

Постановка задачи о разработке автоматизированной системы обработки технических данных конечного интернет-пользователя республики Эквадор.

77-48211/640684

10, октябрь 2013

Гонсалес Х. К., Галкин В. А.

УДК 681.5

Россия, МГТУ им. Н.Э. Баумана

jcgonzalezgusev@gmail.com

galkin@bmstu.ru

1. Введение.

Доступ в Интернет постоянно растет во всем мире. Эквадор занимает десятое место в американском континенте по числу пользователей Интернета (около 9 млн. пользователей в 2012 г.) [4].



Рисунок 1. Интернет - пользователи в Северной и Южной Америке [4].

Между тем, в июне 2013 года в Эквадоре были учтены 4'487 384 Интернет - пользователей, что составляет 29,96% от общего населения в 15,6 млн. [3]. Распределение пользователей состоит из фиксированного доступа (21,25%) и доступа через мобильные линии (78,75%).

Главное Управление Контроля Телекоммуникационных Сервисов Республики Эквадор (SUPERTEL) является контролирующим органом для надлежащего предоставления телекоммуникационных услуг в стране.

Главной задачей этой организации является наблюдение, аудит, интервенция и технический контроль предоставления телекоммуникационных услуг в Эквадоре с целью обеспечить ответственность, прозрачность, эффективность и качество услуг, предоставляемых провайдерами конечным Интернет пользователям. В настоящее время одна из основных целей организации заключается в создании «культуры требования прав», которая позволяет пользователям телекоммуникационных услуг понять и требовать соблюдения надлежащего предоставления услуг. Надо сказать, что в Эквадоре пользователь редко знает свои права после найма службы: каковы должны быть технические качества услуг и какие правила или регламенты защищают его интересы?

2. Контролируемые параметры услуги.

В целях обеспечения качества услуг доступа в Интернет, действующее законодательство в Эквадоре позволяет контролировать параметры указанные в таблице 1.

Параметр «Среднее время для решения технического сбоя» включает в себя следующие сбои в услуге, предоставляемой клиенту:

- отсутствие доступа.
- обрыв доступа.
- ухудшение качества обслуживания.
- ограничения на использование службы без согласия клиента.

Время решения проблемы рассчитывается от момента, когда пользователь информирует оператора об этом случае, до конечного решения проблемы.

Таблица 1.

Технические параметры качества, доступа в Интернет, контролируемые в Эквадоре
SUPERTEL.

Параметр	Тип измерения	Период измерения	Целевое значение
Взаимоотношение с клиентами	Опрос «Удовлетворенности клиентов», выполняемый оператором, на 5 баллов	Полугодовой	≥ 3
Процент жалоб	Процент жалоб, измеренный по количеству клиентов каждого оператора	Ежемесячный	$\leq 2\%$
Максимальное время на решение жалоб	Среднее время в часах, в которых жалобы разрешаются, за исключением жалоб «биллинга»	Ежемесячный	Максимальный срок: 7 дней для 98% жалоб
Процент жалоб «биллинга»	Процент жалоб клиента за счета-фактуры, измеренный по количеству генерируемых фактур	Ежемесячный	$\leq 2\%$
Среднее время для решения технического сбоя	Среднее время в часах, для решения технических сбоев, которые сообщают клиенты	Ежемесячный	$\leq 24h$.
Процент жалоб по скорости канала доступа	Процент жалоб на скорость канала доступа, когда клиент считает что она меньше нанятой	Ежемесячный	$\leq 2\%$

3. Текущая проблема.

Правила и нормы в Эквадоре в отношении качества предоставления Интернет - доступа, нацелены на предоставление надлежащей услуги конечному потребителю, при определении параметров и целевых значений, которые должны быть выполнены поставщиками услуг - ISP (Internet Service Provider - Интернет-провайдер). Жалобы от клиентов представляют собой единственный источник обнаружения проблем по обслуживанию. Поэтому сегодня важно создать настоящую «культуру требования прав" в стране.

Однако, при такой форме контроля, Интернет-провайдер (ISP) несет полную ответственность за обработку жалоб и доклад SUPERTEL, практически соответствующий критерию "добросовестности", не отражает фактические технические результаты обслуживания абонентов, которые важны с точки зрения статистики и выявления существующих недостатков в сетях ISP.

Таким образом, данные, которые служат для анализа предоставления услуги, в настоящее время сосредоточены в жалобах, введенных клиентами, а не в реальных технических данных поведения службы.

4. Цель.

Глобальная цель проекта состоит в создании автоматизированной системы обработки технических данных конечного Интернет - пользователя, обеспечивающей контроль качества услуги, оговоренной соглашением с пользователем, и статистического анализа специфических показателей..

Ниже представлены показатели или флаги, которые по мнению авторов могут быть рассмотрены для специального технического контроля и которые позволили бы иметь реальную техническую и статистическую картину состояния обслуживания, с точки зрения пользователя:

- отсутствие / обрыв доступа: отмечает время, когда доступ к услуге не возможен. Должен быть обнаружен с помощью сообщений конкретных веб - протоколов доступа в Интернет. Необходимо получить время начала и конца недоступности, и представленная сетевая ошибка.
- скорость передачи данных (бит/сек): может рассматриваться в зависимости от используемого протокола (FTP, HTTP и т.д.) и направления передачи (вверх/вниз).
- задержка: измерение задержки к определенным серверам (мсек), с помощью тестов на ICMP протоколе.
- общие ошибки сети, влияющие на качество услуг, представляемых пользователю.

