электронный научно-технический журнал
ИНЖЕНЕРНЫЙ ВЕСТНИК
издатель ФГБОУ ВПО «Московский государственный технический университет им. Н.Э. Баумана»

Система мониторинга и анализа GPRS/EDGE/3G -трафика

77-48211/425144

06, июнь 2012

Ермаков А. В., Сурков Л. В.

УДК 004.738.5.031

Россия, МГТУ им. Н.Э. Баумана
ermakov.av@gmail.com.
srk@bmstu.ru

Для компаний операторов сотовой связи существует технологическая проблема, заключающаяся в необходимости выполнения оперативной настройки мобильных терминалов абонентов для предоставления им услуги доступа к сети Интернет с использованием технологий WAP или GPRS, а также услуги отправки мультимедийных сообщений. Проблема заключается в том, что по разным причинам (ошибки в ПО, недостаточный объем памяти терминала, человеческий фактор и др.) мобильные терминалы абонентов оказываются ненастроенными и абоненты не могут воспользоваться предлагаемыми им услугами. Для устранения данной проблемы в компании «МобильныеТеле Системы» была внедрена технологическая платформа коррекции абонентского трафика [1]. Абонентский трафик, проходя через такую платформу, модифицируется, что в последствии позволяет абонентам получить доступ к указанным услугам в «прозрачном» режиме, т.е. без изменения настроек на мобильных терминалах.

Однако данная платформа не предусматривает выполнение оценки эффективности ее работы и контроля предоставления услуг абонентам. Для решения этих задач была разработана система мониторинга и анализа GPRS/EDGE/3G-трафика (далее система СМАТ), решающая следующие поставленные бизнес задачи операторов сотовой связи (далее ОСС).

- Предоставление сотрудникам ОСС, ответственным за обслуживание абонентов, информации, необходимой для разбора претензий абонентов, связанных с работой технологической платформы коррекции абонентского трафика.
- Предоставление сотрудникам ОСС, ответственным за эксплуатацию платформы, информации, необходимой для анализа текущей работоспособности платформы коррекции абонентского трафика.
- Предоставление сотрудникам ОСС, ответственным за расчет экономической эффективности работы платформы, информации, необходимой для подготовки отчетов об экономической эффективности.

- Предоставление сотрудникам ОСС общей статистической информации об объемах скорректированного трафика по услугам WAP, Internet, MMS.

На этапе принятия решения в процессе анализа предметной области и выработки функциональных требований к системе мониторинга и анализа GPRS/EDGE/3G-трафика, участниками рабочей группы, осуществлявшей внедрении платформы коррекции абонентского трафика предпринимались меры по минимизации затрат на разработку системы за счет использования средств мониторинга и анализа уже эксплуатируемых в сети оператора связи. Однако по мере детализации требований и анализа технических реализаций существующих систем мониторинга участниками рабочей группы была выявлена невозможность использования их функций для решения вышепоставленных задач. Анализ систем мониторинга активности абонентов, предлагаемых на рынке в этой области, показал, что такие системы, которые бы удовлетворяли поставленным функциональным требованиям, отсутствуют. Большинство систем (например такие так Soleranetworks DS 1200, или Endace EndaceSensors), адаптация которых могла бы быть осуществлена, предполагают сохранение «сырого» трафика для последующего анализа и поиска необходимой информации, что влечет за собой значительные расходы на содержание систем хранения данных и низкую эффективность их использования.

На этапе анализа информационного базиса системы СМАТ (т.е. того минимально необходимого состава исходных данных, необходимого для решения бизнес задач) были определены категории информации, стандарты используемых технологий и модули информационной системы, обрабатывающие исходные данные. Информационный базис включает в себя три источника информации:

– лог-файлы узлов SGSN (англ. Serving GPRS Support Node, узел обслуживания абонентов GPRS),
– данные предбиллинговой платформы,
– сетевой трафик.

Наиболее сложным процессом обработки является процесс анализа сетевого трафика, заключающийся в съеме сетевого трафика с использованием сигнатурного анализа, его загрузка в базу данных (БД), декодирование, агрегирование и анализ.

