

特別研究報告

題目

シグナリング処理負荷を考慮したノード仮想化及びプレーン分離を
適用したモバイルコアネットワークの性能評価

指導教員

松岡 茂登 教授

報告者

阿部 修也

平成 28 年 2 月 16 日

大阪大学 基礎工学部 情報科学科

シグナリング処理負荷を考慮したノード仮想化及びプレーン分離を
適用したモバイルコアネットワークの性能評価

阿部 修也

内容梗概

携帯電話加入者数の増加や高機能なスマートフォン等の普及により、第3世代移動通信システムや Long Term Evolution などのモバイルネットワークにおいて、輻輳への対応が課題となっている。加えて、モバイルネットワークの利用形態として Machine-to-Machine (M2M) や Internet of Things (IoT) に関わる通信を行う端末の、モバイルネットワークへの接続が注目されている。M2M/IoT 通信は、従来端末とはその通信特性が大きく異なり、通信データ量は小さいが、端末数が膨大になる。そのため、M2M/IoT 通信を行う端末を従来の携帯電話端末と同じ方式でモバイルネットワークに接続すると、データ通信に必要なシグナリングに関する通信や処理を行うコントロールプレーンの輻輳が悪化すると考えられる。

本報告では、今後拡大が予想される M2M/IoT 通信を収容するためのモバイルコアネットワークアーキテクチャとして、モバイルコアネットワークを構成するノードのデータプレーンとコントロールプレーンを分離し、一方、あるいは双方をクラウド環境へ設置するものを探り上げ、その性能評価を行った。さらに、クラウド環境へ設置したノードの処理性能をノード間で融通する場合や、モバイルコアネットワーク内で通常は端末毎に設定されるトンネルを集約する場合の性能評価を行った。具体的には、モバイルコアネットワークの各ノードにおけるシグナリング処理の重みを、各処理に関連するシミュレーションソフトウェア及び実アプリケーションソフトウェアのコード量から推定し、待ち行列理論に基づき、ノードにおけるシグナリング処理時間を推定することによって、端末がデータ通信を開始する際に必要となるベアラ確立にかかる時間を算出した。さらに、解析結果の数値例を基に、モバイルコアネットワークに収容可能な M2M/IoT 端末数の評価を行い、プレーン分離や処理機能のクラウド環境への設置、処理能力の融通、及び通信集約がモバイルコアネットワークの容量に与える効果を明らかにした。評価の結果、プレーン分離や処理機能のクラウド環境への設置、処理能力の融通及び通信集約によって M2M/IoT 端末の収容可能台数が最大 200% 増加することを示した。

主な用語

モバイルコアネットワーク, M2M/IoT 通信, Software Defined Networks (SDN), プレーン分離, 通信集約

目次

1	はじめに	5
2	モバイルコアネットワークモデル	7
2.1	モデル 1: MME をクラウドネットワーク内に設置したモデル	7
2.2	モデル 2: SGW/PGW の制御プレーンの機能をクラウドネットワーク内に設置したモデル	12
2.3	モデル 3: SGW/PGW の GTP モジュールをクラウドネットワーク内に設置したモデル	16
2.4	通信集約を行うモデル	20
2.4.1	モデル 3-A: UE が用いる GW 間ベアラを事前に決定するモデル	20
2.4.2	モデル 3-B: UE が用いる GW 間ベアラをオンデマンドに決定するモデル	24
3	性能解析	27
3.1	変数定義	27
3.2	ベアラ確立時間の導出	27
3.3	ノードにおけるシグナリング処理時間の導出	29
4	性能評価	30
4.1	パラメータ設定	30
4.2	評価結果と考察	31
5	まとめと今後の課題	39
	謝辞	40
	参考文献	41

目 次

1	モデル1のネットワーク構成	9
2	LENA を基にしたモデル1のシグナリング手順と処理命令文数	10
3	OAI を基にしたモデル1のシグナリング手順と処理命令文数	11
4	モデル2のネットワーク構成	13
5	LENA を基にしたモデル2のシグナリング手順と処理命令文数	14
6	OAI を基にしたモデル2のシグナリング手順と処理命令文数	15
7	モデル3のネットワーク構成	17
8	LENA を基にしたモデル3のシグナリング手順と処理命令文数	18
9	OAI を基にしたモデル3のシグナリング手順と処理命令文数	19
10	LENA を基にしたモデル3-Aのシグナリング手順と処理命令文数	22
11	OAI を基にしたモデル3-Aのシグナリング手順と処理命令文数	23
12	LENA を基にしたモデル3-Bのシグナリング手順と処理命令文数	25
13	OAI を基にしたモデル3-Bのシグナリング手順と処理命令文数	26
14	各モデルにおける収容 UE 数とベアラ確立時間の関係	33
15	LENA に基づく場合の通信集約の効果	35
16	OAI に基づく場合の通信集約の効果	36
17	通信集約手法と処理性能融通の組み合わせの効果	38

1 はじめに

モバイルネットワークに接続される携帯電話やスマートフォン等の端末数は著しく増加し、モバイルネットワークにおいて、輻輳への対応が課題となっている。さらに、モバイルネットワークの利用形態として M2M/IoT 端末 [1] の接続が着目されており、モバイルネットワークに接続される端末数は今後も増加の一途を辿ると考えられている。M2M/IoT 通信は、携帯電話やスマートフォンなどの従来端末の通信とはその特性が大きく異なり、通信データ量は小さいものの、周期性や間欠性をもち、さらに端末数が膨大になる。さらに、M2M/IoT 端末の多くは移動性を持たず、データを送信するが受信することは少ないという特徴を持つ。さらに、M2M/IoT 端末をモバイルコアネットワークに収容する場合の Average Revenue Per User (ARPU) は、従来の端末に比べて小さくなることを見込まれている [2]。これらのことから、M2M/IoT 端末を従来の携帯電話端末と同じ方式でモバイルネットワークに接続すると、データの送信に必要なシグナリングに関する通信や処理を行う制御プレーンの輻輳が悪化すると考えられる。

このような問題に対し、Serving Gateway (SGW)、PDN Gateway (PGW) 間に確立されるモバイルコアネットワークベアラを端末固有に設定するのではなく、グループで共用することでモバイルコアネットワーク内の資源利用を抑える方法 [3, 4] や、即時性を有するユーザの通信に比べて M2M/IoT 通信の遅延制約が緩いという特徴を用いて、各ノードの処理負荷を軽減するなどの対策 [5] が考えられている。さらに、文献 [6–9] では、モバイルコアネットワークのノードを仮想化 [10]、及び Software Defined Networks (SDN) [11] 化し、制御プレーンをクラウドネットワーク内に設置することでノードの処理負荷を減少させる方法について検討されている。これらの研究では、モバイルコアネットワークのノードの仮想化や SDN 化によってサーバ資源の利用効率の向上や低コスト化が可能であるとされているが、その具体的な評価は行われていない。また、仮想化及び SDN 化を行うことによって、モバイルコアネットワークにおける従来のシグナリング処理に加えて、仮想化された機能モジュール間のシグナリング処理や、SDN 制御のためのシグナリング処理が増加することが考えられる。特に、今後膨大な数の M2M/IoT 端末をモバイルコアネットワークに収容し、端末からの周期的な通信が集中的に発生すると、仮想化及び SDN 化によって増加したシグナリングオーバーヘッドが原因となり、収容可能な端末数の減少や、通信に発生する遅延時間が増大することが考えられる。

文献 [12] においては、SDN 化されたモバイルコアネットワークを対象とし、M2M/IoT 端末を収容する際のシグナリングオーバーヘッドやノード負荷を解析的に評価しているが、各ノードにおけるシグナリング処理負荷を、各ノードが送受信するメッセージ数のみに基づいて評価している。また、モバイルコアネットワークのノードの一つである Mobility Management

Entity (MME) ノードの仮想化効果を評価した文献 [6] においても, MME が送受信するメッセージ数のみを用いて評価を行っている. また文献 [13] では, メッセージ数ではなく送受信されるメッセージサイズを基に, MME におけるシグナリング負荷の評価を行っている. しかし実際には, シグナリング処理の種類によって異なる処理負荷が発生するため, これらの検討によって得られた結果は必ずしも妥当であるとはいえない.

そこで本報告では, モバイルコアネットワークの各ノードにおけるシグナリング処理負荷を考慮した, モバイルコアネットワークの性能評価を行う. 具体的には, シグナリング処理負荷の重みを, LTE/EPC ネットワークのシミュレーションソフトウェアである LENA [14], 及び実アプリケーションソフトウェアである Open Air Interface (OAI) [15] における, 各処理に関連する実装コードの命令文数から決定し, その重みをもとに, シグナリング処理によって発生する遅延時間を決定することにより, より精度の高い性能評価を行う. そのために, まず 2 章では, 評価の対象とする仮想化かつ SDN 化されたモバイルコアネットワークアーキテクチャについて複数のモデルを検討する. また, 2 つの異なる集約手法を用いて通信集約を行うモデルについても検討する. 3 章では, 2 章で検討したネットワークモデルに基づき, 実装コードの命令文数から, 待ち行列理論を用いた数学的解析によって各モデルにおけるシグナリングメッセージの総伝搬遅延時間及び各ノードにおける処理遅延時間を導出し, ベアラ確立時間を算出する解析的手法を示す. 4 章では解析結果を用いて, 各モデルにおいて収容可能な M2M/IoT 端末数を評価することによって, 各モデルの性能評価を行い, プレーン分離や処理性能の融通, また通信集約の有効性を示す. 最後に, 5 章でまとめと今後の課題について述べる.

2 モバイルコアネットワークモデル

文献 [12] において評価対象とされているモデルを参考に、本報告で性能評価を行うモバイルコアネットワークのモデルを検討する。また、User Equipment (UE) と呼ばれる M2M/IoT 端末が通信を開始する際に発生するシグナリング処理の違いや、仮想化及び SDN 化によって新たに発生する SDN 制御メッセージについて、それぞれのモデルを基に検討する。なお、LTE のシグナリング手順については文献 [16] に基づいており、SDN 化を行った際の仮想ネットワーク制御には OpenFlow を用いることを想定している。

また、各モデルにおいて UE の通信開始の際にノード間で発生するシグナリングメッセージや、各ノード及び UE がシグナリングメッセージを処理するために実行するプログラムの命令文数は、C++ で書かれた LTE/EPC ネットワークのシミュレータである LENA、及び C 言語で書かれた LTE/EPC ネットワークの実装の 1 つである OAI のソースコードとその命令文数を基にしている。これは、異なる言語やコーディング方針に基づいて記述された LTE/EPC ネットワークのソースコードを用いて評価を行うことで、それらが与える影響を明らかにするためである。また、ネットワークモデルの構築、及びシグナリング手順と処理負荷の決定については、文献 [17–20] を参考にした。

2.1 モデル 1: MME をクラウドネットワーク内に設置したモデル

図 1 に、制御プレーンのノードである MME、Home Subscriber Server (HSS)、Policy and Charging Rules Function (PCRF) をクラウドネットワーク内に設置した、モデル 1 のネットワーク構成を示す。従来の LTE/EPC ネットワーク構成は本モデルのようになっていると考えられる。

また、図 2 及び図 3 に、それぞれ LENA 及び OAI に基づく、モデル 1 における、UE がアイドル状態からアクティブ状態になり、通信開始要求を行う際のシグナリング手順とシグナリング経路、各ノード及び UE がシグナリングメッセージを処理するために実行されるプログラムの命令文数を示す。図 2 及び図 3 のシグナリング処理部分に記した数値は、その処理を行う部分に該当するソースコードの命令文数を表している。なお、図 3 において、UE で行われる Non-Access Stratum (NAS) に関するシグナリング処理については、MME で行われる NAS に関するシグナリング処理と同程度の処理が発生すると考え、同じ値を設定している。また、同様に、MME-SGW/PGW 間で行われるシグナリングメッセージの送受信に対するシグナリング処理に関しては、両ノードに共通して実装されている、GTPc と呼ばれるモジュールが処理を行っているため、送信時及び受信時の処理について、それぞれ一定の命令文数を設定している。また、本報告においては、LENA 及び OAI のネットワー

ク構成に従い, SGW 及び PGW についてはこれらを一体化した SGW/PGW という単一のノードとしてモデル化を行っている。ただし, OAI では UE-eNodeB 間に設定される Radio resource control (RRC) 接続に関しては省略されている。

