

Distr.
GENERAL

CES/AC.71/2004/23/Rev.1
23 April 2004

RUSSIAN ONLY

**СТАТИСТИЧЕСКАЯ КОМИССИЯ и
ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ
КОМИССИЯ
ОРГАНИЗАЦИИ ОБЪЕДИНЕННЫХ НАЦИЙ
КОНФЕРЕНЦИЯ ЕВРОПЕЙСКИХ
СТАТИСТИКОВ**

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

**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 II: Разработка стратегий ИТ в статистических управлениях

**Совершенствование архитектуры информационно-вычислительной системы Госкомстата
России**

Вспомогательный документ

Представлен Госкомстатом Российской Федерации¹

1. Развитие системы государственной статистики в Российской Федерации предполагает обеспечение своевременности, качества и полноты статистических данных и их эффективного использования при решении задач управления и прогнозирования во всех сферах, включая анализ состояния и динамики развития национальной экономики, финансовой системы, демографической и социальной ситуации. В современных условиях выполнение этих задач невозможно без использования высокопроизводительных вычислительных и телекоммуникационных средств, интегрированных в единую информационно-вычислительную систему федерального масштаба.

2. Основное назначение корпоративной информационно-вычислительной системы (ИВС) государственной статистики заключается в поддержке ее деятельности на федеральном, региональном и районном уровнях. К основным технологическим процессам ИВС относятся:

- производственная деятельность Госкомстата России, определяемая Федеральной программой статистических работ и Производственным планом статистических работ, а именно: сбор, обработка, анализ и доставка официальной статистической информации органам государственной власти и управления;

¹ Автор > Юрий Голованов (golovanov@gks.ru)

- предоставление официальной статистической информации в международные организации, местные администрации, физическим лицам;
- выполнение статистических исследований по заказам органов государственной власти и управления;
- оказание организациями, входящими в структуру Госкомстата России, платных информационно- вычислительных и других услуг заказчикам;
- управление внутренней деятельностью Госкомстата России.

3. Проведение работ по модернизации ИВС Госкомстата России преследует следующие основные цели:

- повышение качества статистической информации, предоставляемой в органы государственной власти и управления, в международные организации и местным потребителям;
- сокращение сроков обработки и анализа статистической информации;
- сокращение сроков выполнения статистических исследований по заказам органов государственной власти и управления;
- снижение бюджетных расходов на содержание системы государственной статистики, в том числе за счет сокращения персонала, занятого малопроизводительным трудом;
- увеличение доходов за счет расширения объемов услуг, оказываемых за плату организациями, входящими в структуру Госкомстата России.

4. Принципиальным элементом постановки является необходимость не просто перевода действующей системы на новую технику, а создания качественно новой системы, которая обеспечит выполнение Госкомстатом России стоящих перед ним задач в условиях продолжающихся социально- экономических реформ.

5. Важнейшими принципами реализации новой ИВС являются:

- позиционирование ИВС как основы для создания единого информационного пространства государственной статистики, ее использование как для решения задач, определяемых Федеральной программой статистических работ, так и для выполнения работ в интересах региональных заказчиков;
- применение современных информационных технологий и средств на всех этапах технологического процесса – от планирования статистических наблюдений, сбора данных до их анализа и доставки потребителю;
- возможность постоянного эволюционного развития.

ЕДИНОЕ ИНФОРМАЦИОННОЕ ПРОСТРАНСТВО

6. Статистические ресурсы страны включают разнообразную информацию о деятельности предприятий, жизни населения, имеющимся и используемым природным ресурсам, финансам, интеллектуальной собственности. Эта информация сосредоточена в информационных фондах Госкомстата России и организаций, входящих в его структуру, информационных хранилищах, органов власти и управления на федеральном уровне и на уровне субъектов федерации, местных администраций, отдельных предприятий, библиотек, учебных и научно-исследовательских организаций. В общие статистические ресурсы необходимо должна быть включены все методологические, нормативные и учебные материалы в данной области, а также информация из зарубежных источников, получаемая на основе обмена с национальными и международными органами статистики и другими организациями.

