## APPENDIX A
## PROOF OF LEMMA 1

**Lemma 1.** *Given any feasible task replication and power allocation, the Lyapunov drift-plus-penalty function $B$ in our formulation (32) is upper bounded by the following expression.*

$$B \geq \frac{1}{2}E\left[\left(\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t)\right)^2 + (\mu_b t_a)^2\right] + \frac{1}{2}E\left[\left(\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t)\right)^2 + (1_{Q(t)>0}\mu_c t_a)^2\right],$$

*where $R_{i,b}(t)$ denotes the total amount of computing from TaV $i$ in the time slot t, $1_{Q(t)>0}$ represents a characteristic function, which only takes the value of 1 when $Q(t) > 0$.*

*Proof.* By squaring both sides of the equation, we can get the expression below.

$$\begin{aligned} Q(t+1)^2 &= \left(\max\left\{Q(t) + \sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t) - \mu_b t_a, 0\right\}\right)^2 \\ &\leq Q(t)^2 + (\mu_b t_a)^2 + \left(\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t)\right)^2 \\ &\quad - 2\mu_b t_a Q(t) - 2\mu_b t_a \sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t) \\ &\quad + 2\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t)\mu_b t_a. \end{aligned}$$

The following inequality can be obtained as follows.

$$\frac{1}{2}[Q(t+1)^2 - Q(t)^2] \leq \frac{1}{2}\left[\left(\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t)\right)^2 + (\mu_b t_a)^2 + 2Q(t)\left(\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t) - \mu_b t_a\right)\right].$$

Similarly, we can obtain the following expression.

$$\begin{aligned} &\frac{1}{2}[H(t+1)^2 - H(t)^2] \leq \frac{1}{2}\left(\left(\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t)\right)^2 \right. \\ &+(1_{Q(t)>0}\mu_c t_a)^2 + 2\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t)H(t) \\ &\left. -2\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t)1_{Q(t)>0}\mu_c t_a - 2H(t)1_{Q(t)>0}\mu_c t_a\right) \\ &\leq \frac{1}{2}\left[(1_{Q(t)>0}\mu_c t_a)^2 + \left(\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t)\right)^2 \right. \\ &\left. +2H(t)\left(\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t) - 1_{Q(t)>0}\mu_c t_a\right)\right]. \end{aligned}$$

According to the definition of Lyapunov drift function, we can get the expression below.

$$L(\Theta(t)) = \frac{1}{2}\left(Q(t)^2 + H(t)^2\right),$$

$$\triangle L(t) = E[L(\Theta(t+1)) - L(\Theta(t))].$$

Then, we can get a bound on $\triangle L(t)$.

$$\triangle L(t) \leq B + \left[Q(t)\left(\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t) - \mu_b t_a\right) + H(t)\left(\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t) - 1_{Q(t)>0}\mu_c t_a\right)\right],$$

where $B$ is the upper bound by the expression as below.

$$B \geq \frac{1}{2}E\left[\left(\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t)\right)^2 + (\mu_b t_a)^2\right] + \frac{1}{2}E\left[\left(\sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t)\right)^2 + (1_{Q(t)>0}\mu_c t_a)^2\right],$$

□

## APPENDIX B
## PROOF OF LEMMA 2

**Lemma 2.** *An algorithm that can effectively control $Q(t)$ and $H(t)$ within prescribed limits ($Q(t) \leq Q_{\max}$, $H(t) \leq H_{\max}$) will ensure a capped worst-case delay. The worst-case delay for the task transmission to the server queue is bounded by the constant $T_{\max}$ defined as follows:*

$$T_{\max} = \lceil \frac{Q_{\max} + H_{\max}}{\mu_b t_a} \rceil,$$

*where $\lceil\cdot\rceil$ denotes the smallest integer that is greater than or equal to the value.*

