complete probability formula, we can express above understanding mathematically as follows.

$$\begin{aligned}
&P\{A_T(X+1) = j\} \\
&\quad = P\{A_T(X+1) = j | I \leq T_{AtoF}\} P_{AtoF} \\
&\quad + P\{A_T(X+1) = j | T_{AtoF} < I \leq T_{AtoF} + T_F\} P_{FW} \\
&\quad + P\{A_T(X+1) = j |\ T_{AtoF} + T_F < I \leq T_{AtoF} + T_F \\
&\qquad + T_{FtoD}\} P_{FtoD} \\
&\quad + P\{A_T(X+1) = j | T_{AtoF} + T_F + T_{FtoD} < I\} P_{DS}
\end{aligned} \tag{4}$$

wherein P_{AtoF} represents the condition that the frame $X+1$ arrives in the process of $AtoF$. Moreover, because the frames arrive by following the Poisson process, I is an exponential random variable with parameter λ and we have

$$P_{AtoF} \triangleq P(I \leq T_{AtoF}) = 1 - e^{-\lambda T_{AtoF}} \tag{5}$$

Furthermore, the Dual-Mode EEE port will enter into the process of $FtoA$ after the finishing of the $AtoF$ process. Thus,

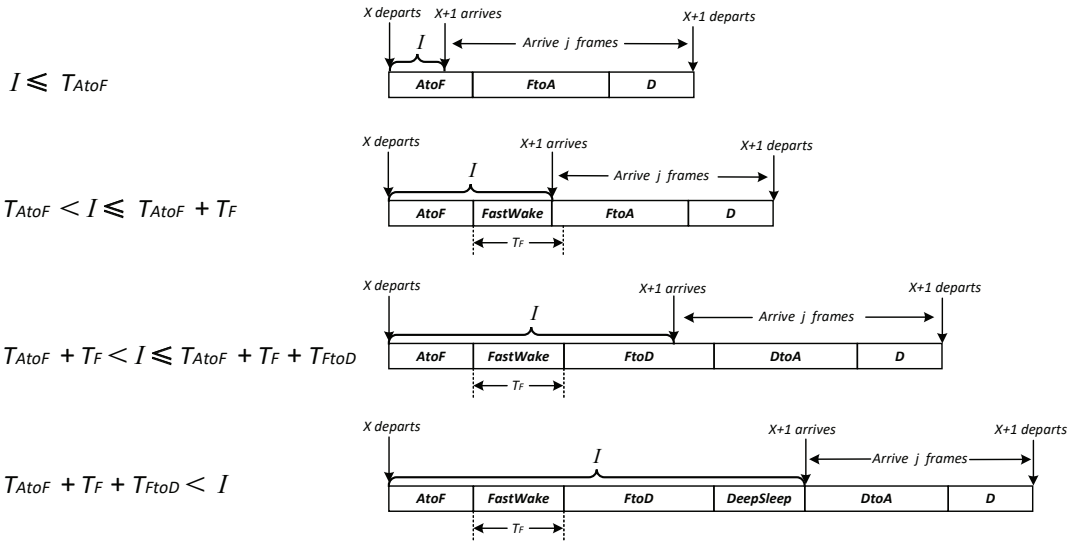

Fig. 2: Different Conditions due to different Arrival Time of the Frame $X+1$

recording I as I_1 in this condition, we have

$$P\{A_T(X+1)=j|I\le T_{AtoF}\}=P\{j\ frames\ arrive\ during\ T_{AtoF}+T_{FtoA}+D-I_1\}\triangleq b_j \tag{6}$$

Similarly, P_{FW} denotes the probability that the frame $X+1$ arrives at the $FastWake$ state and

$$P_{FW}\triangleq P(T_{AtoF}<I\le T_{AtoF}+T_F)=e^{-\lambda T_{AtoF}}(1-e^{-\lambda T_F}) \tag{7}$$

In this condition, the Dual-Mode EEE port will enter into the process of $FtoA$ after the arrival of the frame $X+1$. Hence, we have

$$P\{A_T(X+1)=j|T_{AtoF}<I\le T_{AtoF}+T_F\}=P\{j\ frames\ arrive\ during\ T_{FtoA}+D\}\triangleq c_j \tag{8}$$

Similarly, P_{FtoD} denotes the probability that the frame $X+1$ arrives in the process of $FtoD$ and

$$\begin{aligned}P_{FtoD}&\triangleq P(T_{AtoF}+T_F<I\le T_{AtoF}+T_F+T_{FtoD})\\&=e^{-\lambda(T_{AtoF}+T_F)}(1-e^{-\lambda T_{FtoD}})\end{aligned} \tag{9}$$

In this condition, the Dual-Mode EEE port will enter into the process of $DtoA$ after the finishing of the process $FtoD$. Hence, recording I as I_2, we have

$$P\{A_T(X+1)=j|\ T_{AtoF}+T_F<I\le T_{AtoF}+T_F+T_{FtoD}\}=P\{j\ frames\ arrive\ during\ T_{AtoF}+T_F+T_{FtoD}+T_{DtoA}+D-I_2\}\triangleq d_j \tag{10}$$

Similarly, P_{DS} denotes the probability that the frame $X+1$ arrives in the $DeepSleep$ state and

$$\begin{aligned}P_{DS}&\triangleq P(I>T_{AtoF}+T_F+T_{FtoD})\\&=e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}\end{aligned} \tag{11}$$

In this condition, the Dual-Mode EEE port will enter into the process of $DtoA$ after the arriving of the frame $X+1$. Therefore, we have,

$$P\{A_T(X+1)=j|T_{AtoF}+T_F+T_{FtoD}<I\}=P\{j\ frames\ arrive\ during\ T_{DtoA}+D\}\triangleq e_j \tag{12}$$

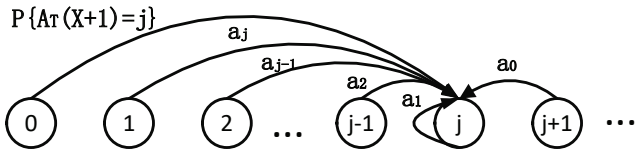

Fig. 3: Transformation about any state j in the Markov chain

Substituting equations (5)-(12) into equation (4), there is

$$\begin{aligned}&P\{A_T(X+1)=j\}\\&=b_jP_{AtoF}+c_jP_{FW}+d_jP_{FtoD}+e_jP_{DS}\\&=b_j(1-e^{-\lambda T_{AtoF}})+c_je^{-\lambda T_{AtoF}}(1-e^{-\lambda T_F})\\&+d_je^{-\lambda(T_{AtoF}+T_F)}(1-e^{-\lambda T_{FtoD}})\\&+e_je^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}\end{aligned} \tag{13}$$