Одна из наиболее сложных задач, с которой пришлось столкнуться в процессе разработки – это нарушение порядка поступления и загрузки исходных данных в БД о событиях, фиксируемых в сети. В частности нарушался порядок поступления пакетов в сетевой интерфейс системы СМАТ. Причина этого нарушения обусловлена технологическими особенностями механизмов зеркалирования сетевого трафика, которые используются для получения исходных данных, а также задержек обработки этого трафика в технологических узлах сети GPRS/UMTS–комплекса (BTS, SGSN, GGSN, SDP, FW).

Рисунок 1 иллюстрирует схему интеграции системы СМАТ в сеть оператора сотовой связи. На схеме указаны логические точки съема (получения) исходной информации и не детализировано расположение коммутационного оборудования. На этой схеме, отражающей логическую структуру сети, видно, что каждый запрос абонента последовательно фиксируется (обрабатывается) на узлах GPRS/UMTS–комплекса. Учитывая, что каждый узел при обработке запроса абонента создает некоторую задержку

(не одинаковую для разных узлов), а также учитывая задержку механизмов протоколирования и доставки данных, становятся понятны причины нарушения порядка поступления данных в систему CMAT.

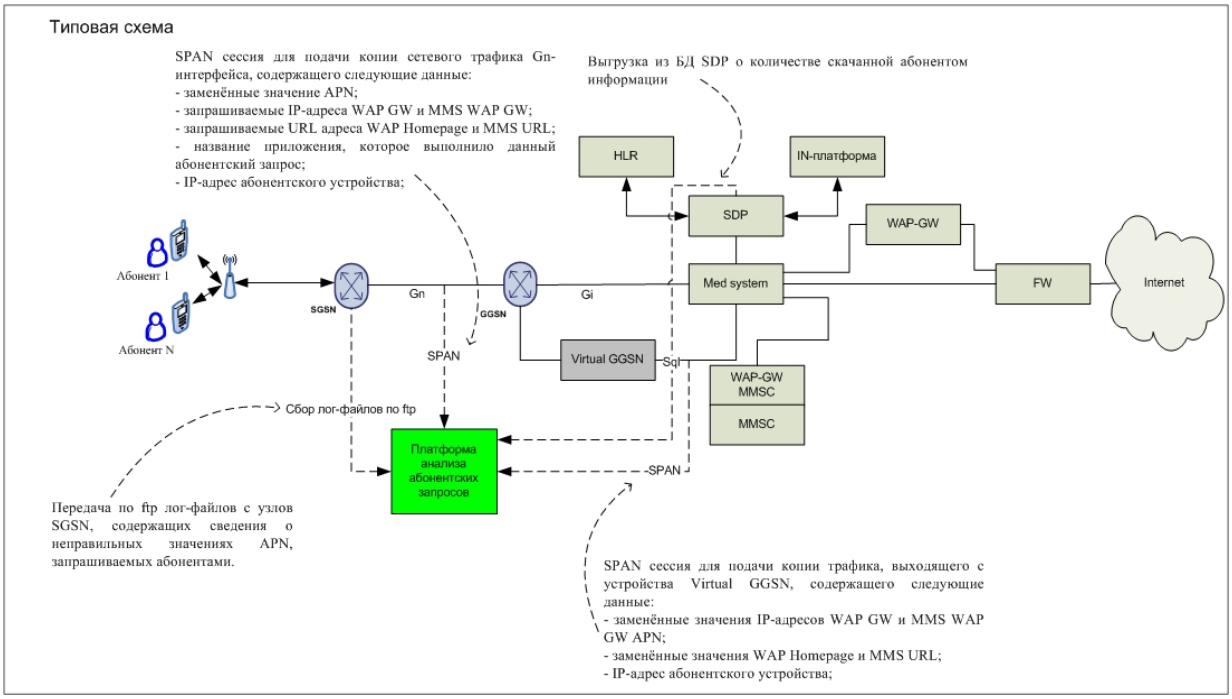


Рисунок 1. Типовая схема интеграции системы CMAT на сети МТС

Для решения данной проблемы и преодоления указанных трудностей предложен порядок обработки сетевого трафика и информации, содержащейся в лог-файлах узлов SGSN, иллюстрируемый рисунками 2 - 4.

