A Metering Infrastructure for Heterogenous Mobile Networks[‡]

Andreas Monger*, Marc Fouquet[‡], Christian Hoene*, Georg Carle[‡] and Morten Schläger[†]

[‡] Network Architectures and Services, Technical University of Munich
* Computer Networks and Internet, Wilhelm Schickard Institute
for Computer Science, University of Tübingen, Germany
[†] Nokia Siemens Networks, Berlin

***Abstract*—The next generation of mobile access networks will support different access technologies like UMTS, WLAN, WiMax and LTE. Optimal handover decisions between different access networks have to take various types of information into account, such as the load of possible target cells. However, collecting this information and transporting it to mobility management decision engines is costly in terms of bandwidth. With our flexible *Generic Metering Infrastructure* (GMI) we are able to collect state information from the core and access networks efficiently. Using the GMI, network operators will be able to control handover decisions for their users.**

I. Introduction

Today many operators already run WLAN access points to supplement their GSM and UMTS networks. In the future they will add new radio access technologies such as Long Term Evolution (LTE) [1] or WiMax [2] to improve user experience and offer new services and to increase the capacity of their access networks. To support multiple access technologies, new network control- and management functionalities have to be introduced, which support heterogeneous handovers and resource management to make sure that mobile devices are optimally connected to one or more of the currently available access networks.

Future mobile networks should be able to offer seamless handovers between different access technologies. As the management of heterogeneous networks is a major topic in current research a new term named *Always Best Connected* (ABC) has evolved in recent years [3]. It is a catchphrase that involves both sides in mobile communication to be considered: On the one hand, users would like to be connected to "best" available network in terms of e.g. signal quality and bandwidth. On the other hand the network provider wants to share the available resources between millions of users in a fair or privilege-based fashion.

State of the art mobile devices support multiple access technologies and the handover between them. Generally, either the mobile devices or their users decide which access technology to choose. But a mobile-driven handover decision is not always desirable because the following reasons: First, this kind of network selection poses a burden on the end user, if it involves user interaction. To hide this complexity from users, smart decision functionalities should be provided that do not involve the user directly. Second, each node has to scan for neighbouring networks, which consumes a considerable amount of power. To prolong the battery run time of mobile devices, radio transceiver that are not needed shall be turned off most of the time. Third, as each mobile node optimises only its own connection without considering the impact of its decisions on other nodes, the decision leads to potentially suboptimal solutions. If, for example, one mobile node connects to a WLAN access network while having a bad link and slow modulation, this can significantly worsen the connection quality of other nodes that are also connected to the same access point [4]. Instead it would be better if the node connected to another access technology in which it would not degrade the connection quality of adjacent fellow nodes.

The "Wizard of Oz" view on the network allows for a decision functionality on the network side that can take numerous users and base stations of heterogeneous networks and network-side parameters like the utilisation of certain base stations into account. Thereby it can provide better decisions than a single user may make, which promises better performance and allows for a network-controlled sharing of resources in heterogeneous networks.

Currently, a lot of effort in 3GPP [5], [6] and IEEE [7] is spent on integrating multiple existing and future access technologies. Our work contributes to these efforts and provides an architecture that enables the management of distributed, heterogeneous networks. To provide heterogeneous handovers we follow an approach that separates data collection, decision making and execution of handovers. Previous work (such as [3], [8], [9]) has focused on decision making, when to conduct the handovers. These approaches require state information from the access network. For example, Staehle [9] defines how to get this data from an UMTS network. IEEE 802.21 [7] defines a media-independent interface to support handover decision on the mobile device. Our approach differs from the related work as it proposes an abstraction layer between data collection and decision making (similar to IEEE 802.21) but allows for technology-independent support of heterogeneous network-controlled and network-assisted handovers.

We present the *Generic Metering Infrastructure* (GMI) that

This work has been developed in cooperation with Nokia Siemens Networks within the BmBF ScaleNet-SYMPATHIE project.

is able to provide decision making entities with the desired information. In Chapter II we describe the tradeoffs that we needed to consider during the design of the GMI. The following benefits can be provided by the GMI:

- Sections III-B, III-B8 and V will show that we can significantly reduce the number of signaling messages that contain information on the state of the networks by generating optimised information delivery paths and combining multiple aggregation techniques.
- Towards the decision-making entities, the GMI offers an interface that generalises the configuration of measurement tasks and supports the most common types of access networks (section III-B9).
- We offer an information collection and delivery service that may serve clients in (soft) real-time (see III-B8) and enable a faster and more precise decision making.

We have implemented the components that make up the Generic Metering Infrastructure. Chapter IV focuses on this subject. The Implementation enabled us to start evaluating our approach in an emulated environment. Chapter V will show our first results. Finally, we have a look at related work in chapter VI and provide a summary in chapter VII.