Defining $a_j\triangleq P\{j\ frames\ arrive\ during\ D\}$, we can represent above analysis of the transition from other states to state j by Fig. 3. Accordingly, the following transition probability matrix of the Markov chain can be obtained.

$$\begin{aligned}\pi_j&=\sum_{i=1}^{j+1}\pi_ia_{j-i+1}+\pi_0P\{A_T(X+1)=j\}\\&=\sum_{i=1}^{j+1}\pi_ia_{j-i+1}+\pi_0[b_j(1-e^{-\lambda T_{AtoF}})+\\&\quad c_je^{-\lambda T_{AtoF}}(1-e^{-\lambda T_F})+d_je^{-\lambda(T_{AtoF}+T_F)}\\&\quad(1-e^{-\lambda T_{FtoD}})+e_je^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}]\end{aligned} \tag{14}$$

In the right of equation (14), the first sum denotes the condition that Dual-Mode EEE is on Active state, and the rest items associated with π_0 represent the condition that there are no frames in Dual-Mode EEE and the system waits for the arrival of frame $X+1$ by entering into either the $FastWake$ state or the $DeepSleep$ state, or in the process of $AtoF$ or $FtoD$.

Substituting equation (14) into equation (3), the Z-transform of N can be expressed as

$$\begin{aligned}\widehat{N}(z) = &\sum_{j=0}^{\infty}(\sum_{i=1}^{j+1}\pi_i a_{j-i+1})z^j + \sum_{j=0}^{\infty}\pi_0(1-e^{-\lambda T_{AtoF}})b_j z^j \\ &+ \sum_{j=0}^{\infty}\pi_0 e^{-\lambda T_{AtoF}}(1-e^{-\lambda T_F})c_j z^j \\ &+ \sum_{j=0}^{\infty}\pi_0 e^{-\lambda(T_{AtoF}+T_F)}(1-e^{-\lambda T_{FtoD}})d_j z^j \\ &+ \sum_{j=0}^{\infty}\pi_0 e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}e_j z^j\end{aligned} \tag{15}$$

Let $T_2 \triangleq T_{AtoF}+T_F+T_{FtoD}-I_2$ and $T_1 \triangleq T_{AtoF}-I_1$. Similar to the definition of A_T, let A_D, $A_{T_1+T_{FtoA}+D}$, $A_{T_{FtoA}+D}$, $A_{T_2+T_{DtoA}+D}$, and $A_{T_{DtoA}+D}$ be the random variables respectively representing the number of frames, which arrive within the times denoted by the corresponding subscripts. Obviously, the Z-transform of these random variables are respectively related to a_j, b_j, c_j, d_j and e_j. With these definitions, we can rewrite equation (15) as

$$\begin{aligned}\widehat{N}(z) = &\frac{1}{z}[\widehat{N}(z)-\pi_0]\widehat{A}_D(z) \\ &+ \pi_0(1-e^{-\lambda T_{AtoF}})\widehat{A}_{T_1+T_{FtoA}+D}(z) \\ &+ \pi_0 e^{-\lambda T_{AtoF}}(1-e^{-\lambda T_F})\widehat{A}_{T_{FtoA}+D}(z) \\ &+ \pi_0 e^{-\lambda(T_{AtoF}+T_F)}(1-e^{-\lambda T_{FtoD}})\widehat{A}_{T_2+T_{DtoA}+D}(z) \\ &+ \pi_0 e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}\widehat{A}_{T_{DtoA}+D}(z)\end{aligned} \tag{16}$$

Therefore, we can solve $\widehat{N}(z)$ from equation (16) and have

$$\begin{aligned}\widehat{N}(z) = &\frac{\pi_0}{z-\widehat{A}_D(z)}[z(1-e^{-\lambda T_{AtoF}})\widehat{A}_{T_1+T_{FtoA}+D}(z) \\ &+ ze^{-\lambda T_{AtoF}}(1-e^{-\lambda T_F})\widehat{A}_{T_{FtoA}+D}(z) - \widehat{A}_D(z) \\ &+ ze^{-\lambda(T_{AtoF}+T_F)}(1-e^{-\lambda T_{FtoD}})\widehat{A}_{T_2+T_{DtoA}+D}(z) \\ &+ ze^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}\widehat{A}_{T_{DtoA}+D}(z)]\end{aligned} \tag{17}$$

C. Latency Distribution Model

After obtaining the expression (17) of $\widehat{N}(z)$ based on the Markov Chain, we can calculate the probability density function of the latency of frames now. Substituting equation (17) into equation (2) and making variable substitution $z = \frac{\lambda-s}{\lambda}$, we can get the Laplace transform of $\widetilde{T}(s)$.

$$\begin{aligned}\widetilde{T}(s) = &\frac{\pi_0\widetilde{D}(s)}{\lambda-s-\lambda\widetilde{D}(s)}[(1-e^{-\lambda T_{AtoF}})(\lambda-s)\widetilde{T}_1(s)\widetilde{T}_{FtoA}(s) \\ &+ e^{-\lambda T_{AtoF}}(1-e^{-\lambda T_F})(\lambda-s)\widetilde{T}_{FtoA}(s) - \lambda \\ &+ e^{-\lambda(T_{AtoF}+T_F)}(1-e^{-\lambda T_{FtoD}})(\lambda-s)\widetilde{T}_2(s)\widetilde{T}_{DtoA}(s) \\ &+ e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}(\lambda-s)\widetilde{T}_{DtoA}(s)]\end{aligned} \tag{18}$$

Wherein $\widetilde{T}_{FtoA}(s) = e^{-sT_{FtoA}}$ and $\widetilde{T}_{DtoA}(s) = e^{-sT_{DtoA}}$. Moreover, we have the following proposition.