Количество-временная диаграмма является отображением последовательности выполнения циклических операций, предусмотренных логикой CMAT, и характеристикой полноты данных на данной итерации (цикле) обработки этих данных. Блоки данных определяют принадлежность данных тому или иному источнику (сетевой трафик, лог-файлы) и логические отношения этих данных внутри блока. В блоках данных указано по два набора записей (соответствующих двум циклам работы CMAT) и записи «выколотые» (или пропущенные) в соответствующем наборе. Из обозначения наборов записей (т.е. количества записей одного блока данных и полученных на одном цикле работы) и выколотых записей видно, что отсутствуют условные ограничения (т.е. нет необходимости выполнения условия связи «один к одному») на количество записей в одном наборе. Также на схеме определена последовательность выполнения операций объединения блоков данных, а повторение отдельных блоков данных указывает на то, что циклические операции сдвинуты во времени (начало следующего цикла обработки может быть раньше завершения предыдущего).

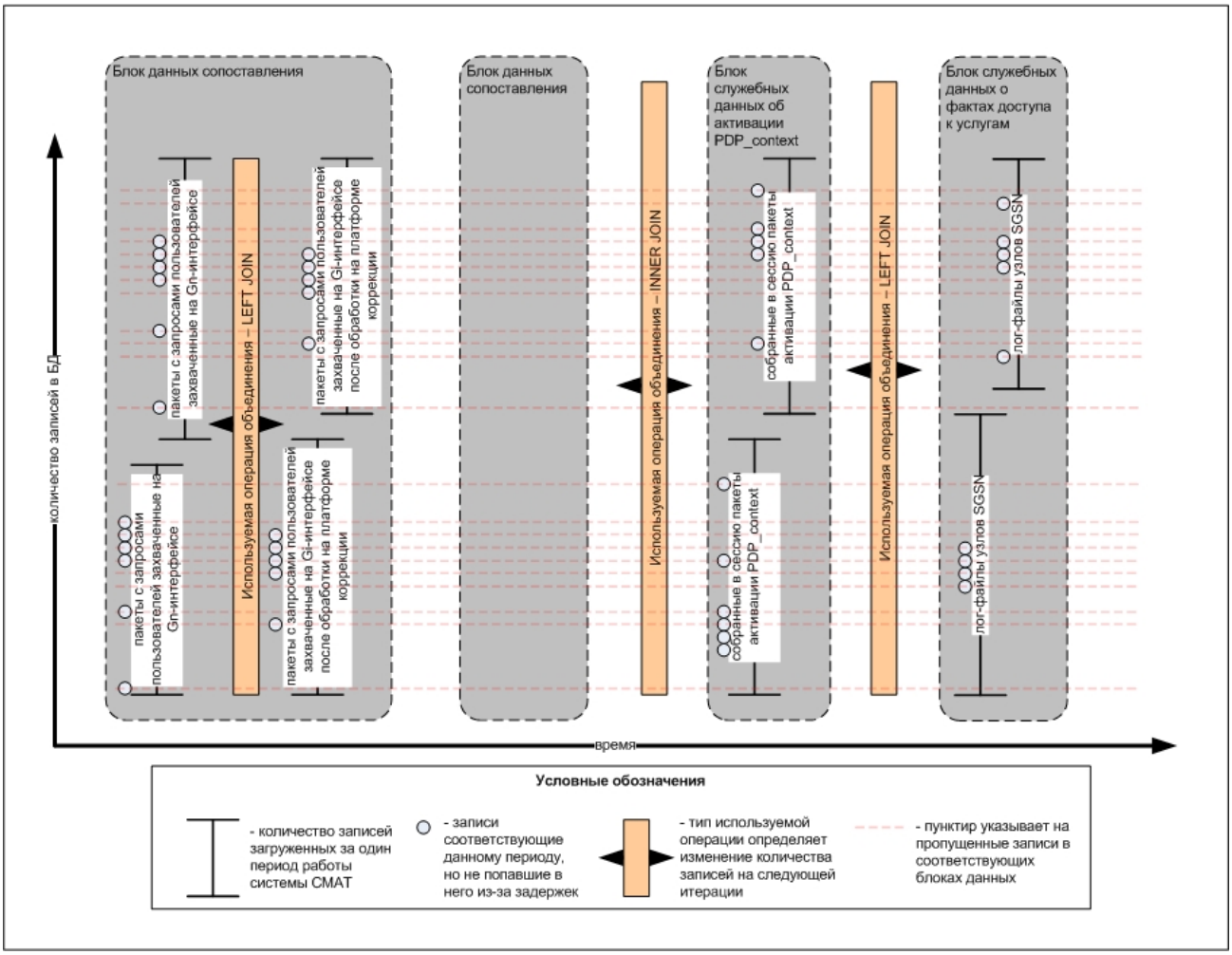


Рисунок 2. Количество-временная диаграмма первого и второго цикла

Идея разработанной логики работы заключается в предварительной обработке «сырых» данных, в результате которой происходит отсеивание избыточной и не используемой информации перед началом логической обработке нормализованных данных и организации механизма посткоррекции Посткоррекция- процедура, которая выполняется с некоторой задержкой от основной процедуры сопоставления данных, при которой часть данных становиться доступной пользователям системы в режиме, приближенном к реальному времени, и затем дополняется недостающей информацией. Основная логическая обработка «сырых» данных (отсеивание) выполняется в соответствующих модулях. Описанный циклический подход к обработке данных выполняется в результирующей БД. Результирующие данные сохраняются в структурированном виде для последующего анализа и их использования при построении разнообразных статистических отчетов.