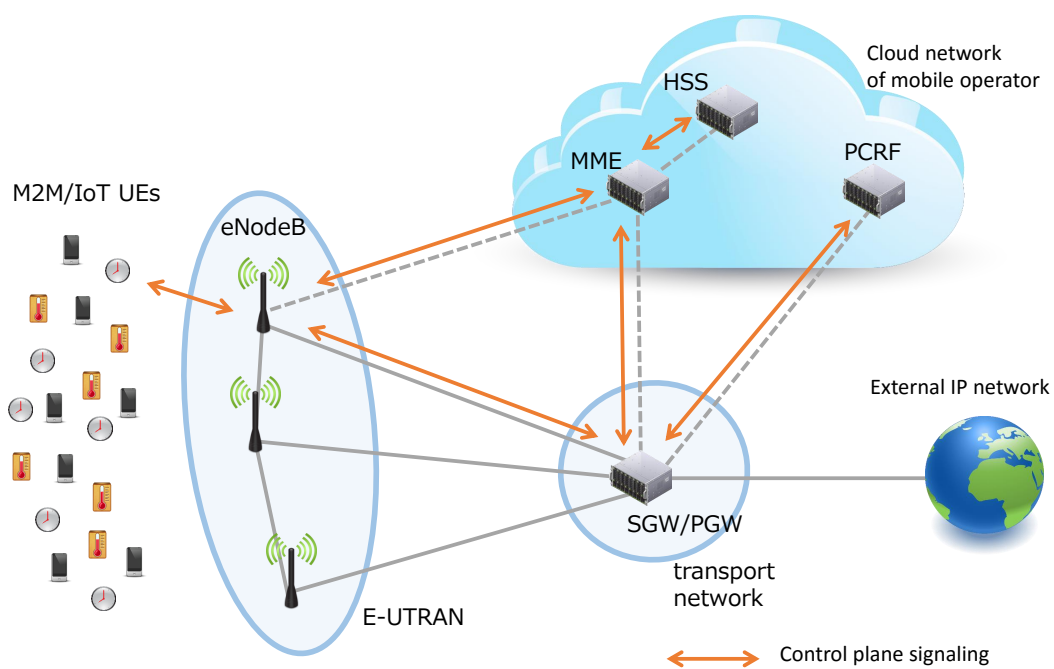


図 1: モデル1のネットワーク構成

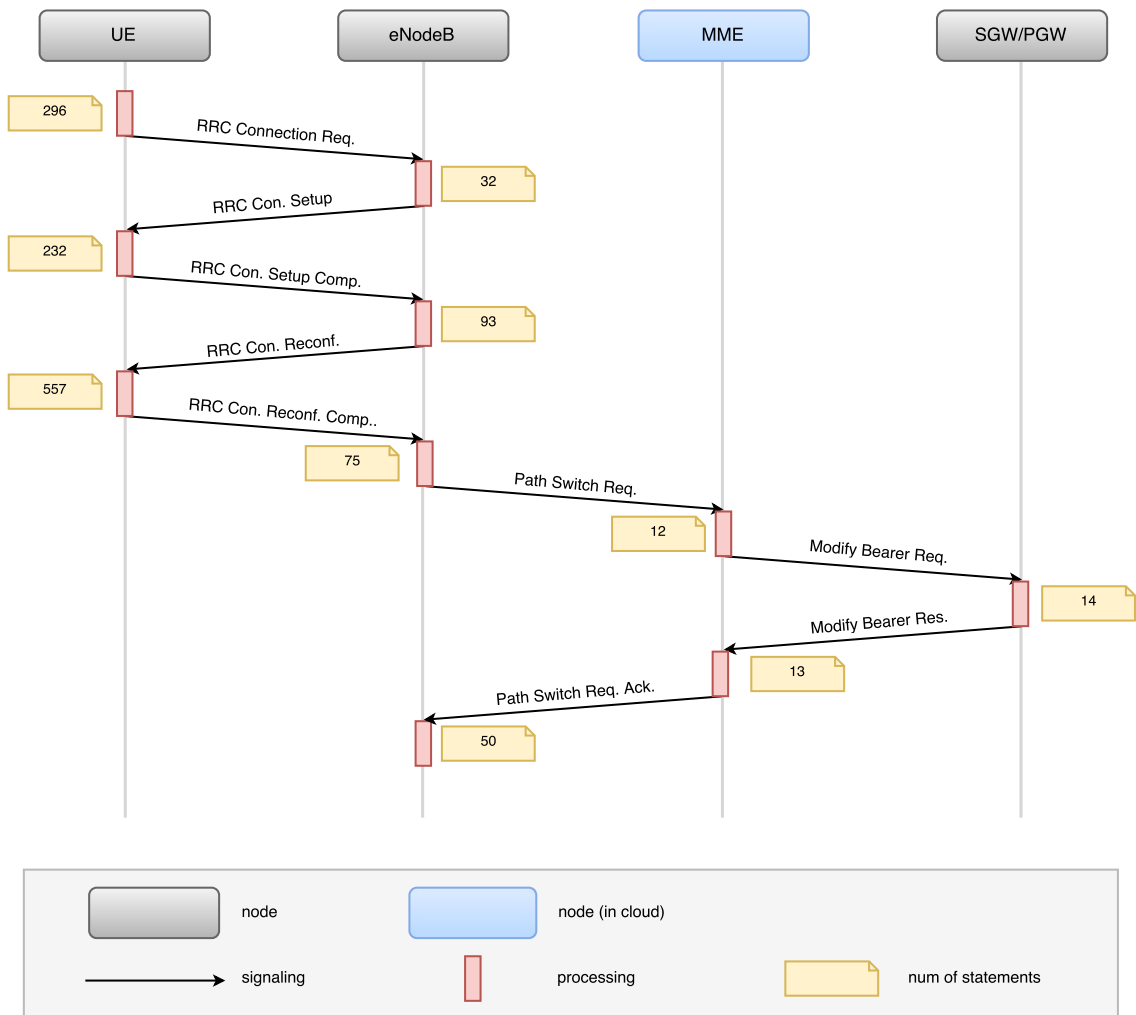


図 2: LENA を基にしたモデル 1 のシグナリング手順と処理命令文数

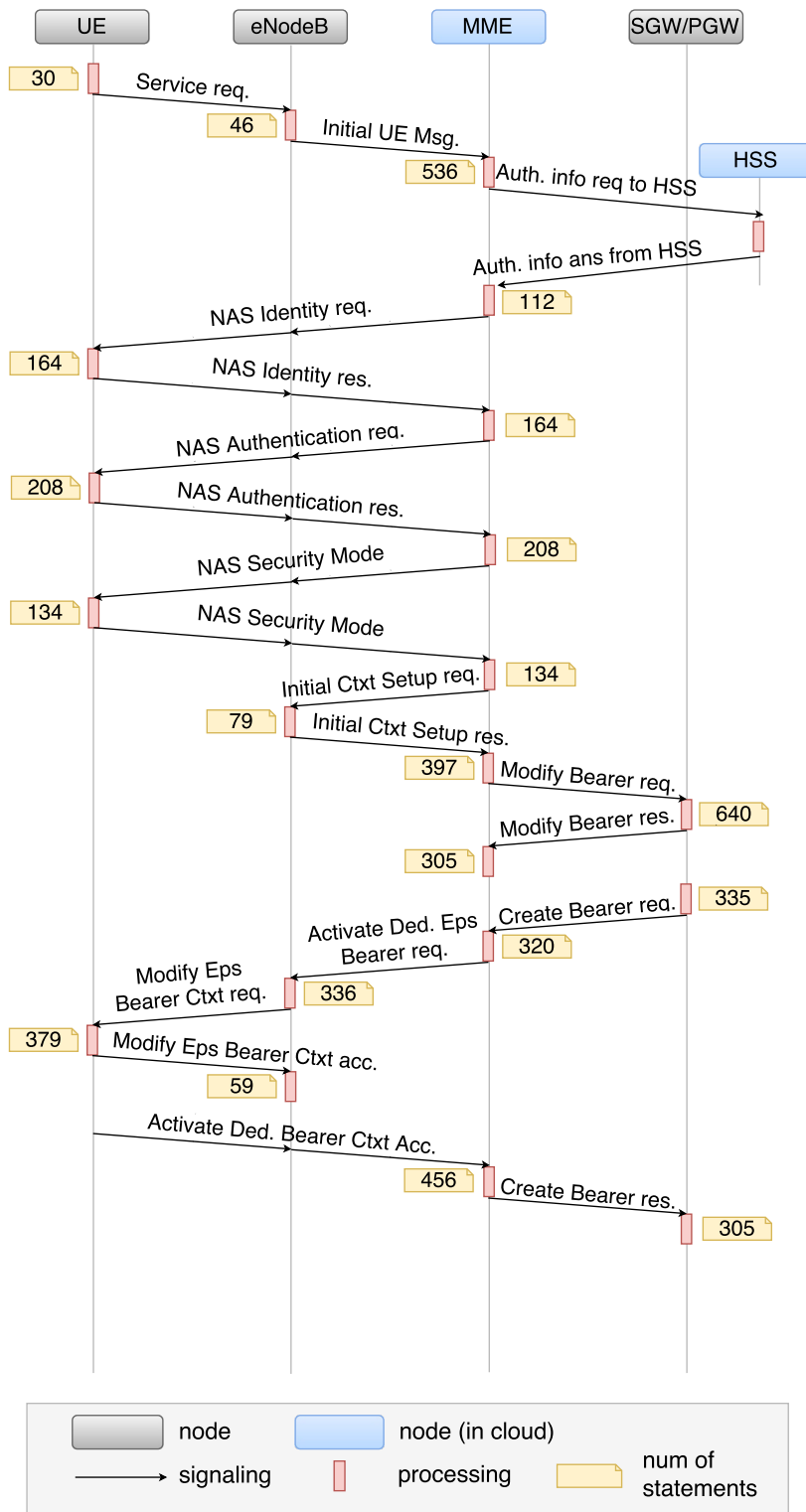


図 3: OAI を基にしたモデル 1 のシグナリング手順と処理命令文数

2.2 モデル 2: SGW/PGW の制御プレーンの機能をクラウドネットワーク内に設置したモデル

図 4 に、SGW/PGW を制御プレーンの機能を持つ SGW/PGWc とデータプレーンの機能を持つ SGW/PGWd に分離し、SGW/PGWc をクラウドネットワーク内に設置したモデル 2 のネットワーク構成を示す。このプレーン分離により、ベアラ確立時の制御メッセージの処理に関して MME との伝搬遅延時間が短縮されることが考えられる。

本モデルにおいては、ベアラを用いた通信に必要な GPRS Tunneling Protocol (GTP) ベアラのマッチング機能 (GTP モジュール) はクラウドネットワーク内に設置されている。その理由は、現在の OpenFlow の仕様では GTP のヘッダフィールドを用いたパケットのマッチングを行えないため、その機能をコントローラ側で実装することが考えられるためである [12]。そのため、GTP マッチングの処理を行う場合でも SGW/PGWc との間には遅延が発生しない。

しかし、GTP モジュールがクラウドネットワーク内に設置されていることにより、ベアラ確立が完了した後のデータ通信において、データパケットはすべてクラウドネットワーク内を通過する必要がある。そのため、データパケットの通過時には遅延時間の増加やスループットの低下が考えられることから、通信データ量が大きいスマートフォンや携帯電話端末との通信には適さず、M2M/IoT 端末のように、通信データ量に比べてベアラ確立のためのオーバーヘッドが比較的大きい端末の収容に適しているモデルであると言える。

また、図 5 及び図 6 に、それぞれ LENA 及び OAI に基づく、モデル 2 におけるシグナリング手順とシグナリング経路、各ノード及び UE がシグナリングメッセージを処理するために実行されるプログラムの命令文数を示す。図 2 及び図 3 と比較すると、ベアラ設定のためのシグナリング処理が完了した後に、SGW/PGWd に対する OpenFlow のメッセージ交換が追加されている。なお、このメッセージ処理に必要な命令文数は “Modify Bearer Req. / Res.” を参考に設定している。図 5 及び図 6 より、モデル 1 では、MME と SGW/PGW の間で行われていたシグナリングメッセージ交換が、クラウドネットワーク内で行われていることがわかる。そのため、モデル 2 はモデル 1 よりもシグナリングメッセージの総伝搬遅延時間が小さくなると考えられる。