Proposition 1. *The Laplace transform of $\widetilde{T}_1(s)$ and $\widetilde{T}_2(s)$ are respectively*

$$\widetilde{T}_1(s) = \frac{\lambda(e^{-sT_{AtoF}} - e^{-\lambda T_{AtoF}})}{(\lambda-s)(1-e^{-\lambda T_{AtoF}})} \tag{19}$$

$$\widetilde{T}_2(s) = \frac{\lambda(e^{-sT_{FtoD}} - e^{-\lambda T_{FtoD}})}{(\lambda-s)(1-e^{-\lambda T_{FtoD}})} \tag{20}$$

Proof. Because I is an exponential random variable with parameter λ, i.e.,

$$I \sim f_I(t) = \lambda e^{-\lambda t}, 0 < t$$

we can obtain the following probability density function according to the definitions of I_1 and I_2

$$I_1 \sim f_{I_1}(t) = \frac{\lambda e^{-\lambda t}}{1-e^{-\lambda T_{AtoF}}}, 0 < t \le T_{AtoF} \tag{21}$$

$$\begin{aligned}I_2 \sim f_{I_2}(t) = &\frac{\lambda e^{-\lambda t}}{e^{-\lambda(T_{AtoF}+T_F)} - e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}}, \\ &T_{AtoF}+T_F < t \le T_{AtoF}+T_F+T_{FtoD}\end{aligned} \tag{22}$$

Hence, the distribution functions $F_{I_1}(t)$ and $F_{I_2}(t)$ of I_1 and I_2 are

$$P(I_1 \le t) = \begin{cases}\frac{e^{-\lambda t}-1}{e^{-\lambda T_{AtoF}}-1}, & 0 < t \le T_{AtoF} \\ 1, & T_{AtoF} < t\end{cases}$$

$$P(I_2 \le t) = \begin{cases}0, & 0 < t \le T_{AtoF}+T_F \\ \frac{e^{-\lambda t} - e^{-\lambda(T_{AtoF}+T_F)}}{e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})} - e^{-\lambda(T_{AtoF}+T_F)}}, & T_{AtoF}+T_F < t \le T_{AtoF}+T_F+T_{FtoD} \\ 1, & T_{AtoF}+T_F+T_{FtoD} < t\end{cases}$$

Because $T_1 = T_{AtoF} - I_1$ and $T_2 = T_{AtoF}+T_F+T_{FtoD}-I_2$, we have

$$\begin{aligned}P(T_1 \le t) &= P(T_{AtoF} - I_1 \le t) = 1 - P(I_1 < T_{AtoF} - t) \\ P(T_2 \le t) &= P(T_{AtoF}+T_F+T_{FtoD} - I_2 \le t) \\ &= 1 - P(I_2 < T_{AtoF}+T_F+T_{FtoD} - t)\end{aligned}$$

Accordingly, we can obtain the following distribution function of T_1 and T_2.

$$F_{T_1}(t) = \begin{cases}\frac{e^{-\lambda T_{AtoF}} - e^{-\lambda(T_{AtoF}-t)}}{e^{-\lambda T_{AtoF}}-1}, & 0 < t \le T_{AtoF} \\ 1, & T_{AtoF} < t\end{cases}$$

$$F_{T_2}(t) = \begin{cases}\frac{e^{-\lambda T_{FtoD}} - e^{-\lambda(T_{FtoD}-t)}}{e^{-\lambda T_{FtoD}}-1}, & 0 < t \le T_{FtoD} \\ 1, & T_{FtoD} < t\end{cases}$$

Then the probability function of T_1 and T_2 can be calculated by conducting the derivative operation.

$$T_1 \sim f_{T_1}(t) = \begin{cases}\frac{\lambda e^{-\lambda(T_{AtoF}-t)}}{1-e^{-\lambda T_{AtoF}}}, & 0 < t \le T_{AtoF} \\ 0, & T_{AtoF} < t\end{cases}$$

$$T_2 \sim f_{T_2}(t) = \begin{cases} \frac{\lambda e^{-\lambda(T_{FtoD}-t)}}{1-e^{-\lambda T_{FtoD}}}, & 0 < t \le T_{FtoD} \\ 0, & T_{FtoD} < t \end{cases}$$

Finally, according to the definition of the Laplace transform, the expression of $\widetilde{T}_1(s), \widetilde{T}_2(s)$ are respectively as follows.

$$\widetilde{T}_1(s) = \int_0^{T_{AtoF}} f_{T_1}(t)e^{-st}dt = \frac{\lambda(e^{-sT_{AtoF}} - e^{-\lambda T_{AtoF}})}{(\lambda - s)(1 - e^{-\lambda T_{AtoF}})}$$

$$\widetilde{T}_2(s) = \int_0^{T_{FtoD}} f_{T_2}(t)e^{-st}dt = \frac{\lambda(e^{-sT_{FtoD}} - e^{-\lambda T_{FtoD}})}{(\lambda - s)(1 - e^{-\lambda T_{FtoD}})}$$

□

Substituting these results into equation (18) and referring to equation (1), we have the following theorem.

Theorem 1. *The latency distribution model of the 40-100Gbps Dual-Mode EEE is*

$$\begin{aligned} \widetilde{T}_Q(s) = \frac{\pi_0}{\lambda - s - \lambda\widetilde{D}(s)} \{ & \lambda e^{-s(T_{AtoF}+T_{FtoA})} \\ & + e^{-\lambda T_{AtoF}}[e^{-\lambda T_F}(s-\lambda) - s]e^{-sT_{FtoA}} \\ & + e^{-\lambda(T_{AtoF}+T_F)}\lambda e^{-s(T_{FtoD}+T_{DtoA})} \\ & - e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})} s e^{-sT_{DtoA}} - \lambda \} \end{aligned} \tag{23}$$

The rest unknown variables in equation (23) are $\widetilde{D}(s)$ and π_0, wherein $\widetilde{D}(s)$ denotes the Laplace transform of service time of frames and π_0 denotes the probability of no frames in Dual-Mode EEE when a frame departs. Once $\widetilde{D}(s)$ is specified, π_0 can be solved out from equation $\widehat{N}(1) = 1$ referring to (17). Therefore, given any general distribution of the service time D, the probability density function of T_Q, i.e., the queueing time of frames in Dual-Mode EEE, can be obtained by taking inverse Laplace transform at both sides of equation (23). In other words, equation (23) is the desired M/G/1 model.

According to equation (23), the probability density function of latency of frames is directly associated with the incoming rate λ, the service time D of frames, the constants $T_{AtoF}, T_{FtoA}, T_{FtoD}$ and T_{DtoA}, as well as the threshold of timer T_F in the 40-100Gbps Dual-Mode EEE. It means that equation (23) can provide direct guidelines to the parameter configuration of the 40-100Gbps Dual-Mode EEE under given traffic load.

IV. Analysis of Latency Distribution

Subsequently, we analyze the latency distribution of frames based on the above M/G/1 model by assuming that the service time D follows an exponential distribution with an average value μ. Under this condition, $\widetilde{D}(s) = \frac{\mu}{\mu+s}$ and $E(A_D) = \lambda E(D) = \frac{\lambda}{\mu}$. When the service time D follows other probability distribution functions, we can also obtain similar results though the computation process becomes more complex. Due to the limited space, we omit the complex computation about the general distribution of service time D.

A. Expression of π_0

To analyze the latency distribution of frames, we should firstly deduce the expression of π_0 by specifying $z = 1$ in equation (17). Obviously, according to the definition of Z-transform, there is $\widehat{N}(1) = \sum_{j=0}^{\infty} \pi_j = 1$ and $\widehat{N}'(1) = E(N)$.