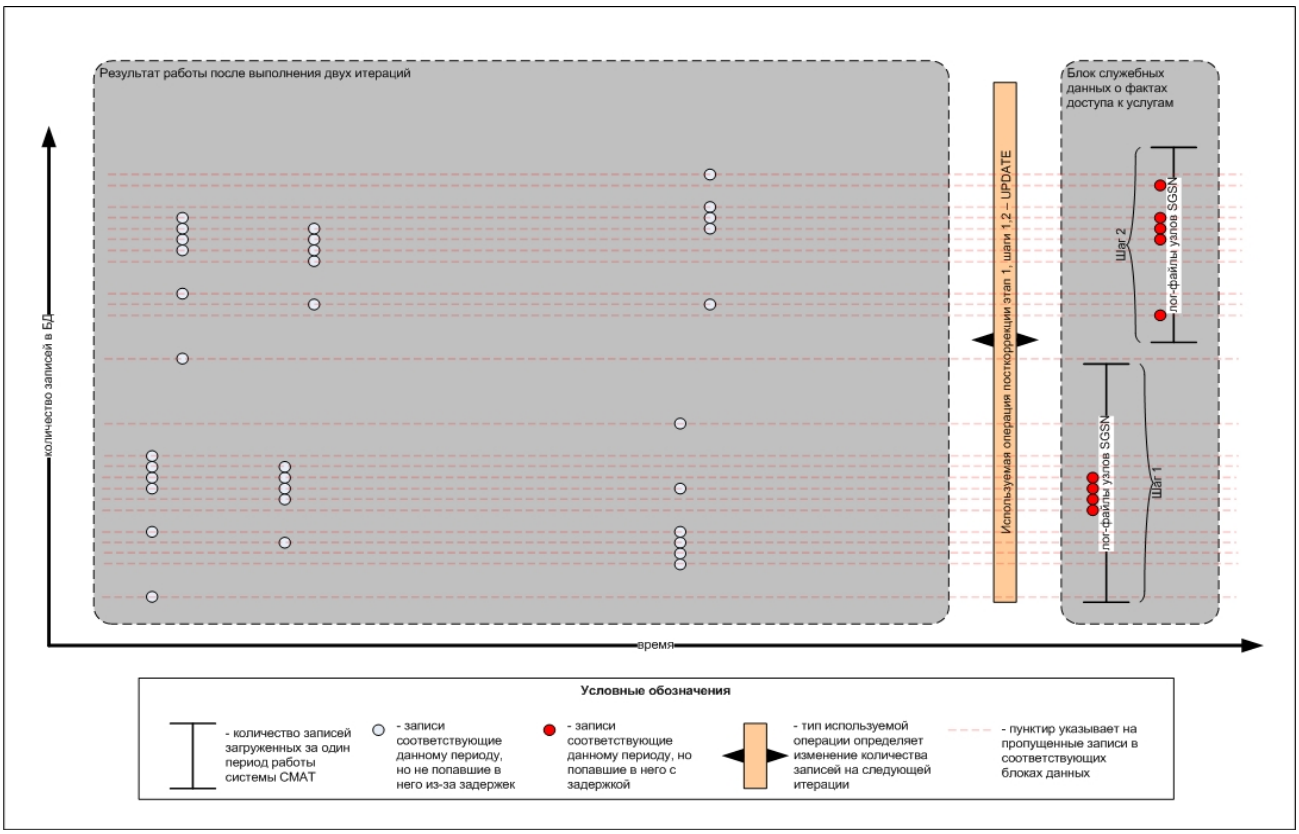


Рисунок 3. Количество-временная диаграмма первого этапа посткоррекции.