С помощью этих показателей, мы получаем техническую информацию предоставления услуг доступа в Интернет и статистику в реальном времени, с точки зрения пользователя. Вариант контроля представлен на рисунке 2.

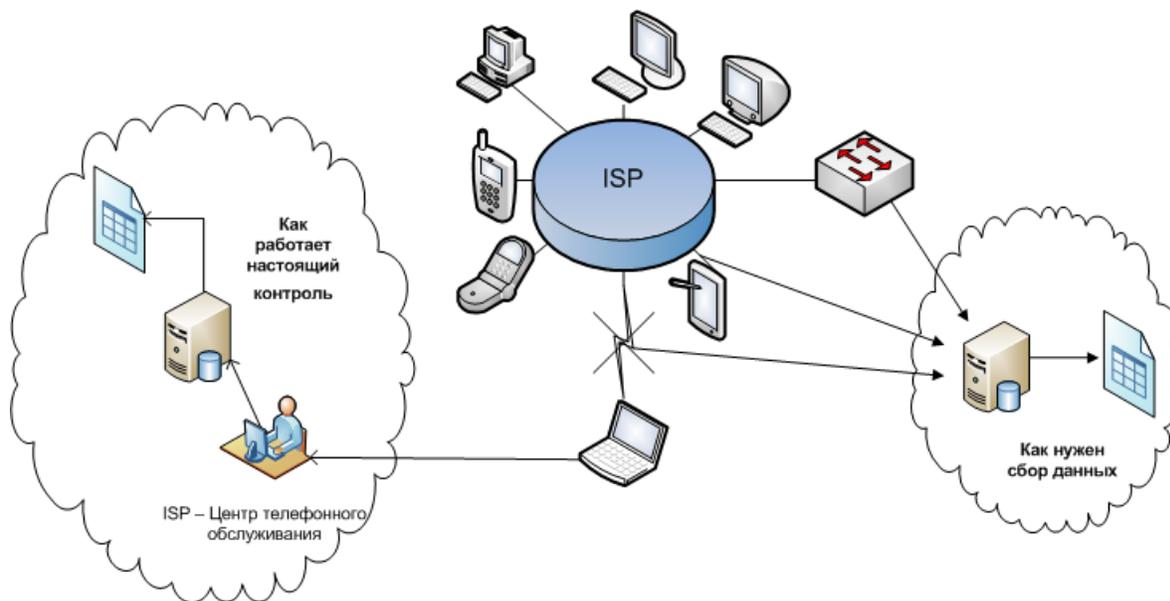


Рисунок 2. Существующий и предлагаемый варианты контроля сервиса.

5. Предлагаемое решение (структура).

После того как определилась текущая проблема и необходимости проекта, устанавливается предлагаемое решение. Система должна содержать следующие этапы:

- сбор данных (конкретного приложения на компьютере).
 - передача и прием данных.
 - обработка данных.
 - статистические отчеты и мониторинг.
- i. Этап сбора данных терминала пользователя поясняет рисунок 3 . Сбор данных должен быть разработан с учетом следующих характеристиками:
- модуль сбора данных должен быть установлен после запроса пользователя.
 - собирать, по крайней мере технические данные, определенные в предыдущем пункте настоящей статьи.
 - нужно произвести анализ последствия установки модуля сбора данных с точки зрения ресурсопотребления. Модуль не должен оказывать заметного влияния на работу приложений пользователя.
 - генерирующиеся файлы должны быть временно записаны на устройстве клиента, и удалены после их передачи.

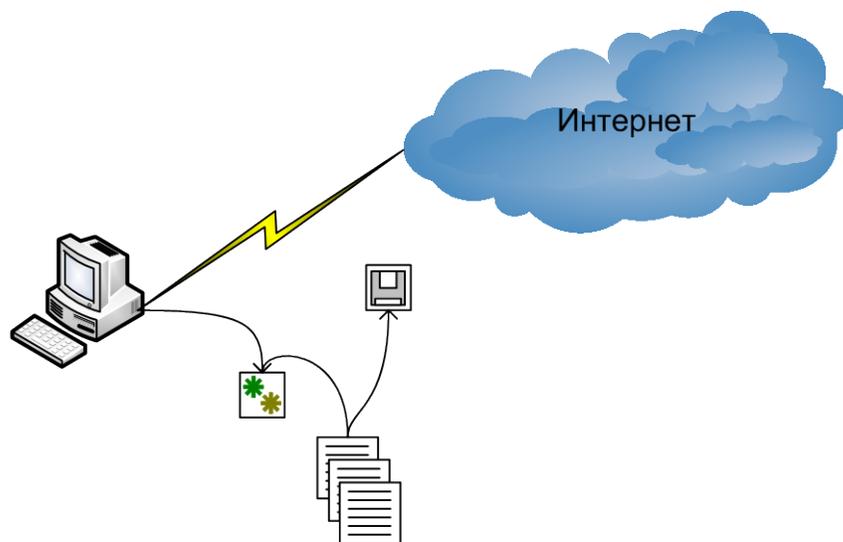


Рисунок 3. Модуль сбора выполняет выбор необходимых технических данных, и сохраняет информацию до ее передачи.