Similarly,

$$\widehat{A}'_D(1) = E(A_D) = \lambda/\mu$$
$$\widehat{A}'_{T_{FtoA}+D}(1) = E(A_{T_{FtoA}+D}) = \lambda(T_{FtoA} + 1/\mu)$$
$$\widehat{A}'_{T_{DtoA}+D}(1) = E(A_{T_{DtoA}+D}) = \lambda(T_{DtoA} + 1/\mu)$$

Moreover,

$$\begin{aligned} \widehat{A}'_{T_{AtoF}+T_{FtoA}+D-I_1}(1) &= E(A_{T_{AtoF}+T_{FtoA}+D-I_1}) \\ &= \lambda[T_{AtoF} + T_{FtoA} + 1/\mu - E(I_1)] \end{aligned}$$

$$\begin{aligned} \widehat{A}'_{T_{AtoF}+T_F+T_{FtoD}+T_{DtoA}+D-I_2}(1) \\ = E(A_{T_{AtoF}+T_F+T_{FtoD}+T_{DtoA}+D-I_2}) \\ = \lambda[T_{AtoF} + T_F + T_{FtoD} + T_{DtoA} + 1/\mu - E(I_2)] \end{aligned}$$

where

$$E(I_1) = \int_0^{T_{AtoF}} t f_{I_1}(t)dt = \frac{T_{AtoF}e^{-\lambda T_{AtoF}}}{e^{-\lambda T_{AtoF}} - 1} + \frac{1}{\lambda}$$

and

$$E(I_2) = \int_{T_{AtoF}+T_F}^{T_{AtoF}+T_F+T_{FtoD}} t f_{I_2}(t)dt = \frac{1}{\lambda} + \frac{(T_{AtoF} + T_F + T_{FtoD})e^{-\lambda T_{FtoD}} - (T_{AtoF} + T_F)}{e^{-\lambda T_{FtoD}} - 1}$$

Note that the definitions of $f_{I_1}(t)$ and $f_{I_2}(t)$ are respectively equations (21) and (22), as shown in the proof of Proposition 1. With the above results, we can have the following deduction by utilizing L'Hospital's rule.

$$\begin{aligned} 1 &= \widehat{N}(1) \\ &= \pi_0 \lim_{z\to 1} \{ \frac{(1 - e^{-\lambda T_{AtoF}})\widehat{A}_{T_1+T_{FtoA}+D}(z)}{1 - \widehat{A}'_D(z)} \\ &+ \frac{(1 - e^{-\lambda T_{AtoF}})z\widehat{A}'_{T_1+T_{FtoA}+D}(z)}{1 - \widehat{A}'_D(z)} - \frac{\widehat{A}'_D(z)}{1 - \widehat{A}'_D(z)} \\ &+ \frac{e^{-\lambda T_{AtoF}}(1 - e^{-\lambda T_F})[\widehat{A}_{T_{FtoA}+D}(z) + z\widehat{A}'_{T_{FtoA}+D}(z)]}{1 - \widehat{A}'_D(z)} \\ &+ \frac{e^{-\lambda T_{AtoF}}e^{-\lambda T_F}(1 - e^{-\lambda T_{FtoD}})\widehat{A}_{T_2+T_{DtoA}+D}(z)}{1 - \widehat{A}'_D(z)} \\ &+ \frac{e^{-\lambda T_{AtoF}}e^{-\lambda T_F}(1 - e^{-\lambda T_{FtoD}})z\widehat{A}'_{T_2+T_{DtoA}+D}(z)}{1 - \widehat{A}'_D(z)} \\ &+ \frac{e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}[\widehat{A}_{T_{DtoA}+D}(z) + z\widehat{A}'_{T_{DtoA}+D}(z)]}{1 - \widehat{A}'_D(z)} \} \\ &= \frac{\pi_0}{\mu - \lambda}[\mu\lambda(T_{AtoF} + T_{FtoA}) + \mu e^{-\lambda T_{AtoF}}(1 - e^{-\lambda T_F}) \\ &+ \mu\lambda e^{-\lambda(T_{AtoF}+T_F)}(T_{FtoD} + T_{DtoA} - T_{FtoA}) \\ &+ \mu e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}] \end{aligned} \tag{24}$$

Therefore, the expression of π_0 is

$$\pi_0 = \frac{\mu - \lambda}{B} \tag{25}$$

where the expression of B is as follows,

$$\begin{aligned} B = {} & \mu\lambda(T_{AtoF} + T_{FtoA}) + \mu e^{-\lambda T_{AtoF}}(1 - e^{-\lambda T_F}) \\ & + \mu\lambda e^{-\lambda(T_{AtoF}+T_F)}(T_{FtoD} + T_{DtoA} - T_{FtoA}) \\ & + \mu e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})} \end{aligned} \tag{26}$$

B. Latency Distribution of Frames

With π_0 and $\widetilde{D}(s)$, we can deduce the probability density function $f_{T_Q}(t)$ of latency distribution via inverse Laplace transform, and then obtain the latency distribution function, which is the integral of $f_{T_Q}(t)$. Substituting the expressions of π_0 and $\widetilde{D}(s)$ into equation (23), we have

$$\begin{aligned} \widetilde{T_Q}(s) = {} & \frac{\pi_0(\mu + s)}{-s^2 + (\lambda - \mu)s}\{\lambda e^{-s(T_{AtoF}+T_{FtoA})} - \lambda \\ & + [e^{-\lambda(T_{AtoF}+T_F)}(s - \lambda) - e^{-\lambda T_{AtoF}}s]e^{-sT_{FtoA}} \\ & + e^{-\lambda(T_{AtoF}+T_F)}\lambda e^{-s(T_{FtoD}+T_{DtoA})} \\ & - se^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}e^{-sT_{DtoA}}\} \end{aligned} \tag{27}$$

After the deduction of the inverse Laplace transform of equation (27) shown in the Appendix, we can obtain the probability density function $f_{T_Q}(t)$, which is composed by five parts as shown in equations (37)-(42). Subsequently, we can deduce the corresponding probability distribution function $F_{T_Q}(t)$ by integration.

$$F_{T_Q}(t) = \begin{cases} F^1_{T_Q}(t), & 0 \le t < T_{FtoA} \\ F^2_{T_Q}(t), & T_{FtoA} \le t < T_{AtoF} + T_{FtoA} \\ F^3_{T_Q}(t), & T_{AtoF} + T_{FtoA} \le t < T_{DtoA} \\ F^4_{T_Q}(t), & T_{DtoA} \le t < T_{DtoA} + T_{FtoD} \\ F^5_{T_Q}(t), & T_{DtoA} + T_{FtoD} \le t \end{cases} \tag{28}$$

where $F^1_{T_Q}(t)$, $F^2_{T_Q}(t)$, $F^3_{T_Q}(t)$, $F^4_{T_Q}(t)$ and $F^5_{T_Q}(t)$ are respectively defined as follows.