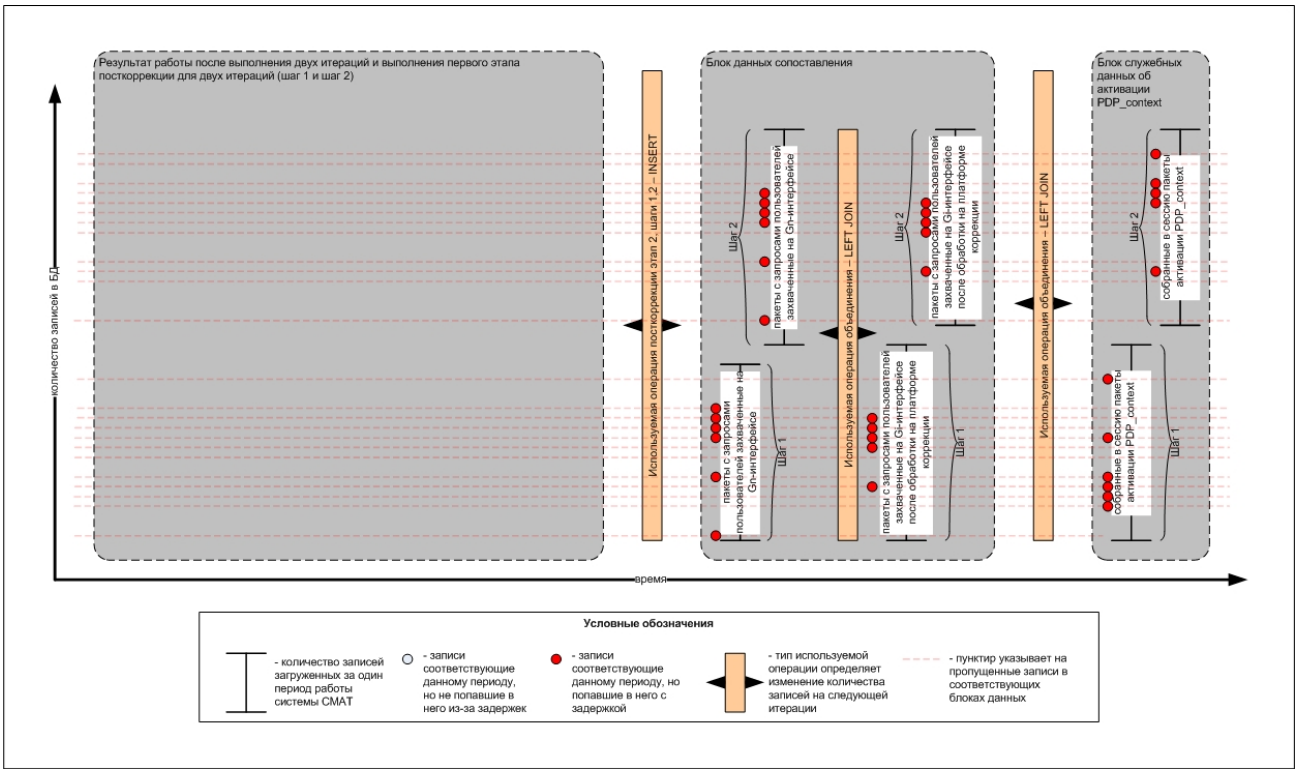


Рисунок 4. Количество-временная диаграмма этапа посткоррекции для первого и второго цикла.