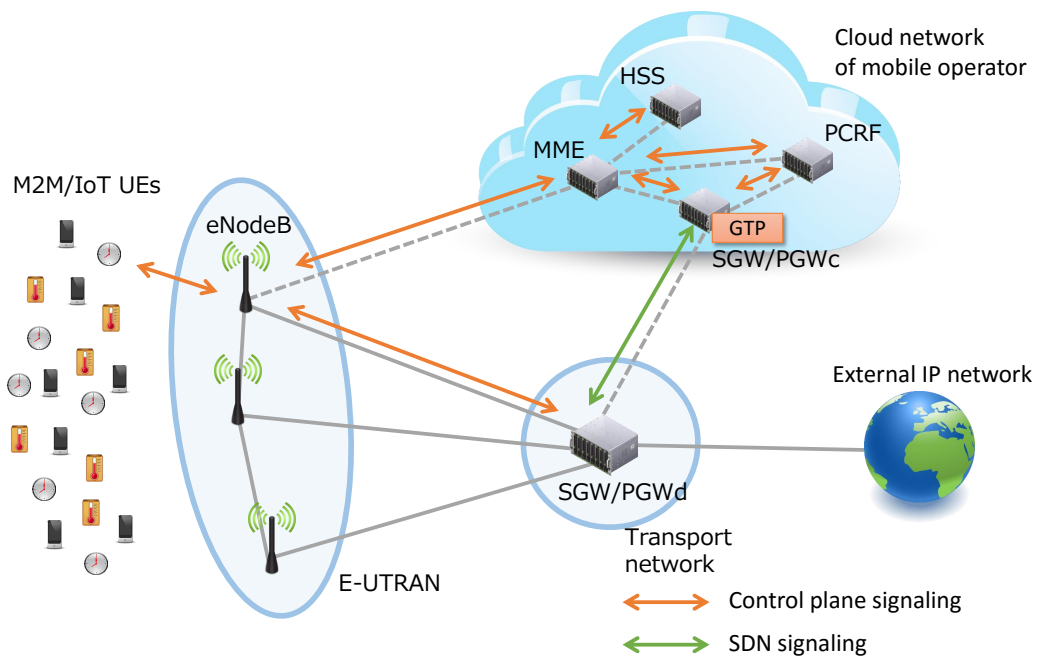


図 4: モデル 2 のネットワーク構成

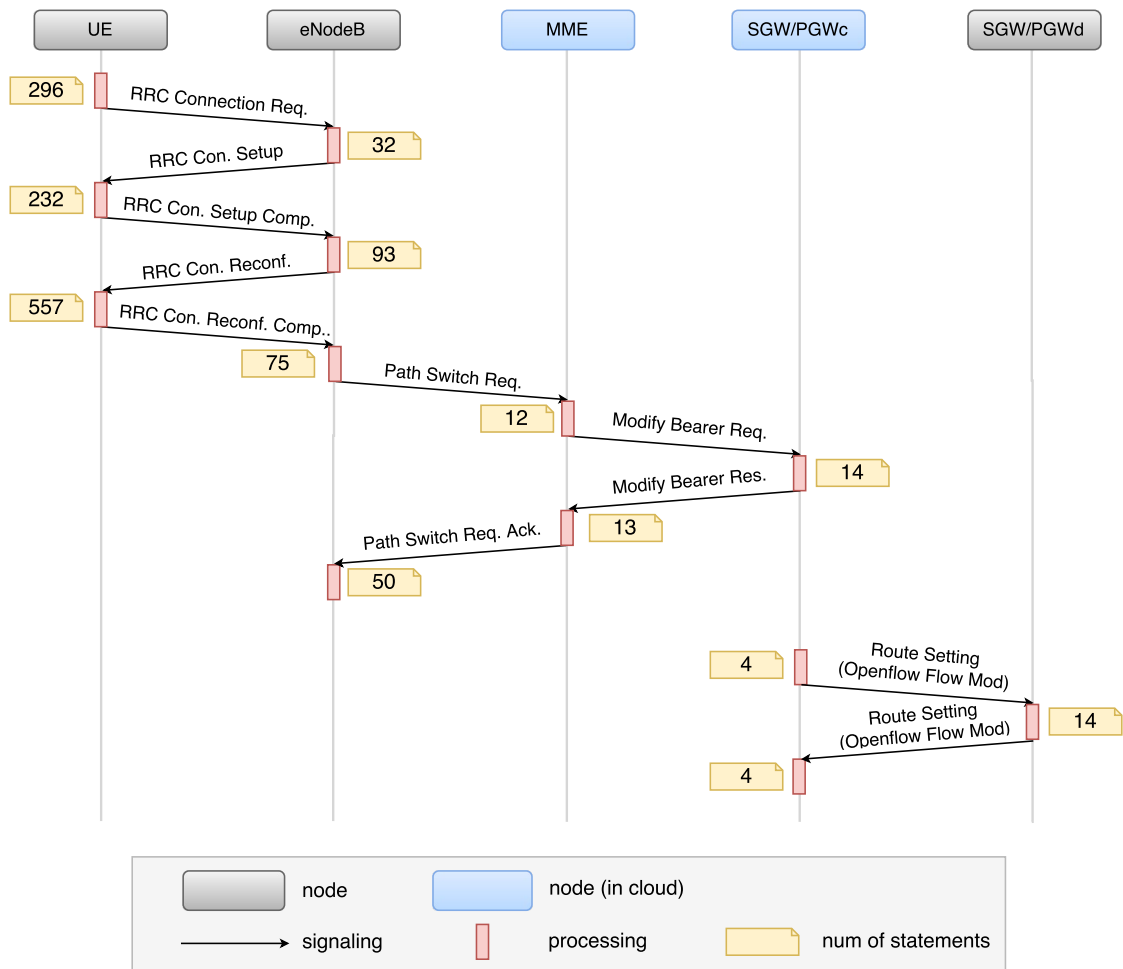


図 5: LENA を基にしたモデル 2 のシグナリング手順と処理命令文数

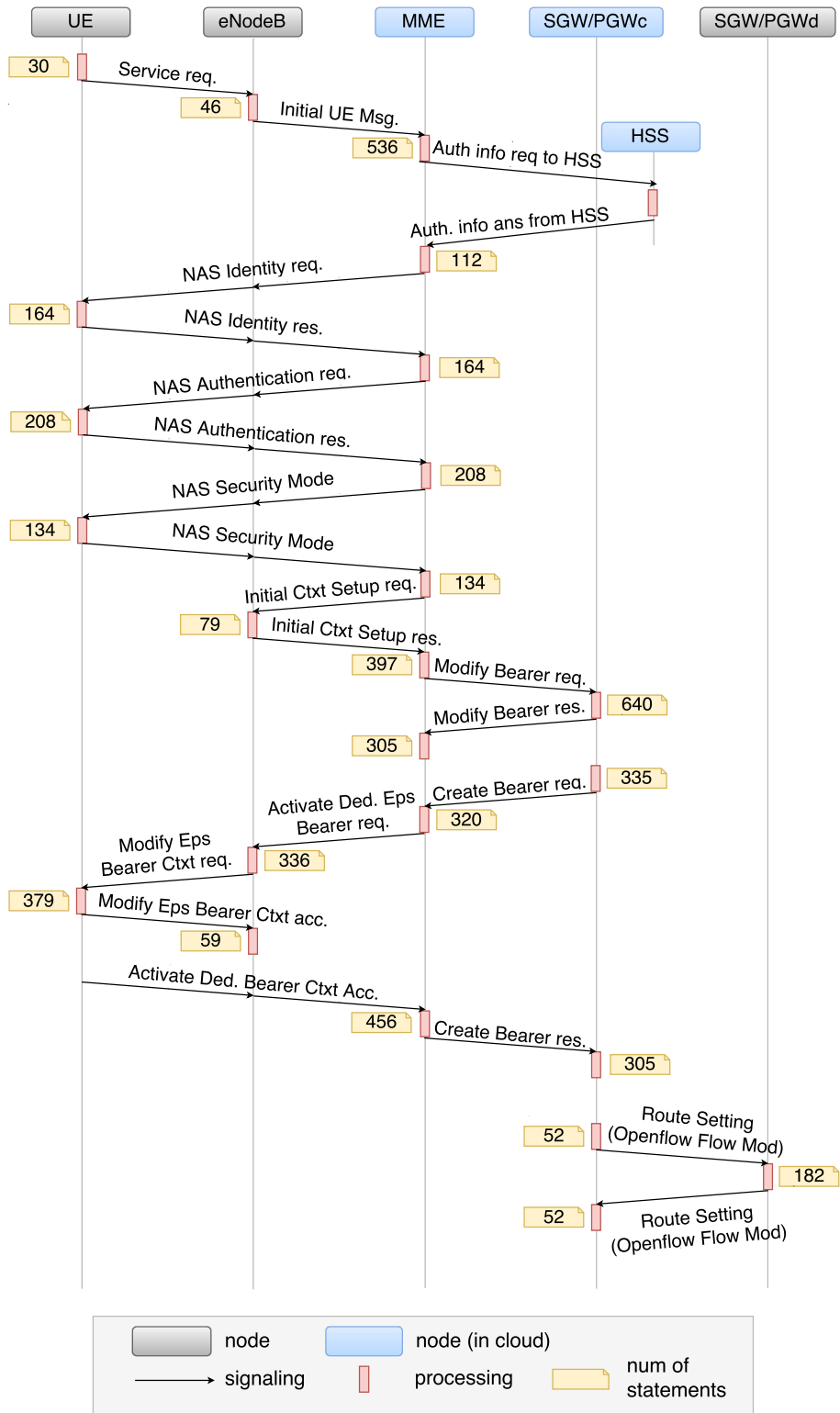


図 6: OAI を基にしたモデル 2 のシグナリング手順と処理命令文数

2.3 モデル3: SGW/PGW の GTP モジュールをクラウドネットワーク内に設置したモデル

図7に、GTP モジュールをハードウェアとして実装、あるいは OpenFlow スイッチにプログラム性を持たせることによりソフトウェアとして実装することを想定した、モデル3のネットワーク構成を示す。このモデル3は、トランスポートネットワークに設置した GTP モジュールがマッチング処理を行うため、モデル2で懸念されていたようなデータパケットの伝搬遅延時間の増加やスループットの低下は避けることができると考えられる。

また、図8及び図9に、それぞれ LENA 及び OAI に基づく、モデル3におけるシグナリング手順とシグナリング経路、各ノード及び UE がシグナリングメッセージを処理するために実行されるプログラムの命令文数を示す。モデル2と比較すると、SGW/PGWd に対する OpenFlow のメッセージ交換の後に、GTP モジュールへマッチングのための情報を投入するためのメッセージ交換が追加されている。なお、この処理に必要な命令文数は、SGW/PGWd に対する OpenFlow メッセージの処理と同等であると仮定することで決定している。