7. Таким образом, ставится задача интеграции статистических информационных ресурсов в масштабах страны для эффективного использования при решении задач управления и

прогнозирования во всех сферах и на всех уровнях, от федерального до муниципального. В частности, создание развитой информационной инфраструктуры в регионах и на муниципальном уровне позволит осуществлять комплексный анализ совокупности происходящих в этих образованиях социальных, экономических и технологических процессов, необходимых для своевременного принятия решений. В ряде регионов уже начата реализация комплексных территориальных информационных систем, с включением статистических ресурсов как базовых элементов.

8. Создание интегрированных информационного ресурсов позволит не только повысить качество информации, предоставляемой органам власти различных уровней, но и проводить единую политику в области информационного обеспечения потребителей, в первую очередь органов власти. Под этим подразумевается не только техническая и организационная возможность создания интегрированного информационного ресурса, но и целенаправленное формирование потребности у различных потребителей в определенном виде информации.

Важнейшее значение приобретает интеграция информационных ресурсов Госкомстата России с ресурсами других министерств и ведомств, обеспечивающих сбор и накопление ведомственной статистики, а также осуществляющих регистрационные и лицензионные функции. В первую очередь в эту группу входят МНС России, ГТК России, МВД Российской Федерации, Министерство экономического развития и торговли.

9. Практическая реализация данного принципа предполагает:

- переход на сквозную технологию сбора, передачи, обработки, накопления и представления официальной статистической информации;
- разработку и внедрение типовой технологии и стандартов информационного обмена;
- использование совместимой системы метаданных, включающей как общероссийские классификаторы, так и единый Каталог статистических показателей

ВЛИЯНИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

10. Произошедшие за прошедшее время изменения в информационных технологиях, как в производительности отдельных вычислительных комплексов, так и в телекоммуникациях, позволяют проводить перестройку существующих процессов сбора, обработки и распространения информации.

11. В частности, для повышения эффективности работы ИВС предполагается существенно улучшить автоматизацию наиболее трудоемкого этапа – первичного ввода данных – за счет использования электронных методов сбора отчетности непосредственно с предприятий и использования автоматизированного сканирования.

12. Существенное снижение показателя цена/производительность для серверов баз данных, аппаратных вычислительных платформ и устройств хранения данных позволяют осуществить повсеместный переход к банковской технологии хранения и обработки информации, внедрению архитектур клиент-сервер.

13. Продолжающееся развитие Интернет и его доступность в большинстве регионов России делают возможным позиционирование Интернет в качестве одного из основных средств взаимодействия ИВС с внешними поставщиками и потребителями статистической информации. Использование Интернет -технологий в рамках корпоративной ИВС при условии организации защиты информации или глобальной сети передачи данных с ограниченным доступом будет являться определяющим при взаимодействии организаций Госкомстата России между собой и для межведомственного обмена информацией.

ВОЗМОЖНОСТИ РАЗВИТИЯ

14. Одним из очевидных свидетельств развития в области ИТ является продолжающееся снижение значения показателя «цена/производительность» вычислительных средств. Таким образом, практически во всех случаях целесообразно осуществлять ориентацию на применение стандартных программно-технических компонент, а не на разработку собственных решений, как это, в частности, отмечается в рекомендациях Европейской Статистической комиссии. Другой важный фактор связан с резкими изменениями на рынке готовых компонент. В этой связи, развитие ИВС следует ориентироваться, в первую очередь, на использование стандартных (промышленных) интерфейсов взаимодействия между этими компонентами. Во всех случаях замена одной из компонент системы, поддерживающей стандартный интерфейс, на более совершенную версию будет проще, чем замена нестандартного средства, которое реализует собственный интерфейс взаимодействия.

15. Таким образом архитектура ИВС должна базироваться преимущественном использовании стандартных интерфейсов, подсистем и типов компонент с хорошо определенной функциональностью. Такой подход должен обеспечить необходимую стабильность и адаптивность ИВС под меняющиеся внешние условия.

