

Distr.  
GENERAL

CES/AC.71/2004/18  
1 March 2004

RUSSIAN ONLY

**СТАТИСТИЧЕСКАЯ КОМИССИЯ и  
ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ КОМИССИЯ  
ОРГАНИЗАЦИИ ОБЪЕДИНЕННЫХ НАЦИЙ**

**ЕВРОПЕЙСКАЯ КОМИССИЯ  
СТАТИСТИЧЕСКОЕ УПРАВЛЕНИЕ  
ЕВРОПЕЙСКИХ СООБЩЕСТВ  
(ЕВРОСТАТ)**

**КОНФЕРЕНЦИЯ ЕВРОПЕЙСКИХ СТАТИСТИКОВ**

**ОРГАНИЗАЦИЯ ЭКОНОМИЧЕСКОГО  
СОТРУДНИЧЕСТВА И РАЗВИТИЯ (ОЭСР)  
СТАТИСТИЧЕСКИЙ ДИРЕКТОРАТ**

**Совместное совещание ЕЭК/Евростата/ОЭСР по вопросам управления  
статистическими информационными системами (УСИС)  
(Женева, 17-19 мая 2004 года)**

Тема III: Использование ПО с открытым исходным кодом (open source) в статистике

**Open Sources Software и управление информационными технологиями учреждения  
статистики**

**Вспомогательный документ**

Представлен Государственным Комитетом по Статистике Азербайджана<sup>1</sup>

**1. ВВЕДЕНИЕ**

1. Движение Open Sources Software это направление коллективной разработки программ, при котором программы распространяются свободно всем желающим вместе с исходными кодами. В отличие от классического закрытого метода разработки, метод Open Sources – это такой способ, при котором:

- Программное обеспечение создается без единого подробно разработанного и утвержденного проекта;
- Программное обеспечение создается виртуальным коллективом свободных программистов, открытым для любого желающего;
- Документы разработки и тексты программ полностью открыты всем вообще, не только участникам разработки.

2. Целью данного документа не является анализ движения Open Sources, его психологических, моральных, социальных и коммерческих аспектов. Об этом написано

---

<sup>1</sup> Автор> Файг Жалиов (ssc@azstat.org)

достаточно много, в частности такими лидерами движения, как Richard Stallman, Eric Raymond, Tim O'Reilly и другими.

3. Ограничимся лишь вопросами использования Open Sources Software в статистических учреждениях и перспективами применения самого метода для разработки прикладных программ обработки статистической информации.

## **2. ОПЫТ ИСПОЛЬЗОВАНИЯ У НАС OPEN SOURCES SOFTWARE**

4. В настоящее время методом Open Sources разработано большое количество продуктов. В основном – это сетевые операционные системы, сервера для различных сетевых сервисов, а также большое число утилит и программ, выполняющих различные функции. В настоящее время используя только Open Sources Software можно создать полноценную компьютерную сеть любого размера с очень широкими функциональными возможностями.

5. С начала 2000 года мы взяли курс на широкое использование Open Sources Software для построения корпоративной сети Госкомстата Азербайджана. Такое решение было принято в первую очередь из экономических соображений. До этого времени мы использовали коммерческие сетевые операционные системы фирм Novell и Microsoft. Любое расширение корпоративной сети, добавление новых функциональных возможностей и обновление устаревшего программного обеспечения требовало значительных финансовых затрат. Ограниченность финансирования всегда являлась основным тормозом развития информационных технологий. Часто мы были вынуждены использовать устаревшие версии программных продуктов только потому, что не в состоянии были приобрести новые. Это значительно снижало функциональность корпоративной сети и ее безопасность. Поэтому мы приняли для себя концепцию, которую вкратце можно сформулировать так:

- Применять продукцию Open Sources Software всюду, где это экономит средства и не наносит ущерба функциональности информационных технологий.
- Закупать и применять коммерческую продукцию только в том случае, если соответствующих аналогов среди Open Sources Software нет, либо эти аналоги уступают коммерческим по своим характеристикам.

6. Схема действующей в настоящее время корпоративной сети Госкомстата Азербайджана, построенная согласно этой концепции изображена на рисунке 1.

