

SIP ネットワークにおける EAPEC(Environment-Adaptive Personal Communications) の性能評価に関する検討

上岡英史 山田茂樹
国立情報学研究所

次世代モバイルコンピューティングネットワークに向けたシームレス通信アーキテクチャ「環境適応型パーソナル通信 (EAPEC : Environment-Adaptive Personal Communications)」は、SIP (Session Initiation Protocol) と 3GPP の all-IP ネットワークアーキテクチャを基本にいくつかの機能モジュールを追加することによって実現される。これらのアプリケーションサーバを経由したエンドエンド間の通信遅延はリアルタイム性が要求されるテレコム系サービスの品質に大きく影響する。本論文では、EAPEC におけるエンドエンド間のネットワークアーキテクチャをモデル化し、シミュレーションによって得られる呼設定遅延時間を PSTN の基準値と比較することによって、SIP ネットワークにおける EAPEC の性能評価に関する検討を行う。

Study on Performance Evaluation of Environment-Adaptive Personal Communications, EAPEC in SIP Networks

Eiji Kamioka and Shigeki Yamada
National Institute of Informatics

This paper describes the modeling and performance evaluation of a seamless communication architecture, called EAPEC (Environment-Adaptive Personal Communications), for next generation wide-area mobile computing networks. The EAPEC lays its foundation on the 3GPP-based all-IP network architecture and is implemented by adding functional components to the 3GPP-based all-IP network. To evaluate performance of the EAPEC architecture, it is modeled by integrating a queuing network model for application layer processing delay with a statistical model for IP network delay. Call setup delay, a performance measure important for EAPEC realtime applications, is obtained by mathematical analysis and simulation of the integrated model.

1. はじめに

次世代のモバイルコンピューティング・ネットワークにおいては、「いつでも、どこでも、誰でも」コンピュータを使うことができるユビキタスコンピューティングを念頭に入れたシステムが求められると考えられる。しかしながら、ユビキタスコンピューティングの重要な概念である「コンピュータの不可視性（コンピュータが人間の日常生活に自然に

溶け込み、事実上ユーザの視点からはコンピュータが見えなくなる）」を実現するためには、コンピューティングと通信に関する気遣いをユーザから開放し、ユーザによって通信端末が無意識のうちに利用される環境を作り上げる必要がある。我々は、このような目的を達成するためのアプローチとして、「環境適応型パーソナル通信 (EAPEC : Environment-Adaptive Personal Communications)」の提案を行ってきた [1]。

EAPEC は 3GPP (Third Generation Partnership Project) の all-IP ネットワークアーキテクチャに基づいて設計されており、テレコム系サービスの呼設定プロトコルとして SIP (Session Initiation Protocol) [2] を用いている。リアルタイム性が要求されるテレコム系サービスでは、エンドエンド間の遅延がサービスの品質を左右する重要な要素であり、その値が PSTN (Public Switched Telephone Network) で達成されている基準値と比較して大きく下回することは、ユーザのニーズを満たすことができないだけでなく、電話網の IP 化の発展に対する妨げにもなりかねない。

このような背景のもとに、我々は EAPEC のテレコム系サービスにおける性能評価としてエンドエンド間の呼設定遅延に着目し、PSTN の基準値と比較することによってどのような要因が 3GPP の all-IP ネットワークに影響を及ぼすことになるか、そして、all-IP ネットワークを構築する際に注意しなければならない点に関する検討を行った。呼設定遅延時間の取得に関しては、EAPEC におけるエンドエンド間のネットワークアーキテクチャをモデル化し、ネットワーク性能評価シミュレーションツール [3] を用いてシミュレーションを行った。

以降の章では、EAPEC の概要および 3GPP の all-IP SIP ネットワークアーキテクチャを示し、性能評価方法および EAPEC のモデル化について述べ、得られた結果について議論する。