АРХИТЕКТУРА ИВС

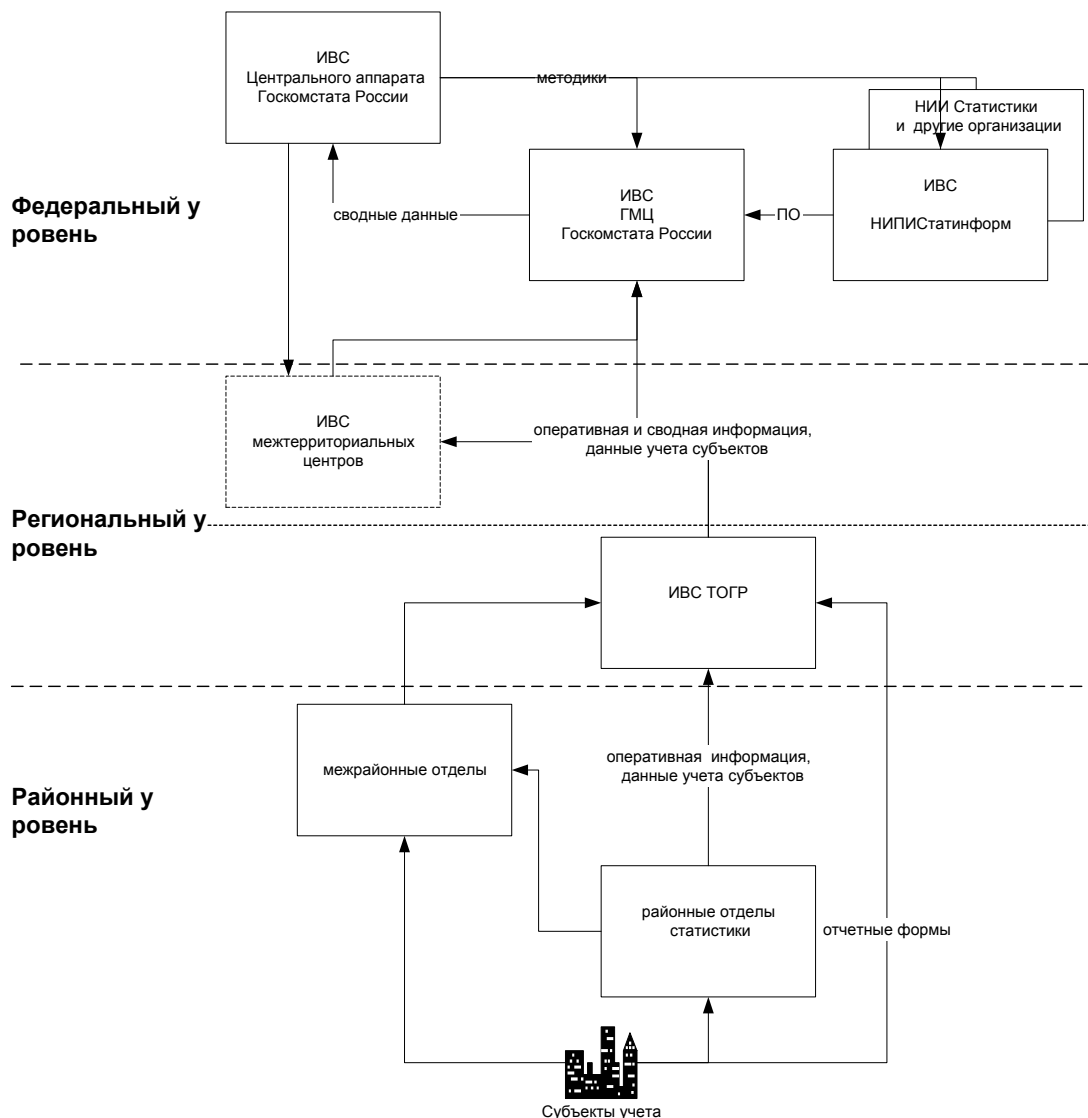
16. ИВС должна включать ряд компонент (подсистем), которые могут быть сгруппированы по организационно-иерархическому или функциональному принципам.