II. Problem Analysis

Although there already is a way of monitoring and managing current provider networks it lacks of real-time capability [10]. In UMTS networks, for example, mechanisms for performance management exist which allow for requesting various information for individual cells. However, UMTS NodeBs are usually connected to the Core Network by links with limited capacity. In this case, for Operation, Administration and Maintenance (OAM) a channel of approximately 64 kBit/s, is available. The load of a specific cell can be requested using these interfaces. However, the usual OAM interface is realised by uploading ASN.1-files via FTP in intervals of 30 minutes for bandwidth reasons. This kind of interface is not suitable for quick mobility decisions, so alternatives need to be explored.

Furthermore, the signalling information between the User Equipments (UEs) and the decision making entities on the network's side reduces a user's bandwidth, since this information has to be sent over the same link as the user's traffic.

Therefore, the design of the GMI must keep the load on sources of information low and save bandwidth and computing resources, especially at bandwidth bottlenecks in the RANs. Redundant transmission of data must be avoided and the number of measurement reports should be kept as low as possible. On the other hand the decision making entities need current information to make good decisions, in some situations it is even necessary to retrieve measurement data on the instant (i.e. in a request/reply fashion).

Based on research of our project partners at DAI-Laboratories [8] four exemplary categories of data have been identified that are required for their Network Resource Management (N-RM) decisions.

- *Load information* is required for prevention or handling of overload situations. Thus in case of an overloaded cell or access network, UEs can be moved to different networks.
- *Signal quality* gives the N-RM the chance to move users to a different access network, if the radio conditions are insufficient.
- *Mobility and location information* about a user helps the N-RN to estimate, which alternative networks or cells are available at the user's location.
- *Perceived Quality of Service (QoS)* can be a general indicator that tells the N-RM about a customer whose service quality is degrading. The reason for this will usually be either network overload, an inappropriate access selection or bad signal quality - or a combination of these factors.

Having examined the information categories we will focus on the question of how the data should be reported to client applications of the GMI (e.g. the N-RM). We distinguish three types of information delivery mechanisms that address different requirements to serve a client's needs:

- *Periodic reports* keep the N-RM constantly informed about the network's state. By monitoring certain measurements the N-RM may be able to act proactively on changing conditions before critical situations occur.
- If nevertheless a critical situation occurs, reports should be *triggered* immediately, so the N-RM can react as quickly as possible.
- Additionally, it should be possible for N-RM to get direct access to a value in a request/reply fashion. Unlike the former methods that require a previous announcement of interest in measured data, such a request is only sent *once* and is answered immediately.

III. Design

Major goals of our design are flexibility in configuration and a reduced load on the transport network to save resources.

Most of today's protocols for network management (for example SNMP [11]) are based on the client-/server-paradigm that relies on a closely time-constrained request-/reply-message architecture. But this approach does not suit well for our field of operation. A basic assumption that has already been stated by T. Bandh [12] is that information is of most interest in critical situations. If e.g. the load in a cell rises to a critical level, an N-RM will surely want to be informed immediately and regularly, but if the level of load reduces to a "normal" level the information is of less importance. This observation led to the design of a modified Publish/Subscribe System (P/S-System).

A. Publish/Subscribe Systems

Publish/Subscribe Systems are event-driven and decouple senders and receivers in two dimensions: Instead of polling the source of information regularly, an interested party registers for events only once (decoupling of time). There might be more than one interested party for some information and the source of the information may not want to or not even be able to send a copy of the information to each recipient. Again it

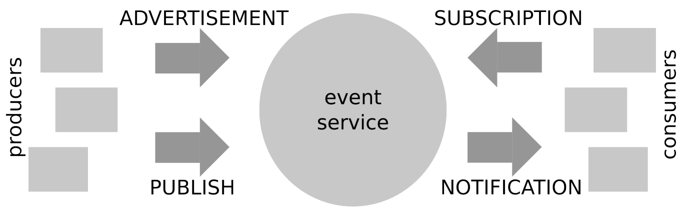

Fig. 1. Basic concept of an Event Service.

would be desirable to publish the information only once for all recipients (decoupling of space). This concept leads to highly asynchronous communication.

Figure 1 illustrates the basic concept of P/S-Systems. As can be observed the interaction is entirely information-driven, thus the source of information and the receiver of information are unaware of each other. Objects of interest or *producers* may *advertise* events to the event-broker system in form of a topic or type that specifies the sort of information they may *publish*. On the other hand, interested parties or *consumers* may *subscribe* at the P/S-System for certain sorts of information they're interested in. As soon as a source of information *publishes* some "news" the event-broker-system starts to dispatch this message and *notifies* all interested parties of the occurred event.

