Credit Pre-Reservation Mechanism for Mobile Prepaid Service

Jyoti Bhosale
Department of Computer Science Engineering
VDF school of Engg,Latur,Maharashtra, India

Priyanka Pawar
Department of Computer Science Engineering
M.S.Bidve College of Engg,Latur,Maharashtra, India

Abstract— Online Charging System (OCS) supports multiple prepaid and postpaid sessions simultaneously. Through credit reservation, the OCS assigns some credit units to a session. These credit units are decremented based on the traffic volume or the duration time. If the assigned credit units are consumed before the session is completed, reserve units (RU) operation is executed to obtain more credit units from the OCS. If the credit at the OCS is depleted, the prepaid session is forced to terminate. During the RU operation, packet delivery is suspended until extra credit units are granted from the OCS. To avoid session suspension during credit reservation, we propose the credit pre-reservation mechanism (CPM) that reserves credit earlier before the credit at the GGSN is actually depleted. Analysis and simulation experiments are conducted to investigate the performance of the mechanism. Our study indicates that the CPM can significantly improve the performance of the OCS prepaid mechanism.

Keywords— Credit reservation, diameter protocol, online charging, prepaid service, Universal Mobile Telecommunications System.

I. INTRODUCTION

Pricing, charging and billing are important activities in telecommunications [1], [2], [3]. Advanced mobile telecommunications operation incorporates data applications (specifically, mobile Internet applications [4], [5]) with real time control and management, which can be archived by a convergent and flexible *Online Charging System* (OCS) [6], [7], [8]. Such convergence is essential to mitigate fraud and credit risks, and provide more personalized advice to users about charges and credit limit controls. The OCS allows simultaneous prepaid and postpaid sessions to be charged in real-time. This feature is important for a telecom operator to deliver multiple sessions simultaneously. Through online charging, the operator can ensure that credit limits are enforced and resources are authorized on a per-transaction basis.

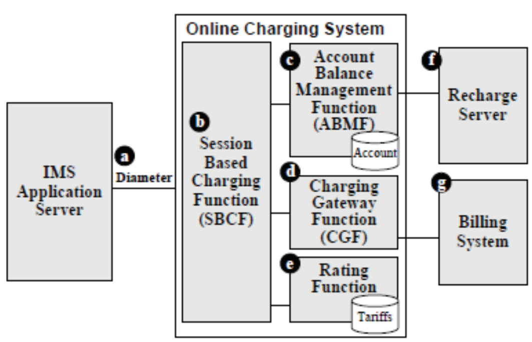


Fig. 1. OCS architecture for mobile telecommunications services

By merging the prepaid and postpaid methods, the OCS proposed in Universal Mobile Telecommunications System (UMTS) provides two-way communications between network nodes and the OCS to transfer information about rating, billing and accounting illustrates how UMTS integrates with the OCS [9]. Fig. 1 shows the OCS architecture for IMS services [2].In this architecture, online charging for the IMS services is performed by using the Diameter Credit Control (DCC).

Protocol (see Fig. 1 (a)) [1]. The OCS provides the *Session Based Charging Function* (SBCF; Fig. 1 (b)) responsible for online charging of network bearer and user sessions. In the OCS, the *Account Balance Management Function* (ABMF; Fig. 1 (c)) keeps a user's balance and other account data. When the prepaid user's credit depletes, the ABMF connects the Recharge Server (Fig. 1 (f)) to trigger the recharge account function. The SBCF interacts with the Rating Function (Fig. 1 (e)) to determine the price of the requested service. The rating function handles a wide variety of ratable instances, such as data volume, session connection time and event service (e.g. for web content charging). The SBCF The paper is organized as follows. Section 2 describes the Diameter credit control mechanism exercised at the interface. Section 3 proposes the CPM. Section 4 develops an analytical model for the CPM. Section 5 uses numerical examples to investigate the performance of the CPM. Section 6 gives our conclusions.