17. С точки зрения организационной структуры, в состав ИВС должны войти компоненты, реализующие ту или иную функциональность соответственно на федеральном, региональном или местном уровнях с учетом возможного наличия промежуточных уровней – межтерриториальных центров обработки и межрайонных отделов.(рис.1)

18. Корпоративная ИВС федерального уровня включает:

- (i) ИВС Центрального аппарата Госкомстата России, на базе которой ведутся, в частности, информационные фонды статистических показателей и готовых документов;
- (ii) ИВС ГМЦ Госкомстата России, осуществляющего обработку статистической информации и ведение статистических регистров;
- (iii) ИВС ЦИГС НИПИСтатинформ, который обеспечивает проектирование, разработку и сопровождение программно-технологических решений.

19. На региональном уровне в состав ИВС включены ИВС территориальных органов Госкомстата России (ТОГР), которые обеспечивают сбор, предварительную обработку статистической информации в регионах, предоставление информации для региональных администраций и других потребителей, ведение части регистров и классификаторов по соответствующему региону. На районном уровне в состав ИВС должны быть включены ЛВС и рабочие станции районных и межрайонных (в тех регионах, где они получили распространение) отделений территориальных органов Госкомстата России.



1. Рис. 1

20. Существующая организационная структура и схема сбора и обработки данных построена по схеме "звезда" и предполагает управление из одного центра 89 объектами. С точки зрения общих принципов управления, системного подхода, значительной разнородности территорий (различия достигают сотен раз) по объемам обрабатываемых данных, отечественного и зарубежного опыта и конкретных шагов по созданию представительств Президента РФ, данную схему и существующий порядок нельзя считать оптимальным. Поэтому в предстоящий период предусматривается возможность перераспределения функций ТОГР в будущем.

21. На базе некоторого количества (от 7 до 10) ТОГР могут быть созданы межтерриториальные центры обработки данных (ЦОД). При этом на эти центры будут возложены функции сбора, предварительной обработки (контроля) всей поступающей информации в пределах зоны, охватываемой центром, передача ее на ГМЦ Госкомстата России, обеспечение статистических услуг для пользователей. Остальные ТОГР в зоне такого центра будут в этом случае выполнять только минимально необходимые функции по передаче информации с районного уровня в центр (там, где это необходимо), взаимодействию с местными органами

власти, оказанию технической поддержки, и т.п. Предполагается, что данная реорганизация будет проводиться постепенно на основании анализа пилотных проектов в ряде регионов. При этом в проектных решениях ИВС предусматривается обеспечение инвариантности по отношению к таким изменениям. На практике это означает, что система должна будет обеспечить одновременную работу по новой технологии части территорий по новой схеме, части территорий - по старой.

Создание Центров обработки данных

22. Выбор ИТ-архитектуры закладывает тот фундамент, на котором будут эксплуатироваться прикладные системы для решения основных задач, стоящих перед заказчиком. Любой просчет на этой стадии впоследствии сложно исправить, что грозит обернуться серьезными потерями.

23. Цель проекта, начатого в декабре 2002 г. состояла в модернизации и реорганизации технологического процесса и технической базы информационной системы Госкомстата России. Для главного статистического ведомства страны требовалось создать общий интегрированный информационный ресурс и инструментарий для проведения единой политики по оперативному обеспечению аналитической информацией своих пользователей.

24. В структуре Государственного комитета РФ по статистике имеются три основных уровня: районный, региональный и федеральный. Сбор данных осуществляется на всех уровнях, а консолидация и обработка огромных массивов информации ведется на региональном и федеральном уровнях. Такая иерархическая организация ведомства определила структуру его информационно-вычислительного комплекса. Для построения инфраструктуры ИС на федеральном уровне и в нескольких опорных региональных узлах в качестве ключевой была выбрана технологии построения Центров Обработки Данных (ЦОД). Северо-Западный регион с центром в Санкт-Петербурге было решено сделать опорным, в настоящее время там близится к завершению пилотный проект по созданию первого Центра Обработки Данных, к которому будут подключены объекты Северо-Западного региона: Калининград, Вологда, Архангельск и Псков.