- ii. Передача результатов работы модуля сбора поясняется рисунком 4 и должна быть выполнена со следующими характеристиками:
- передавать данные через уже открытый Интернет доступ. Нужно выполнить технический анализ способности сети для каждого сервера с учетом приема данных от каждого клиента.
 - для передачи данных следует рассмотреть параметры безопасности на сервере и клиентах (открыть тоннель передачи после прохождения соответствующего процесса аутентификации).
 - должна присутствовать проверка целостности данных.
 - в первоначальной этапе проекта можно разработать одностороннюю связь (от клиента к серверу), но не следует исключать в будущем расширение двусторонней связи.



Рисунок 4. Данные должны быть переданы на центральный сервер в защищенной среде.

iii. Центральный сервер в соответствии с рисунком 5 должен содержать программу обработки данных со следующими характеристиками:

- поддерживать объем входящей информации от всех конечных пользователей.
- обеспечивать автоматическую обработку информации (прием, обработка, хранение).
- накапливать и обрабатывать информацию в специализированную базу данных, структура которой должна быть оптимизирована с точки зрения информационной избыточности.

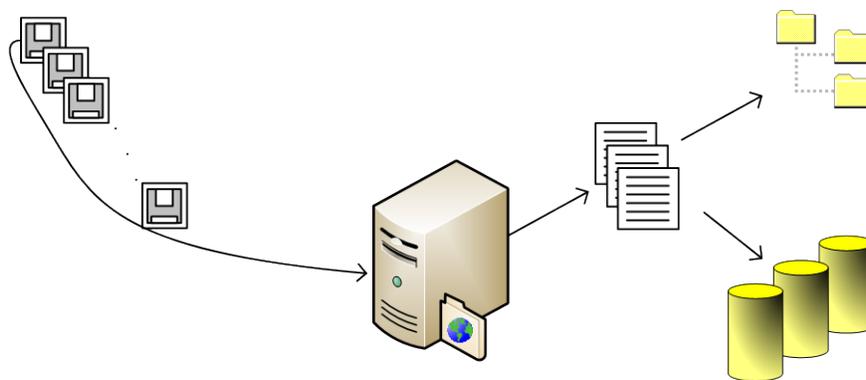


Рисунок 5. Обработка собранной информации производится на центральном сервере. Информация сохраняется в определенных папках и вводится в определенную базу данных

iv. Сервер должен содержать в соответствии с рисунком 6 модуль статистики и отчетности, учитывая:

- мониторинг пользователей, которые введены в систему, состояние последней передачи.
- статистические отчеты системы.
- статистические отчеты собранной технической информации, включая технические характеристики оборудования (тип и модель устройства, провайдер, технология и т.п).

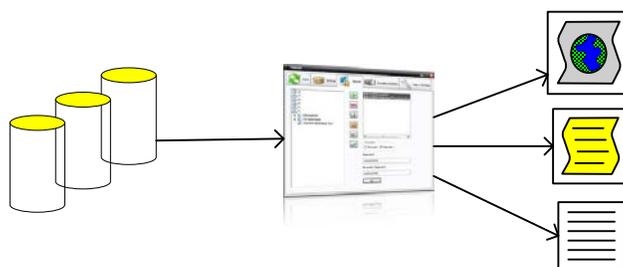


Рисунок 6. Автоматизированная система имеет приложение для создания отчетов и анализа статистических данных.

Таки образом, мы получаем систему, которая показывает технические данные предоставления услуги Интернет - доступа в режиме реального времени, для статистических и контрольных целей, будучи также:

- Интерактивной и интегрируемой модульной системой.
- Собственной разработкой.
- Масштабируемой.
- С распределенным сбором данных в организованной среде.
- Формирующей статистику и отчеты по мониторингу.
- В безопасной среде и с проверкой целостности.

6. Оценка функционирования системы.

Чтобы получить оценку эффективности функционирования системы, построим модель, используя для этого язык программирования GPSS (*General Purpose Simulation System*) и метод аналитико-имитационного моделирования [1, 2].

6.1. Функциональная модель.

Для описания функциональной структуры системы (создание функциональной модели), используется методология IDEF0 [5], в виде набора взаимодействующих и взаимосвязанных блоков.

Для этого, определяем глобальный процесс, связанный с ролью нашей системы: «Контролировать качественную Интернет - услугу», со всеми данными при входе, выходе, управлениями, и.т.д., как показано на рисунке 7.