$$F^1_{T_Q}(t) \triangleq P(T_Q \le t) = \pi_0\{\frac{\mu\lambda t}{\mu - \lambda} + \frac{\lambda^2[e^{(\lambda-\mu)t} - 1]}{(\mu - \lambda)^2}\} \tag{29}$$

with $0 \le t < T_{FtoA}$.

$$\begin{aligned} F^2_{T_Q}(t) \triangleq {} & F^1_{T_Q}(t) + \pi_0\{e^{-\lambda T_{AtoF}}(1 - e^{-\lambda T_F}) \\ & + \frac{\mu\lambda e^{-\lambda(T_{AtoF}+T_F)}(t - T_{FtoA})}{\mu - \lambda} \\ & - \frac{\lambda e^{-\lambda T_{AtoF}}[e^{(\lambda-\mu)(t-T_{FtoA})} - 1][(1 - e^{-\lambda T_F})\mu - \lambda]}{(\mu - \lambda)^2}\} \end{aligned} \tag{30}$$

with $T_{FtoA} \le t < T_{AtoF} + T_{FtoA}$.

$$\begin{aligned} F^3_{T_Q}(t) \triangleq {} & F^2_{T_Q}(t) - \frac{\pi_0\lambda}{\mu - \lambda}[\mu(t - T_{AtoF} - T_{FtoA}) \\ & + \frac{\lambda e^{(\lambda-\mu)(t-T_{AtoF}-T_{FtoA})} - \lambda}{\mu - \lambda}] \end{aligned} \tag{31}$$

with $T_{AtoF} + T_{FtoA} \le t < T_{DtoA}$.

$$\begin{aligned} F^4_{T_Q}(t) \triangleq {} & F^3_{T_Q}(t) \\ & + \pi_0 e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}\frac{\mu - \lambda e^{(\lambda-\mu)(t-T_{DtoA})}}{\mu - \lambda} \end{aligned} \tag{32}$$

with $T_{DtoA} \le t < T_{DtoA} + T_{FtoD}$.

$$\begin{aligned} F^5_{T_Q}(t) \triangleq {} & F^4_{T_Q}(t) - \frac{\pi_0\lambda e^{-\lambda(T_{AtoF}+T_F)}}{(\mu - \lambda)}[\mu(t - T_{FtoD} \\ & - T_{DtoA}) + \frac{\lambda e^{(\lambda-\mu)(t-T_{FtoD}-T_{DtoA})} - \lambda}{\mu - \lambda}] \end{aligned} \tag{33}$$

with $T_{DtoA} + T_{FtoD} \le t$.

The latency distribution function $F_{T_Q}(t)$ is segmented. For example, the latency of the first frame, which arrives at the $FastWake$ state or in the process of $AtoF$, is mainly involved in the term $F^2_{T_Q}(t)$. The latency of the first frame, which arrives at the $DeepSleep$ state or in the process of $FtoD$, is mainly involved in the term $F^4_{T_Q}(t)$. Note that the latency of frame isn't solely composed by the waiting time for state transitions and the service time, but also consists of the queuing time in the Dual-Mode EEE.

Although the expression of $F_{T_Q}(t)$ seems complex, there are only three variables T_F, λ and μ in equation (28), and all the other coefficients are constants. Therefore, the probability, where the waiting time T_Q of frames is smaller than a given value t, depends on all of the traffic pattern λ, the service rate μ and the configuration T_F of the Dual-Mode EEE. In other word, given the traffic pattern and the desired 99^{th} tail latency t, we can find whether there exists a suitable parameter T_F from equation (28).

V. VERIFICATION

Subsequently, we will verify the above analytical results by showing that their degenerated results are consistent with many existing works.

At first, we focus on the M/G/1 model (23). When $T_{AtoF} = 0, T_{FtoA} = 0$ and $T_F \to +\infty$, there is $e^{-\lambda T_{AtoF}} = 1, e^{-\lambda T_F} = 0$ and the model (23) would degenerate to a general M/G/1. Under this condition, these is $\widetilde{T_Q}(s) = \pi_0 s/[\lambda\widetilde{D}(s) + s - \lambda]$, which is the same as the latency distribution model of general M/G/1 shown in [14]. It verifies the correctness of model (23).

Secondly, the probability distribution function (28) will degenerate to be the same as that in [18], if we consider T_{FtoA} as the sum of the timer and the wake-up time in the 1-10Gbps EEE, and set $T_{AtoF} \to 0$, $T_F \to +\infty$. Under this situation, we have $e^{-\lambda T_{AtoF}} = 1, e^{-\lambda T_F} = 0$ and the expression of π_0 and $F_{T_Q}(t)$ respectively degenerate to be

$$\pi_0 = \frac{\mu - \lambda}{\mu(\lambda T_{FtoA} + 1)} \tag{34}$$

$$F_{T_Q}(t) = \begin{cases} \pi_0\{\dfrac{\mu\lambda t}{\mu-\lambda} + \dfrac{\lambda^2[e^{(\lambda-\mu)t}-1]}{(\mu-\lambda)^2}\}, \\ 0 \le t < T_{FtoA} \\ 1-\pi_0[\dfrac{\lambda\mu e^{(\lambda-\mu)(t-T_{FtoA})}}{(\mu-\lambda)^2} - \dfrac{\lambda^2 e^{(\lambda-\mu)t}}{(\mu-\lambda)^2}], \\ T_{FtoA} \le t \end{cases} \tag{35}$$

This expression is exactly the same as the equation (9) of [18].

Finally, after obtaining the probability density function $f_{T_Q}(t)$, we can also deduce the expectation of the random variable T_Q to get the following mean latency of frames for the 40-100Gbps Dual-Mode EEE.

$$\begin{aligned} E(T_Q) = {} & \frac{-\pi_0}{2(\lambda-\mu)^2}\{2\lambda e^{-\lambda T_{AtoF}}(e^{-\lambda T_F}-1) \\ & + \lambda(T_{AtoF}+T_{FtoA})[\mu(\lambda-\mu)(T_{AtoF}+T_{FtoA}) - 2\lambda] \\ & + 2T_{FtoA}(\lambda-\mu)e^{-\lambda T_{AtoF}}[\mu+(\lambda-\mu)e^{-\lambda T_F}] \\ & - \lambda\mu(\lambda-\mu)e^{-\lambda(T_{AtoF}+T_F)}T_{FtoA}^2 \\ & - 2\lambda^2(T_{FtoD}+T_{DtoA})e^{-\lambda(T_{AtoF}+T_F)} \\ & + \lambda\mu(\lambda-\mu)(T_{FtoD}+T_{DtoA})^2 e^{-\lambda(T_{AtoF}+T_F)} \\ & + e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}[2\mu(\lambda-\mu)T_{DtoA} - 2\lambda]\} \end{aligned} \tag{36}$$

This result is the same as that of [10], confirming the correctness of model.