Результатом разработки является решение CMAT, полностью реализующее предъявляемые к системе функциональные требования. Доступ к аккумулируемой информации осуществляется через web-интерфейс, где предусмотрены механизмы построения разнообразных отчетов, экспорта данных, протоколирования действий пользователей, разграничение прав пользователей и другие функциональные возможности. Кроме того выполнена интеграция системы CMAT со службой каталогов Windows, что увеличивает удобство использования системы пользователями и снижает затраты на администрирование, так как в этом случае пользователи используют свои персональные учетные записи.

Система CMAT состоит из нескольких компонентов, каждый из которых выполняет отведенную ему функцию.

- Модуль загрузки лог-файлов узлов SGSN.

  Лог-файлы узлов SGSN выкладываются на специализированный FTP-сервер .Данные файлов с заданной периодичностью анализируются модулем загрузки лог-файлов узлов SGSN. После чтения файлов их содержимое построчно загружается в базу данных для последующего анализа.

- Модуль съема и декодирования трафика.

  Одной из наиболее значимых, для функционирование системы CMAT, является информация, получаемая благодаря пассивному анализу сетевого трафика. Выполнение этого анализа осуществляется последовательно, в несколько этапов: на первом этапе трафик попадает на сетевой интерфейс системы CMAT, далее каждый пакет, попавший на сетевой интерфейс проходит сигнатурную проверку, в случае успешного сравнения с помощью сигнатуры содержимое пакета декодируется и представляется в виде лог-файла заданного формата. На последнем этапе указанный лог-файл читается загрузчиком для записи полученной информации в БД.

- Модуль взаимодействия с предбиллинговыми платформами.

  В зону ответственности данного модуля входит получение и загрузка информации о количестве переданного за время абонентских сессий объемов трафика и загрузка этой информации в результирующую БД.

- Модуль отображения информации.

  Модуль отображения информации реализован в виде WEB-интерфейса, откуда осуществляется доступ пользователей к данным, содержащимся в результирующей БД, а так же управление системой. Заметим, что модуль отображения информации проинтегрирован со службой каталогов AD Windows, что позволяет пользователям корпоративной сети использовать свои учетные записи.

- Модуль резервирования данных.

  Задачей этого модуля является ежедневное резервирование схемы БД и данных, критических для работы системы CMAT. Резервирование осуществляется на удаленный сервер, используемый только для этой цели.

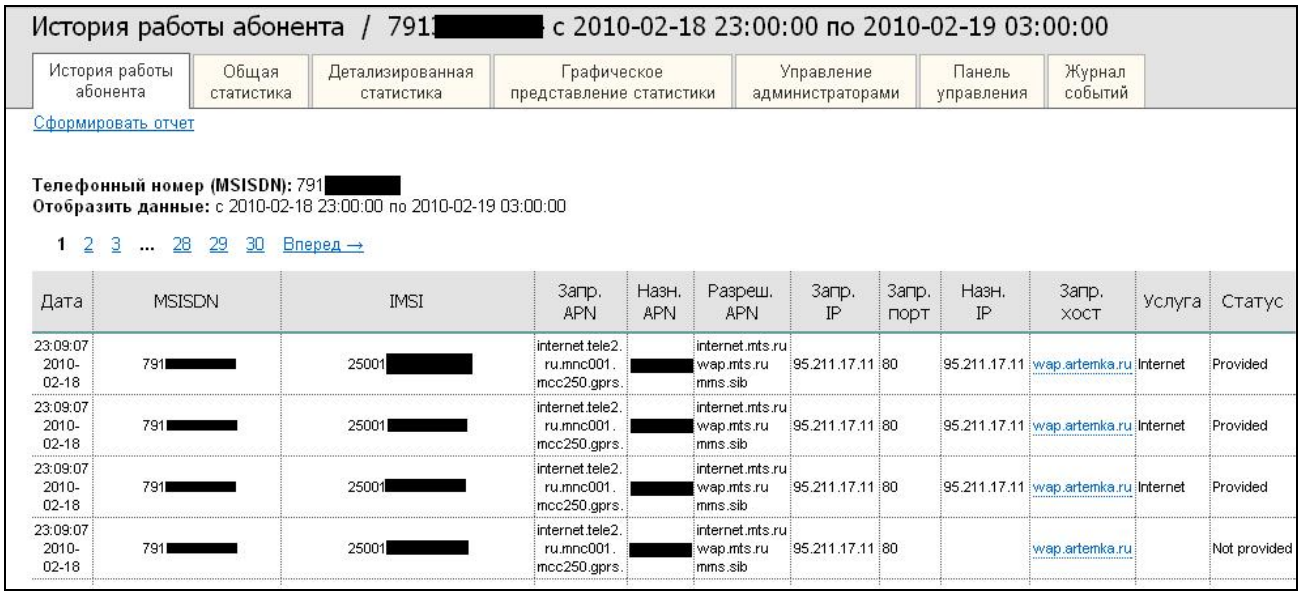
- Модуль мониторинга.