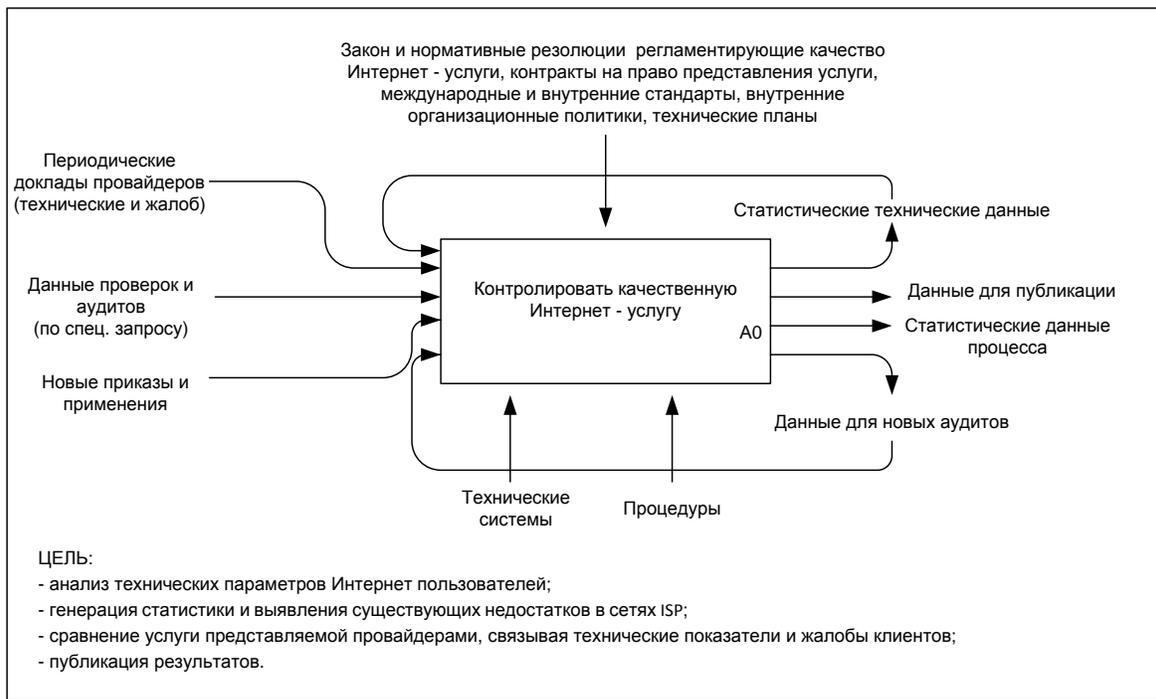


Рисунок 7. Функциональная модель общего процесса «Контролировать качественную Интернет - услугу».

Как показывает рисунок 7, общий процесс осуществляет функцию обеспечения надлежащего предоставления Интернет – услуги, используя для этого разные технические системы и процедуры, на основе внутренних и международных правил, а также согласно плану технических проверок и аудитов. Процесс выделяет данные (технические и статистические), которые используются с целью анализа, публикации информации интересной для пользователя (например, средняя скорость или количество обрывов услуги за месяц), осуществляет поиск недостатков в сетях ISP (замечания, штрафы поставщикам услуг) и др..

Таким образом, автоматизированная система обработки технических данных конечного Интернет – пользователя является частью технических процедур – механизмом общего процесса. Структура функциональной модели данной процедуры показана на рисунке 8.

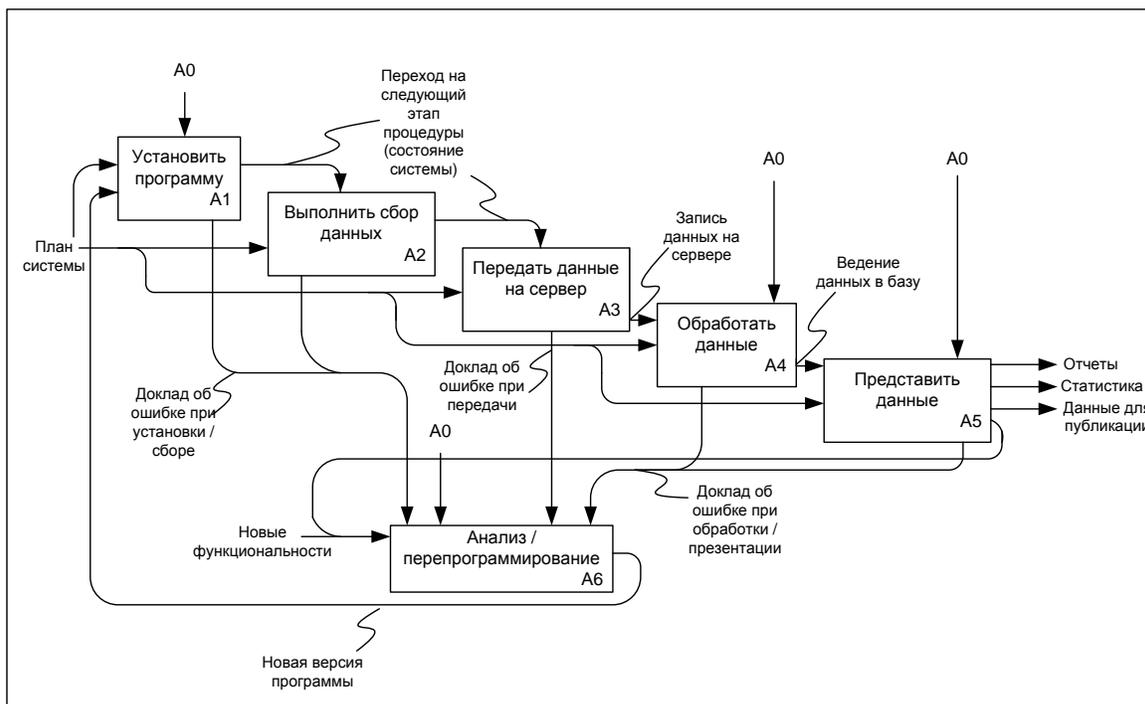


Рисунок 8. Функциональная модель специфической процедуры автоматизированной системы обработки технических данных конечного Интернет – пользователя.