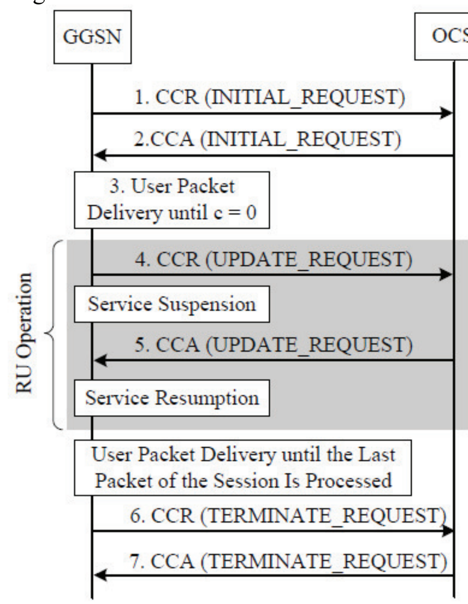


Fig. 2. Message flow of the diameter credit control mechanism.

II. DIAMETER CREDIT CONTROL MECHANISM

In online charging services, credit control is achieved by exchanging Diameter Credit Control Request (CCR) and Credit Control Answer (CCA) messages through the interface defined between the GGSN and the OCS [10]. The message flow of the online charging is illustrated in Fig. 2 with the following Steps:

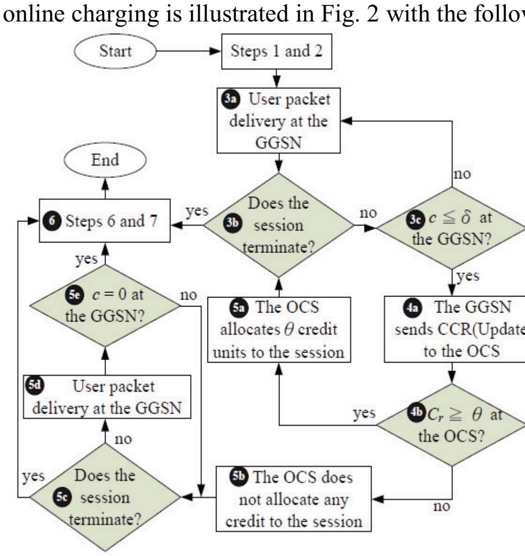


Fig. 3. Flowchart of the CPM (steps 1, 2, 6, and 7 are the same as those in the diameter credit control mechanism).

Step 1. To initiate the service session, the GGSN sends a CCR message with type INITIAL REQUEST to the OCS.

Step 2. Through the interactions among the SBCF, the ABMF and the RF, the OCS determines the tariff of the requested service session and then reserves the amount θ of credit units for the service session if the credit balance Cr at the OCS is no less than θ (i.e., $Cr \geq \theta$). Then the OCS returns a CCA message with type INITIAL REQUEST to the GGSN. This message indicates the amount θ of granted credit to the session.

Step 3. After receiving the CCA message, the GGSN starts to deliver the user packets. Let c be the amount of the remaining credit of the service session. This credit is decremented at the GGSN. At the beginning of this Step,

$c = \theta$. During the session, the granted credit units may be depleted (i.e., $c = 0$). If so, the session is suspended and Step 4 is executed.

Step 4. The GGSN sends a CCR message with type UPDATE REQUEST to the OCS to report the used credit units and request for additional credit units. Packet delivery is suspended, and any newly arriving packets are buffered at the GGSN.

Step 5. The OCS deducts the amount of used credit from the user account and reserves extra amount θ of credit units for the session. Then the OCS acknowledges the GGSN with a CCA message to indicate that θ credit units have been reserved. The session is resumed at the GGSN and the buffered packets are delivered. Note that Steps 3 - 5 may repeat many times before the service session is complete. At Steps 2 and 4, if $Cr < \theta$ at the OCS, then the session is forced to terminate. For the discussion purpose, Steps 4 and 5 are called an *RU operation*.

Step 6. When the session is completed, the GGSN sends

a CCR message with type TERMINATE REQUEST to terminate the session and report the amount of used credit.

Step 7. The OCS deducts the amount of the used credit from the account. Then the OCS acknowledges the reception of the CCR message by sending a CCA message with type TERMINATE REQUEST.

In the above procedure, when the credit units are depleted at the GGSN, an RU operation is executed to request for additional credit units. The session must be suspended until extra credit units are granted from the OCS. Delayed processing of user packets at the GGSN may seriously degrade the quality of service. To resolve this issue, we propose the credit prereservation mechanism in the next section.