  Модуль мониторинга с заданным расписанием выполняет проверку аппаратных ресурсов и контроль работоспособности модулей системы CMAT: наличие

свободного места на жестких дисках, наличие трафика на сетевом интерфейсе, поступление «свежих» данных в результирующую БД и др. В случае аварии ответственному специалисту высылается уведомление об аварии на его адрес электронной почты.

На рисунках 4,5 представлены примеры web-форм, которые используются для построения разнообразных отчетов:

- История запросов абонента, использовавшего неправильные параметры доступа, представленная в виде отчета сформированного через WEB-интерфейс, рисунок 5.



История работы абонента / 791[redacted] с 2010-02-18 23:00:00 по 2010-02-19 03:00:00

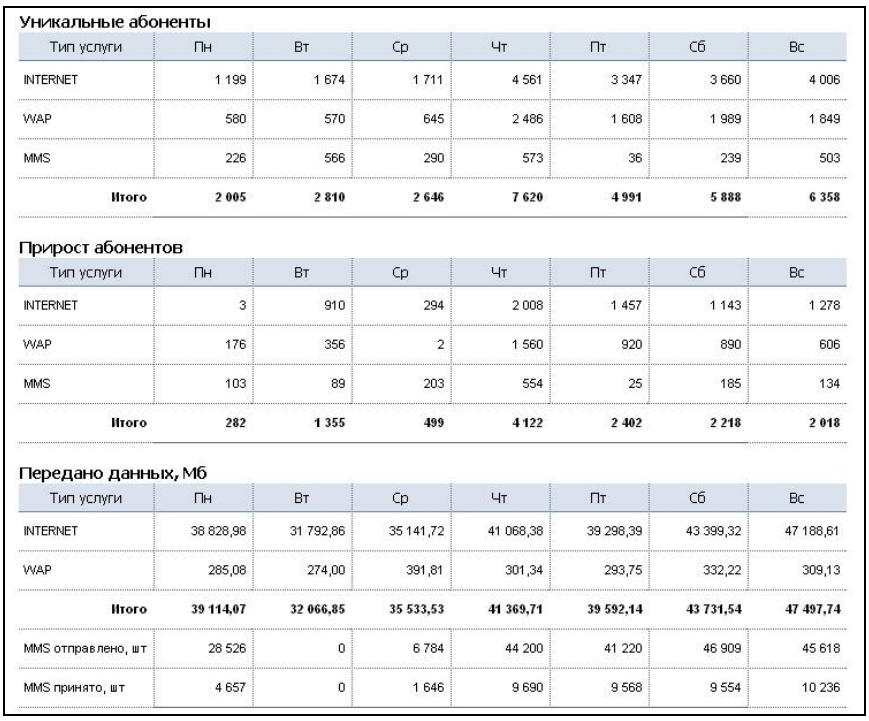
Телефонный номер (MSISDN): 791[redacted]
Отобразить данные: с 2010-02-18 23:00:00 по 2010-02-19 03:00:00