2. EAPEC の概要

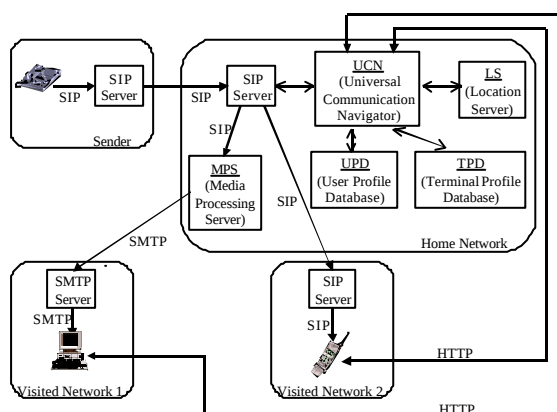


図1 EAPEC のネットワークアーキテクチャ

図1は SIP ネットワーク上に実装された EAPEC のアーキテクチャを表す。EAPEC は、3つのデータ

ベースモジュール、LS (Location Server)、UPD (User Profile Database) および TPD (Terminal Profile Database) と、2つの制御モジュール、UCN (Universal Communication Navigator) および MPS (Media Processing Server) で構成される。

LS はユーザの位置情報を管理し、UPD はユーザに関する認証情報、課金情報およびプリファレンス情報を管理する。また、TPD はユーザが利用する通信端末の情報を管理する。UCN はこれら3つのデータベースを参照し、ユーザのネットワーク利用環境に適した通信端末および通信メディアを選択し、その制御を行う。MPS は送信者から受信者に送られるコンテンツをインターセプトし、UCN の制御にしたがって通信メディア変換または通信サービス変換を行い、受信端末へそのコンテンツを転送する。各モジュールの役割、機能は文献 [1] に詳しい。

3. 3GPP の all-IP SIP ネットワーク

3.1 SIP サーバの構成

図1においては、SIP を用いた呼設定を制御するサーバを単に SIP Server と記述してあるが、3GPP の all-IP ネットワークはその仕様は複雑化されており、機能の異なる4種類の SIP サーバ (HSS, P-CSCF, S-CSCF, I-CSCF) によって構成される [4]。それらの機能は以下のとおりである。

1. HSS (Home Subscriber Server)
ユーザの位置情報を管理する LS (Location Server) とユーザに関する情報を記憶する PDB (Profile Database) からなるサーバである [5]。
2. P-CSCF (Proxy Call State Control Function)
ユーザ端末が最初にコンタクトするサーバで、呼設定要求 INVITE リクエストや位置登録要求 REGISTER リクエストなど、端末からのすべてのリクエストは、まず P-CSCF に送られる。
3. S-CSCF (Serving CSCF)
個々の呼の状態を管理するとともに LS に自分自身 (S-CSCF 名) を登録する機能も持つ。ユーザ名とそのユーザを管理する S-CSCF 名の情報ペアが LS に記憶される。
4. CSCF (Interrogating CSCF)
ユーザからの REGISTER リクエストに対して呼の状態を管理する S-CSCF を新規に割り当て、INVITE リクエストに対して LS に位置情

報 (S-CSCF 名) を問い合わせる機能を持つ。

以下、これらの SIP サーバを利用した呼設定手順について説明する。

3.2 呼設定要求手順

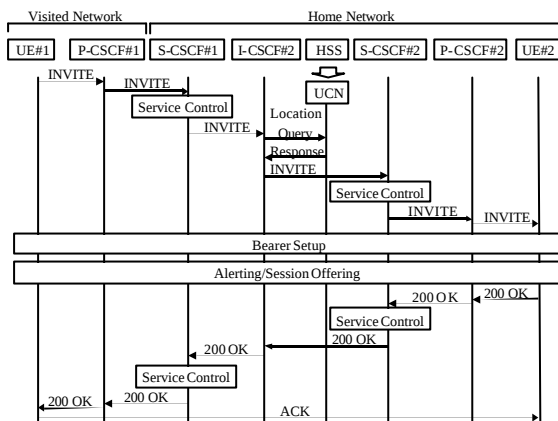


図2 3GPP の all-IP ネットワークアーキテクチャに基づく呼設定フロー

図2は送信者が Home Network にいて、受信者が Visited Network にいる場合の SIP サーバを経由した呼設定フローを示す (なお、この呼設定フローは 2000 年 10 月時点で 3GPP が Draft として提示したものである [6])。UE#1 はあるユーザが Visited Network (移動先ネットワーク) 内で使用する通信端末で、UE#2 は Home Network 内の通信端末を示す。この図では、受信者のアドレスを管理する S-CSCF#2 は Home Network 内にある場合を示した。送信端末からの INVITE リクエストが送信側 P-CSCF (P-CSCF#1)、送信側 S-CSCF (S-CSCF#1)、受信側 I-CSCF (I-CSCF#1)、受信側 HSS、受信側 S-CSCF (S-CSCF#2)、受信側 P-CSCF (P-CSCF#2) を経由して受信端末に到達し、これと全く逆の経路をたどって受信端末が送信端末に 200 OK レスポンスを返すことによって呼が確立される。

