

ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНАЯ СИСТЕМА ВЦ РАН

С 1998 г.

Г.М. Михайлов, М.А. Копытов, М.В. Кулагин Ю.П. Рогов, С.В. Сорокин

Принципы построения информационно-вычислительной системы ВЦ РАН (ИВС ВЦ РАН) и ее базы, локальной вычислительной сети (ЛВС), изложены в [1]. Анализ состояния ИВС по завершении определенного этапа ее развития, главным образом связанного с развитием инфраструктуры ЛВС, дан в [2]. Содержание данной статьи является логическим продолжением упомянутых монографий.

В ВЦ РАН в период, предшествующий развертыванию ИВС, уже был накоплен достаточно большой опыт использования систем коллективного пользования, построенных на программно-аппаратных средствах ранних поколений отечественных ЭВМ (БЭСМ-6, ЕС 1066). Для тех систем с точки зрения предоставляемых системных функций была характерна ориентация на пользователя-программиста, специалиста в своей области, использующего вычислительные ресурсы и соответствующее системное программное обеспечение для того, чтобы реализовывать собственные алгоритмы или производить расчеты по готовым программам.

Иными словами, программист использовал вычислительную систему, чтобы "считать" на ней.

Главные функции старых систем: предоставление вычислительных ресурсов, организация эффективного их использования в различных режимах доступа к системе (мультипрограммирование, пакетная обработка, разделение времени и ресурсов и т.п.), предоставление средств автоматизации процесса программирования (текстовые редакторы, компиляторы, библиотеки стандартных программ, загрузчики, отладчики и т.п.). Использование вычислительных ресурсов для предоставления информационных услуг было минимальным.

Состав пользователей вычислительных систем был достаточно однородным. Однако с появлением персональных компьютеров и в особенности в связи с развитием систем управления базами данных, с появлением локальных и глобальных сетей ЭВМ, с переходом от работы с текстовой и битовой информацией к широкому использованию графической, аудио- и видеоинформации начинается новая жизнь вычислительных систем. Однородность пользователей нарушается, появляются новые типы пользователей.

Если все множество пользователей вычислительных систем разбить на две группы, то одну группу составят пользователи-программисты, которых мы охарактеризовали как специалистов, "считающих" на компьютерах, а другую - все остальные, использующие вычислительные ресурсы и специальное программное обеспечение для удовлетворения собственных информационных запросов и обмена разнообразной информацией с другими источниками (например, электронная почта).

Такое разделение пользователей достаточно условно. В настоящее время все пользователи, в том числе и пользователи-программисты, естественно, активно используют вычислительные средства и в информационных целях. Так, при использовании старых систем в ВЦ РАН соотношение между пользователями-программистами и остальными было примерно 99:1, т.е. практически все пользователи "считали" на ЭВМ.

Этот факт безусловно повлиял на особенности реализации новой ИВС ВЦ РАН, которую мы начали развертывать в 1993 г. Правда, уже в 1993 г. в специальной литературе сообщалось, что в мире соотношение между пользователями-программистами и прочими пользователями составляет 1:5, т.е. только 20% пользователей используют вычислительные системы для целей программирования. Следует признать, мы относились к этой статистике скептически, полагая, что для пользователей ИВС ВЦ РАН она не в полной мере справедлива. Однако в процессе пятилетней эксплуатации ИВС ВЦ РАН мы сами смогли убедиться в существенном изменении структуры современных пользователей вычислительных систем. В настоящее время коллектив пользователей ИВС ВЦ РАН составляет 300 человек. И соотношение между указанными группами пользователей в ВЦ РАН - 1:10.

Безусловно, что эти статистические данные помимо общей мировой тенденции сокращения пользователей-программистов отражают и субъективный фактор, а именно то, что современного пользователя-программиста, в частности пользователя ВЦ РАН, не удовлетворяют предоставляемые вычислительные ресурсы. Ему нужны более мощные вычислительные системы. О перманентном требовании пользователей-программистов к повышению производительности используемых вычислительных серверов нам известно достаточно хорошо.

Тем не менее при совершенствовании работы ИВС уже нельзя игнорировать факт, что информационные запросы пользователей выходят на первое место. Это значит, что при реализации информационных функций ИВС важно рассматривать вопросы повышения их эффективности. Современная ИВС должна не только обеспечивать для пользователей-программистов эффективную и надежную работу вычислительных систем, но и комфортность, надежность и безопасность при работе с информационными системами.

Все системные функции (традиционные и информационные) реализуются в мультипрограммной и мультипроцессорной работе множества системных процессов, исполняемых вместе с программами пользователей на множестве физических серверов ЛВС. В условиях сетевого взаимодействия программ многие системные функции реализуются по технологии "клиент-сервер", что увеличивает общую нагрузку на физические ресурсы ИВС. Повысить эффективность исполняемых функций - это значит не только совершенствовать сами системные программы, реализующие эти функции, но и правильно их распределять по всему полю физических серверов ЛВС.

В результате почти пятилетнего периода эксплуатации и развития ИВС ВЦ РАН мы пришли к выводу, что необходимо разделить все множество физических хостов ЛВС ВЦ РАН на функциональные группы или сегменты, каждый из которых с наибольшим эффектом должен обеспечивать конкретные системные функции. Речь, естественно, идет о локализации серверных программ. Клиентские системные программы, реализующие запросы к серверным процессам, в соответствии со здравым смыслом, должны иметь возможность становиться активными в каждой физической точке ЛВС.