*Proof.* The task arrives at slot $t$ can depart the queue at the slot $t+1$ earliest. We prove that these tasks can depart the queue by time $t + T_{\max}$ by contradiction. It must be that $Q(t_x) > 0$ for all $t_x \in \{t+1, \cdots, t+T_{\max}\}$. We have:

$$H(t_x + 1) = \max[H(t_x) + \sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t) - \mu_b t_a, 0],$$

and hence for all $t_x \in \{t+1, \cdots, t+T_{\max}\}$:

$$H(t_x + 1) \geq H(t_x) + \sum_{i=1}^{I} x_{i,b}(t)R_{i,b}(t) - \mu_b t_a.$$

Summing the above over $t_x \in \{t+1, \cdots, t+T_{\max}\}$ yields:

$$\begin{aligned} &H(t + T_{\max} + 1) - H(t+1) \\ &\quad \geq \mu_b t_a T_{\max} - \sum_{t_x=t+1}^{t+T_{\max}} \sum_{i=1}^{I} x_{i,b}(t_x)R_{i,b}(t_x). \end{aligned}$$

Using the fact that $H(t+T_{\max}+1) \leq T_{\max}$ and $H(t+1) \leq 0$, we have:

$$\mu_b t_a T_{\max} - H_{\max} \leq \sum_{t_x=t+1}^{t+W_{\max}} \sum_{i=1}^{I} x_{i,b}(t_x) R_{i,b}(t_x).$$

Because we have assumed that not all of the tasks arrive at $t$ slot depart by time $t+W_{\max}$, we must have:

$$\sum_{t_x=t+1}^{t+T_{\max}} \sum_{i=1}^{I} x_{i,b}(t_x) R_{i,b}(t_x) < Q(t) \leq Q_{\max}.$$

Combining above equations, we can get:

$$\mu_b t_a T_{\max} - H_{\max} \leq Q_{\max}.$$

Therefore:

$$T_{\max} \leq \frac{Q_{\max} + H_{\max}}{\mu_b t_a}.$$

This contradicts the definition we have made before. □

## APPENDIX C
## PROOF OF LEMMA 3

**Lemma 3.** *The BER $\Omega_{i,b}(t)$ between TaV $i$ and the edge cloud and the BER $\Omega_{i,j}(t)$ between TaV $i$ and SeV $j$ under Rayleigh fading conditions have the approximation as below.*

$$\Omega_{i,b}(t) \approx \frac{(e^{R_{i,b}(t)} - 1)\left(\sum_{k=i+1}^{I} |h_{k,b}(t)|^2 \alpha_i^n(t) p_{k,b}(t) + \sigma^2\right)}{p_{i,b}(t)\,|h_{i,b}(t)|^2},$$

$$\Omega_{i,j}(t) \approx \frac{(e^{\frac{K_{i,j}}{L_{i,j}}} - 1)\sigma^2}{p_{i,j}(t)\,|h_{i,j}(t)|^2}.$$

*Proof.* According to Taylor's Expanded Form, we can get the approximations of BER as below.

$$\begin{aligned}\Omega_{i,b}(t) \approx & 1 - k_1 \gamma_{i,b}(t) \left(1 + \frac{k_2}{\gamma_{i,b}(t)} + \frac{k_2^2}{2{\gamma_{i,b}}^2(t)}\right.\\ & \left. -1 - \frac{k_3}{\gamma_{i,b}(t)} - \frac{k_3^2}{2{\gamma_{i,b}}^2(t)}\right),\end{aligned}$$

where $k_1 = \sqrt{\frac{L_{i,b}}{2\pi(e^{2\frac{K_{i,b}}{L_{i,b}}}-1)}}$, $k_2 = 1 - e^{\frac{K_{i,b}}{L_{i,b}}} + \sqrt{\frac{\pi(e^{2\frac{K_{i,b}}{L_{i,b}}}-1)}{2L_{i,b}}}$, and $k_3 = 1 - e^{\frac{K_{i,b}}{L_{i,b}}} - \sqrt{\frac{\pi(e^{2\frac{K_{i,b}}{L_{i,b}}}-1)}{2L_{i,b}}}$.