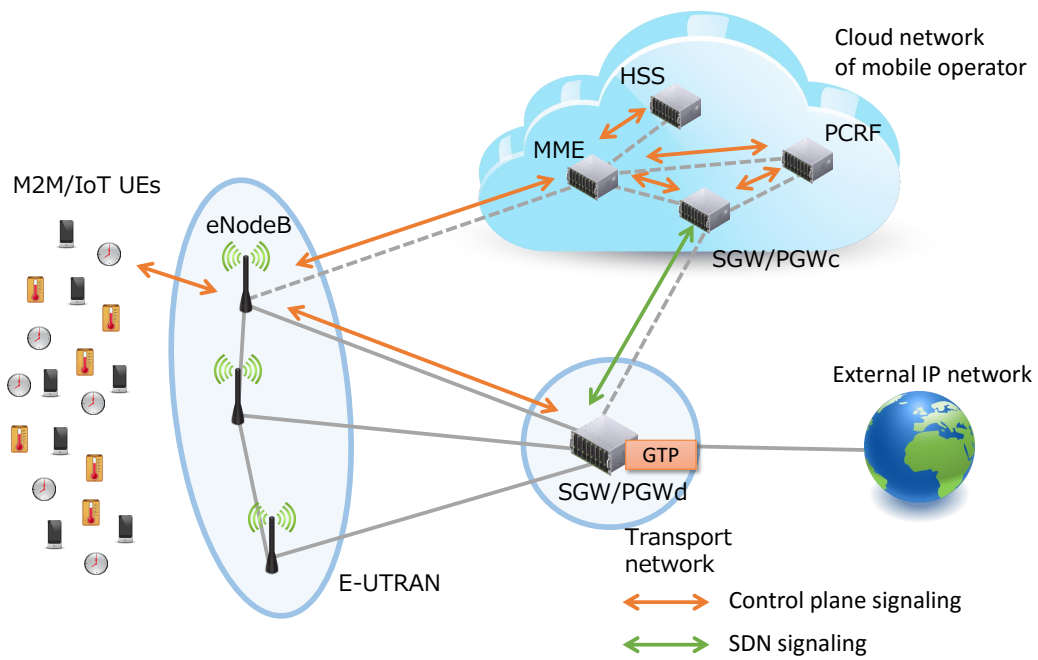


図 7: モデル 3 のネットワーク構成

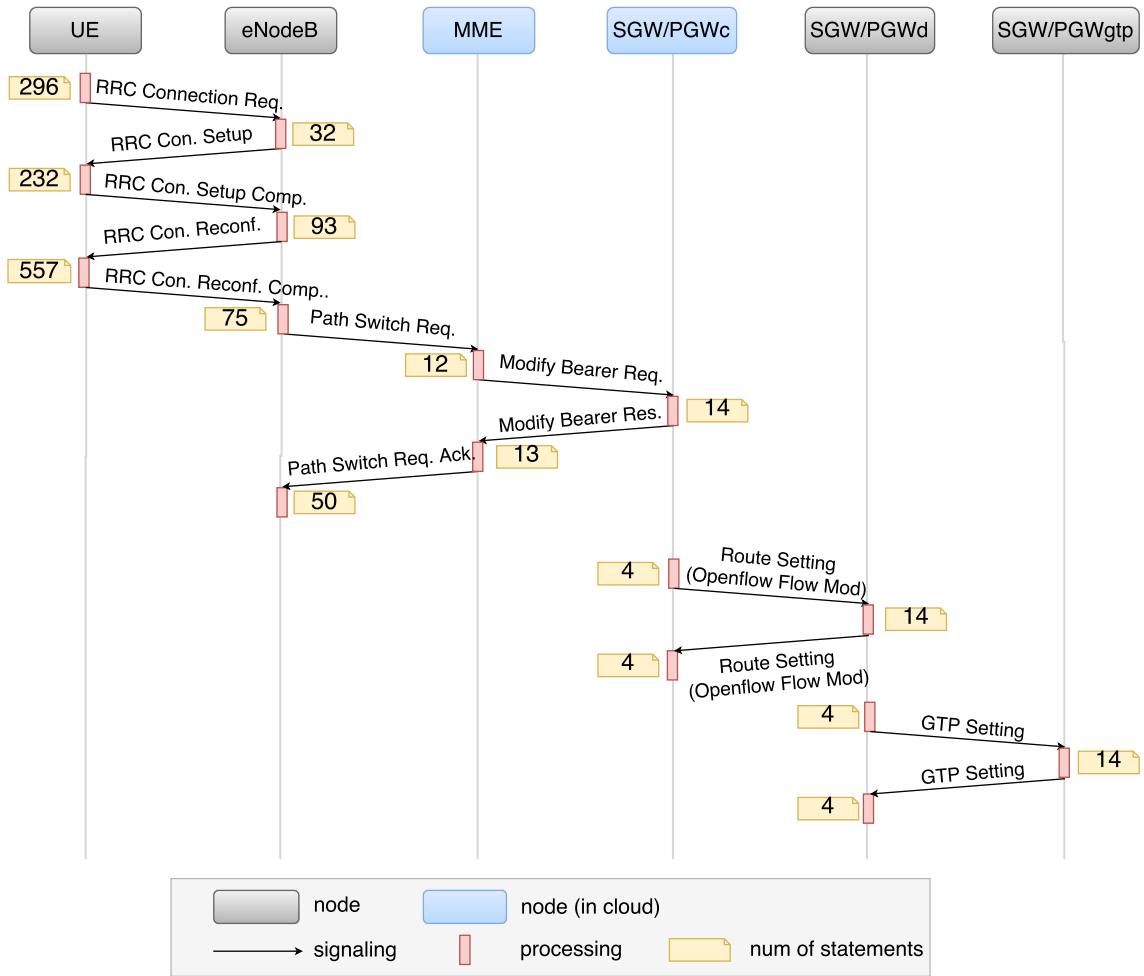


図 8: LENA を基にしたモデル 3 のシグナリング手順と処理命令文数

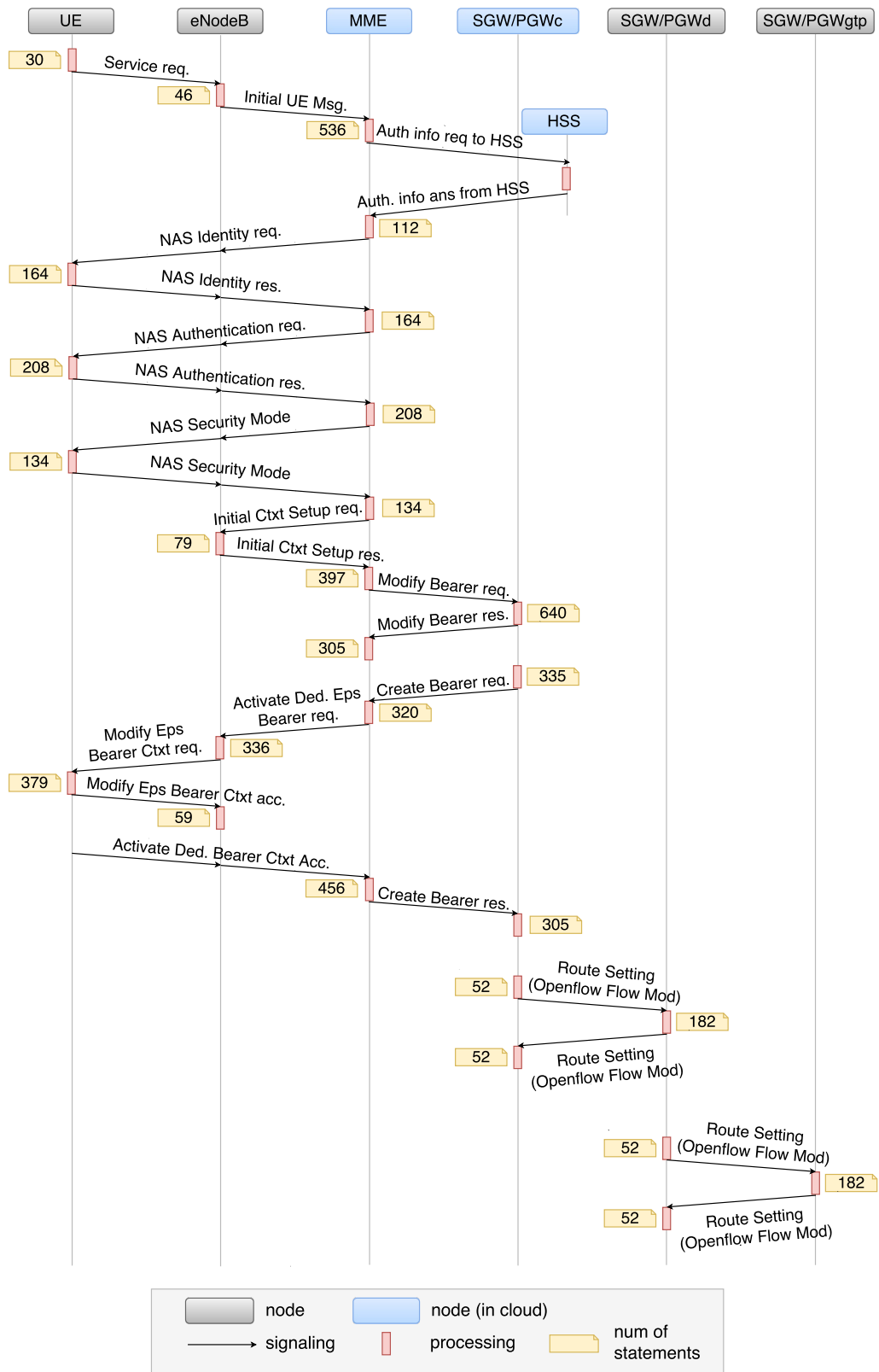


図 9: OAI を基にしたモデル 3 のシグナリング手順と処理命令文数

2.4 通信集約を行うモデル

本報告では、SGW/PGW において UE からの通信を集約する手法 [9, 12] の効果についても評価を行う。通信集約は、ベアラ確立時にシグナリング手順そのものを変更することによって実現されることを前提としているため、上述したすべてのネットワークモデルに適用可能であるが、本報告ではモデル 3 に対して適用したものを評価対象とする。

通常、GW 間ベアラと UE は、UE がモバイルネットワークにアタッチされる際に、UE 固有の International Mobile Subscriber Identity (IMSI) によって対応付けられ、MME が IMSI と GW 間ベアラの対応を管理する。文献 [21] においては、MME による IMSI の管理負荷を低減するために、複数の UE で IMSI を共有する手法が提案されている。本報告では、その手法を参考に、以下のような通信集約手法を想定し、評価対象とした。

本報告における通信集約手法においては、集約対象となる複数の UE に対して設定する、GW 間ベアラで用いる仮想的な IMSI と UE 固有の IMSI とを対応付けるための対応表を MME が管理するものとし、ベアラ確立の際に行われるシグナリング処理において、GW 間ベアラを設定する部分において、UE に対応する仮想 IMSI を用いるものとする。このようにすることで、通信集約を実現するために必要となるシグナリング手順の変更を小さくできると考えられる。

以下では、通信集約において、1 つの GW 間ベアラに集約する UE の台数、すなわち集約度を K と表す。また、通信集約を行う場合は MME が仮想 IMSI のプールを持つことを想定しており、 K 台の UE に対して仮想 IMSI を 1 つ払い出す際に、MME は IMSI の対応テーブルを更新し、PCRF へ通知するための新たなシグナリング処理が必要となる。

以降、本報告において検討した、2 種類の通信集約手法をモデル 3 に対して適用したモデルを説明する。

2.4.1 モデル 3-A: UE が用いる GW 間ベアラを事前に決定するモデル

モデル 3-A では、UE に電源が入る時に発生するアタッチ処理の際に、UE 固有の IMSI と仮想 IMSI の対応を決定すると仮定する。そのため、UE 固有の IMSI と仮想 IMSI の対応付けに関する PCRF への通知処理はアタッチ時に発生する。

図 10 及び図 11 に、それぞれ LENA 及び OAI に基づく、モデル 3-A におけるシグナリング手順とシグナリング経路、各ノード及び UE がシグナリングメッセージを処理するために実行されるプログラムの命令文数を示す。本モデルにおいては、図 10 及び図 11 に示す通り、ある UE が通信を開始する際、その UE と対応付けられた仮想 IMSI に対応する GW 間ベアラが確立されていなければ通常のシグナリングと同様の処理を行い、ベアラをアクティブートして用いる。一方、ある UE が通信を開始する際、UE が対応付けられた仮想 IMSI に対

応する GW 間ベアラが既に確立されていた場合は、GW 間ベアラの確立に関するシグナリング処理を省略できる。これにより、通信集約による GW のシグナリング負荷やベアラ維持負荷を低減することができる。

ただし、本モデルにおいては、MME が持つ仮想 IMSI のプールや IMSI の対応表の容量は、仮想 IMSI 数及び収容端末数にそれぞれ比例すると考えられる。また、GW 間ベアラに対してアタッチ時にネットワーク帯域の確保を行う場合、各ベアラに割り当てるネットワーク帯域を実際に通信を行う UE 数に応じて動的に変更することが困難となる。そのため、実際に通信を行っている UE に対してではなく、アタッチ処理を行った UE に対してネットワーク帯域の確保を行うため、UE の通信特性によってはネットワーク帯域が有効に利用できないことが懸念される。