Переход системы из одного состояния в другое происходит лишь после успешного завершения предыдущего, для функций A1 – A3. При ошибке в любой из них, вызывается функция для анализа и перепрограммирования приложения сбора данных (A6), с последующей переустановкой. A6 так же вызывается общим процессом A0 для ввода изменений, а также в результате новых функциональностей выдающихся по разным причинам (новые технические показатели, применения, регулирование, и т.д.). Функции обработки данных (A4) и представления данных (A5) вызываются параллельно общим процессом A0, и используют результаты предыдущих функций. На выходе система выдает все технические данные (презентации, графы, статистика, показатели), ради которых создана система контроля, и вызывает новые процессы организации (информирование клиентов, доклады, сообщения для поставщиков).

6.2. Параметризация системы.

Для параметризации, система рассматривается как СМО (система массового обслуживания), со следующими характеристиками¹:

- один **канал** связи (**Chn**), Интернет - канал на сервере, для получения данных собранных у клиента. Тем не менее потому, как будет использован протокол для передачи данных (FTP, например [6]), возможна параллельная обработка заявок, поэтому на входе система рассчитывается как многоканальная СМО (**n**)².
- **заявки (Z)** представляют собой каждый запрос (IP соединение) со стороны клиента на передачу собранной информации (файл), и последовательная передача данных. Тут нужно отметить важность выбора адекватной пропускной способности на сервере. Действует **накопитель – буфер на этапе передачи (B_{tx})**, и образуется **очередь на передачу (Q_{tx})**. **Длительность обслуживания заявок при передаче V(τ)**, или **время передачи** каждого файла = размер файла (**F_{sz}** - кбайт) / пропускная способность канала (**R_{ch}** - битов/сек) / количество файлов одновременно передающихся (n, количество допущенных соединений – сокетов).
- поток заявок образуют все пользователи или **клиенты (CIt)** вошедшие в проект (имеют установленную программу). Интервал времени генерации одного файла на каждом клиенте – 30 минут. Таким образом, заявки (**Z**) поступают, каждую секунду в количестве C/1800. Если рассчитываем секунду как единица времени, то **интенсивность λ = C/1800**. Интенсивность будет зависеть от количества клиентов, но в нормальном состоянии считается неизменной – поток является стационарным. Так как используется передача данных на сервер с **n** допущенных соединений одновременно, то считается, что в каждый момент времени может появиться только одна заявка, которая обслуживается очередным сокетом, деля для этого пропускную способность. Поток является простейшим. **Интервал времени между соседними заявками A(τ) = 1 – e^{-λτ} = 1 – e^{-C/1800τ}** (в секундах). В этапе передачи система не содержит накопителя, является СМО с отказами (отказ происходит при достижении [n+1]-й заявки на передачу данных).
- после этого, действует накопитель - буфер на этапе обработки (**Bpr**) для файлов достигших сервера (прошли первый этап передачи), но еще не

¹ Описывается со стороны сервера, при получении файлов накопленных на каждом клиенте.

² Количество одновременных соединений ограничивается в целях безопасности.

обработанных. Так же, генерируется **очередь на обработку (Q_{pr})**. Длину очереди представляет количество не обработанных файлов. **Обработать** файл, значит переписать его данные в общую базу на сервере. Емкость накопителя в данной модели предполагается неограниченной (переполнение накопителя очень маловероятно, рассчитывая на средний размер файлов). Обзор методов оптимизации вычислительных ресурсов для достижения наилучшей производительности системы в разных СУБД и операционных системах, представленный в [7], [8], [9]³, позволяет в качестве критерия выбрать длительность **обслуживания заявок на этапе обработки (P_{rc})** в секундах.

- заявки обрабатываются без приоритета, в порядке поступления (FIFO – *Fist In First Out*). Это касается дисциплин буферизации и обслуживания.

Диаграмма системной модели представлена на рисунке 9.

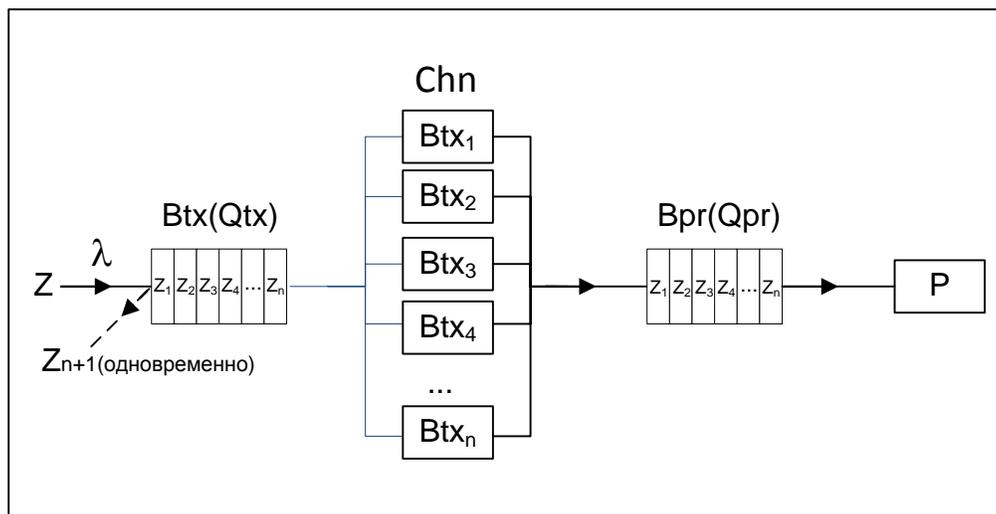


Рисунок 9. Диаграмма модели системы.