Filter Models: As not all published information is relevant to each of the consumers, there must be a way to reduce the amount of information that will be delivered to a single client. After a consumer submitted a subscription, the event broker is in charge of deciding whether a notification is of interest for a consumer or not. This issue is solved by the introduction of *filters*. A filter is a boolean function that can be applied on all notifications and evaluates to *true*, if the consumer is interested in the notification, or *false*, if it is not [13].

The filter model determines the degree of flexibility a P/S-System achieves. Basically there are two categories of approaches for filter models. The first category defines a fixed set of topics. The producer of information may decide under which topic it publishes its information.

Channel-based filters offer a predefined set of topics to publish information. The flexibility for classifying the messages is limited by the amount of existing channels, thereby additional filtering of information on the client's side is often needed. The subject-based model organises notification topics (or subjects) in a tree-structure. Again the object of interest chooses the subject to publish its reports. A subscribing consumer may specify a single leaf of the tree or an intermediate node. After subscribing to an intermediate node a consumer receives all reports that are published at any leaf of the corresponding sub-tree.

These two approaches are simple and easy to implement, but they hinder changes. If the topic-assignment for a type of notification is changed, both producer and consumer have to switch topics simultaneously, to avoid losing information.

The second category of filter models enables subscriptions which refer to the actual content of a message (content-based filters [14], [15]). A producer does not have to categorise its notifications anymore. The event system is responsible for deciding whether the information contained inside the message is relevant for each subscriber or not. Although this approach is more flexible in terms of specifying subscriptions, it is also much more complex to realise and places a heavy burden on the event system.

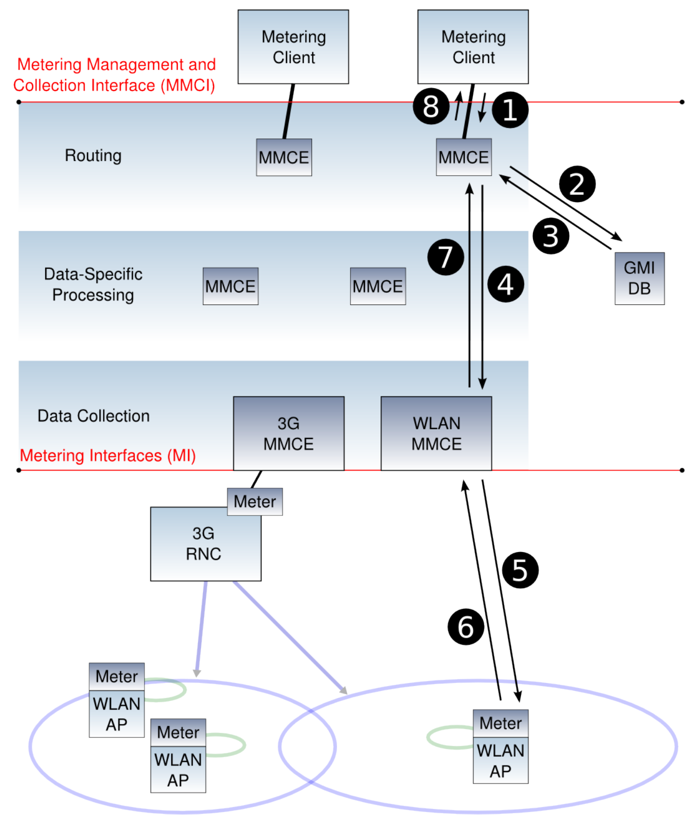

Fig. 2. GMI signalling overview.

B. GMI Event-Service

The GMI adapts and modifies the concept of P/S-Systems. Its event-brokers are called Metering Management and Collection Entities (MMCE).

1) Overview and signalling: An overview of the GMI can be seen in Figure 2. At the top we have the interested partys, which are called metering clients in our terminology. At the bottom we find the actual meters that produce the measurement data, in mobile networks they could be placed at network nodes like RNCs or WLAN APs.

The GMI itself is split into 3 sublayers. The MMCEs that interface directly with the metering clients are primarily meant to route requests to the correct lower MMCE. Here it should be noted that MMCEs are logical functions that don't necessarily have to be "physical boxes", an MMCE could also be a process on the Metering Client's machine.

Data-specific MMCEs offer additional value-added services to the clients, examples are given in section III-B6. On this layer we can also build multicast-like distribution trees if multiple clients are interested in the same data.