VI. Numerical Analysis and Simulation

A. Numerical Analysis

Subsequently, we conduct numerical analysis on the analytical results to guide the configuration of the 40-100Gbps Dual-Mode EEE.

At first, μ is set to be the link capacity $40Gbps$ and $\lambda = \mu/2$. The latency distribution of frames against time is shown in Fig. 4, with kinds of different values of parameter T_F. Obviously, as the probability distribution function, $F_{T_Q}(t)$ always converges to 1 with the increase of t. Moreover, although $F_{T_Q}(t)$ is segmented, it's continuous at all the segmented points $T_{FtoA} = 0.34\mu s$, $T_{AtoF} + T_{FtoA} = 0.52\mu s$, $T_{DtoA} = 5.5\mu s$ and $T_{DtoA} + T_{FtoD} = 6.22\mu s$. In addition, when T_F is set as the default value $3.5\mu s$, $F_{T_Q}(t)$ approaches to 1 before t reaches $5.5\mu s$, indicating that the Dual-Mode EEE seldom enters into the $DeepSleep$ state with relatively large T_F. This is also why the curve of $T_F = 10\mu s$ is totally covered by the curve of $T_F = 3.5\mu s$, and the variation of $F_{T_Q}(t)$ becomes slow after $t \ge 0.52\mu s = T_{AtoF} + T_{FtoA}$.

Next, the impacts of the arrival speed λ and parameter T_F on the 99^{th} tail latency of frames are shown in Fig. 5. In details, we calculate each t_v such that $F_{T_Q}(t_v) = 99\%$ for every given T_F and λ. Similar to above results, the 99^{th} tail latency drops from about $6.22\mu s$ suddenly, as the Dual-Mode EEE seldom enters into the $DeepSleep$ state, when T_F becomes larger than a cliff point. Moreover, the larger the λ, the smaller the cliff point of T_F, but the 99^{th} tail latency of frames stays at a larger value after the cliff point of T_F. This is because the queueing latency of frames becomes significant

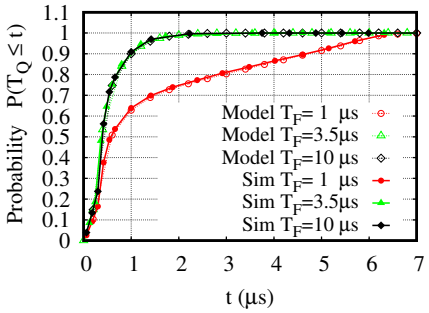

Fig. 4: Distribution Function

with the increase of traffic load. This is also why the 99^{th} tail latency drops from a value larger than $6.22\mu s$ under the conditions of $\lambda = 0.9\mu$ and $\lambda = 0.7\mu$. Similar conclusion is also suitable to the mean latency of frames, which is plotted according to the equation (36) generated by our model, as shown in Fig. 6. These results indicate that T_F should be set large enough according to the load for small 99^{th} tail latency and mean latency of frames in the Dual-Mode EEE. In this way, the unnecessary large wake up time for the $DeepSleep$ mode is always avoided.

Besides the latency of frames, however, we should also consider the impact of T_F on power consumption. Therefore, we show the corresponding power consumption ratio under above conditions in Fig. 7, based on the equations (3), (4) and (5) in [9]. Consisting with the intuition, the larger the value of λ, the larger the power consumption of EEE. Moreover, larger T_F leads to smaller power consumption, especially when the value of λ is small. Note that the smallest power consumption ratio is still as high as 80% of the normal states, because the power consumption of the $FastWake$ state is 70%. With the Dual-Mode strategy, the $DeepSleep$ state helps a little on power saving due to the following two reasons. First, the 40-100Gbps EEE can't stay at the $DeepSleep$ state for a long time as it will be woken up once any frame arrives. Second, the power consumption is 100% during the large wake-up time T_{DtoA}. For better power saving, strategy like frame coalescing may be employed at the $DeepSleep$ state [11], because the overhead of latency $T_{DtoA} = 5.5\mu s$ has been spent after the choice of the $DeepSleep$ state.

B. Simulations

Although the correctness of our theoretical results has been verified by its consistency with existing works, we conduct simulations based on NS3 to show the accuracy of the theoretical results here as a step further. Specifically, we implement the 40-100Gbps Dual-Mode EEE by extending the codes of [17] in NS3. Moreover, to simulate the Poisson arrival of frames, we let the arrival interval of frames follow the exponential distribution with parameter λ. Furthermore, the frame size is drawn from an exponential distribution with average value 1KB such that the service time of frames follows an exponential distribution. For each result, we repeat the

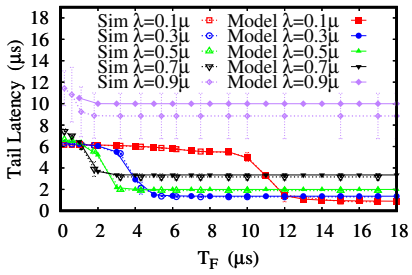

Fig. 5: 99^{th} Tail Latency of Frames

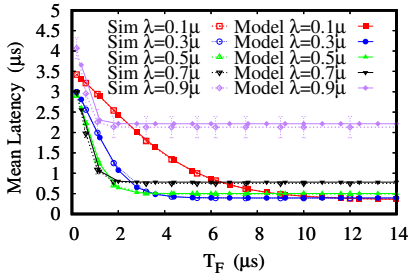

Fig. 6: Mean Latency of Frames

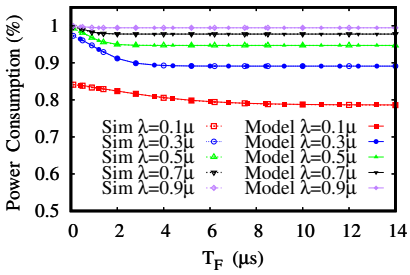

Fig. 7: The Power Consumption

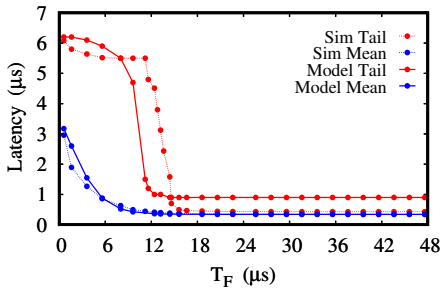

Fig. 8: Trace-Driven Simulation

corresponding simulation 15 times and show their average and standard deviation.