このように呼設定フローを考える場合、図2に示すように、EAPEC の UCN 部分は 3GPP の仕様でいう HSS の拡張モジュールとして all-IP ネットワーク内に実装される。すなわち、EAPEC が提供するサービスは 3GPP の SIP ネットワークに円滑に導入することができることを示す。なお、図2に示す UCN は、それがアクセスする 3 つのデータベース、LS、UPD および TPD をも含めて表している。

4. 性能評価方法と EAPEC のモデル化

4.1 性能評価方法

EAPEC をネットワーク的な視点から性能評価を行う場合、ネットワーク・スループット、エンド・ツー・エンド遅延、パケット損失等いくつかの評価項目が存在するが、このうちエンド・ツー・エンド遅延はリアルタイム・テレコム系サービスにおける呼設定遅延時間に大きな影響を及ぼす。ここでは EAPEC における呼設定遅延時間に着目し、PSTN (Public Switched Telephone Network) におけるそれと比較することによって性能評価を行う。

PSTN における電話サービスの場合、呼設定遅延時間は送信者が最後の数字をダイヤル (プッシュ) してからリングバックトーンを受けるまでの時間と定義されている。ITU-E.721 では、呼設定遅延時間の平均値は市内呼、市外呼、国際呼でそれぞれ 3.0 秒、5.0 秒、8.0 秒以内で、95 パーセンタイル値はそれぞれ、6.0 秒、8.0 秒、11.0 秒以内であることが推奨されている。EAPEC における呼設定遅延時間は、送信端末から INVITE リクエストが送り出されてから、Ringing レスポンスを受信するまでの時間に対応する。以下、これらのネットワーク内における振る舞いをモデル化する方法を示し、理論計算およびシミュレーションによって得られる呼設定遅延時間の結果について議論する。

4.2 EAPEC のモデル化

EAPEC における呼設定遅延時間は、サーバにおけるメッセージの処理待ち合わせに伴う遅延と IP ネットワーク遅延の組み合わせによって表すことができる。前者は SIP サーバ内で INVITE リクエストと Ringing レスポンスを処理するためのアプリケーション層レベルの遅延、後者は IP パケットがルータ、SIP サーバ、回線等を経由することによって生じるトランスポート層以下の遅延である。

各 SIP サーバと UCN における SIP メッセージの処理遅延は、その挙動を待ち行列モデルで表現することによって遅延時間を求めることができる。一方、IP ネットワーク遅延に関しては、混雑していないインターネットにおける片方向遅延時間分布は対数正規分布あるいは正規分布に近いことが実トラヒックを用いた統計的分析によって明らかにされている [7]。したがって、これらのモデルを用いて EAPEC における呼設定遅延時間の評価を行う。