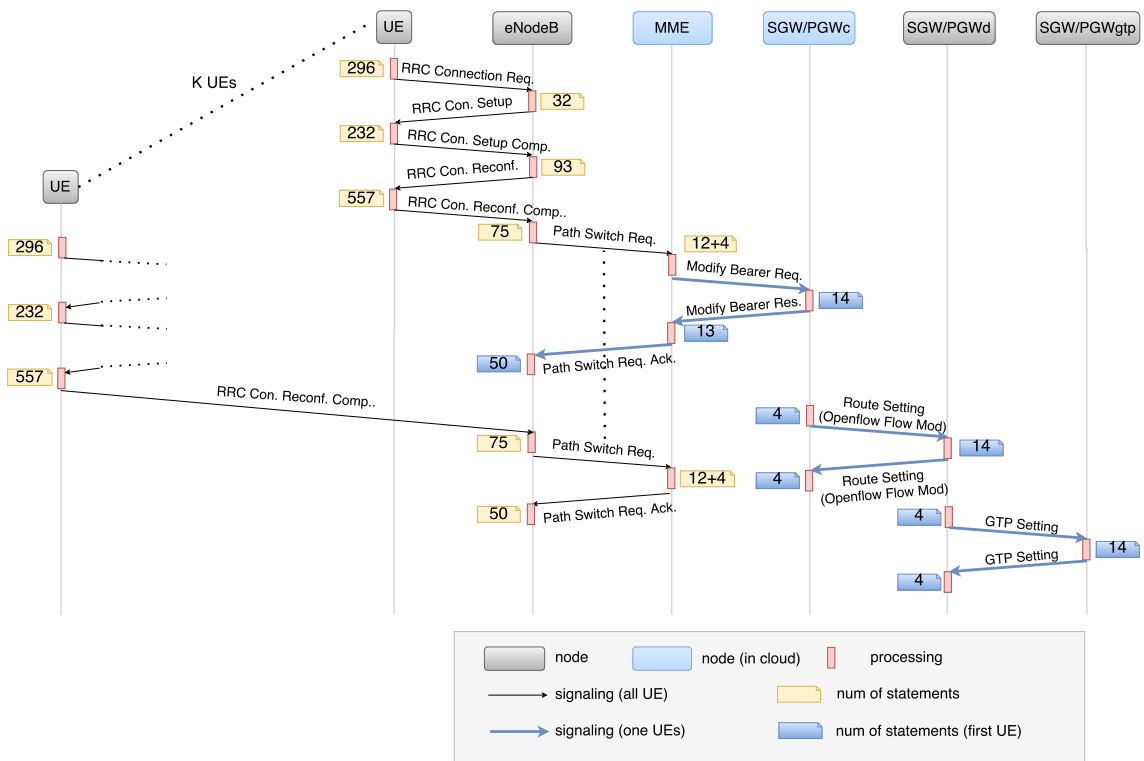


図 10: LENA を基にしたモデル 3-A のシグナリング手順と処理命令文数

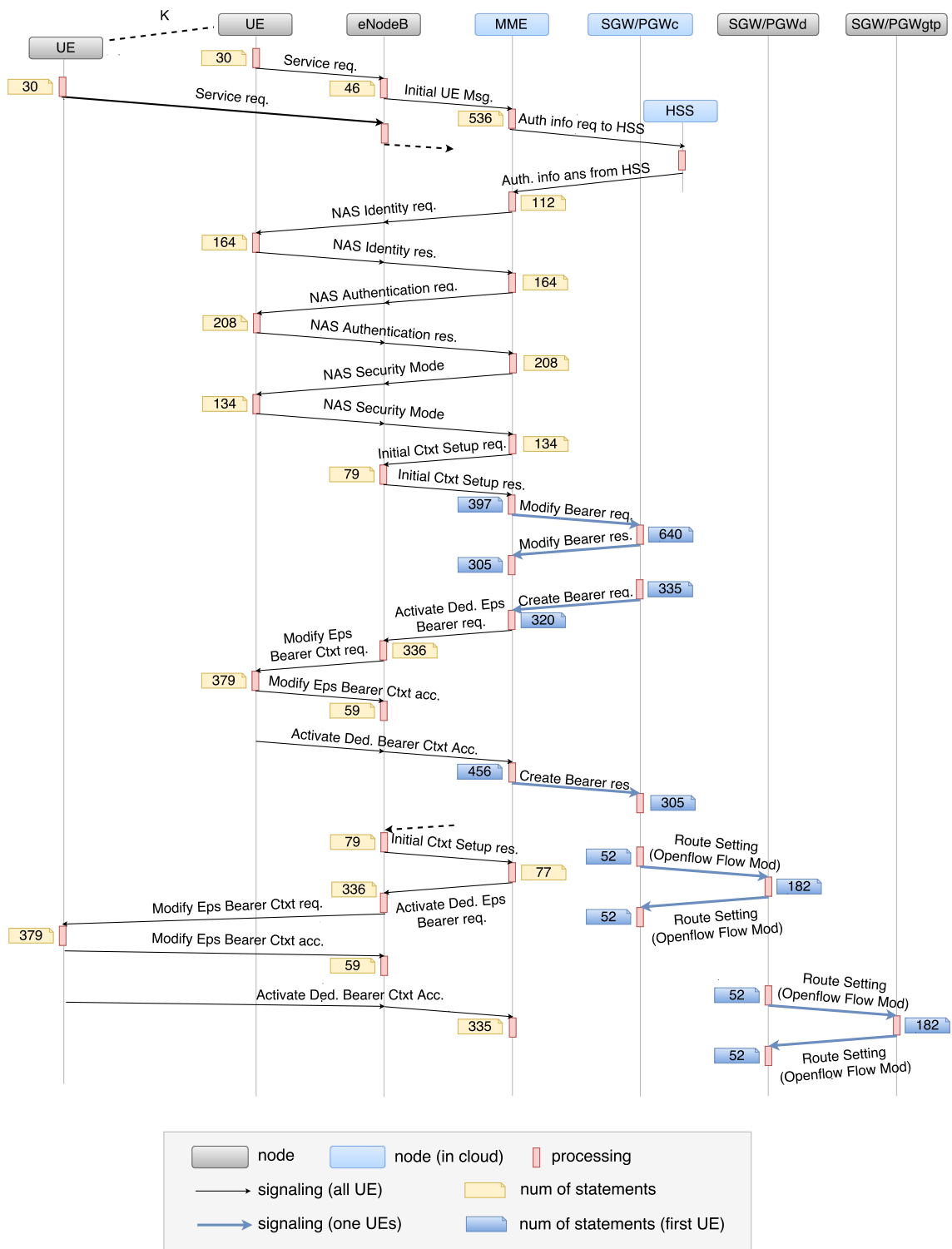


図 11: OAI を基にしたモデル 3-A のシグナリング手順と処理命令文数

2.4.2 モデル 3-B: UE が用いる GW 間ベアラをオンデマンドに決定するモデル

モデル 3-B においては，図 12 及び 13 に示す通り，UE からの通信開始要求が MME に到着した際，MME から GW 間ベアラを設定するためのシグナリング処理を中断し， K 台の UE からの通信開始要求が到着するまで待機する．この待機時間は集約度 K に比例し，かつ収容する端末数に反比例する．ここで，UE のシグナリングを待機させることによって発生し得るタイムアウト処理等は，本報告では考慮していない． K 台からの通信開始要求が到着した時点で， K 台の UE が通信時に利用する GW 間ベアラと，それに対応する仮想 IMSI を決定し，GW 間ベアラを確立するためのシグナリング処理を行う．本モデルにおいては，モデル 3-A と異なり，UE 固有の IMSI と仮想 IMSI の対応付けに関する PCRF への通知処理は通信開始時に発生する．なお，図 12 及び図 13 において，PCRF への通知処理を行う部分は，メッセージの送受信処理であることを考慮し，“Modify Bearer Req. / Res.” を参考に命令文数を設定している．

本モデルにおいては，通信要求が発生した UE に対してのみ GW 間ベアラが確立されるため，確保されたネットワーク帯域に応じた通信量が必ずしも見込めないモデル 3-A と比較すると，ネットワーク帯域の利用効率が高いことが期待される．さらに，MME が持つ仮想 IMSI のプールや IMSI の対応表のサイズは通信を行う UE 数に比例すると考えられるが，それぞれ仮想 IMSI 数及び収容端末数に比例するモデル 3-A と比較すると，その容量は小さくなると考えられる．一方，本モデルにおいては， K 台からの通信要求が到着するまでシグナリングを中断するため，ベアラ確立時間が増大することが考えられる．

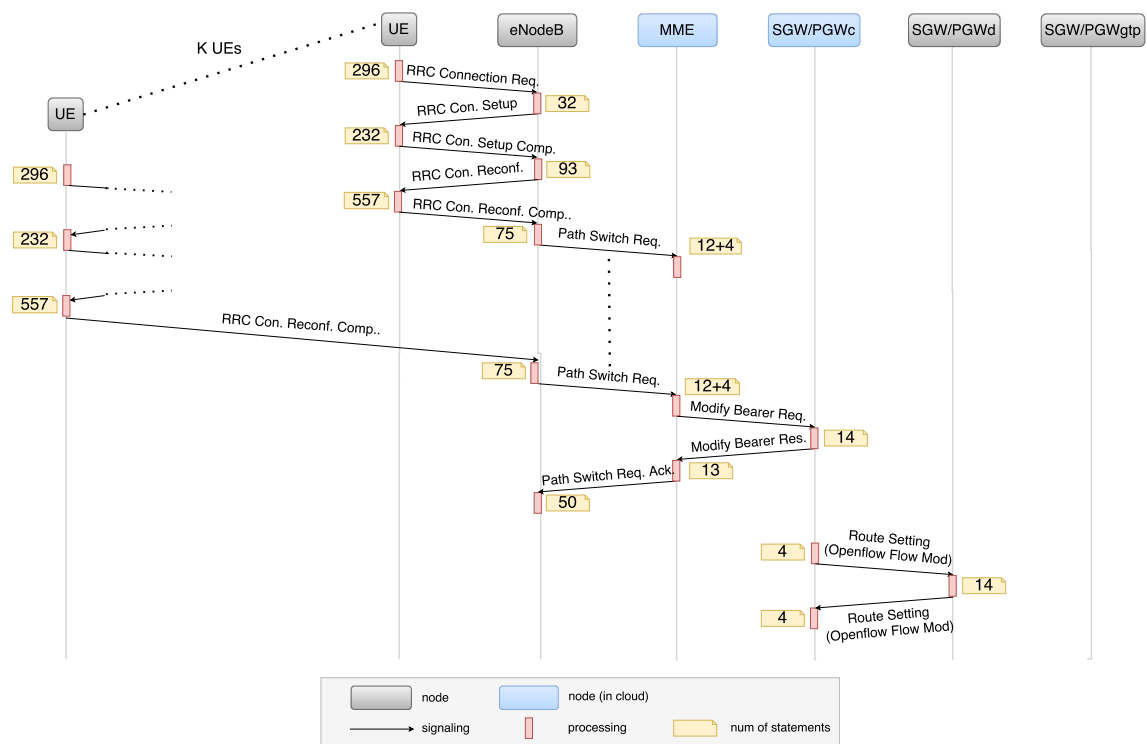


図 12: LENA を基にしたモデル 3-B のシグナリング手順と処理命令文数

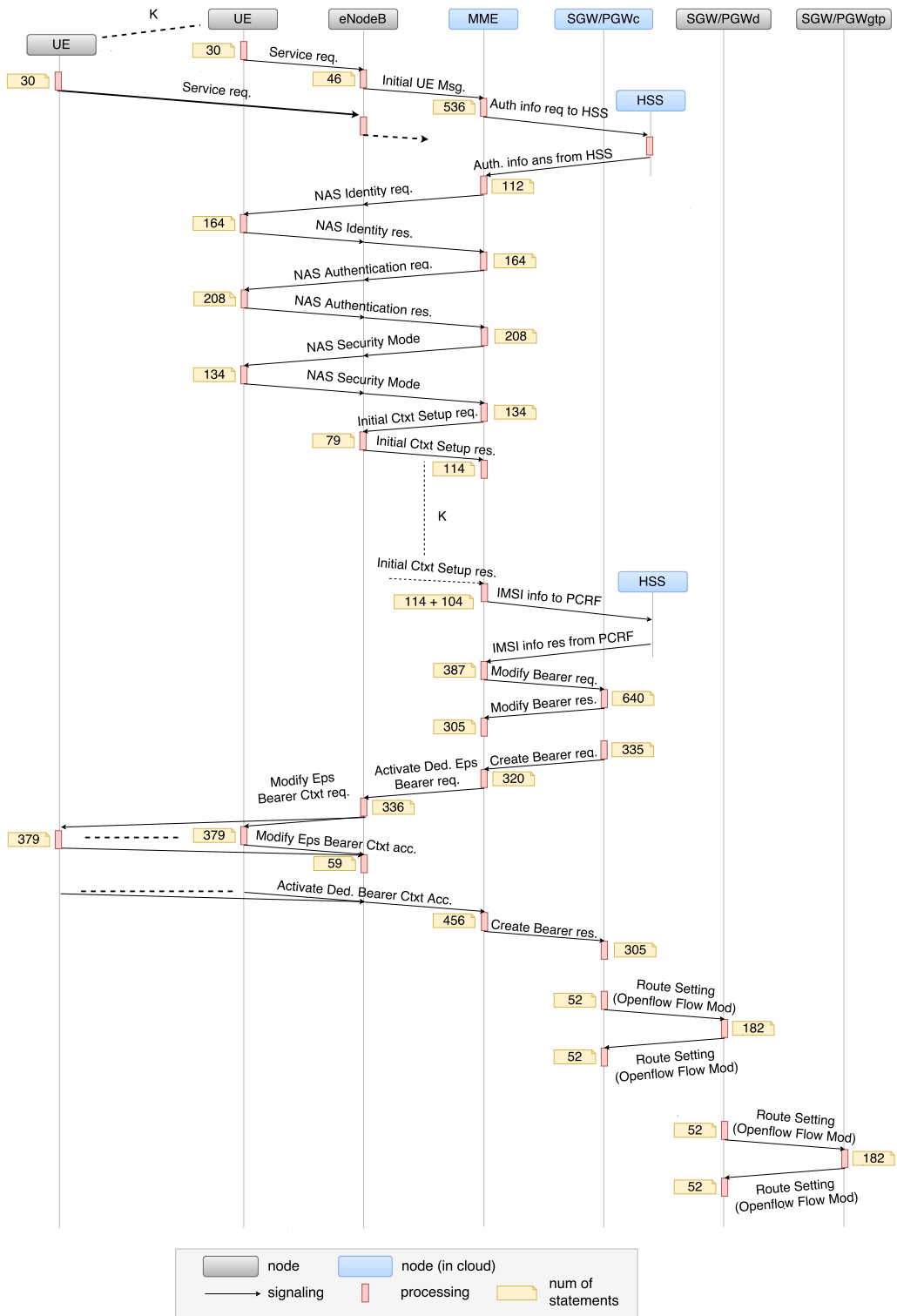


図 13: OAI を基にしたモデル 3-B のシグナリング手順と処理命令文数

3 性能解析

本章では，2章において検討したモデル及びシグナリング手順を用いて，UEに通信要求が発生してから，すべてのシグナリング処理が終了するまでにかかる時間である，ベアラ確立時間を解析的に導出する．

3.1 変数定義

解析で用いる変数を以下のように定義する．ただし，UE，eNodeB，MME及びSGW/PGWをそれぞれ U ， B ， M ，及び G と略記し，それらを総称するためのノードを N と表記する． N_c ， N_d ，及び N_g はそれぞれノード N に対する制御プレーンノード，データプレーンノード，及びGTPモジュールを表す．また， $\tau_{N1,N2}$ をノード $N1$ ， $N2$ 間の伝搬遅延時間とし， t_N をノード N におけるシグナリングメッセージの平均処理時間， $C_{N1,N2}$ をノード $N1$ ， $N2$ 間のシグナリング伝搬回数， P_N をノード N におけるシグナリング処理回数とする．ただし， N ， $N1$ ， $N2$ はそれぞれ U ， B ， M ， G ， G_c ， G_d ， B_c ， B_d ， G_g 及び B_g のいずれかである．

3.2 ベアラ確立時間の導出

2章で定義したモデル i ($i = 1, 2, 3, 3-A, 3-B$)において，ベアラ確立時間を T_i ，シグナリングの過程で発生するすべてのノードにおける処理遅延時間 t の総和を t_i ，伝搬遅延時間 τ の総和を τ_i とすると， T_i は以下のように表せる．

$$T_i = \tau_i + t_i$$

構成ノードがUE，eNodeB，MME，SGW/PGWであるモデル1のベアラ確立時間 T_1 は，図3より以下の式で表される．

$$T_1 = (C_{U,B}\tau_{U,B} + C_{B,M}\tau_{B,M} + C_{M,G}\tau_{M,G}) \\ + (P_U t_U + P_B t_B + P_M t_M + P_G t_G)$$