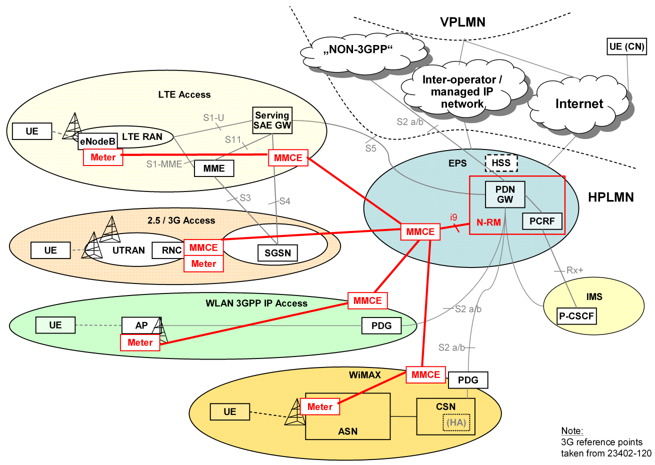

Fig. 3. Mapping of the GMI to the SAE network architecture.

MMCEs on the data collection layer directly interact with the meters. As the meters may require different protocols for configuration and data delivery, the primary purpose of these MMCEs is translation between GMI messages and the Meter's protocols. Of course one could also develop native GMI meters that do not require this step.

Figure 2 also shows some example signalling. This is the simplest case without distribution trees or data-specific MMCEs - a single Client requests data from one single Meter.

The Metering Client is interested in some WLAN data and sends a CREATE message to his local MMCE (1). This MMCE performs a lookup in the DNS-like GMI-database (2) to find the source of the requested data. Having receiving a reply (3), the MMCE forwards the CREATE message to the MMCE assigned for that meter (4). This meter-assigned MMCE is in charge of configuring the meter for this measurement task (5). As soon as new data is available, the Meter sends a report to its assigned MMCE (6). The message is translated into a GMI PUBLISH message which is forwarded to the Client subsequently (7), (8).

2) Positioning of the MMCEs: Figure 3 shows, how the GMI could be deployed in 3GPP's System Architecture Evolution networks. Here a single Network Resource Management (N-RM) instance is the only metering client. In a real network, multiple resource management decision engines and possibly other management system would obtain their information from the GMI.

All network elements shown in Figure 3 produce metering data that is of interest to the clients of the GMI. The MMCEs are placed as close to the meters as possible, but above bandwidth bottlenecks. This is easy to accomplish with UMTS networks, as RNCs are central entities which possess the required information for hundreds of cells. With HSDPA this situation partly changes, as scheduling and radio resource management have been moved to the NodeB, so in this case a Meter on the NodeB is required.

LTE networks have no RNC anymore, the radio resources are controlled by the eNodeB, so a Meter is required there. With WLAN and WiMax, information about the radio links is also available at the actual base stations.

3) Late Duplication: The MMCEs form an acyclic network of nodes that allows for building up distribution paths for published information that are similar to multicast trees. Thereby the so-called "late-duplication" is applied here. This reduces the bandwidth consumption of the event-system as the messages that have multiple recipients are duplicated as close to the recipient as possible. C. Chalmers [16] investigated the benefit of using multicast trees compared to unicast communication. The general advantage of multicast trees is hard to predict because it depends on multiple factors, such as the breadth and height of the tree and the number of receivers. But it can be observed that the number of messages sent grows logarithmically, if any of the named factors grows linearly. This aspect is especially important in huge networks. By reducing the number of sent duplicate messages to a minimum, the system's complexity and costs are kept low.

4) Addressing: As already indicated in section III-B1 the GMI is a modified Publish/Subscribe System. The major difference lies in the decoupling of senders and receivers of notifications. The decoupling in space would prevent the metering clients from directly assigning metering tasks to specific meters. This aspect has been overcome by introducing a scheme that combines the addressing of meters and the filter model of the GMI's event service.

This approach introduces a DNS-addressing scheme that is based on the 3GPP TS TS23.003 [17] (Annex C and D). This standard proposes DNS-like addressing of network functions such as a Serving GPRS Support Node (SGSN). An SGSN can be addressed by appending its name and identifier (e.g.: 1B34) to an operator's top-level domain:

```
sgsn1B34.<mobile national code>.<mobile
    country code>.3gppnetwork.org
```

The 3GPP only defines DNS-names for those network nodes that are addressable by IP. But as we want to be able to address meters like UMTS NodeBs that are not capable of IP, we propose an easy extension of the naming scheme to these entities.

The introduction of a new MMCE-domain below the operators top-level domain spans a new overlay network of Metering Management and Collection Entities where each meter forms its own DNS-domain within the mmce-domain. That way a self-explaining addressing scheme can be deployed, for example data regarding an UMTS NodeB could be found under the address

```
nodeB0123.mmce
.mnc123.mcc123.3gppnetwork.org
```

Obviously this scheme introduces an intended indirection. The DNS-names do not refer to the meters themselves but to their correspondent MMCEs. They are responsible for administering metering tasks and forwarding measurement reports.

5) GMI Subject-Tree: The GMI uses the subject-based filter model that has been briefly introduced in section III-A.