III. CREDIT PRE-RESERVATION MECHANISM

In the *credit pre-reservation mechanism* (CPM), we define a threshold δ. When the amount c of the remaining credit for the session is not larger than δ (i.e., $c \leq \delta$), the GGSN conducts an RU operation to request extra credit from the OCS. During the RU operation, the GGSN continues to process the user packets. Hopefully, the GGSN will receive the extra amount θ of credit units before $c = 0$, and therefore the user packets need not be buffered (i.e., they are not suspended for processing) at the GGSN. Fig. 3 illustrates the flowchart of CPM, which modifies Steps 3-5 in Fig. 2 as follows:

Step 3a. The GGSN delivers the user packets and deducts the reserved credit units.

Step 3b. If the processed packet is the last one of the service session, then Step 6 is executed to terminate the session(the session is successfully completed).Otherwise, the execution proceeds to Step 3c.

Step 3c. Let δ be the CPM threshold. If $c \leq \delta$, Step 4a is

executed. Otherwise, the execution proceeds to Step 3a.

Step 4a. The GGSN sends a CCR message with type UPDATE REQUEST to request for additional credit. During the RU operation, if $c > 0$, the user packets are continuously delivered at the GGSN. When $c = 0$, the session is suspended and the newly arriving packets are buffered.

Step 4b. If the OCS does not have enough credit units (i.e., $Cr < \theta$), Step 5b is executed. Otherwise, Step 5a is executed.

Step 5a. The OCS sends the CCA message to the GGSN to indicate that extra amount θ of credit units have been reserved for the session. Then the execution proceeds to Step 3b. If the last packet arrives during the RU operation, the termination operation (Steps 3b and 6) is executed after the GGSN has received the CCA message. (This is called *delayed termination*) In this case, the session is successfully completed.

Step 5b. The OCS sends the CCA message to the GGSN.

This message indicates that no credit is reserved for the session.

Step 5c. If the previously processed packet is the last one of the session, then the session is successfully completed.

Step 6 is executed.

Step 5d. The GGSN continues to deliver the user packets.

Step 5e. If $c = 0$, then the session is forced to terminate, and Step 6 is executed. Otherwise, the execution proceeds to Step 5c.

Step 6. The session terminates by executing Steps 6 and 7 in Fig. 2. In the CPM, if δ is set too small, then the credit units for a session are likely to be depleted and the session must be suspended during the RU operation. On the other hand, if is set too large, many credit units are reserved in the active sessions, and the credit in the OCS is consumed fast. In this case, an incoming session has less chance to be served, and an in-progress session is likely to be force-terminated. Therefore, it is important to select an appropriate δ value to "optimize" the CPM performance.

IV. ANALYTIC MODEL FOR THE CPM

We describe an analytic model to investigate the CPM performance. We assume that the prepaid session arrivals for a user form a Poisson process with rate γ. The inter-arrival time between two packet arrivals has the mean $1/\lambda$. The round-trip transmission delay for the RU operation (i.e., the round-trip message delay for the CCR and CCA message pair) has the mean $1/\mu$. An arrival packet is the last one of the session with probability α; in other words, the session continues with probability $1 - \alpha$, and the expected number of packets delivered in a session is $1/\alpha$ Initially, a user has C credit units at the prepaid account in the OCS. Without loss of generality, we assume that each user packet consumes one credit unit. Define a *low credit* (LC) *period* as an interval such that during this interval, $c \leq \delta$ for a session. At the beginning of an LC period, the session initiates an RU operation. If more than θ packets arrive during this RU operation, then $\theta - \delta$ packets will be buffered at the GGSN.

Consequently, at the end of the RU operation, another RU operation must be issued to obtain more credit units to absorb the buffered packets and to ensure that $c > \delta$ after the buffer is empty. Before an LC period ends, the RU operation may be executed for several times until the session has reserved more than δ credit units. The output measures investigated in our study are listed below.