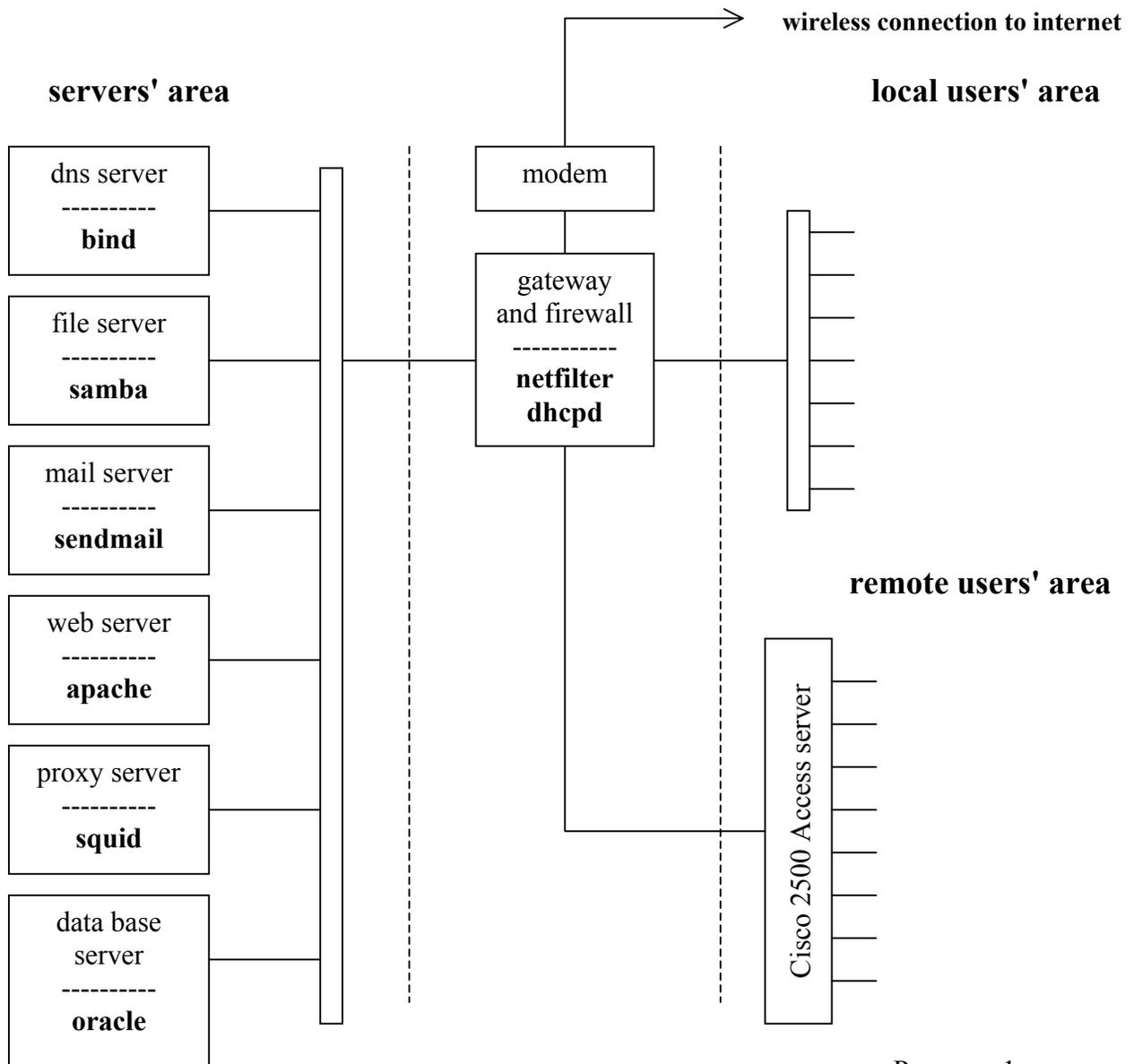


Рисунок 1

7. Все сервера, приведенные на схеме работают под управлением операционной системы Linux и все услуги Intranet-Internet полностью реализованы на различных свободно распространяемых программных продуктах Open Sources Software. Основные из них – это:

- служба DNS домена, реализованная на базе программы BIND;
- файловый сервер, реализованный на сервере Samba;
- Web-сервер Apache интегрированный с Perl, PHP и MySQL (Наш сайт [www.azstat.org](http://www.azstat.org) распространяется с этого сервера);
- почтовая служба на базе почтового сервера Sendmail;
- служба группового доступа к внешним интернет-ресурсам на базе Proxy-сервера Squid;
- шлюз и файрвол на базе netfilter.

8. Наряду с продуктами Open Sources Software используются и коммерческое программное обеспечение различных производителей. Основные из них – это:

- система управления базами данных Oracle. Пока еще не существует и вряд ли появится в ближайшее время свободно распространяемый продукт такого класса.
- практически на всех рабочих станциях пользователей используется MS Windows. Переход здесь на Linux связан со значительными затратами по переучиванию всего персонала и перепрограммированию существующих прикладных задач.