Because the previously defined domain-names already imply a hierarchical structure, they can directly be mapped into a subject-tree. This tree does not only allow addressing of meters but it also contains addresses for the data that is measured there. Each measurement forms a new subdomain within its meter's domain. That way, a measurement can be addressed by appending its name to the address of the meter.

For instance the topic *FailOutintraNodeB* (number of failed outgoing intra-NodeB hard handovers) can be measured at each NodeB [18]. Because there can be multiple reasons for a handover-failure this topic is split up into several reasons or sub-measurements. Among others these can be: *NoReply, ProtocolError, RadioLinkFailure* or *sum*. The resulting name of the measurement *NoReply* would be:

```
NoReply.FailOutintraNodeB.nodeB0123
.mmce.mnc123.mcc123.3gppnetwork.org
```

It shall be understood that the other parameters can be addressed accordingly.

These capabilities can be advertised on start-up of the network to the attached MMCE of a meter. Thereby each MMCE has to cope only with the information of its attached meter and doesn't have to be aware of the information that can be found on other Metering Management and Collection Entities distributed in the network.

6) Tracking mobile sources of data: As mentioned before, subject-trees do have their drawbacks, especially when source of certain information changes its location. For example, an N-RM application may want to be kept informed about the signal quality of a single user. The information resides at the base station the user is currently attached to. But the source of information may change as the user switches to a different base station. If the metering client is unaware of that, it will not receive updates anymore.

This means that the proposed approach has to be extended to meet the requirement of keeping track of changing sources for the same data. The introduction of "*hooks*" is meant to address this problem. Event producers with dynamically changing subjects of information may advertise *hooks* to announce predefined sorts of information, that contain variable content.

A hook is defined by a template and its available instances. The template defines a sub-tree structure containing measurements that can be found at each instance. Each meter that can provide dynamically changing data must use hooks to announce its changing capabilities. An MMCE that is attached to such a meter advertises a template for its hook and its current instances instead of only specifying a static configuration. An example for such a hook is

```
DownlinkCQI.imsi012345.uehook.nodeB0123
.mmce.mnc123.mcc123.3gppnetwork.org
```

A data specific MMCE (DS-MMCE) as introduced in section III-B1 can be used to hide this mobility. Such an MMCE can subscribe to data under a hook and to a trigger that fires when the data changes its location, i.e. a handover-tigger regarding a UE. On a change, the MMCE would re-subscribe to the data at the new location. So the DS-MMCE would make this data available to other MMCEs and clients at a constant location.

7) Generic Measurements: Another feature of our system are so-called "generic measurement tasks". Meters can contain plug-in modules for specific tasks, i.e. flow-based QoS measurements. A configuration for such a plug-in may be sent to a meter using normal GMI mechanisms, which means that it is encapsulated in a GMI-subscription. At the meter, the configuration is forwarded to the plug-in. The results of the measurements are assigned to an identifier (i.e. a flow ID) and published in the P/S tree at a special hook.

```
AvgDelay.flow0123456.qosplugin.ggsn0123
.mmce.mnc123.mcc123.3gppnetwork.org
```

This is useful when a metering task requires configuration that is too complex to be encoded in a GMI address.

8) Measurements and Events: As already mentioned in section II, a basic requirement of our architecture is the support of a flexible set of measurement tasks.

To emphasise that a subscription in the GMI's sense differs from the classical sense of a P/S-System, we replaced the *SUBSCRIBE* message with a *CREATE* message. Its basic functionality remains the same (declare the interest in a certain type of information). But a CREATE-message may also cause new measurement tasks to be created at the meter.

Periodic measurements: A metering client can subscribe to periodic metering tasks. In this case the subscriber specifies a desired report period and the measurement value it wants to stay informed about. If there already is a subscription that matches the desired measurement, the sender of the *CREATE* message is appended to the already existing list of receivers. If there is no matching subscription, a new metering task will be started. Thereby the GMI ensures that only one periodic measurement task with the given report period is active at a time.

Triggers: It is also possible to set triggers for measurements. Such a subscription sets one or multiple thresholds for a metered value. If the value rises above or falls below the given threshold the metering client is informed immediately. Trigger subscriptions contain a hysteresis parameter to make sure that a value which oscillates around the threshold does not cause an unnecessarily large amount of messages. A trigger notification consists of two values: the former value and the currently measured value that caused at least one trigger to fire. This enables an implicit aggregation of triggered measurement reports. On the one hand the meter has to send only one notification; on the other hand each intermediate node can decide which subscribers of triggers have to be informed. This enables "late duplication" according to the classical P/S-scheme: Messages that are of interest for multiple recipients are duplicated on their way to the destinations as late as possible to avoid redundant message transmissions over the same links. Additionally, a client may subscribe to classical events like handovers and connection losses. These events are not associated with a numerical measurement value inside the meter, but stand on their own.