Так же, формируются следующие показатели эффективности системы:

- число заявок в системе, в очереди и на обслуживании (N_z) = количество заявок в системе.
- вероятность потери заявок (P_n) = количество заявок не допустивших в систему (отказ - O_t) / общее количество заявок [%].
- интенсивность потока обслуженных заявок (λ_z) = количество обслуженных заявок / время моделирования [заявок/сек].

³ Методы будут подробно анализированы на этапе проектирования.

- среднее время передачи заявки (T_{tx}), среднее время обработки файла (T_{pr}), включая ожидание в очередях [сек], и среднее время пребывания заявок в системе (T_c).

6.3. Имитационная модель.

Текст программы на языке GPSS (General Purpose Simulaton System [1], [2]) представлен на рисунке 10.

```

Clт      EQU      5000      ;пользователи вошедшие в проект
Z        EQU      1800/Clт  ;время генерации заявки каждого
                               ;пользователя (в секундах)
Btx_n    STORAGE  5        ;количество допущенных одновременных
                               ;соединений
Qtx_n    EQU      5        ;очередь на этапе передачи = Btx_n
Rch      EQU      512000   ;пропускная способность канала (б/сек)
Fsz      EQU      Uniform(2,10,100) ;размер генерирующегося файла
Prc      EQU      Uniform(3,0.5,2) ;длительность обслуживания заявок на
                               ;этапе обработки (в секундах)
*****
      GENERATE      (Exponential(4,0,Z)) ;генерация заявок
      TEST L        Q$Qtx,Qtx_n,Full  ;контроль достигших одновременных
                               ;отказ обслуживания
      QUEUE        Total_time        ;контроль пребывания заявок в системе
      QUEUE        Qtx                ;счет очереди канала передачи
      ENTER        Btx_n              ;вход в очередной канал передачи
      ADVANCE      (TxTime())         ;вызов процедуры вычисления времени
                               ;передачи
      LEAVE        Btx_n              ;выход из канала передачи
      DEPART       Qtx                ;выход из очереди канала передачи
      QUEUE        Qpr                ;счет очереди на этапе обработки файла
      SEIZE        Qpr                ;вход в блок обработки файла
      ADVANCE      Prc                ;вычисление времени обработки файла
      RELEASE      Qpr                ;выход из блока обработки файла
      DEPART       Qpr                ;выход из очереди на этапе обработки
                               ;файла
      TERMINATE    1                  ;завершение модели в желательной
                               ;последовательности
      DEPART       Total_time        ;контроль пребывания заявок в системе
Full     TERMINATE  1                  ;завершение модели в случае отказа
                               ;соединения
*****
PROCEDURE TxTime() BEGIN
  TEMPORARY Arg1, Time;
  IF (Q$Qtx = 0) THEN BEGIN
    Time = (Fsz # 8 # 1024) / Rch;
  END;
  ELSE BEGIN
    Time = (Fsz # 8 # 1024 # Q$Qtx) / Rch;
  END;
  RETURN (Time);
END;

```

Рисунок 10. Модель системы на языке GPSS.

Модель использует для генерации заявок экспоненциальный закон распределения, зависимый от числа пользователей (учитывая, что каждый файл генерируется через каждые 30 минут). Образуется очередь в первоначальной этапе передачи данных, зависящая от количества допущенных одновременных соединений (отказ обслуживания при достижении $[n+1]$ -й заявки). Время передачи рассчитывается PLUS процедурой и непосредственно связано с пропускной способностью канала передачи, размера генерирующегося файла (в виде равномерно распределенной случайной величины), и очередью на этапе передачи. Время обслуживания на этапе обработки рассчитывается как равномерно распределенная функция. Рассматривается так же очередь на этапе обработки.

Имитационное моделирование системы проведено в течение одной недели непрерывной работы (604800 секунд). Получены результаты для следующих значений варьируемых параметров модели:

- а.** R_{ch} – 512 Кбит/сек. Fsz – 10 и 100 Кбайт. Prc – 0,5 и 2 сек.
- б.** R_{ch} – 1 Мбит/сек. Fsz – 10 и 100 Кбайт. Prc – 0,5 и 2 сек.
- в.** R_{ch} – 1 Мбит/сек. Fsz – 4 и 20 Кбайт. Prc – 0,3 и 1,5 сек.

Система с пропускной способностью 512 Кбайт/сек, стабильная на этапе передачи данных до достижения примерно 2000 пользователей, после чего вероятность потери заявок ($Pп$) резко возрастает. При пропускной способности 1 Мбит/сек, этапа передачи остается стабильной и до 3000 пользователей ($Pп < 5\%$), и достигает лучших результатов ($Pп < 3\%$) если выбрать наибольшее количество одновременных соединений ($B_{tx} = 15$ или 20). Эти зависимости представлены на рисунке 11.