図3は 3GPP の IP ネットワークアーキテクチャ

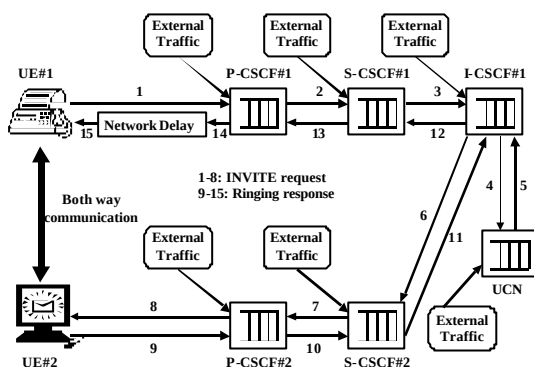


図3 EAPCにおける呼設定遅延モデル

に基づく EAPC の呼設定フローの評価モデルを示す．呼設定フローを考える場合，図2の HSS にあたる部分は 図3の UCN (Universal Communication Navigator [1]) の処理とみなすことができる．UE#1 から送信された INVITE リクエストは P-CSCF#1, S-CSCF#1 を経由して I-CSCF#1 に転送され，I-CSCF#1 は UCN へ 転送先アドレスを問い合わせる．I-CSCF#1 は UCN からそのリプライを受け取ると，INVITE リクエストは次段 SIP サーバである S-CSCF#2 へ 転送され，P-CSCF#2 を経由して UE#2 へ到着する．UE#2 は Ringing レスポンスを生成し，P-CSCF#2, S-CSCF#2, I-CSCF#1, S-CSCF#1, P-CSCF#1 を経由して UE#1 へと転送される．INVITE リクエストは UCN を経由して UE#2 に送られるが Ringing レスポンスはその必要がないため，UCN を経由しない．図3において，手順 1 から 15 にかかる時間が EAPC における呼設定遅延時間となる．

各 SIP サーバと UCN における待ち行列モデルにおいては，呼設定のための SIP メッセージ (INVITE リクエストおよび Ringing レスポンス) に伴うトラヒックと，処理遅延を引き起こす外部トラヒックを用意する．これら 2 つのトラヒックはポアソン分布に基づいて生成される．また，サーバにおけるサービス時間分布に関しては，指数分布モデルと一定分布モデルの 2 つを仮定する．一方，ネットワーク遅延に関しては，IP ネットワークがあまり混雑していないと仮定し，文献 [7] に示されるように片方向遅延分布は正規分布を用いる．それゆえ，それが足し合わされた往復の遅延分布も正規分布になることから，ネットワーク遅延は図3に示すように集約した形で表すことができる．

図3からわかるように，各 SIP サーバおよび UCN は論理的に直列に配置されることから，上記 2 種類

のサービス時間分布を用いた呼設定遅延モデルをそれぞれ M/M/1 タンデムモデル，M/D/1 タンデムモデルと呼ぶことにする．

M/M/1 タンデムモデルにおいては，各サーバは独立な M/M/1 待ち行列モデルとして取り扱うことが可能であり，全サーバにおける平均遅延時間は個々のサーバにおける平均待ち時間の和として容易に計算することができる．一方，M/D/1 タンデムモデルに関しては，ポアソン生起の外部トラヒックが呼設定用 SIP メッセージに伴うトラヒックに比べて十分多ければ，各サーバは独立な M/D/1 待ち行列モデルとして近似することができる．しかしながら，トラヒックの一般性を損なわずに全サーバにおける平均遅延時間を扱うためには，M/M/1 タンデムモデルのように単純ではないのでシミュレーションを用いることにする．また，全サーバにおける遅延時間の 95 パーセンタイル値を算出するためには，たまたみこみの複雑な計算を必要とするため，これに関しては M/M/1 タンデムモデル，M/D/1 タンデムモデルともにシミュレーションを用いることにする．

5. 性能評価結果

シミュレーションにおいて，片方向 IP ネットワーク遅延時間の平均値および標準偏差は公開されている実測ベースの値をもとに，それぞれ 50 ms, 7 ms とする [8]．また，各 SIP サーバおよび UCN は常に一定の使用率になるように負荷をかけ，ネットワークはトラヒック的にバランスのとれた状態を仮定する．

5.1 サーバの使用率と呼設定遅延の関係

図4は シミュレーションによって得られた M/M/1 タンデムモデルにおけるサーバの使用率と呼設定遅延時間との関係を表す．サーバの平均処理時間は 10 ms と 100 ms を設定し [9]，各々の呼設定遅延時間の平均値と 95 パーセンタイル値を示す．サーバの平均処理時間が 100 ms の場合，サーバ使用率が 0.7 を超えると呼設定遅延時間の平均値および 95 パーセンタイル値はともに E.721 で推奨される値 (3 秒 および 6 秒) を満たすことができなくなっている．しかしながら，平均処理時間が 10 ms と高性能なサーバとなると，サーバの使用率が 0.9 を超えても E.721 の推奨値を満足している．これは，3GPP の all-IP ネットワークアーキテクチャが多段の SIP サーバによって構成されていること