$$\begin{aligned}\Omega_{i,j}(t) \approx & 1 - g_1 \gamma_{i,j}(t) \left(1 + \frac{g_2}{\gamma_{i,j}(t)} + \frac{g_2^2}{2\gamma_{i,j}^2(t)}\right.\\ & \left. -1 - \frac{g_3}{\gamma_{i,j}(t)} - \frac{g_3^2}{2{\gamma_{i,j}}^2(t)}\right),\end{aligned}$$

where $g_1 = \sqrt{\frac{L_{i,j}}{2\pi(e^{2\frac{K_{i,j}}{L_{i,j}}}-1)}}$, $g_2 = 1 - e^{\frac{K_{i,j}}{L_{i,j}}} + \sqrt{\frac{\pi(e^{2\frac{K_{i,j}}{L_{i,j}}}-1)}{2L_{i,j}}}$, and $g_3 = 1 - e^{\frac{K_{i,j}}{L_{i,j}}} - \sqrt{\frac{\pi(e^{2\frac{K_{i,j}}{L_{i,j}}}-1)}{2L_{i,j}}}$.

After simplification, we can obtain the below expressions.

$$\Omega_{i,b}(t) \approx \frac{(e^{\frac{K_{i,b}}{L_{i,b}}} - 1)\left(\sum_{k=i+1}^{I} |h_{k,b}(t)|^2 p_{k,b}(t) + \sigma^2\right)}{p_{i,b}(t)\,|h_{i,b}(t)|^2},$$

$$\Omega_{i,j}(t) \approx \frac{(e^{\frac{K_{i,j}}{L_{i,j}}} - 1)\sigma^2}{p_{i,j}(t)\,|h_{i,j}(t)|^2}.$$

□

## APPENDIX D
## PROOF OF LEMMA 4

**Lemma 4.** *The expected service reliability incurred by PA-TRLF could achieve the effect that has a gap with the maximal service reliability, which can be described below.*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t_x=0}^{T-1} E\{\Upsilon(t_x)|\Theta(t_x)\} \leq \Upsilon^{\sup}(t) - \frac{B}{V},$$

*where $\Upsilon^{\sup}(t)$ is the supremum of the average service reliability achieved by any joint strategy under the required constraints, and $B$ is mentioned in **Lemma 1**.*

*Proof.* From **Lemma 2**, we can get the below expression.

$$\begin{aligned}&\triangle L(t) - VE\{\Upsilon(t)|\Theta(t)\} \leq B + Q(t)E\{\sum_{i=1}^{I} x_{i,b}(t) R_{i,b}(t)\\ &- \mu_b t_a |\Theta(t)\} + H(t)E\{\sum_{i=1}^{I} x_{i,b}(t) R_{i,b}(t)\\ &- 1_{Q(t)>0}\mu_c t_a |\Theta(t)\} - VE\{\Upsilon(t)|\Theta(t)\}.\end{aligned}$$

$B$ is the upper bound, which is mentioned in **Lemma 1**. From **Section 3.1.4** in the study [16], we have the following inequality.

$$B - V\Upsilon^{\sup}(t) \leq \triangle L(t) - VE\{\Upsilon(t)|\Theta(t)\}.$$

By adding $t = 0, 1, \ldots, T-1$ cumulatively, we get:

$$\begin{aligned}TB - TV\Upsilon^{\sup}(t) \leq & E\{L(\Theta(t)) - L(\Theta(0))\}\\ & - V\sum_{t=0}^{T-1} E\{\Upsilon(t)|\Theta(t)\}.\end{aligned}$$

Since $L(\Theta(t)) \geq 0, L(\Theta(0)) = 0$, both sides divide by the same $-TV$, when $t \to +\infty$, we can obtain that:

$$\lim \sup_{t\to+\infty} \frac{1}{t} \sum_{t=0}^{T-1} E\{\Upsilon(t)\} \leq \Upsilon^{\sup}(t) - \frac{B}{V}.$$

With Jensen's inequality, we can get:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} E(\Upsilon(t)) \leq \Upsilon^{\sup}(t) - \frac{B}{V}.$$

□

## Appendix E
## Proof of Lemma 5

**Lemma 5.** *The virtual task queue in the edge cloud and the virtual delay queue for all TaVs in NOMA-enabled VEC are stable, which can be presented as follows.*

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t_x=0}^{T-1}\left(E\{Q(t_x)\}+E\{H(t_x)\}\right)\le\frac{V(\hat{\Upsilon}(t)-\check{\Upsilon}(t))-B}{\varsigma},$$

*where $\hat{\Upsilon}(t)$ and $\check{\Upsilon}(t)$ are defined as the maximum and minimum service reliability among all strategies. $B$ is mentioned in **Lemma 1**.*

*Proof.* We now prove that $Q(t)$ and $G_i(t)$ are mean rate stable under Algorithm 3.