As shown in Fig. 7, our simulation results are consistent with the model deduced in [9]. It confirms the correctness of our implementation of the 40-100Gbps Dual-Mode EEE. Furthermore, as shown in Figures 4, 5 and 6, all the simulation results fit our model well under all kinds of different conditions, verifying the correctness of our analytical results. When $\lambda = 0.9\mu$, the standard deviation of the simulation results becomes large, as the random queueing latency becomes more significant than the fixed state transitions latency, as shown in Figures 5 and 6. In addition, there is a slight difference between the average value of the simulation results and the models, possibly due to the limitation of frame sizes in NS3, which cuts off the exponential distribution of the service time of packets. Therefore, the simulation results also suggest to set T_F large enough such that the $FastWake$ state can be utilized to complement the shortage of the $DeepSleep$ mode on large wake up time. In this way, the Dual-Mode EEE can achieve small mean and tail latency of frames by avoiding the unnecessary large wake up time of the $DeepSleep$ mode.

Finally, we conduct simulations driven by the real trace provided by [17]. As the trace is obtained from 10Gbps links, we shorten each packet arriving interval such that the load of the 40Gbps EEE is about 10%. As a comparison, we produce the corresponding results of our model (28) under this load. The mean latency of our model matches well with the mean latency of the trace-driven simulation, as shown in Fig. 8. Moreover, all the curves are of the same trends, i.e., both the 99^{th} tail latency and the mean latency drops greatly after T_F is large enough similar to above results, even if the trace doesn't follow the assumption of the Poisson traffic pattern. It confirms that our model is useful for guiding the parameter configuration of the 40-100Gbps Dual-Mode EEE.

VII. Conclusion

In this paper, we establish the M/G/1 model for the latency distribution of frames in the 40-100Gbps Dual-Mode EEE. This model is powerful to generate many of existing results under different simplified scenarios. Based on this model, we can deduce out both the mean latency and the tail latency of frames, and accordingly utilize it to guide the configuration of Dual-Mode EEE. Numerical analysis of the results shows that the parameter T_F of Dual-Mode EEE should be set large enough such that the $FastWake$ state is fully utilized to complement the shortage of the $DeepSleep$ mode on large wake up time. In other words, the analytical results reveal the cliff points of the tail and mean latency of frames under different traffic load for the configuration of parameter T_F. Simulations confirm the accuracy of our analytical results.

References

[1] IEEE P802.3ba, *40Gb/s and 100Gb/s Ethernet Task Force*, http://www.ieee802.org/3/ba/index.html, 2010.

[2] P. Patel-Predd, *Energy-efficient Ethernet*, IEEE Spectr., May 2008.

[3] IEEE P802.3az, *Energy Efficient Ethernet Task Force*, Available at http://grouper.ieee.org/groups/802/3/az/public/index.html.

[4] Stephen M. Rumble, Diego Ongaro, Ryan Stutsman, Mendel Rosenblum and John K. Ousterhout, *Its Time for Low Latency*, HotOS, 2011

[5] M. Mostowfi, *A Simulation Study of Energy-Efficient Ethernet With Two Modes of Low-Power Operation,* IEEE Communications Letters Vol. 19, no.10, Oct. 2015

[6] H. Barrass, *Options for EEE in 100G*, Jan. 2012, https://www.ieee802.org/3/bj/public/jan12/barrass_01a_0112.pdf

[7] P. Reviriego, K. Christensen, J. Rabanillo and J. A. Maestro, *An initial evaluation of energy efficient ethernet,* IEEE Commun. Lett. vol. 15, no. 5, pp. 578-580, May 2011.

[8] IEEE Computer Society, *IEEE 802.3bj-2014 Amendment 2: physical layer specifications and management parameters for 100 Gb/s operation over backplanes and copper cables,* June 2014.

[9] M. Mostowfi and K. Shafie, *An analytical model for the power consumption of Dual-Mode EEE,* Electronics Letters, Vol. 52, no. 15, Jul. 2016.

[10] M. Mostowfi and K. Shafie, *Average packet delay in Dual-Mode EEE: An analytical model,* Electronics Letters, Vol. 52, no. 21, Oct 2016.

[11] S. Herrerłaalonso, M. Rodrłguezprez, M. Fernndezveiga, C. Lpezgarcła, *Frame coalescing in Dual-Mode EEE,* 2015, https://arxiv.org/pdf/1510.03694.pdf.

[12] M. Mostowfi and K. Shafie, *Dual-Mode Energy Efficient Ethernet with Packet Coalescing: Analysis and Simulation,* Sustainable Computing Informatics & Systems, 2017.

[13] Mor Harchol-Balter, *Performance Modeling and Design of Computer Systems: Queueing Theory in Action,* Cambridge University Press, Feb. 2013.
[14] Jinli Meng, Fengyuan Ren, Chuang Lin. *Modeling and understanding burst transmission for energy efficient ethernet,* Computer Networks, July 2017.
[15] Xiaodan Pan, Tong Ye, Tony T. Lee, Weisheng Hu, *Power Efficiency and Delay Tradeoff of 10GBase-T Energy Efficient Ethernet Protocol*, IEEE/ACM Transactions on Networking, Oct. 2017.
[16] Sergio Herreria-Alonso, Miguel Rodriguez-Perez, Manuel Fernandez-Veiga, Candido Lopez-Garcia, *A GI/G/1 Model for 10Gb/s Energy Efficient Ethernet Links*, IEEE Trans. on Communications, Nov. 2012.
[17] Angelos Chatzipapas and Vincenzo Mancuso, *An M/G/1 Model for Gigabit Energy Efficient Ethernet Links With Coalescing and Real-Trace-Based Evaluation*, IEEE/ACM Trans. on Networking, Oct. 2016.
[18] Nail Akar, *Delay Analysis of Timer-Based Frame Coalescing in Energy Efficient Ethernet,* IEEE Communications Letters, July 2013.
[19] J. Dean and L. A. Barroso, *The Tail At Scale*, Communications of the ACM, Volumn 56:74-80, 2013
[20] Giuseppe DeCandia et. al., *Dynamo: Amazons Highly Available Key-value Store*, SOSP, 2007.
[21] Farhan Khan, *The Cost of Latency*, 2015, https://www.digitalrealty.com/blog/the-cost-of-latency.
[22] Ramin, *How Latency Can Make or Break Your Sales on Cyber Monday*, 2017, https://www.cdnetworks.com/web-performance/how-latency-can-make-or-break-your-sales-on-cyber-monday/