25. Информационная система должна обладать адекватной для решения столь масштабных задач функциональностью, достаточной производительностью и высокой доступностью (непрерывностью технологического процесса). Эти критерии подразумевают создание надежных технических средств путем дублирования компонентов и технологических процессов информационной системы, а также внедрение политики безопасности. Всем этим требованиям наиболее адекватно отвечает технология построения Центра Обработки Данных – мощного инструмента для управления и интеграции информационных ресурсов.

26. Технология построения Центра Обработки Данных – это новый виток в развитии традиционной концепции Вычислительного Центра, которая берет свои истоки в организации вычислительных процессов на базе централизованной модели вычислений. Современные ЦОД больше сфокусированы на решение бизнес-задач путем предоставления услуг в виде информационных сервисов, нежели прежние ВЦ, ориентированные на предоставление услуг по поддержанию вычислений. Общей для обеих технологий остается базовая модель централизованных вычислений, подразумевающая консолидацию ресурсов и организацию процесса эксплуатации с целью поддержания уровня качества услуг.

27. Какие преимущества предоставляет такая модель? Прежде всего, централизованный подход позволяет упростить управление вычислительным комплексом и более эффективно использовать его ресурсы путем их динамического перераспределения за счет того, что все ресурсы сосредоточены в одном месте. Очевидно, что политика обеспечения высокого уровня доступности и безопасности, обеспечивающая оперативный доступ к информационным ресурсам,

централизованно осуществляется гораздо проще. Оптимально низкая общая стоимость затрат на владение единой информационной структурой – один из главных аргументов в пользу выбора консолидированного подхода. Все перечисленные преимущества полностью согласуются с требованиями, предъявляемыми к новой информационной структуре Государственного комитета РФ по статистике, на базе которой будут решаться стоящие перед статистическим ведомством стратегические задачи.

28. Технология ЦОД, основанная на централизованной модели вычислений, включает в себя комплекс взаимосвязанных организационно-технических решений, направленных на поддержание высокого уровня доступности информационных ресурсов. Неотъемлемой частью ЦОД являются отказоустойчивые серверные комплексы, построенные на базе кластерных технологий. Технологические процессы, как правило, дублируются Резервным Центром, основное предназначение которого – перевод вычислительного процесса в РЦ с целью предоставления информационных услуг в непрерывном режиме в случае выхода из строя основного серверного комплекса. Этот перевод также может осуществляться путем применения кластерных технологий между основным и Резервным Центрами. Для организации автоматического перевода критичных информационных сервисов на вычислительные мощности РЦ используются кластерные технологии, объединяющие серверы ЦОД и РЦ в единый серверный комплекс с удаленными узлами. В качестве обязательного решения по обеспечению непрерывности бизнеса в ЦОД применяется централизованная система резервного копирования. Эффективным средством предупреждения простоев служит проактивный мониторинг состояния вычислительных средств и программного обеспечения ИС – составная часть системы эксплуатации ЦОД. Все перечисленные технологии ЦОД направлены на сокращение времени возможного простоя. Для ликвидации причины неполадок служит план восстановления после аварий, осуществление которого возможно только при условии соблюдения должного уровня сервисного обслуживания.

29. Концепция ЦОД для Госкомстата России была разработана в качестве типового решения с целью дальнейшего тиражирования ее на региональных объектах. Вопросы повышения уровня доступности ценной статистической информации потребовали организации основного федерального центра обработки данных и дублирующего его резервного центра. Центральные узлы сети размещены в Центральном аппарате Госкомстата России и ГМЦ (главном межрегиональном вычислительном центре). Поэтому основной центр планируется внедрить на базе ГМЦ, а в Центральном аппарате -- организовать резервный центр. Прямой волоконно-оптический канал длиной 14 км соединит оба центра, где будут установлены симметричные системы хранения данных общей емкостью 5 Тбайт с возможностью масштабирования. В рамках пилотного проекта типовое решение по созданию регионального ЦОД будет первоначально опробовано на Северо-Западном объекте регионального уровня – в Санкт-Петербурге. В ЦОД Санкт-Петербурга предусмотрена система хранения емкостью 1 Тбайт. Резервный центр здесь пока не планируется, но все процедуры резервного копирования и восстановления после сбоев являются стандартными и обязательными к исполнению.