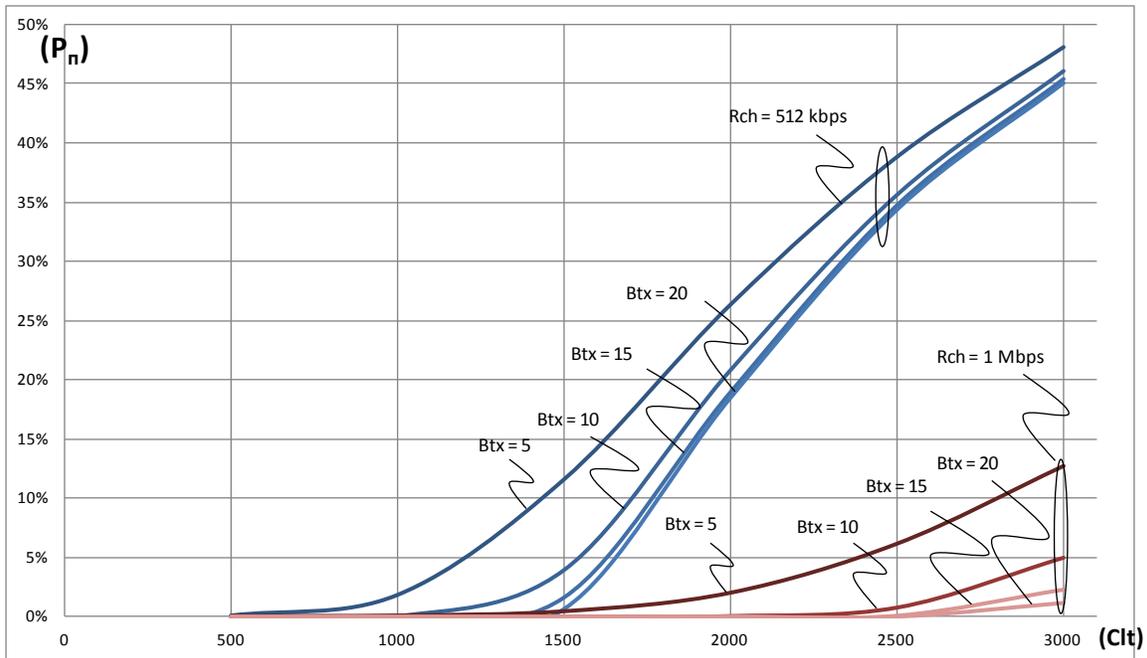
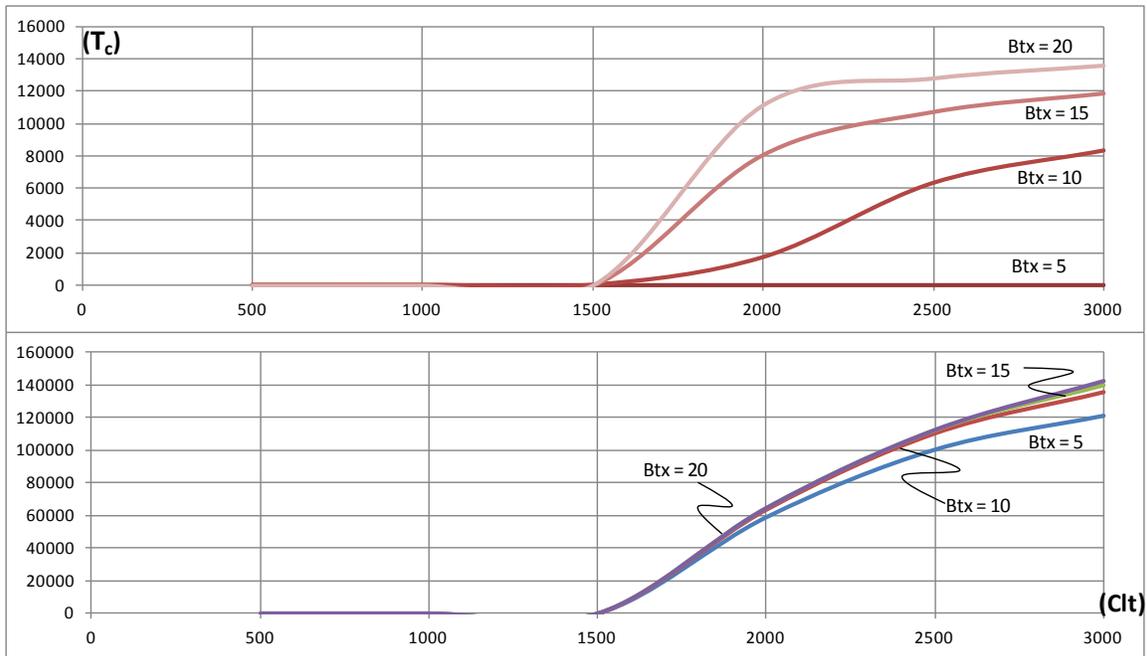
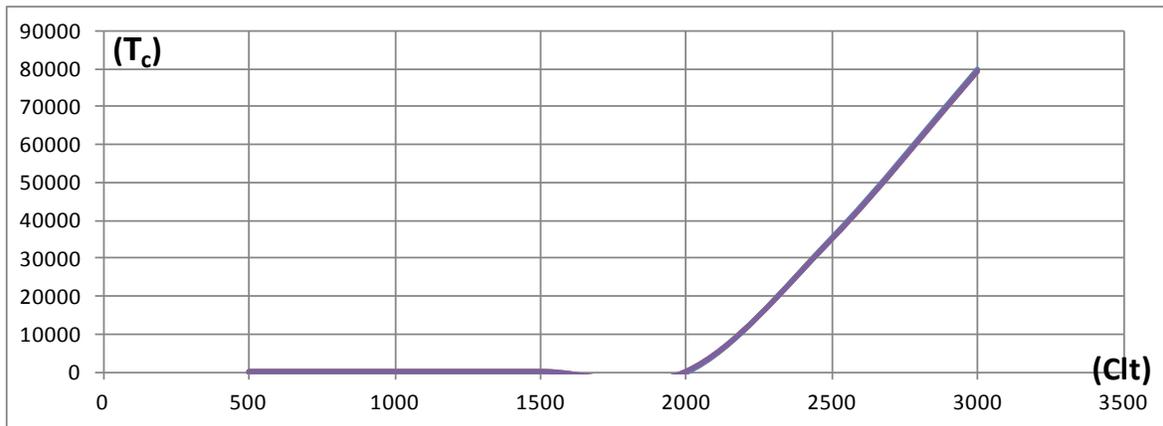