Request/Reply: The last type of reporting is an immediate response to a request of a metering client for a certain value of data. This notification is not an event in the classical sense of a P/S-System. In this case the metering client simply sends a request for the value (which is a message similar to a subscription) and receives a reply containing the value (which is handled like a notification). In this case no aggregation is possible as the message is only sent to a single client. However caching of values can limit the number of requests to the meters, if the cached information is still current enough.

9) Interface towards the Clients: The GMI is a service that acts as a middleware between metering clients and the meters. A metering client is expected to connect to only one MMCE that serves as its access point to the GMI. This MMCE is assigned by the network operator.

The GMI provides an abstraction layer that allows a metering client to create measurement tasks for every meter within the network the same way. The message format for different meters (e.g. a WLAN access point and an SGSN) has the same structure although the actual configuration of measurement tasks at these meters may be very different since vendor- and implementation-specific aspects often have to be considered. Here the lowest MMCE, which directly interfaces with the meters, is in charge of translating the requests according to the meters' specification. Thereby a metering client does not have to worry about device-specific aspects of different meters.

IV. Implementation

The key components of the envisioned system have been implemented in our laboratory. Our MMCE instances maintain the subscriptions in a tree. In our current implementation all messages are XML-based and also interpreted as tree-like data structures. Every piece of metered information must be advertised in advance, so the system can add it to the subject tree. Resources can be dynamically added and removed at runtime. The MMCEs that receive these advertise-messages use them to store routing information as annotations in the subject-tree. The routing information can be distributed among the MMCEs if needed, so each MMCE only holds the routing information that is relevant for its own operation.

Received "PUBLISH" messares are interpreted as a subset of the subject-tree. The forwarding decisions made at each intermediate MMCE are based on an algorithm that traverses the received tree node by node and matches it to the subject-tree. This easily allows to determine the receivers of an event.

We have implemented our concept in the Python programming language. Our MMCEs are individual applications that communicate via TCP sockets and can be run on a single machine or be distributed on several ones.

As the GMI is middleware, it needs data sources and metering clients to run. As a data source we have implemented a small network emulation application, which allows simulated users to move on a map, connect to different radio cells and start sessions. Each cell maintains a capacity counter. The model is simple, but it allows for homogeneous and heterogeneous handovers and is sufficient for initial tests of the GMI. Figure 4 shows a screenshot of our simulator.

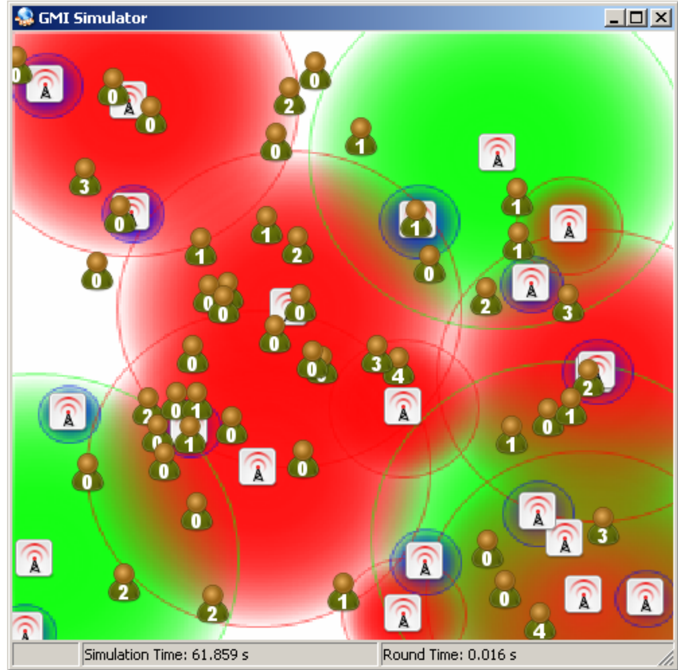

Fig. 4. Screenshot of the Simulator that generates input data for the GMI. Here three different access technologies (UMTS, WLAN, LTE) and 50 mobile users are shown. The number printed on the users are the currently active sessions.

V. Evaluation

The evaluation of the GMI is a difficult task, since the software is middleware. The behavior and the performance of the GMI greatly depend on the underlying meters and - especially - on the subscriptions of the metering clients. In this work we present some preliminary evaluation results that give some insights on the applicability of the GMI's set of measurement tasks (see section III-B8).

In the evaluation, the load of one cell of our simulator is measured. The load value is updated at the meter in intervals of one second. A light-weight metering client sends a subscription for the data and compares the received result with the expected original curve using the $L2$ norm. Basically a subsampling of the original curve is applied, for example each 10th value is transmitted if the metering task is set to "periodic, 10 seconds".