9. Для разработки прикладных задач используются как свободно распространяемые средства, так и коммерческие. Например, для разработки прикладных клиент-серверных статистических задач на стороне сервера используется MySQL, а на стороне клиента Borland C++ Builder и Delphi. Для локальных приложений широко применяется MS Visual FoxPro. В этих задачах применяется также большое количество свободно распространяемых программ и модулей. Удачное совмещение коммерческих продуктов с продуктами Open Sources Software дает наиболее оптимальный результат.

10. Использование продуктов Open Sources Software имеет свои особенности, отличные от коммерческих продуктов. Эти особенности всегда надо учитывать, если мы хотим успешно использовать для своих целей эти продукты:

- В отличие от коммерческих продуктов, новые версии программ Open Sources Software появляются очень часто. Наряду с версиями, считающимися стабильными всегда существуют новейшие, которые не прошли достаточного тестирования. Некоторые из них работают тоже стабильно, но они могут содержать и ошибки. При первоначальной установке продукта желательно выбирать последнюю стабильную версию, рекомендованную автором. Новейшие недостаточно проверенные версии программ можно использовать лишь тогда, когда они содержат новые функции существенно необходимые нам. Конечно, это связано с определенным риском. В таких случаях можно оценить относительную стабильность продукта, если проанализировать списки замеченных и исправленных ошибок на сайте разработчика.
- Новые версии продуктов Open Sources Software всегда содержат большое число новых возможностей и функций, недоступных ранее. Если эти функции необходимы нам, то мы удаляем старые и устанавливаем новые версии программ. Однако при этом надо всегда иметь в виду, что некоторые авторы практически не заботятся об обратной совместимости программ, удаляя из кода устаревшие функции. С одной стороны это очень хорошо, так как упрощает программу и улучшает ее эксплуатационные характеристики. Я тоже считаю, что все устаревшее надо решительно выбрасывать. Чрезмерное увлечение поддержкой обратной совместимости в конце концов приводит к тому, что полезные функции программы тонут в куче старого мусора (что в частности происходит с большинством коммерческих продуктов). Однако может оказаться, что некоторые наши старые конфигурационные файлы и даже некоторые связанные с этим продуктом другие программы перестанут работать. Поэтому обновляя версии, необходимо прежде всего внимательно изучить всю сопровождающую документацию, а после установки любого продукта проверить работу всех связанных с ним программ и если необходимо, то внести в них исправления.

- При построении корпоративной сети на базе продуктов Open Sources Software всегда следует следить за взаимной совместимостью совместно работающих продуктов. Это следует всегда учитывать при подборе и обновлении версий продуктов. Например, такие продукты как Apache WEB-сервер, PHP и MySQL часто используются совместно. Однако, мы сталкивались с ситуацией, когда рекомендованные последние стабильные версии этих продуктов не поддерживали совместную работу. Это часто происходит из-за того, что один продукт изменяется, а авторы другого еще не успели внести соответствующие изменения у себя. В таких случаях приходится долго разбираться с сопровождающей документацией продуктов, проводить эксперименты и выбирать наиболее оптимальное сочетание версий продуктов.
- Свободное распространение исходного кода продуктов Open Sources Software – это главная идея и движущая сила этого движения. Лицензионные условия этих продуктов достаточно мягкие и часто они позволяют любому лицу вносить изменения и распространять измененный продукт дальше. Поэтому всегда существует определенный риск того, что мы приобретем продукт с внесенными в него ошибками или другими нежелательными изменениями. Поэтому приобретать продукт можно только из надежных источников. Это может быть сайт разработчика или рекомендованные разработчиком зеркальные сайты, либо диски с дистрибутивом программ, распространяемые известными фирмами.
- Продукты Open Sources Software обычно можно устанавливать как трансляцией исходного кода, так и с заранее подготовленного кем-то грм-бинарника. Опытные программисты предпочитают обычно первый вариант, так как в этом случае достигается наиболее точная настройка рабочего продукта и оптимизация его эксплуатационных характеристик. Однако это требует высокой квалификации персонала системных программистов и затрат рабочего времени.

11. Из всего вышеизложенного следует, что использование продукции Open Sources Software в реальных системах требует наличия высококвалифицированного персонала системных программистов. Одновременно нагрузка и значимость этих программистов значительно возрастает. Однако персоналу работать с этими продуктами значительно интереснее, чем с коммерческими. Поэтому люди проявляют энтузиазм и их профессионализм быстро растет.

12. Первоначально мы перешли на продукцию Open Sources Software только из экономических соображений, однако в процессе их эксплуатации поняли, что цена здесь не самое главное. Главное то, что практически все системные продукты Open Sources Software имеют очень высокие эксплуатационные характеристики. Здесь имеется в виду надежность, быстродействие, идеологическая новизна, безопасность, экономия ресурсов вычислительной техники и многое другое.