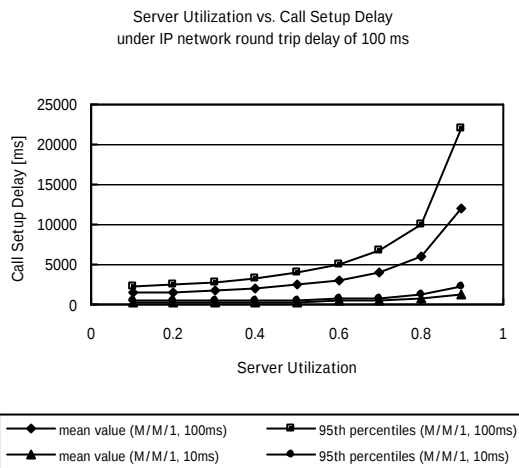


図4 サーバにおける平均処理時間を 100 ms と 10 ms に設定したときの M/M/1 タンデムモデルの呼設定遅延時間。

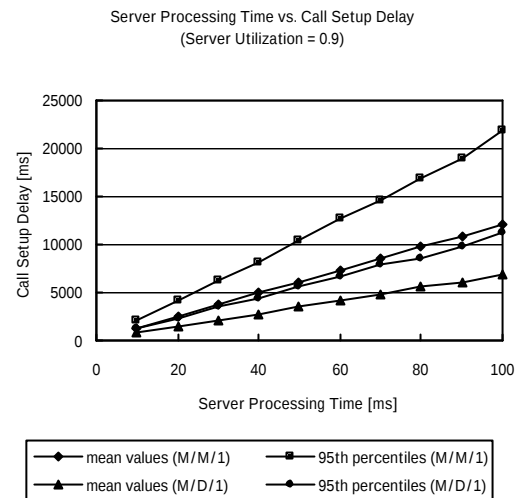


図6 サーバ使用率を 0.9 に設定したときの M/M/1 および M/D/1 タンデムモデルのサーバ処理時間と呼設定遅延時間の関係

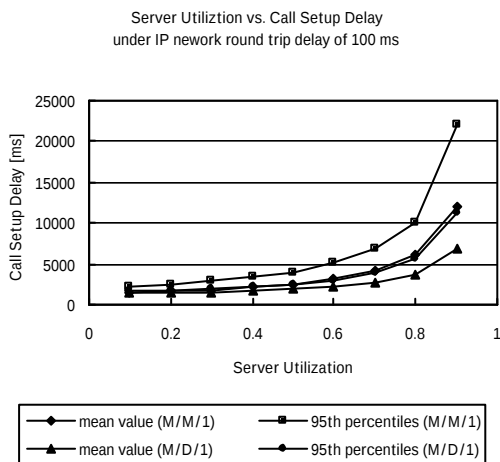


図5 サーバにおける平均処理時間を 100 ms としたときの M/M/1 および M/D/1 タンデムモデルの呼設定遅延時間

が大きく影響している。すなわち、たとえ 1 段のサーバにおける処理遅延が小さかったとしても、このモデルのように呼設定が確立するまでに 12 段ものアプリケーションサーバで同様の処理遅延が蓄積されるため、結果としてエンドエンド間では大きな処理遅延を生み出してしまう。