また，SGW/PGWをプレーン分離し，制御プレーンをクラウドネットワーク内に置いたモデル2において，FlowModによる経路設定にかかるメッセージの往復回数を C_{FM} とすると， T_2 は，図6より次式で表される．

$$T_2 = T_1 + (2C_{FM}\tau_{G_c,G_d} + (C_{FM} + 1)t_{G_c} + C_{FM}t_{G_d})$$

モデル3においては、SGW/PGWdとGTPモジュールとの間でGTPマッチングのための設定を行うシグナリングが発生する。マッチング設定のために必要なメッセージの往復回数を C_{GTP} 、あるノード N においてGTPモジュールで行われるGTP設定のために必要となる処理時間を t_{Ng} とすると、 T_3 は、図9より以下のように表される。

$$T_3 = T_2 + (2C_{GTP}\tau_{G_d, G_g} + (C_{GTP} + 1)t_{G_d} + C_{GTP}t_{G_g})$$

通信集約を行うモデル3-A及び3-Bでは、 K 台のUEに対して仮想IMSIを1つ払い出す際に、MMEはIMSIの対応テーブルを更新し、PCRFへ通知するための新たなシグナリング処理が必要となる。モデル3-Aにおいては、このPCRFに対するシグナリング処理は、UEに電源が入るときに発生するアタッチ処理の際に行われる。そのため、通信要求が発生する際には、MMEがIMSIの対応テーブルを検索する処理のみが追加される。また、モデル3-Aにおいては、UEの通信要求がMMEに到着した際に、対応するGW間ベアラが既に確立されているか否かによって、ベアラ確立時間が変化する。

一方、モデル3-Bでは、PCRFに対するシグナリング処理は、通信要求発生時に行われる。ただし本解析では、このシグナリング処理によるMMEの処理負荷及びMME-PCRF間の伝搬遅延時間は考慮しているが、PCRFにおける処理遅延時間は無視している。さらに、モデル3-Bでは、 K 台のUEからの通信開始要求がMMEに到着するまでの待機時間が存在し、この平均待機時間は集約度 K と収容端末数に依存する。さらに、通信集約を行うことによって、ベアラ確立後のデータパケットの送信時に発生する遅延時間が大きくなるのが考えられる。この点の評価については、本解析においては対象外としているが、今後の課題として挙げられる。

以上のことから、モデル3-A及び3-Bにおけるベアラ確立時間は、ノード $N1, N2$ 間において集約対象ではないシグナリングメッセージの伝搬回数 $C'_{N1, N2}$ 、ノード N における集約対象ではないシグナリング処理回数 P'_N 、UEの台数 n_N 、通信周期 D 、及びMMEにおいてIMSIの対応を取得する時間 t'_M を用いて以下の式で表される。ただし、*case1*はGW間ベアラが既に確立されている場合、*case2*はGW間ベアラが確立されていない場合を表す。

$$T_{3-A} = \begin{cases} (C_{U,B}\tau_{U,B} + C_{B,M}\tau_{B,M} + C'_{M,G}\tau_{M,G}) \\ + (P_U t_U + P_B t_B + P'_M t_M + t'_M) & (case1) \\ T_3 + t'_M & (case2) \end{cases}$$

$$T_{3-B} = T_3 + t'_M + \frac{KD}{2n_U} \quad (1)$$

3.3 ノードにおけるシグナリング処理時間の導出

ノード N において必要となる処理時間 t_N は，サーバの並列数が r である M/G/1/PS 待ち行列モデルを用いて導出する．M/G/1/PS 待ち行列モデルにおいて，ジョブの到着率を λ ，ワークロード分布を $S(x)$ ，その平均を $E[S]$ ，システム利用率を $\rho = \lambda \cdot E[S]$ とすると，リクエストがサーバに到着してから，サービスが終了するまでの平均系内時間 $E[R]$ は，以下のように表される．

$$E[R] = \frac{\rho^r}{1-\rho} \frac{E[S^2]}{2E[S]} + \frac{1-\rho^r}{1-\rho} E[S]$$

t_N の導出にあたっては，ジョブの到着率として，ノード N のシグナリング処理頻度を用いる．また，ワークロード分布にはノード N のシグナリング処理命令文数分布を用いる．シグナリング処理命令文数分布は，ノードにおいて発生するシグナリング処理に関連するコードの命令文数の分布である．従って，その平均 $E[S]$ はノードにおけるすべてのシグナリング処理に関連するコードの命令文数の平均値となる．従って，ノード N の台数を n_N ，単位時間あたりに処理できる命令文数で表現される処理能力を A_N ，UE の台数を n_U ，通信周期を D ，ノード N で行われるシグナリング処理の平均命令文数を V_N とすると，ノード N における λ 及び $E[S]$ は以下のように導かれる．

$$\lambda = \frac{P_N n_U}{D n_N}, E[S] = \frac{V_N}{A_N}$$

また，通信集約を行う場合には，SGW/PGW 間のベアラを設定するために必要となるシグナリングが，集約を行わない場合に比べて $1/K$ 倍となる．これは，上式における P_N が集約度に応じて小さくなることを意味するため， λ が小さくなり，シグナリングに関わるノードの負荷が低下する．

4 性能評価

本章では，3における解析の数値例を示すことによって，2で検討したネットワークアーキテクチャの性能評価を行う．

4.1 パラメータ設定

UEの通信周期 D は600秒とし，各ノード間の伝搬遅延時間は以下のように決定する．ただし，eUTRANは図1中のeNodeBのみからなるネットワークを，トランスポートネットワークはSGW/PGWで構成されたネットワークを表しており，クラウドネットワーク内ノードはモデル1におけるMME，モデル2，モデル3，モデル3-A，及びモデル3-BにおけるMMEとSGW/PGWcのように，クラウドネットワーク内に設置されたノードを表す．

- UE-eUTRAN間: 20 msec
- eUTRAN-トランスポートネットワーク間: 7.5 msec
- eUTRAN-クラウドネットワーク内ノード間: 10 msec
- トランスポートネットワーク-クラウドネットワーク内ノード間: 10 msec
- クラウドネットワーク内ノード-クラウドネットワーク内ノード間: 1 msec
- GTP モジュール-データプレーンノード間: 1 msec

また，文献 [12] 中のノードのシグナリング処理時間の係数及び最大収容端末数から，LENAを基にしたモデルにおける各ノードのシグナリング処理能力を以下のように決定する．

- UE: 3,000 命令文/sec
- eNodeB: 3,000 命令文/sec
- MME: 30,000 命令文/sec
- SGW/PGW: 30,000 命令文/sec
- GTP モジュール: 30,000 命令文/sec

同じように，OAIを基にしたモデルにおける各ノードのシグナリング処理能力を以下のように決定する．

- UE: 3,000 命令文/sec

- eNodeB: 6,000 命令文/sec
- MME: 3,000,000 命令文/sec
- SGW/PGW: 3,000,000 命令文/sec
- GTP モジュール: 600,000 命令文/sec

また、SGW/PGW に関して、制御プレーンとデータプレーンを分離した場合は、それぞれの処理能力は分離前と等しいものとする。

ノードを仮想化してクラウドネットワーク内に設置することで、ノードの処理負荷に応じて処理能力を適応的に増減させることが可能になる。本解析では、モデル 3 において SGW/PGWc と MME の間で、処理性能をその総和を保ったまま融通できると仮定して評価を行う。具体的には、処理性能の融通可能なノードに関して、処理性能の初期値の合計を維持したまま、ノードの使用率がほぼ等しくなるように処理性能を配分する。

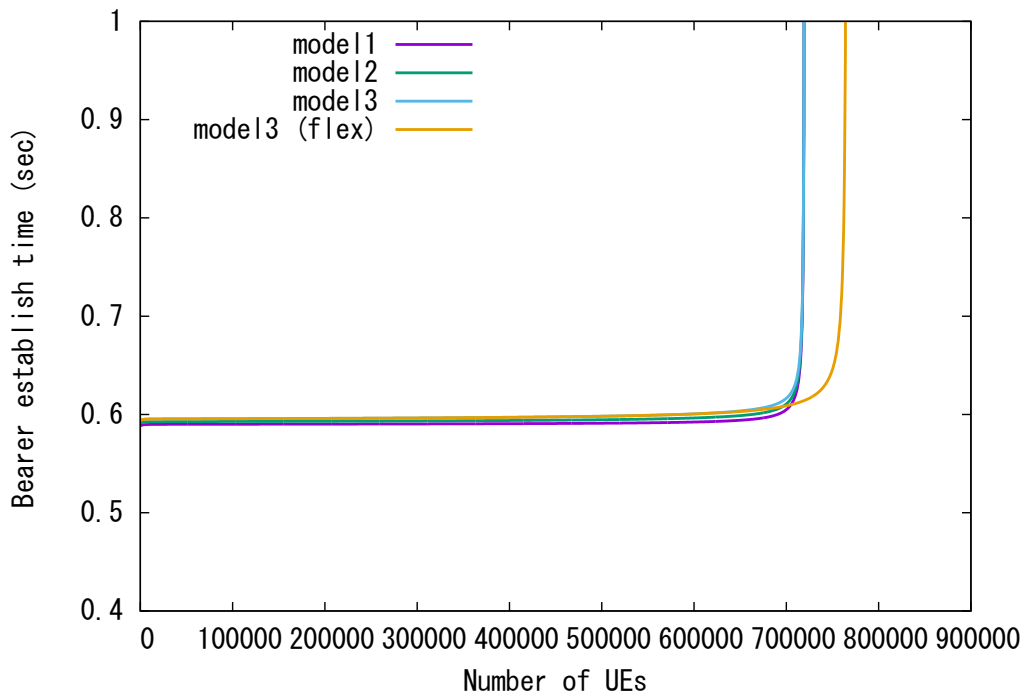
4.2 評価結果と考察

図 14 に、通信集約を行わないモデル 1 (図中の model 1)、モデル 2 (model 2)、モデル 3 (model 3)、及びモデル 3 (model 3 (flex)) において処理性能の融通を行った場合の、収容する UE 数とベアラ確立時間の関係を示す。図より、どの場合においても、UE 数が増加すると、ベアラ確立時間が急激に増加することがわかる。これは、LTE/EPC ノードのいずれかの利用率が高くなり、ノードの処理遅延時間が急増するためである。以降では、ベアラ確立時間が急激に大きくなる直前の UE 数を、そのネットワークにおける収容可能端末数と見なす。

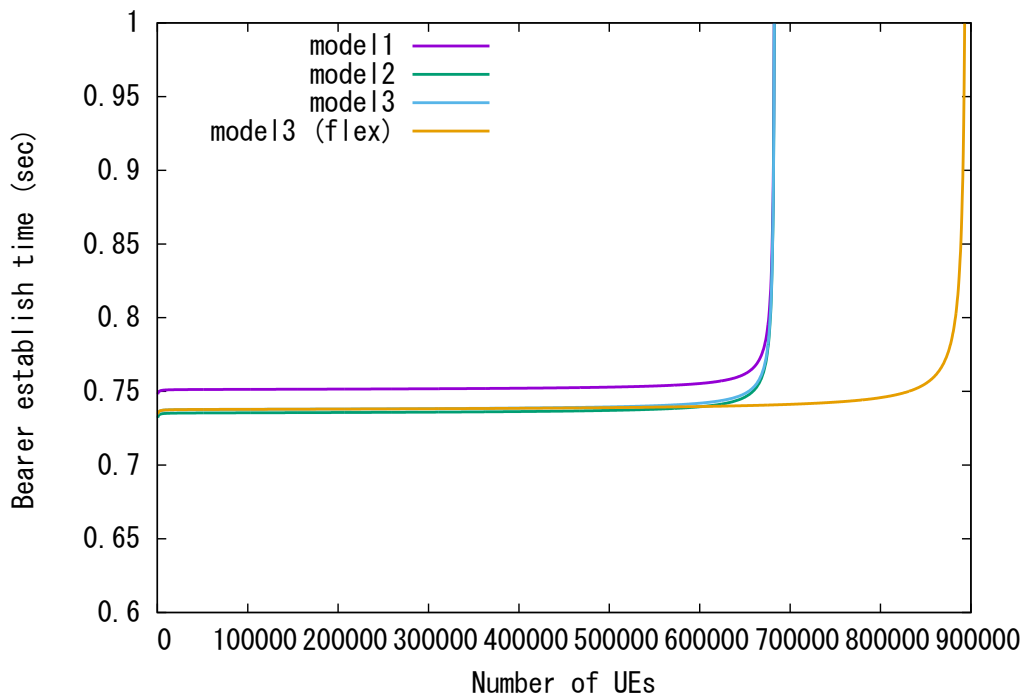
図 14 より、LENA 及び OAI のいずれを基にしてシグナリング手順とシグナリング経路を決定した場合においても、処理性能の融通を行わない場合には、収容可能端末数にモデル間の違いがないことがわかる。これは、モデル 1、モデル 2、及びモデル 3 において、収容する UE 数が増加した際に性能のボトルネックとなるのは MME であり、これら 3 つのモデルにおける MME の処理負荷は等しいためである。

一方、処理性能の融通を行うことで収容可能端末数が増加することがわかる。モデル 3 においては、モデル 1 のシグナリング手順に加え、図 8 及び 9 に示す通り、SDN 化による追加的なシグナリング手順が発生する。そのため、SGW/PGWc における処理負荷はモデル 1 における SGW/PGW の処理負荷よりも大きい。ボトルネックとなる MME に SGW/PGWc の処理性能の余剰分を融通することによって、MME の処理負荷が軽減されるため、収容可能端末数が増加する。また、図 14(a) と図 14(b) の比較から、処理能力の融通の効果は、OAI