### **3. ПЕРСПЕКТИВЫ РАЗРАБОТКИ ПРИКЛАДНЫХ ПРОГРАММ ДЛЯ СТАТИСТИКИ МЕТОДОМ OPEN SOURCES**

13. Автоматизация статистических задач решается обычно тремя путями одновременно:

- Использование стандартных пакетов;
- Разработка прикладных программ своим персоналом программистов;
- Заказ прикладных программ сторонним фирмам;

14. Стандартные пакеты – это готовые, разработанные различными фирмами, программные продукты, которые имеют развитый интерфейс конечного пользователя и большой выбор стандартных статистических функций. Используя их, сам пользователь путем некоторых настроек может автоматизировать часть своих работ и выполнить весь цикл записи и обработки данных на своем настольном компьютере. К ним относятся, например, MS Excel, SPSS, STATA и множество других продуктов. Стандартные пакеты обычно пригодны для класса задач с небольшим объемом данных и несложным алгоритмом расчета. Применение стандартных пакетов значительно снижает нагрузку на персонал разработчиков программ. Поэтому наиболее верный подход – это применять стандартные пакеты всюду где это возможно и где это дает полноценный результат.

15. Однако не все статистические задачи поддаются автоматизации стандартными пакетами. В таких случаях приходится разрабатывать специальное программное обеспечение. Применяя универсальные алгоритмические языки программирования можно разработать приложения для автоматизации любой задачи. При этом могут быть реализованы любые специфические для конкретной задачи тонкости, которые невозможно реализовать применяя пакеты в силу существующих в них ограничений. К таким задачам, например, относятся статистические задачи требующие применения клиент-серверных технологий.

16. Обычно персонал программистов учреждения статистики невелик и не способен охватить разработку всех задач учреждения. Мы сталкиваемся с проблемой, когда с одной стороны существует большое количество разнообразных задач требующих разработки специальных программных продуктов, с другой – ограниченный контингент программистов, явно недостаточный для выполнения этой работы. Приходится либо увеличивать состав разработчиков, либо заказывать часть задач сторонним фирмам на договорной основе. То и другое требует значительных затрат.

17. Естественно возникают вопросы. Можно ли придумать какие-либо альтернативные пути разработки статистических задач без дополнительных затрат? Можно ли разрабатывать задачи обработки статистической информации методом Open Sources? Если да, то какой класс задач и каким образом организовать работу? Открытый метод разработки задач путем привлечения виртуального коллектива единомышленников мог бы частично решить проблему. Во всяком случае можно порассуждать на эту тему.

18. Рассмотрим некоторые из сопутствующих вопросов:

- **Класс задач.** Задачи, предлагаемые для открытой разработки должны быть достаточно универсальны, иметь широкую сферу применения и предлагать оригинальные и интересные решения. Это необходимо для привлечения большого коллектива свободных программистов для ее реализации. Следует иметь в виду, что программисты подключаются к проекту вовсе не из альтруистических соображений. Это потенциальные пользователи, которые также как и мы заинтересованы в конечном результате. Нельзя рассчитывать на привлечение

большого виртуального коллектива, если предлагаемая задача имеет очень узкую сферу применения. Не случайно, подавляющее большинство задач, разработанных методом Open Sources – это системное программное обеспечение. Это происходит потому, что эти программы имеют очень большую аудиторию пользователей а сами пользователи являются одновременно и программистами.

- **Управление проектом.** Проект может быть объявлен, на рабочем сайте. Но нельзя начинать его с пустого предложения. Важно с самого начала предложить разработанную программу с исходными текстами. Она может быть на начальной стадии разработки, содержать ошибки и плохо работать. Это не важно. Главное идея. Проект должен управляться высококвалифицированным специалистом учреждения, который имеет качества общения с коллективом.

#### **4. РАЗРАБОТКА ПРИКЛАДНЫХ ПРОГРАММ ДЛЯ СТАТИСТИКИ КОРПОРАТИВНЫМ МЕТОДОМ**

19. Это метод совместной разработки прикладных задач несколькими организациями, заинтересованными в ее результатах. Такая работа может быть организована путем договоренности между этими организациями. Задача может быть поделена на функционально независимые части каждая из которых может быть выполнена одной из организаций. Исходные тексты задач могут быть открыты для облегчения тестирования и исправления ошибок. Таким методом могут разрабатываться задачи, которые не являются достаточно универсальными и имеют узкую отраслевую сферу применения. Поэтому они не могут рассчитывать на привлечение большого виртуального коллектива свободных программистов. Однако они интересны для нескольких статистических учреждений, которые и могут заняться реализацией совместно.

- - - - -