B**:** the average number of packets buffered during an RU operation

W**:** the average packet waiting time

P_r**:** the probability that during an LC period, two or more RU operations are executed

P_{nc}**:** the probability that a session is not completely served; i.e., the probability that a new session request is blocked or an in-progress session is forced to terminate Xs**:** the average number of the RU operations performed in a session To derive Pr and B, we first consider the case where α=0; i.e., a session is never terminated. Let K be the number of packets arriving in one RU operation (excluding the first packet arrival that triggers the RU operation). It is clear that

$$Pr = \Pr[K \geq \theta] \qquad (1)$$

We assume that an RU operation delay has the Erlang density function (t) with the shape parameter b = 2 and the scale parameter $h = 1/\mu$. (I.e., t is the summation of two Exponential delays. This assumption will be relaxed, and more general distributions will be considered in the simulation model.) Therefore the Laplace-Stieltjes Transform $f*$ (s) of the RU operation delay is

$$f^*(s) = \left(\frac{\mu}{\mu+s}\right)^2 \qquad (2)$$

For α = 0, the probability that $K = k$ can be calculated as follows:

$$\Pr[K = k, \alpha = 0] = \int_{t=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t}(t)dt \qquad (3)$$

$$= \left(\frac{\lambda^k}{k!}\right)\int_{t=0}^{\infty} t^k e^{-}(t)dt \qquad (4)$$

$$= \left(\frac{\lambda}{k!}\right. \quad (-1 \quad \left[\frac{d^k f^*(}{ds^k}\right. \qquad (5)$$

$$= \left(\frac{\lambda}{k!}\right. \quad (-1 \quad \left[\frac{d^k}{ds^k}\left(\frac{\mu}{\mu+2}\right.\right. \qquad (6)$$

$$= \frac{\lambda^k (k+1}{(\lambda+\mu)^k} \qquad (7)$$

In (3), the RU operation delay is t with the probability $f(t)dt$. During period t, there are k packet arrivals following the Poisson distribution with the rate λ. Eq. (5) is derived from (4) using Rule P.1.1.9 in [11]. Substitute (2) in (5), we obtain (7).

Now we consider the case when $\alpha \geq 0$. If a session is terminated during an RU operation, $\Pr[K = k]$ is derived by considering the following cases:

(I) When k = 0, we have $\Pr[K = 0] = \Pr[K = 0, \alpha = 0]$

(II) When $k > 0$, there are two sub cases:

(II a) There are exactly k packet arrivals during an RU operation (with probability $\Pr[K = k,\ \alpha = 0]$) and the session is not terminated by any of the first $k - 1$ packets (with the probability $(1 - \alpha)^{k-1}$). Note that the k-th packet can be the last one of the session.

(II b) There are more than k packet arrivals during an RU operation if the session is never terminated (with probability $\sum_{i=k+1}^{\infty}\Pr[K = i, \alpha = 0]$) and the session is actually terminated at the k-th packet arrival (with probability $(1 - \alpha)^{-1}\alpha$).

Based on the above cases, $\Pr[K = k]$ is derived as

$$\Pr[K = k] = \Pr[K = 0,\ \alpha = 0], \qquad k=0 \qquad (8)$$

$$\Pr[K = k] = \Pr[K = 0,\ \alpha = 0]\,(1-\alpha)^{k-1} + \sum_{i=k+1}^{\infty}\Pr[K = 0, \alpha = 0]\,(1-\alpha)^{k-1}\alpha \quad k>0 \quad (9)$$

From (7), (8) can be derived as

$$\Pr[K = 0] = \left(\frac{\pi}{\lambda+\pi}\right)^2 \qquad (10)$$

$$\Pr[K = k] = (k + 1)\mu^2 + \alpha\lambda\,[(k + 2)\mu + \lambda] \qquad (11)$$

The expected number B of buffered packets is derived as follows:

When the number k (of packet arrivals during an RU operation) is no less than the threshold δ (i.e., $k \geq \delta$), the session will have $k-\delta$ buffered packets. Therefore, from (11)

$$B = \sum_{k=\delta}^{\infty}(k-\delta)\Pr[K = k]$$

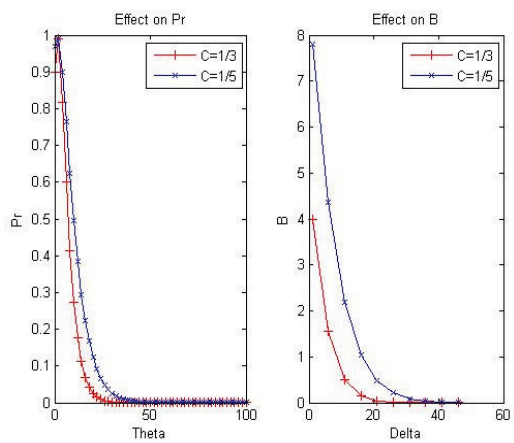
$$= \sum_{k=\delta}^{\infty}(k-\delta)\left[\frac{(1-\alpha)}{(\lambda+\mu)^{k+2}}^{k-1} * \frac{\lambda^k}{1}\right]$$