Дата	MSISDN	IMSI	Запр. APN	Назн. APN	Разреш. APN	Запр. IP	Запр. порт	Назн. IP	Запр. хост	Услуга	Статус
23:09:07 2010-02-18	791[redacted]	25001[redacted]	internet.tele2.ru.mnc001.mcc250.gprs.	[redacted]	internet.mts.ru wap.mts.ru mms.sib	95.211.17.11	80	95.211.17.11	wap.artemka.ru	Internet	Provided
23:09:07 2010-02-18	791[redacted]	25001[redacted]	internet.tele2.ru.mnc001.mcc250.gprs.	[redacted]	internet.mts.ru wap.mts.ru mms.sib	95.211.17.11	80	95.211.17.11	wap.artemka.ru	Internet	Provided
23:09:07 2010-02-18	791[redacted]	25001[redacted]	internet.tele2.ru.mnc001.mcc250.gprs.	[redacted]	internet.mts.ru wap.mts.ru mms.sib	95.211.17.11	80	95.211.17.11	wap.artemka.ru	Internet	Provided
23:09:07 2010-02-18	791[redacted]	25001[redacted]	internet.tele2.ru.mnc001.mcc250.gprs.	[redacted]	internet.mts.ru wap.mts.ru mms.sib	95.211.17.11	80		wap.artemka.ru		Not provided

Рисунок 5. Пример отчета для просмотра истории запроса абонента.

- Общие и детализированные статистические данные о количестве абонентов использовавших платформу коррекции, о количестве модифицированного трафика абонентов и о других данных представленных в разрезе времени и типов сервиса, рисунок 6.

Разработка, апробация и внедрение системы CMAT позволила решить ряд принципиально важных задач в компании МТС, направленных на повышение качества обслуживания абонентов, что крайне важно в условиях конкуренции на рынке операторов сотовой связи. Кроме того, система CMAT позволяет получить информацию, необходимую для подготовки отчетности об экономической эффективности внедренного решения коррекции абонентского трафика.



Уникальные абоненты

Тип услуги	Пн	Вт	Ср	Чт	Пт	Сб	Вс
INTERNET	1 199	1 674	1 711	4 561	3 347	3 660	4 006
WAP	580	570	645	2 486	1 608	1 989	1 849
MMS	226	566	290	573	36	239	503
Итого	**2 005**	**2 810**	**2 646**	**7 620**	**4 991**	**5 888**	**6 358**

Прирост абонентов

Тип услуги	Пн	Вт	Ср	Чт	Пт	Сб	Вс
INTERNET	3	910	294	2 008	1 457	1 143	1 278
WAP	176	356	2	1 560	920	890	606
MMS	103	89	203	554	25	185	134
Итого	**282**	**1 355**	**499**	**4 122**	**2 402**	**2 218**	**2 018**

Передано данных, Мб

Тип услуги	Пн	Вт	Ср	Чт	Пт	Сб	Вс
INTERNET	38 828,98	31 792,86	35 141,72	41 068,38	39 298,39	43 399,32	47 188,61
WAP	285,08	274,00	391,81	301,34	293,75	332,22	309,13
Итого	**39 114,07**	**32 066,85**	**35 533,53**	**41 369,71**	**39 592,14**	**43 731,54**	**47 497,74**
MMS отправлено, шт	28 526	0	6 784	44 200	41 220	46 909	45 618
MMS принято, шт	4 657	0	1 646	9 690	9 568	9 554	10 236

Рисунок 6. Пример отчета детализированной статистической информации.

В заключении отметим, что система CMAT является сетенезависимым решением (т.е. может применять в любых сетях операторов связи), обеспечивающим идентификацию абонентов посредством анализа их сетевого трафика передачи данных, а не путем анализа идентификационной информации Кроме того, система обеспечивает хранение не образцов сетевого трафика, а плоских текстовых файлов, что является многократно более эффективным приемом при организации хранилища данных.

Литература:

1. Услуга «Доступ без настроек» позволяет пользоваться сетью Интернет и отправлять или получать MMS-сообщения с неправильными настройками мобильного телефона. http://www.kras.mts.ru/services/internet/without_setting/
2. Exploiting Commodity Multi-core Systems for Network Traffic Analysis, Luca Deri, ntop.org, IIT/CNR, Pisa, Italy. http://luca.ntop.org/MulticorePacketCapture.pdf
3. Эффективное проектирование приложений Oracle. : Пер. с англ. Томас Кайт – М.: «Лори», 2006. – 635с.: ил.
4. Улучшенный пассивный захват пакетов за устройством опроса (device polling) http://www.fssr.ru/hz.php?name=News&file=article&sid=3377
5. GTP protocol 3GPP TS 29.060 V6.9.0 (2005-06)
6. RFC 2616 http://tools.ietf.org/html/rfc2616
7. WAP 2.0 http://www.wapforum.org/what/technical.htm