Denote $\Phi(t)=TV\{\hat{\Upsilon}(t)\}-V\sum_{t=0}^{T-1}E\left\{\Upsilon(\bar{t})|\Theta(t)\right\}-TB+E\left\{L(\Theta(0))\right\}$.

Based on **Lemma 3**, we can get $E\{L(\Theta(t))\}\le\Phi(t)$, i.e., $E\{Q(t)^2\}\le 2\Phi(t)$ and $E\{H(t)^2\}\le 2\Phi(t)$. Thus, we have the following inequality.

$$0\le E\{Q(t)\}\le\sqrt{2\Phi(t)}.$$
$$0\le E\{H(t)\}\le\sqrt{2\Phi(t)}.$$

Since $\lim_{T\to\infty}\frac{\sqrt{2\Phi(t)}}{T}=0$, we can also acknowledge that

$$\lim_{T\to\infty}\frac{E\{Q(t)\}}{T}=0.$$
$$\lim_{T\to\infty}\frac{E\{H(t)\}}{T}=0.$$

To find the upper bound of the average sum queue length of $E\{Q(t)\}$ and $E\{H(t)\}$, we can obtain the below expression in the proof of **Lemma 3**,

$$\begin{aligned}&L(\Theta(t+1))-L(\Theta(t))-VE\{\Upsilon(t)|\Theta(t)\}\ge\\&B-V\Upsilon^{\sup}(t)+\left[Q(t)\left(\sum_{i=1}^{I}x_{i,b}(t)R_{i,b}(t)-\mu_b t_a\right)\right.\\&\left.+H(t)\left(\sum_{i=1}^{I}x_{i,b}(t)R_{i,b}(t)-1_{Q(t)>0}\mu_c t_a\right)\right],\end{aligned}$$

where $\Upsilon^{\sup}(t)$ is the supremum of the average service reliability achieved by any joint strategy under the required constraints.

Then, summing over $t\in\{0,1,2,\cdots,T-1\}$, we can obtain the expression below.

$$\begin{aligned}&L(\Theta(T))-L(\Theta(0))-\sum_{t=0}^{T-1}VE\{\Upsilon(t)|\Theta(t)\}\ge\\&TB-TV\Upsilon^{\sup}(t)+\varsigma\left(E\{H(t)\}+E\{Q(t)\}\right),\end{aligned}$$

where $\varsigma>0$, and after dividing $T$ and $\varsigma$ by both sides of the above inequality, we can finally get.

$$\begin{aligned}&\frac{1}{T}\left(E\{H(t)\}+E\{Q(t)\}\right)\le\frac{E(L(T))}{\varsigma T}-\frac{E(L(0))}{\varsigma T}\\&-\frac{B-V\Upsilon^{\sup}(t)+\frac{1}{T}\sum_{t=0}^{T-1}VE\{\Upsilon(t)|\Theta(t)\}}{\varsigma},\end{aligned}$$

$$\Upsilon^{\sup}(t)-\frac{1}{T}\sum_{t=0}^{T-1}E\{\Upsilon(t)|\Theta(t)\}\le\hat{\Upsilon}(t)-\check{\Upsilon}(t).$$

From the predefined value $L(0)$, we can get the conclusion.

$$\begin{aligned}&\lim_{T\to\infty}\frac{1}{T}\sum_{t_x=0}^{T-1}\left(E\{H(t_x)\}+E\{Q(t_x)\}\right)\\&\le\frac{V(\hat{\Upsilon}(t)-\check{\Upsilon}(t))-B}{\varsigma}.\end{aligned}$$

□