Besides using simple periodic measurements, we tested a combination of periodic measurements and triggers. The upper hysteresis thresholds of four triggers are set to 50%, 65%, 80% and 90% of the cell's total capacity (the lower thresholds are at 48%, 63%, 78% and 88%). Figure 5 shows a small part of the data sampled with these triggers. Here the filled curve is the original data while the black lines show what the metering client sees.

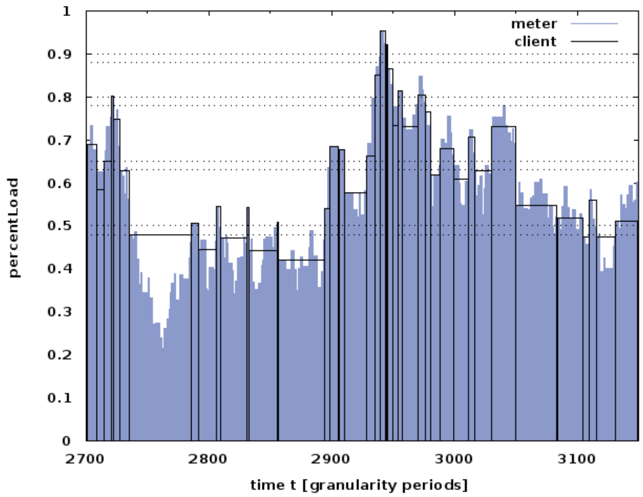

Fig. 5. Example from our evaluation data set. The filled curve is the original load value that appears in a simulated cell (view of the meter). The black curve is the output of a GMI-measurement job using four triggers (view of the GMI's client). Pairs of dashed horizontal lines indicate the upper- and lower thresholds for the triggers.

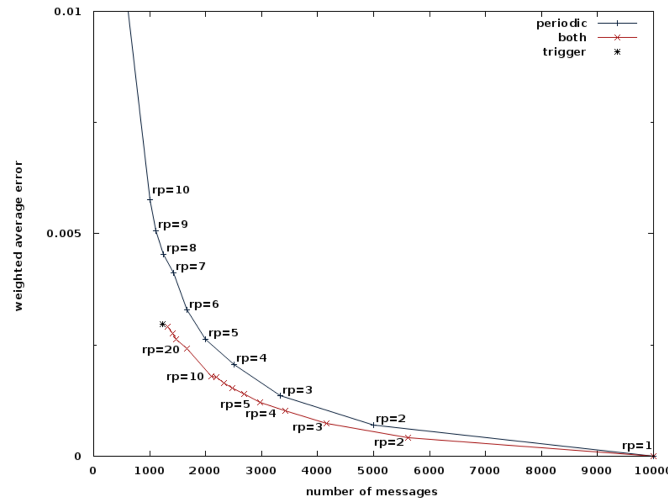

Fig. 6. Evaluation results: Measurement accuracy versus the number of sent messages.

During the 10 000 seconds of simulation time, the load in the cell varies, but it is generally above 50%; some peaks even touch the cell's capacity limit.

Figure 6 shows the result of the evaluation. Here the accuracy of the measurement ($L2$ distance between measured curve and real curve) is drawn against the number of reports that had to be sent to achieve this accuracy - so values that are closer to the origin are better. With the "periodic" curve, one can see the trade-off between accuracy and report period - more reports produce a higher accuracy. The "both" curve shows that adding triggers increases the accuracy. Without triggers, a report period of 5 seconds (which equals to 2000 messages) is needed to achieve an accuracy of 0.0026 in the $L2$ norm. The same accuracy can be reached by activating the triggers as described above and setting the periodic report period to 35 seconds - and with this setting only 1470 messages are necessary.

In our tests, the size of the GMI-XML PUBLISH messages was approximately 370 bytes per message on application layer. However XML is only used in our implementation and not conceptually required, the message size can be reduced by a factor for 10 when using a more compact representation like WBXML ([19]).

For the future we plan a more extensive evaluation with a simulated network resource management as the metering client. Then we will finally be able to show the performance of the GMI in its intended use-case.

VI. Related Work

In recent years, there has been a lot of research on handovers in heterogeneous networks.

The approaches in [3], [8], [20] and [9] use policies to make "vertical" handover decisions between different access technologies. Most of these approaches ([3], [20] and [9]) leave the decision which network to choose to the mobile terminal.

This is a reasonable design concept since the information about signal quality of the surrounding base stations is available there. With network-centric decision engines this information must be transported to the core network over the expensive air-interface. Transporting this data also introduces undesired delay that leads to potentially imprecise values.

On the other hand a management facility inside the network is able to take global information, i.e. on the load situation into account and therefore is able to make better decisions. Additionally, it can help the mobile terminal to find adjacent networks without forcing it to scan for available access points which would deplete its battery power.