したがって、サーバの処理性能は呼設定遅延時間に大きな影響を与えるため、高使用率が見込まれるネットワークにおいては、高速なサーバの導入が不可欠である。

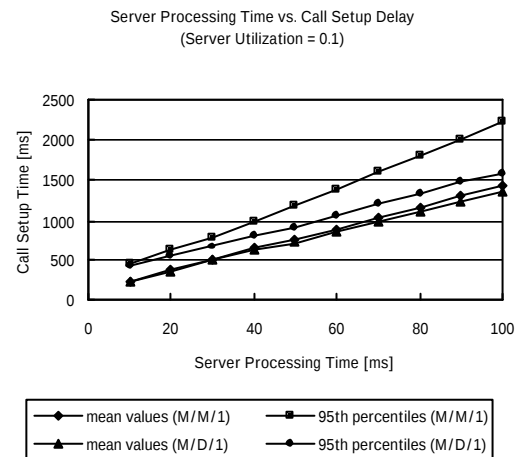


図7 サーバ使用率を 0.1 に設定したときの M/M/1 および M/D/1 タンデムモデルのサーバ処理時間と呼設定遅延時間の関係

5.2 サーバのサービス時間分布の違いによる呼設定遅延への影響

図5はサーバの使用率を 100 ms とした場合の M/M/1 タンデムモデルと M/D/1 タンデムモデルの図4と同様のシミュレーション結果を表す。図5は M/D/1 タンデムモデルが M/M/1 タンデムモデルに比べて効果的であることを示す。すなわち、3GPP の IP ネットワークアーキテクチャに基づくシステムサーバは、一定サービス時間分布に従う処理を実現するようなスケジューリングが有効であることを意

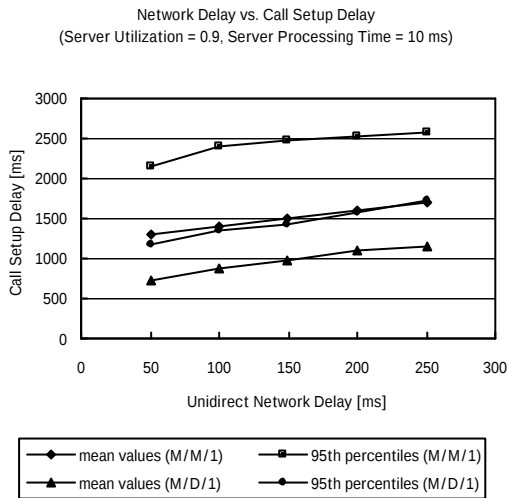


図8 サーバ処理時間を 10 ms、サーバ使用率を 0.9 としたときの M/M/1 および M/D/1 タンデムモデルの IP ネットワーク遅延と呼設定遅延時間の関係

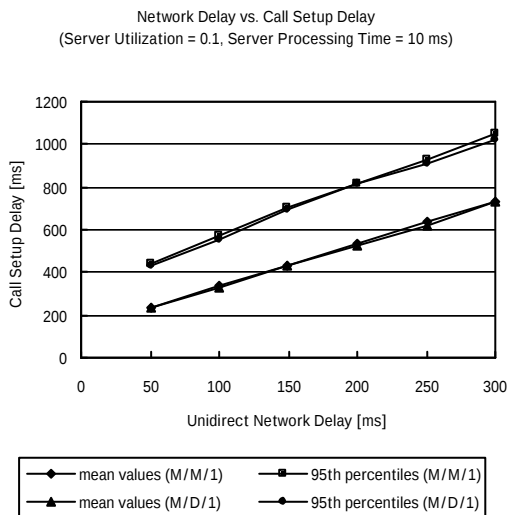


図9 サーバ処理時間を 10 ms、サーバ使用率を 0.1 としたときの M/M/1 および M/D/1 タンデムモデルの IP ネットワーク遅延と呼設定遅延時間の関係

味する。

5.3 サーバの処理性能と呼設定遅延の関係

図6および図7はサーバ使用率が 0.9 と 0.1 の場合の 1 メッセージあたりのサーバ処理時間と呼設定遅延時間の関係を表す。これらの結果は、サーバの使用率にかかわらずサーバの処理時間に対して呼設定遅延時間が直線的に増加していることを示す。また、図6からサーバの使用率が 0.9 と高い場合で

も、M/M/1 タンデムモデルでサーバ処理時間が 30 ms 以下、また、M/D/1 タンデムモデルで 50 ms 以下であれば呼設定遅延時間が平均値および 95 パーセンタイル値とともに E.721 の推奨値を満足することが読み取れる。このようにサーバ処理時間と呼設定遅延時間の関係が直線的であるという結果は、あるネットワークに導入するサーバを選択する際に非常に有用である。