$$= \{(k+1)^2 + \alpha\lambda\,[(k+2) + \lambda]\}$$

$$= \left[\frac{(1-\alpha)\lambda}{\lambda+\mu}\right]^{\delta+1}$$

$$= \left[\frac{\delta\mu^2 + \alpha\delta\lambda\mu + 2\mu^2 + 2\lambda\mu + \alpha\lambda\mu + \alpha}{(1-\alpha)(\mu+\alpha\lambda)^2}\right. \qquad (14)$$

Equations (11) and (14) validate against the simulation experiments as illustrated in Fig. 4. In this figure, the dashed curves represent the simulation results, and the solid curves represent the analytical results. These curves indicate that the analytic and the simulation results are consistent.



(a) Effect on Pr (b) Effect on B

Fig. 4. Validation of simulation and analytical results on Pr and B ($\alpha = 0.01$, $\gamma/\mu = 1/20$, and $\delta = 0.3\theta$; solid curves: analytic results; dashed curves: simulation results).

V. CONCLUSION

This paper investigated the prepaid services for the UMTS network where multiple prepaid and postpaid sessions are simultaneously supported for a user. We described the prepaid network architecture based on

UMTS, and proposed the credit pre-reservation mechanism (CPM) that reserves extra credit at the GGSN is actually depleted.

REFERENCES

[1] P. Lin, H.-Y. Chen, Y. Fang, J.-Y. Jeng, and F.-S. Lu, "A secure mobile electronic payment architecture platform for wireless mobile networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2705-2713, 2008.

[2] J. Lau, and B. Liang, "Optimal pricing for selfish users and prefetching in heterogeneous wireless networks," in *Proc. IEEE International Conference on Communications*, June 2007.

[3] Y.-B. Lin and S.-I. Sou, Charging for Mobile All-IP Tele- communication. Wiley 2008.

[4] D. Wu, T. Hou, and Y.-Q. Zhang, "Transporting real-time video over the Internet: challenges and approaches" ,*Proc. IEEE*, vol. 88, no. 12, pp. 1855-1875, Dec. 2000.

[5] A.-C. Pang and Y.-K. Chen, "A multicast mechanism for mobile multimedia messaging service," *IEEE Trans. Veh. Technol.*, vol. 53, no.6, pp. 1891-1902, Nov. 2004.

[6] S.-T. Sou, H.-N. Hung, Y.-B. Lin, N.-F. Peng, and J.-Y. Jeng, "Modeling credit reservation procedure for UMTS online charging System", *IEEE Trans. Wireless Commun.*, vol. 6, no. 11), pp. 4129-4135, Nov. 2007.

[7] S.-I. Sou, Y.-B. Lin, Q. Wu, and J.-Y. Jeng, "Modeling prepaid Application server of VoIP and messaging services for UMTS",*IEEE Trans.Veh.Technol.*, vol. 56, no. 3, pp. 1434-1441, May 2007.

[8] S.-I. Sou, Y.-B. Lin, and J.-Y. Jeng, "Reducing credit reauthorization cost in UMTS online charging system," *IEEE Trans. Wireless Commun.*, vol. 7, no. 9, pp. 3629-3635, 2008.

[9] 3GPP, 3rd Generation Partnership Project; Technical Specification Group Service and System Aspects; Telecommunication management;Charging management; Online Charging System (OCS): Applications and interfaces (Release 10), 3G TS 32.296 version 10.0.0 (2010-03), 2010.

[10] H. Hakala, L. Mattila, J.-P. Koskinen, M. Stura, and J. Loughney, "Diameter credit control application," IETF RFC 4006, Aug. 2005.

[11] E. J. Watson, *Laplace Transforms and Applications*. Birkhauserk,1981.

[12] P. Lin, Y.-B. Lin, C.-H. Gan, and J.-Y. Jeng, "Credit allocation for UMTS prepaid service," *IEEE Trans. Veh. Technol.*, vol. 55 no. 1, pp. 306- 316,Jan. 2006.

[13] H.Y. Lee and Y.-B. Lin, "The simulation model for credit pre reservation mechanism," Technical Report NCTU-08-01, 2008.