Toker [8] introduced a network-assisted approach that combines both benefits. They're using two decision engines, one of which is located in the core network while the other resides on the mobile terminal. The decision engine inside the core network transfers a list of available networks and a policy to the mobile terminal. The mobile terminal's decision engine can now take the signal quality into account to evaluate the received information and subsequently decides which network to choose. The GMI could be used to provide the core network- and RAN-related information which is required by the decision engine on the network side.

VII. Conclusion

In this work we have presented the Generic Metering Infrastructure, an information distribution system for mobile networks that has been developed with heterogeneous mobility management in mind, but which is general enough to support other applications. The GMI is based on a modified Publish/Subscribe System which employs subject-based addressing using domain names and subject-based filtering.

Optimised information distribution mechanisms are applied to reduce the number of messages that need to be transported.

When using the GMI for heterogeneous access management, measurement data has to be sent over expensive backhaul links. Therefore future work will focus on the development of an optimised transport solution to further reduce the amount of message-data.

ACKNOWLEDGEMENT

The authors would like to thank Tobias Bandh for fruitful discussions and Frank-Uwe Andersen for the basis of Figure 3.

REFERENCES

[1] E. Dahlman, H. Ekstrom, A. Furuskar, Y. Jading, J. Karlsson, M. Lundevall, and S. Parkvall, "The 3G long-term evolution - radio interface concepts and performance evaluation," *Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd*, vol. 1, pp. 137–141, 7-10 May 2006.

[2] IEEE802.16, "Air Interface for Fixed Broadband Wireless Access Systems," IEEE Standard for Local and Metropolitan Area Networks, Oct. 2004.

[3] V. Gazis, N. Alonistioti, and L. Merakos, "Toward a generic "always best connected" capability in integrated WLAN/UMTS cellular mobile networks (and beyond)," *Wireless Communications, IEEE*, vol. 12, pp. 20–29, 2005.

[4] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda, "Performance anomaly of 802.11b," *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE*, vol. 2, pp. 836–843 vol.2, 30 March-3 April 2003.

[5] 3GPP, "TS 23.401 v8.1.0: General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access," 2008.

[6] ——, "TS 23.402 v8.1.1: Architecture enhancements for non-3GPP accesses," 2008.

[7] IEEE802.21, "Media Independent Handover Services," http://www.ieee802.org/21/, 2007.

[8] C. Fan, M. Schläger, A. Udugama, V. Pangboonyanon, A. C. Toker, and G. Coskun, "Managing Heterogeneous Access Networks. Coordinated policy based decision engines for mobility management," in *LCN '07: Proceedings of the 32nd IEEE Conference on Local Computer Networks*. Washington, DC, USA: IEEE Computer Society, 2007, pp. 651–660.

[9] R. Pries, A. Mäder, and D. Staehle, "A Network Architecture for a Policy-Based Handover Across Heterogeneous Networks," in *OPNETWORK 2006*, Washington D.C., USA, Aug 2006.

[10] 3GPP, "TS 32.401 V7.0.0: Telecommunication management; Performance Management (PM); Concept and requirements (Release 7)," 2006.

[11] J. Case, R. Mundy, D. Partain, and B. Stewart, "RFC 3410 - Introduction and Applicability Statements for Internet Standard Management Framework," 2002.

[12] T. Bandh, "Automated Real Time Performance Management for Mobile Networks," Diploma Thesis, University of Tübingen, 2006.

[13] A. Zeidler, "A Distributed Publish/Subscribe Notification Service for Pervasive Environments," PhD Thesis, University of Darmstadt, 2004.

[14] A. Carzaniga, "Architectures for an Event Notification Service scalable to Wide-area Networks," PhD Thesis, Politecnico Di Milano, 1998.

[15] G. Mühl, "Large-Scale Content-Based Publish/Subscribe Systems," PhD Thesis, University of Darmstadt, 2002.

[16] R. Chalmers and K. Almeroth, "Developing a Multicast Metric," in *Proceedings of IEEE Global Internet (GLOBECOM'00)*, San Francisco, California, USA, Nov. 2000.

[17] 3GPP, "TS 23.003 v7.4.0: Numbering, addressing and identification (Release 7)," 2007.

[18] ——, "TS 32.405 V7.4.0: Performance Management (PM); Performance measurements Universal Terrestrial Radio Access Network (UTRAN) (Release 7)," 2007.

[19] B. Martin and B. Jano, "WAP Binary XML Content Format," http://www.w3.org/TR/wbxml/, 1999.

[20] H. J. Wang, R. H. Katz, and J. Giese, "Policy-Enabled Handoffs Across Heterogeneous Wireless Networks," in *WMCSA '99: Proceedings of the Second IEEE Workshop on Mobile Computer Systems and Applications*. Washington, DC, USA: IEEE Computer Society, 1999, p. 51.