5.4 IP ネットワーク遅延と呼設定遅延の関係

図8および図9はサーバ処理時間が 10 ms で、サーバ使用率が 0.9 と 0.1 の場合の片方向 IP ネットワーク遅延と呼設定遅延時間の関係を表す。これらの結果から、IP ネットワーク遅延に対して呼設定遅延時間はゆるやかに直線的に増加することがわかる。すなわち、IP ネットワーク遅延はサーバの使用率やサーバの処理時間と比べると、それほど呼設定遅延時間に影響しないということが出来る。また、図9では M/M/1 タンデムモデルと M/D/1 タンデムモデルとの違いはほとんど見られない。これは、図4に示されるように、サーバの処理時間が 10 ms と速い場合には、M/M/1 タンデムモデルと M/D/1 タンデムモデルとでは、サーバの使用率が低い時にはほとんど差がないことを反映している。

IP ネットワーク遅延がサーバの処理遅延に比べて呼設定遅延時間にあまり効いてこない理由として、やはり SIP サーバの段数が考えられる。すなわち、高々数百 ms の IP ネットワーク遅延に対して、12 段のサーバ処理遅延の影響は非常に大きい。

6. まとめ

EAPEC のテレコム系サービスにおける性能評価として、アプリケーションサーバにおける処理遅延と IP ネットワーク遅延をモデル化し、それらがエンドエンド間の呼設定遅延に与える影響を検討した。

サーバの使用率が高くなるにつれて、呼設定遅延は急激に増加するが、その際、処理速度の速い高性能なサーバを用いることによって、E.721 で推奨される遅延値を満足させることができ、また、サーバの性能がそれほど高くない場合でも、一定サービス時間分布に従う処理を実現するようなスケジューリングが可能であれば、呼設定遅延時間を減少させるのに効果があることを示した。

サーバの処理時間と呼設定遅延の関係がサーバの使用率にかかわらず直線的であるという結果は、あ

るネットワークに対してサーバを導入する際、最も適した性能のマシンを選択するのに有効である。

一方、IP ネットワーク遅延はサーバの使用率やサーバの処理時間と比べると、あまり呼設定遅延時間には影響しないことを示した。

以上のような結論は、3GPP の all-IP ネットワークアーキテクチャでは経由するサーバの段数が多いことが大きな原因になっていることがわかる。

これらの EAPEC に関する性能評価に関する検討は、そのアーキテクチャから 3GPP の all-IP SIP ネットワークにそのまま流用することができる。なぜなら、呼設定手順については、3GPP ネットワークアーキテクチャの HSS にあたる部分がそのまま EAPEC の UCN と置き換えることができるからである。HSS と UCN の内部的な処理の違いはあるにせよ、SIP を用いた呼設定手順としては全く同様である。

今回の検討によって、SIP サーバおよび EAPEC の機能モジュールとして動作するアプリケーションサーバをネットワークに実装する際、注意しなければならない点を明確にすることができ、また、3GPP の all-IP ネットワークを構築するにあたって有益な情報を取得することができた。

[8] T.J. Kostas, M.S. Borella, I. Sidhu, G.M. Schuster, J. Grabiec, and J. Mahler, "Real-Time Voice Over Packet-Switched Networks," IEEE Network, Vol.12, No.1, pp.18-27, January 1998.

[9] T. Evers and H. Schulzrinne, "Predicting Internet Telephony Call Setup Delay," Proc. of the First IP Telephony Workshop (IPTel2000), Berlin, April 2000.

参考文献

[1] 上岡英史, 山田茂樹, "SIP を用いた環境適応型パーソナル通信アーキテクチャの提案," 情報処理学会研究報告, 2001-MBL-16, pp.1-7, 2001.

[2] M. Handley, H. Schulzrinne, E. Schooler and J. Rosenberg, "SIP: Session Initiation Protocol," RFC 2543, March 1999.

[3] <http://opnet.com/home.html>.

[4] Lucent Technologies, "CSCF Descriptions," 3GPP TSG-SA Working Group 2 Draft S2-001710, October 2000.

[5] Ericsson, "Registration and Profile Transfer Message Flows," 3GPP TSG-SA Working Group 2 Draft S2-001106, June 2000.

[6] AT&T, "Call Flow - Mobile Termination, roaming, with home control," 3GPP TSG-SA Working Group 2 Draft S2-001723, 2000.

[7] 藤本康平, 阿多信吾, 村田正幸, "インターネットにおける実測に基づいたパケット転送遅延の統計的分析," 通学技報, IN2000-48, pp.41-48, July 2000.