の方が LENA よりも大きいことがわかる。これは、LENA と OAI との間に存在する、シグナリング経路とシグナリング手順、及び各ノードにおけるシグナリング処理に関連する実装コードの命令文数の違いに起因しており、MME と SGW/PGW_c の使用率に差がある OAI の方が、SGW/PGW_c から MME に対して融通できる処理能力が大きいためである。



(a) LENA に基づくシグナリング手順とシグナリング経路で評価した場合



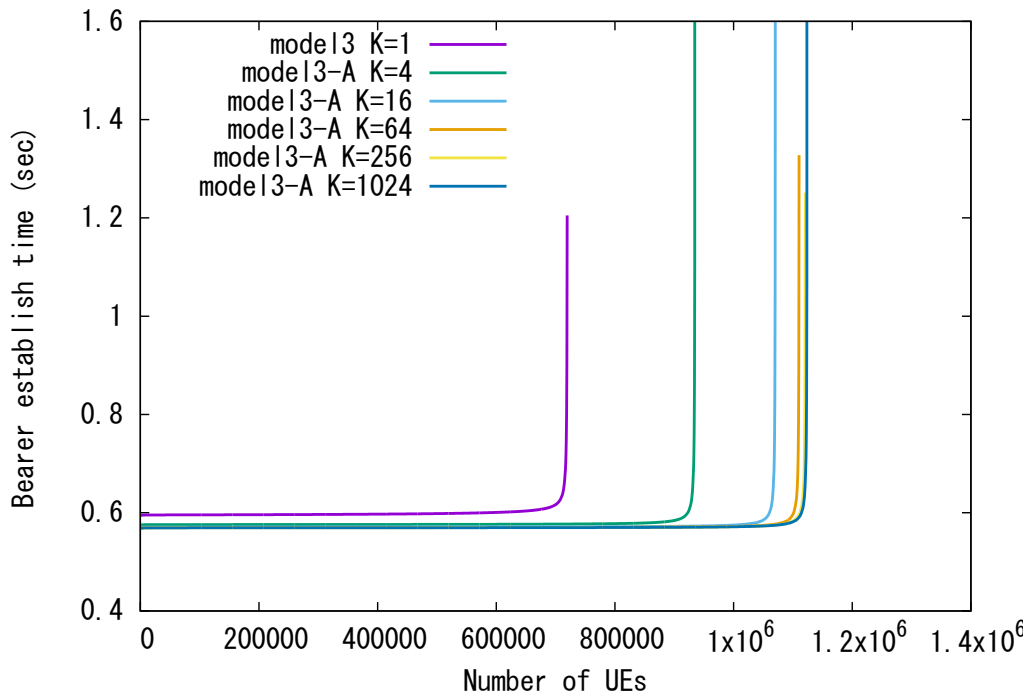
(b) OAI に基づくシグナリング手順とシグナリング経路で評価した場合

図 14: 各モデルにおける収容 UE 数とベアラ確立時間の関係

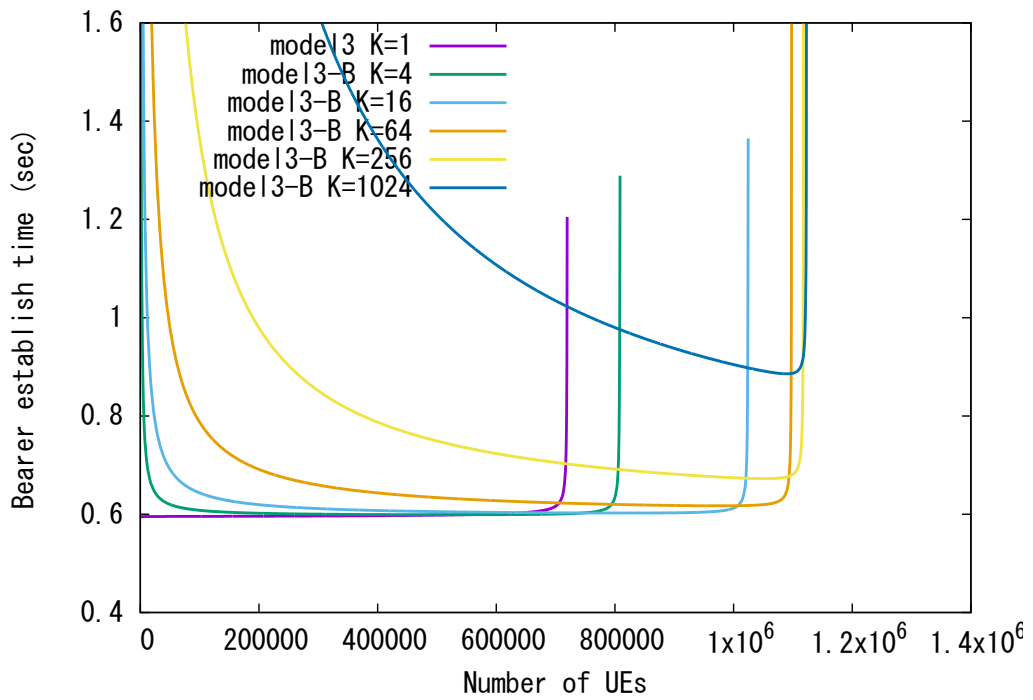
図 15 及び図 16 に、通信集約を行うモデル 3-A、及びモデル 3-B における、集約度を与える影響の評価結果を示す。図より、いずれの集約方法においても、集約を行わないモデル 3 に比べて収容可能端末数が増加していることがわかる。これは、通信集約によって MME 及び SGW/PGW_c のシグナリング頻度が低下することによって、これらのノード負荷が軽減されるためである。また、集約度 K が 64 を超えると収容可能端末数に変化が見られなくなる。これは、集約によって軽減されるシグナリングオーバーヘッドが全体に比べて無視できる程度に小さくなるためであると考えられる。

また、モデル 3-A においては、集約度 K の増加に対して、ベアラ確立時間の減少が確認できる。これは、UE の通信要求が発生した際に、対応する GW 間ベアラが既に確立されていれば、ベアラ確立のためのシグナリング手順が省略されるためである。

一方、モデル 3-B では、集約によってノードの負荷が小さくなることで、ノードの処理遅延時間が僅かに小さくなるものの、総伝搬遅延時間は変化しない。さらに、モデル 3-B では K 台の UE からの通信要求が揃うまでの待機時間が発生する。この待機時間は、式 (1) の第 3 項に示すように、集約度 K に比例し、かつ収容端末数に反比例する。そのため、図 15(b) 及び図 16(b) のいずれにおいても、収容する UE 数が大きくなるにつれて遅延時間が一定値に近づき、ボトルネックとなる MME 及び SGW/PGW_c の使用率が 100% 付近になると遅延時間が急激に増加する。さらに、図 15 と図 16 を比較することによって、通信集約の効果は、OAI の方が LENA よりも大きいことがわかる。これはシグナリング経路とシグナリング手順、及び各ノードにおけるシグナリング処理に関連する実装コードの命令文数の違いに起因しており、OAI の方が、ボトルネックとなる MME と SGW/PGW_c において集約されるシグナリング処理の総命令文数が大きいためである。