Рисунок 11. Влияние пропускной способности и количества одновременных соединений на вероятность потери заявок при увеличении пользователей. Этап передачи данных.

На этапе обработки данных, система является стабильной до 1500 пользователей, учитывая равномерное распределение размера файлов (от 10 до 100 Кбайт) и время их обработки (от 0,5 и 2 секунд). При изменении размера файлов и времени обработки (от 4 до 20 Кбайт и 0,3 и 1,5 секунд соответственно), этап обработки стабилен до 2000 пользователей. Далее, как это следует из рисунка 12, среднее время пребывания заявок в системе возрастает до неприемлемых значений.



а. R_{ch} – 512 Кбит/сек (верхний) и 1 Мбит/сек (нижний). Fsz – между 10 и 100 Кбайт. Prc – между 0,5 и 2 сек.



б. R_{ch} – Мбит/сек. Fsz – между 4 и 20 Кбайт. Prc – между 0,3 и 1,5 сек.

Рисунок 12. Влияние размера и время обработки файлов на среднее время пребывания заявок в системе, при увеличении пользователей. Этапа обработки данных.

7. Заключение.

Отсутствие данных, с помощью которых могли бы быть оценены технические и статистические результаты реального предоставления Интернет - доступа конечному пользователю в Республике Эквадор, делает актуальной постановку задачи о разработке автоматизированной системы для этой цели. В данной статье, в рамках высказанной инновационной идеи реализации автоматизированной системы обработки технических данных конечного пользователя Интернет услуги, определены основная структура и характеристики этой системы, построена функциональная модель и имитационная модель системы, анализируются параметры ее эффективности.

Система может считаться стабильной, и обслуживать, как минимум, 2000 пользователей, с пропускной способностью канала доступа в 1 Мбит/сек и 15 одновременных соединений для передачи данных. При этом необходимо стремиться к генерированию файлов наименьшего размера и к достижению наилучшей производительности сервера на этапе обработки. Настройка и регулировка сервера должны учитываться на этапе проектирования и разработки системы, также как, реальные показатели должны сравниваться с теоретическими результатами, полученными в настоящей статье.

Следующие стадии разработки проекта включают проектирование системы, разработку программного обеспечения для сбора данных, проектирование базы данных и тестирование макета системы.

8. Список литературы.

1. Алиев Т.И. Основы моделирования дискретных систем. Санкт-Петербург: ИТМО, 2009. 363 с.
2. Черненький. В.М. Адаптированное описание системы имитационного моделирования GPSS // Электронная Образовательная Система МГТУ им. Н.Э.Баумана. 43 с. Режим доступа: http://e-learning.bmstu.ru/portal_iu5/pluginfile.php/644/mod_page/content/7/GPSS.pdf (дата обращения 08.10.2013).
3. Источник: Национальный Телекоммуникационный Секретариат (SENATEL) Республики Эквадор. Отчет плотности Интернет – доступа. 2013. . Режим

- доступа: <http://www.regulaciontelecomunicaciones.gob.ec/biblioteca/> (дата обращения 08.10.2013).
4. Организация Internet World Stats. Статистика для всех стран Северной и Южной Америки. 2012. Режим доступа: <http://www.internetworldstats.com/stats2.htm> (дата обращения 08.08.2013).
 5. Госстандарт России. Методология функционального моделирования IDEF0. Москва: ИПК. 2000. 75 с.
 6. Постел Дж., Рейнолдс Дж. Протокол передачи файлов // Рабочее предложение RFC 765. 1985. 69 с. Режим доступа: <http://www.ietf.org/rfc/rfc959.txt> (дата обращения 08.10.2013). [J. Postel, J. Reynolds. File transfer protocol (FTP)].
 7. Сирмон Дж. Стин Х. Параметризация компьютерного обеспечения табличного решения (Эс-Кью-Эл услуги сервера – анализа). // Техническая статья Майкрософт Эс-Кью-Эл Сервер. 2013. 29 с. [John Sirmon, Heidi Steen. Hardware Sizing a Tabular Solution (SQL Server Analysis Services)].
 8. Корпорация Майкрософт. Добавление в базу данных одной или нескольких записей. Режим доступа: <http://office.microsoft.com/ru-ru/access-help/HA010043515.aspx> (дата обращения 08.10.2013).
 9. Корпорация Майкрософт. Планирование и настройка рабочих характеристик хранилища и SQL Server (SharePoint Server 2010). 2011. Режим доступа: <http://technet.microsoft.com/ru-ru/library/cc298801.aspx> (дата обращения 08.10.2013).