30. Благодаря организации новых каналов связи появляется возможность в удаленном режиме организовать ввод данных в БД напрямую. Автоматическое круглосуточное взаимодействие подсистем на основе вычислительных мощностей ЦОД резко сокращает сроки сбора и обработки информации. Благодаря тому, что все данные, справочники, классификаторы, собраны в одном месте, обеспечивается согласованность информации. Все это самым непосредственным образом влияет на улучшение качества информации, получаемой в оперативном режиме. Благодаря консолидации вычислительных мощностей расходы на закупку, эксплуатацию и модернизацию технических средств также заметно снижаются.

31. Основные компоненты типового центра обработки данных включают серверный комплекс, систему хранения данных (СХД), систему эксплуатации для организации эффективного

управления ИТ-ресурсами и компонентами инфраструктуры ЦОД, а также систему информационной безопасности (СИБ), представляющую собой комплекс организационных мер и программно-технических средств защиты.

32. Серверный комплекс ЦОД, который планируется внедрить в рамках пилотного проекта, проектируется с учетом иерархии, в соответствии с которой все серверы можно выделить в четыре группы. Первую группу составляют серверы представления информации: в ЦОД будет реализована технология доступа к приложениям на основе концепции «тонкий клиент», а также доступ к информационным ресурсам через Web-браузер. Вторую группу образуют серверы приложений, которые, собственно и осуществляют выполнение всех бизнес-задач Заказчика. Третья группа – это серверы информационных ресурсов, ответственные за сохранение и предоставление данных серверам приложений. Служебные серверы образуют четвертую группу – они поддерживают работу других подсистем ЦОД, обеспечивая, например, управление системой резервного копирования. Такая архитектура серверного комплекса облегчает возможность его масштабирования, модификации бизнес-логики, позволяет использовать сервисы второго уровня другими приложениями, а также обеспечивает сохранение инвестиций по мере роста обрабатываемых данных за счет перевода старых серверов информационных ресурсов, требующих больших вычислительных мощностей, в категорию серверов приложений.

33. Как уже упоминалось, серверные комплексы ЦОД строятся на кластерных технологиях с целью обеспечения высокого уровня доступности информационных ресурсов или распараллеливания решения сложной вычислительной задачи. Однако высокая готовность самого кластера оправдана только при условии соблюдения высокого уровня доступности данных, которые этот кластер обрабатывает. Высокая готовность данных в ЦОД обеспечивается путем создания Системы Хранения Данных, надежность которой обеспечивает организация процедуры резервного копирования.

34. Полный комплекс Системы Хранения Данных включает устройства хранения данных (дисковые массивы и ленточные библиотеки), причем все дисковое пространство организуется как единый пул хранения, обеспечивающий динамическое выделение ресурсов и поддерживающий разграничение доступа к ним. Для увеличения надежности хранения предусмотрены процедуры тиражирования данных между массивами и создания мгновенных копий. В состав СДХ входит также инфраструктура доступа, которую обеспечивает сеть хранения данных (Storage Area Network, SAN). Компонентом СХД является также программное обеспечение управления хранением данных. В рамках Системы Хранения Данных установлена обязательная процедура резервного копирования и архивирования данных, а также поддерживается система эксплуатации.

35. Все процедуры СХД, включая работу систем резервного копирования, должны быть взаимоувязаны с политикой, проводимой в Резервном Центре. Сеть хранения данных ЦОД также должна быть взаимоувязана с его Резервным Центром. Для организации распределенных кластеров необходим прямой доступ серверов ЦОД к устройствам хранения в РЦ, а это наиболее эффективно решается посредством SAN.

36. Описанная инфраструктура ЦОД обеспечивает гибкость для поддержания новых бизнес задач, которые могут возникнуть у Государственного комитета Российской Федерации по статистике в перспективе.