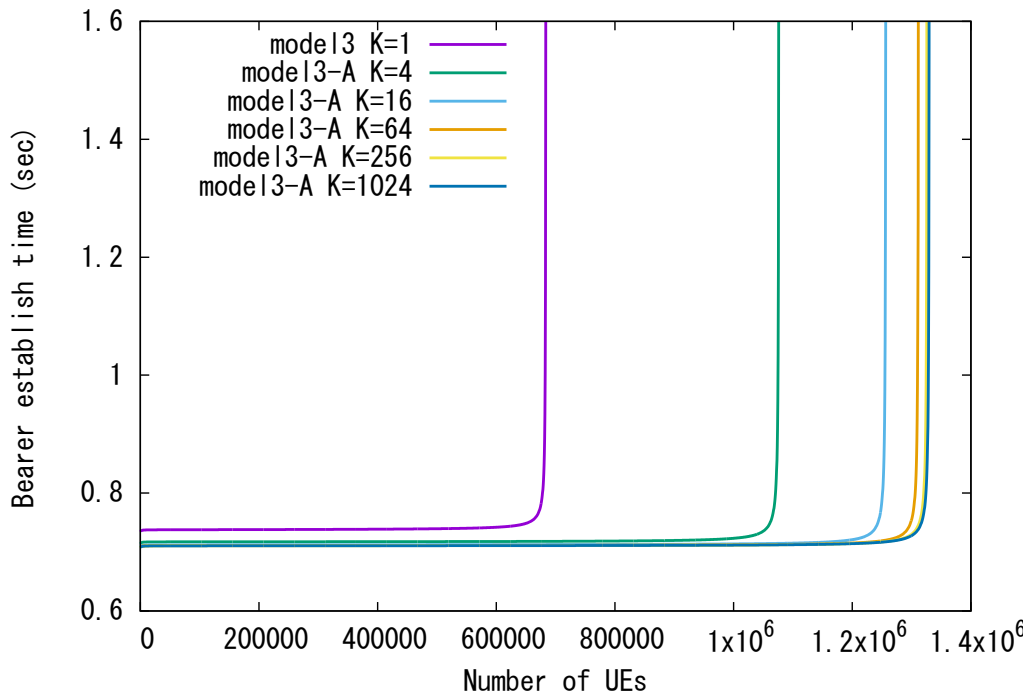


(a) モデル 3-A における通信集約手法の効果

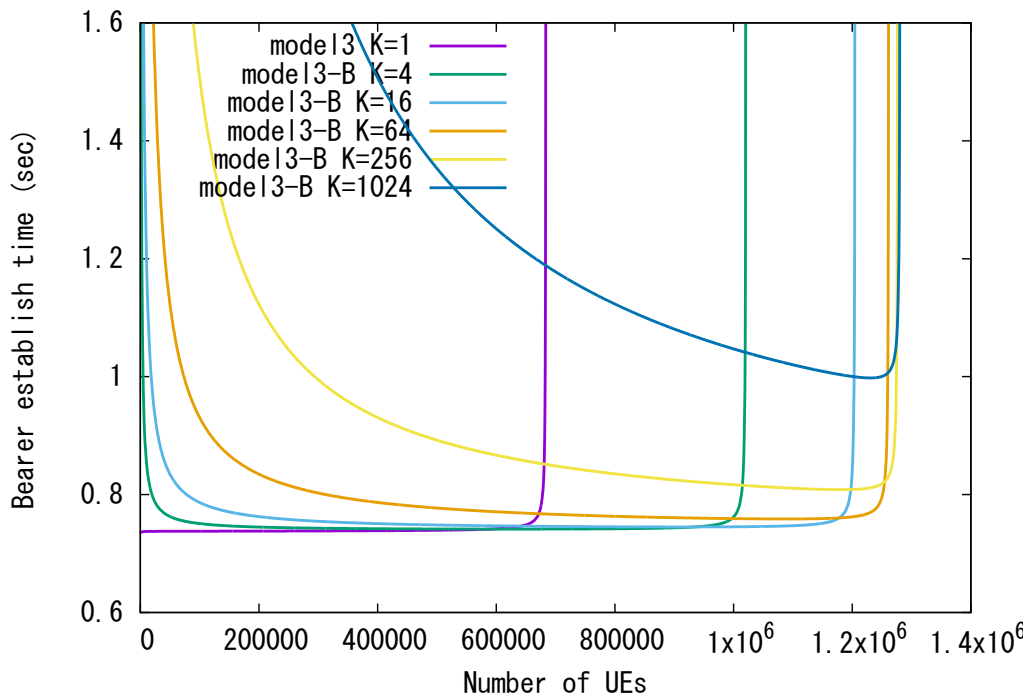


(b) モデル 3-B における通信集約手法の効果

図 15: LENA に基づく場合の通信集約の効果



(a) モデル3-Aにおける通信集約手法の効果

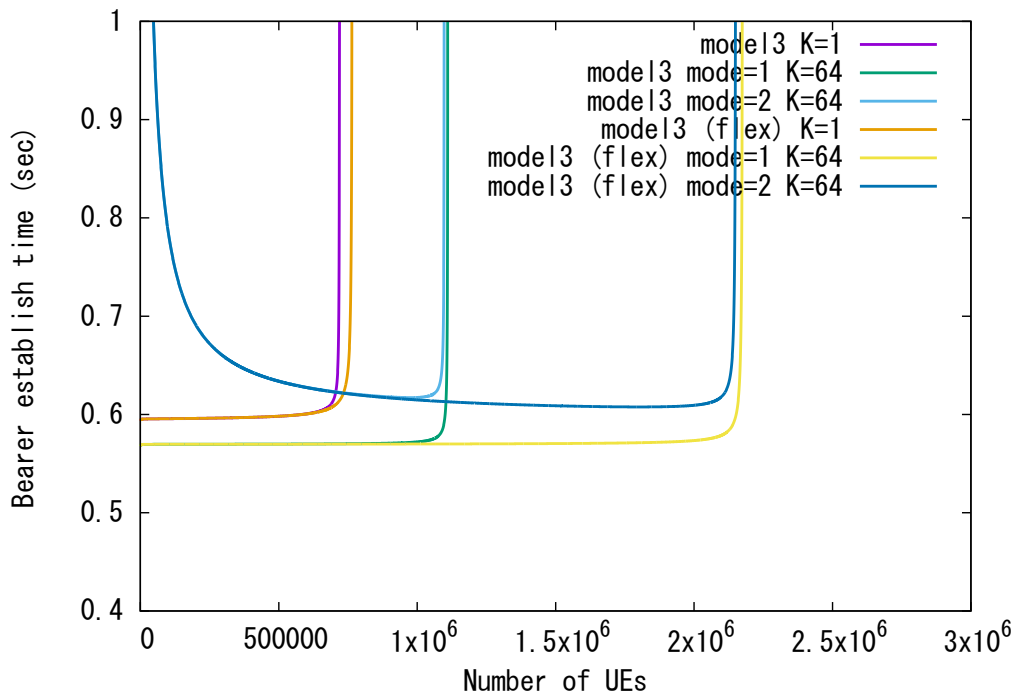


(b) モデル3-Bにおける通信集約手法の効果

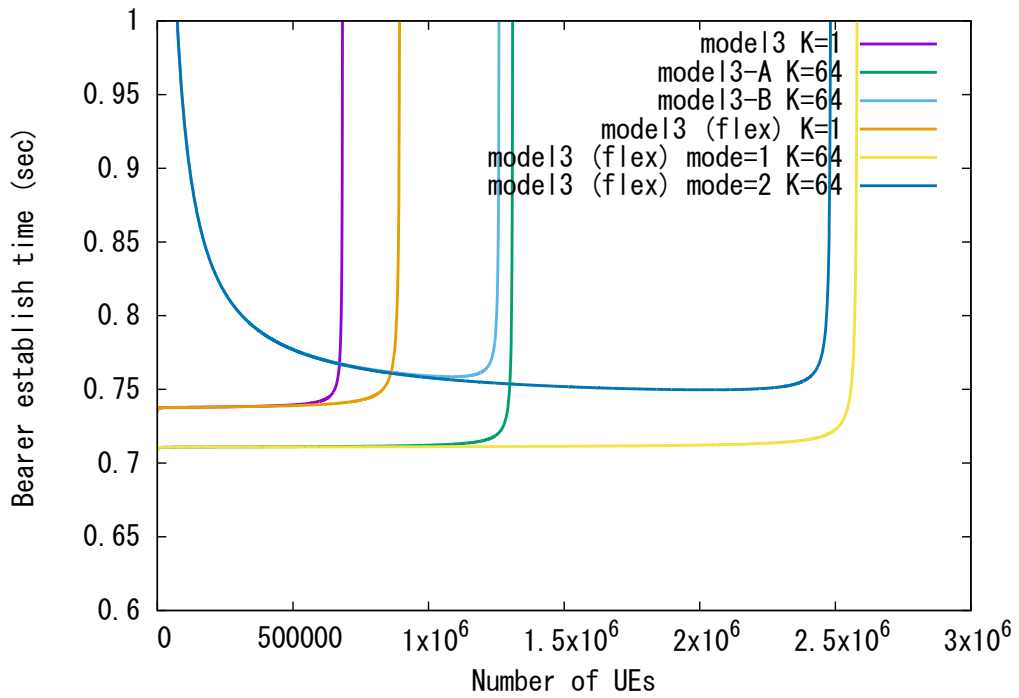
図 16: OAI に基づく場合の通信集約の効果

図 17 に、モデル 3-A、及びモデル 3-B において処理性能の融通を組み合わせた場合の評価結果を示す。図 17 より、いずれの集約方法を用いた場合でも、処理性能の融通を組み合わせることで収容可能端末数が大きく増加していることがわかる。また、この増加量は、処理性能の融通及び通信集約のいずれかを用いた場合の収容可能端末数の増加量の和よりも大きい。この傾向は文献 [12] における結果の傾向と一致している。これは、ノードの仮想化及びプレーン分離だけではノードにおける負荷を大きく減少することはできず、ボトルネックとなる MME に融通できる処理能力は限定的であるが、通信集約との組み合わせにより、各ノードにおける負荷が減少するため、融通できる処理能力の量が増加するためである。

評価の結果、収容可能端末数は、LENA を基にしたシグナリング手順とシグナリング経路で評価した場合はモデル 1 と比較して約 200%増加し、OAI を基にしたシグナリング手順とシグナリング経路で評価した場合は約 280%増加した。増加量に差が存在する理由は、図 14、15、16 で説明したように、処理性能の融通、及び通信集約による効果はいずれも OAI の方が大きいため、その効果を組み合わせることで、より大きな効果が得られたためである。



(a) LENA に基づくシグナリング手順とシグナリング経路で評価した場合



(b) OAI に基づくシグナリング手順とシグナリング経路で評価した場合

図 17: 通信集約手法と処理性能融通の組み合わせの効果

5 まとめと今後の課題

本報告では、モバイルコアネットワークにおける M2M/IoT 通信の収容能力の拡大を目的とし、モバイルコアネットワークを構成するノードのプレーン分離や処理機能のクラウドネットワーク内への設置、処理能力の融通及び通信集約の効果を、数学的解析手法によって評価を行った。評価の結果、従来のモバイルコアネットワークを基にしたモデルにおける M2M/IoT 端末の収容可能台数に比べ、プレーン分離や処理機能のクラウドネットワーク内への設置、処理能力の融通及び通信集約によって、シミュレーションソフトウェアを基にしたモデル、及び実アプリケーションソフトウェアを基にしたモデルにおける収容可能台数が増加することがわかった。また、プレーン分離や処理機能のクラウドネットワーク内への設置を行った上で、処理能力の融通及び通信集約の効果について評価をし、それらの手法を組合せて適用する場合には、それぞれの手法を単一で適用する場合の効果の和よりも大きく効果が得られ、収容可能端末数は、プレーン分離前のモデルと比較すると、シミュレータソフトウェアを基にしたモデルにおいては約 200%、実アプリケーションソフトウェアを基にしたモデルにおいては約 280%増加することがわかった。

今後の課題としては、シグナリング処理負荷のさらに詳細な評価を行いたい。具体的には、実装コードのコンパイル後の命令数による評価や、実機上で EPC ノードを動作させることによる評価を行いたい。また、本報告では単一の EPC ネットワークを対象に評価を行ったが、これを基に、複数の EPC ネットワークから構成される広域モバイルネットワークを対象にした評価を行いたい。

謝辞

本研究を進めるにあたり、多くの方々の御指導を頂きました。指導教員である松岡茂登教授には、初めての研究を進めるにあたり、研究に対する姿勢を教えていただき、また、様々な場面で御支援や御指導をいただきました。ありがとうございました。また、村田正幸教授には、研究を始めた当初より、様々な観点からの鋭い御指摘や御指導を頂きました。私が考えに至っていなかったところを御指摘頂き、改めて自分自信の進めている研究や、その取り組み方を省みる機会を得ることも多く、大変感謝しております。ありがとうございました。そして、日々の研究から本報告の作成に至るまで、長谷川剛准教授には終始丁寧な御指導を頂きました。研究を進めるにあたって至らないところも多く、四苦八苦しながらも、自分自身が進めている研究に対する興味と前向きな気持ちを忘れず、今後の展開にも意欲を持つことができたのは長谷川剛准教授のおかげです。心から感謝しております。ありがとうございました。樽谷優弥助教には、研究から研究室での生活に至るまで様々な質問をさせて頂き、その度に時間を割いて丁寧に、かつ親身になって御指導頂きました。日頃より学生に寄り添い、温かく御指導くださり、大変心強かったです。ありがとうございました。

また、研究に関する御支援や御助言を頂いただけでなく、普段の生活でも様々な場面で松岡研究室の先輩方には御世話になりました。研究室に配属された当初、何もわからない私にも温かく接して下さり、研究室生活を楽しく過ごせたのはひとえに先輩方のおかげだと思っています。ありがとうございました。研究チームで一緒だった村田研究室の坂田航樹君、初めての研究にお互い戸惑いを感じながらも研究の面白さや難しさを共有できたこと、様々な場面で助け合える関係を築けたことを嬉しく思っています。ありがとうございました。最後に、研究チームの垣根を超えて、研究室に配属された当初から良き仲間、良き理解者として私の研究生活になくってはならなかった同期の方々、心より感謝申し上げます。思うように成果が出ないことや、辛いことがあっても皆さんのおかげで乗り越えられました。ありがとうございました。

参考文献

- [1] G. Wu, S. Talwar, K. Johnsson, N. Himayat, and K. Johnson, “M2M: From Mobile to Embedded Internet,” *Communications Magazine*, IEEE, vol. 49, no. 4, pp. 36–43, Apr. 2011.
- [2] A. Daj, C. Samoila, and D. Ursutiu, “Digital Marketing and Regulatory Challenges of Machine-to-Machine (M2M) Communications,” in *Proceedings of Remote Engineering and Virtual Instrumentation (REV), 2012 9th International Conference on*, pp. 1–5, July 2012.
- [3] 岩井 孝法, 長谷川 剛, 若宮 直紀, 本吉 彦, 吉川 隆士, “モバイルコアネットワークにおけるトンネル集約方式,” *電子情報通信学会技術研究報告*, vol. 112, no. 350, pp. 115–120, 2012 年 12 月.
- [4] 櫻井 駿, 長谷川 剛, 若宮 直紀, 岩井 孝法, “モバイルネットワークにおける通信制御コスト削減のための端末集約手法の性能評価,” *電子情報通信学会技術研究報告*, vol. 113, no. 56, pp. 193–198, 2013 年 5 月.
- [5] 岩井 孝法, 長谷川 剛, 若宮 直紀, “モバイルコアネットワークへのアクセス負荷を平滑化するためのバックオフ機構の拡張,” *電子情報通信学会技術研究報告*, vol. 113, pp. 67–72, 2014 年 1 月.
- [6] I. Widjaja, P. Bosch, and H. La Roche, “Comparison of MME Signaling Loads for Long-Term-Evolution Architectures,” in *Proceedings of IEEE Vehicular Technology Conference*, pp. 1–5, IEEE, Sept. 2009.
- [7] A. Basta, W. Kellerer, M. Hoffmann, K. Hoffmann, and E.-D. Schmidt, “A Virtual SDN-Enabled LTE EPC Architecture: A Case Study for S-/P-Gateways Functions,” in *Proceedings of 2013 IEEE SDN for Future Networks and Services (SDN4FNS)*, pp. 8–14b, Nov. 2013.
- [8] A. Basta, W. Kellerer, M. Hoffmann, H. J. Morper, and K. Hoffmann, “Applying NFV and SDN to LTE Mobile Core Gateways, The Functions Placement Problem,” in *Proceedings of the 4th Workshop on All Things Cellular: Operations, Applications, & Challenges*, pp. 33–38, ACM New York, NY, USA, Aug. 2014.

- [9] 長谷川 剛, 村田 正幸, “モバイルコアネットワークにおける M2M 通信集約手法の解析的評価,” 電子情報通信学会技術研究報告, vol. 114, no. 160, pp. 51–56, 2014 年 7 月.
- [10] F. Yousaf, J. Lessmann, P. Loureiro, and S. Schmid, “SoftEPC, Dynamic Instantiation of Mobile Core Network Entities for Efficient Resource Utilization,” in *Proceedings of Communications (ICC), 2013 IEEE International Conference on*, pp. 3602–3606, June 2013.
- [11] X. An, W. Kiess, and D. Perez-Caparras, “Virtualization of Cellular Network EPC Gateways based on a Scalable SDN Architecture,” in *Proceedings of Global Communications Conference (GLOBECOM), 2014 IEEE*, pp. 2295–2301, Dec. 2014.
- [12] 長谷川 剛, 村田 正幸, “M2M 通信収容のための仮想モバイルコアネットワークアーキテクチャに関する一検討,” 電子情報通信学会技術研究報告, vol. 114, no. 477, pp. 493–498, 2015 年 3 月.
- [13] M. R. Sama, S. Ben, H. Said, K. Guillouard, and L. Suci, “Enabling Network Programmability in LTE / EPC Architecture Using OpenFlow,” in *Proceedings of Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2014 12th International Symposium on*, pp. 389–396, IEEE, May 2014.
- [14] “LENA.” available at <http://networks.cttc.es/mobile-networks/software-tools/lena/>.
- [15] “OpenAirInterface.” available at <http://www.openairinterface.org/>.
- [16] V. S. Rao and R. Gajula, “Protocol Signaling Procedures in LTE,” Sept. 2011.
- [17] G. Ralf, Kreher and Karsten, *LTE Signaling: Troubleshooting and Optimization 1st Edition*. Wiley, Jan. 2011.
- [18] C. Christopher, *An Introduction to LTE: LTE, LTE-Advanced, SAE, VoLTE and 4G Mobile Communications: Second Edition*. Wiley, May 2014.
- [19] A. Sassan, *LTE-Advanced: A Practical Systems Approach to Understanding 3GPP LTE Releases 10 and 11 Radio Access Technologies*. Academic Press, Nov. 2013.
- [20] M. Magnus, Olsson and Catherine, *EPC and 4G Packet Networks, Second Edition: Driving the Mobile Broadband Revolution*. Academic Press, Dec. 2012.

- [21] 伊藤 学, 西永 望, 北辻 佳憲, “加入者識別子 (IMSI) の共用によるセルラー IoT 端末群の通信回線集約に関する検討,” 電子情報通信学会技術研究報告, vol. 115, no. 311, pp. 83-88, 2015 年 9 月.