APPENDIX A
THE PROBABILITY DENSITY FUNCTION $f_{T_Q}(t)$

As shown in equation (27), there are five items in the expression of $\widetilde{T}_Q(s)$, i.e., $\widetilde{T}_Q(s) = \pi_0[F_1(s)+F_2(s)+F_3(s)+F_4(s)+F_5(s)]$ where

$$F_1(s) \triangleq \frac{\lambda(\mu+s)e^{-s(T_{AtoF}+T_{FtoA})}}{-s^2+(\lambda-\mu)s}$$
$$F_2(s) \triangleq \frac{[e^{-\lambda T_{AtoF}}e^{-\lambda T_F}(s-\lambda)-e^{-\lambda T_{AtoF}}s](\mu+s)e^{-sT_{FtoA}}}{-s^2+(\lambda-\mu)s}$$
$$F_3(s) \triangleq \frac{e^{-\lambda T_{AtoF}}e^{-\lambda T_F}\lambda(\mu+s)e^{-s(T_{FtoD}+T_{DtoA})}}{-s^2+(\lambda-\mu)s}$$
$$F_4(s) \triangleq -\frac{e^{-\lambda T_{AtoF}}e^{-\lambda T_F}e^{-\lambda T_{FtoD}}s(\mu+s)e^{-sT_{DtoA}}}{-s^2+(\lambda-\mu)s}$$
$$F_5(s) \triangleq -\frac{\lambda(\mu+s)}{-s^2+(\lambda-\mu)s}$$

Assume the inverse Laplace transform of these items are $f_1(t), f_2(t), f_3(t), f_4(t)$ and $f_5(t)$ respectively. We have

$$f_{T_Q}(t) = \pi_0[f_1(t)+f_2(t)+f_3(t)+f_4(t)+f_5(t)] \tag{37}$$

according to the linear property of inverse Laplace transform. Subsequently, we will deduce $f_i(t)$ for all $i = 1, 2, 3, 4, 5$.

A. The expression of $f_1(t)$

Assume the inverse Laplace transform of $H(s)$ is $h(t)$, where $H(s)$ is defined by

$$H(s) = \frac{\lambda(\mu+s)}{-s^2+(\lambda-\mu)s} = -\frac{\lambda}{s+\mu-\lambda} - \frac{\mu\lambda}{\mu-\lambda}\frac{\mu-\lambda}{s(s+\mu-\lambda)}$$

We have $f_1(t) = h(t - T_{AtoF} - T_{FtoA})$ by using the time-domain translation properties of Laplace transform. Looking up the table of Laplace transform, we have $h(t) = [\lambda^2 e^{(\lambda-\mu)t} - \mu\lambda]/(\mu-\lambda)$. Consequently, the expression of $f_1(t)$ is

$$f_1(t) = \frac{\lambda^2 e^{(\lambda-\mu)(t-T_{AtoF}-T_{FtoA})} - \mu\lambda}{\mu-\lambda}, T_{AtoF}+T_{FtoA} \le t \tag{38}$$

B. The expression of $f_2(t)$

Similarly, let

$$R(s) = \frac{[e^{-\lambda T_{AtoF}}e^{-\lambda T_F}(s-\lambda) - e^{-\lambda T_{AtoF}}s](\mu+s)}{-s^2+(\lambda-\mu)s}$$
$$= e^{-\lambda T_{AtoF}}[(1-e^{-\lambda T_F})(1+\frac{\lambda}{s+\mu-\lambda}) - e^{-\lambda T_F}H(s)]$$

Assuming the inverse Laplace transform of $R(s)$ is $r(t)$, we have $f_2(t) = r(t - T_{FtoA})$. Moreover, looking up the table of Laplace transformation, we have

$$r(t) = e^{-\lambda T_{AtoF}}(1-e^{-\lambda T_F})\delta(t) + \frac{e^{-\lambda T_{AtoF}}e^{-\lambda T_F}\mu\lambda}{\mu-\lambda} + \frac{\lambda e^{-\lambda T_{AtoF}}[(1-e^{-\lambda T_F})\mu-\lambda]e^{(\lambda-\mu)t}}{\mu-\lambda}$$

where $\delta(t)$ is the impulse function satisfying

$$\delta(t) = \begin{cases} \infty, t=0 \\ 0, t\neq 0 \end{cases} \quad and \quad \int_{-\infty}^{+\infty}\delta(t)dt = 1$$

Thus, the expression of $f_2(t)$ is

$$f_2(t) = e^{-\lambda T_{AtoF}}(1-e^{-\lambda T_F})\delta(t-T_{FtoA}) + \frac{\lambda e^{-\lambda T_{AtoF}}[(1-e^{-\lambda T_F})\mu-\lambda]e^{(\lambda-\mu)(t-T_{FtoA})}}{\mu-\lambda} + \frac{e^{-\lambda T_{AtoF}}e^{-\lambda T_F}\mu\lambda}{\mu-\lambda}, T_{FtoA} \le t \tag{39}$$

C. The expression of $f_3(t)$

Similar to the deduction of $f_1(t)$, we have

$$f_3(t) = \frac{e^{-\lambda(T_{AtoF}+T_F)}}{\mu-\lambda}[\lambda^2 e^{(\lambda-\mu)(t-T_{FtoD}-T_{DtoA})} - \mu\lambda], \quad T_{FtoD}+T_{DtoA} \le t \tag{40}$$

D. The expression of $f_4(t)$

Define $Q(s)$ as follows.

$$Q(s) = -\frac{e^{-\lambda T_{AtoF}}e^{-\lambda T_F}e^{-\lambda T_{FtoD}}s(\mu+s)}{-s^2+(\lambda-\mu)s}$$
$$= e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})} + \frac{e^{-\lambda T_{AtoF}}e^{-\lambda T_F}e^{-\lambda T_{FtoD}}\lambda}{s+\mu-\lambda}$$

Obviously, $f_4(t) = q(t - T_{DtoA})$, where $q(t)$ is the inverse Laplace transform of Q(t). Looking up the Laplace transform table, we can obtain

$$q(t) = e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}[\delta(t) + \lambda e^{(\lambda-\mu)t}]$$

Accordingly, the expression of $f_4(t)$ is

$$f_4(t) = e^{-\lambda(T_{AtoF}+T_F+T_{FtoD})}[\delta(t-T_{DtoA}) + \lambda e^{(\lambda-\mu)(t-T_{DtoA})}], T_{DtoA} \le t \tag{41}$$

E. The expression of $f_5(t)$

Finally, the expression of $f_5(t)$ as follows,

$$f_5(t) = \frac{\mu\lambda}{\mu-\lambda} - \frac{\lambda^2 e^{(\lambda-\mu)t}}{\mu-\lambda}, 0 \le t \tag{42}$$