Ранее все системные функции реализовывались на всем множестве серверов ЛВС. В результате некоторые серверы были явно перегружены, что отразилось и на эффективности исполнения многих функций, и на надежности работы системы в целом. Например, в реальной эксплуатации была рабочая станция, на которой исполнялись функции файлового сервера, процессы NIS, программы оболочек и протоколов INTERNET в конкурентном мультипрограммном взаимодействии с некоторыми другими системными процессами и процессами пользователей-программистов. По нашим расчетам выделение функциональных сегментов должно повысить надежность ИВС и уровень ее безопасности.

Вычислительный сегмент построен на программно-аппаратных средствах, предназначенных для пользователей-программистов.

Отметим, что хотя и наблюдается некоторый прогресс в оснащении ИВС ВЦ РАН высокопроизводительной вычислительной техникой (ВВТ), в целом уровень оснащенности ВВТ оставляет желать лучшего.

Главная причина - практически полное отсутствие средств на приобретение действительно современной ВВТ. Отчасти этот недостаток компенсируется тем, что некоторые из наших пользователей на соответствующей договорной основе пытаются работать с удаленными серверами ВВТ, благо выход в глобальные сети наша ИВС имеет. Пользователи, использующие эту возможность, реализуют ее с рабочих мест специального сегмента массивно-параллельных приложений.

На вычислительных серверах используются следующие операционные системы (ОС). Это - различные разновидности UNIX (Solaris, OSF, BSD). Естественно, возникает желание использовать последние версии системного программного обеспечения. Однако из-за недостатка средств, а также поскольку некоторые новые версии ОС иногда не поддерживают старые версии средств автоматизации программирования, приходится держать в эксплуатации и старые версии ОС. Например, по этим причинам мы не можем полностью отказаться от Solaris 1 и перейти на работу только с Solaris 2.

Сегмент персональных компьютеров представляет собой совокупность клиентских рабочих мест, расположенных в центральном зале и на рабочих местах сотрудников ВЦ РАН. Каждое рабочее место помимо обычных функций персонального компьютера обеспечивает доступ к общим вычислительным и информационным ресурсам

системы, файловым серверам и выход в INTERNET. Каждый персональный компьютер оснащен специальным адаптером - сетевой платой и связан с центром коммутации через кабельную сеть на витых парах (10BaseT). На персональных компьютерах функционирует каноническая для PC операционная среда, дополненная, если в этом есть необходимость, специальным пакетом PC-NFS для связи с серверами ЛВС. Сейчас в ИВС ВЦ РАН имеется более 70 таких рабочих мест. Включенные в таблицу рабочие станции SUN SPARCStation SLC в настоящее время тоже используются в качестве клиентских рабочих мест.

Файловый сервер теперь функционирует на специальной рабочей станции SUN SPARCStation 10/30. Практически это - главная серверная функция данной рабочей станции. Как уже отмечалось выше, для максимального повышения эффективности работы центральной файловой системы на файловом сервере мультипрограммирование сведено к минимуму. В настоящее время центральная файловая система строится на следующих дисковых устройствах:

- SUN1.05 (1.05 Гб),
- SUN1.3G (1.3 Гб),
- SSI-440 RACa-ray (14 Гб).

Имеются следующие отдельные файловые системы:

- /home/sunny (0.5 Гб),
- /homeb1 (2 Гб),
- /homeb2 (2 Гб),
- /homeb4 (2 Гб),
- /homeb5 (2 Гб),
- /cd-mas (2 Гб).

Все файловые системы через аппарат монтирования и экспортирования доступны в любой точке ЛВС. Кроме центральной файловой системы имеются автономные файловые системы специального назначения, построенные на локальных физических ресурсах отдельных рабочих станций.

Общий объем внешней памяти (без учета локальной памяти сегмента персональных компьютеров) составляет 40 Гб.

Сегмент информационных и сетевых серверов использует рабочую станцию SUN SPARCStation 10/41. На этой станции функционирует web-сервер ВЦ РАН (www.ccas.ru), служба электронной почты Email (sendmail, POP-сервер), доменная система имен (DNS). Рабочая станция SUN SPARCStation 10/51 используется для обслуживания терминального сервера и модемов, а также для лицензионной поддержки программы PC-NFS, установленной на персональных компьютерах, подключенных к ЛВС.

Сейчас, по прошествии более четырех лет, эксплуатации и развития ИВС ВЦ РАН имеет смысл представить более подробное описание информационной составляющей ИВС. Это относительно новая для систем коллективного пользования ВЦ РАН компонента, которой, как мы уже отмечали, не было на старых системах, предшествующих развиваемой ИВС ВЦ РАН.

Необходимость создания информационной составляющей была определена нами сразу же, еще при проектировании ИВС. Подробнее об этом можно прочитать в [1]. Было решено строить информационную систему на базе существующих стандартов и средств, получивших наибольшее распространение в мире и которые стали стандартами де-факто. На период начала проектирования ИВС мы еще не могли точно утверждать, какие конкретно компоненты информационной системы будут взяты нами в качестве базовых. Предстояла работа по тщательной сравнительной оценке этих компонент, после чего надо было сделать окончательный выбор.

Помимо определения состава системного программного обеспечения, необходимого для развертывания информационной системы, предстояло определить, какая конкретная информация должна стать содержимым информационной системы. Необходимо было прояснить, какого рода информация требуется для погружения в базу данных информационной системы.

Кроме этого, необходимо было, чтобы наши ученые, преодолев определенный консерватизм мышления, поняли бы, что поставлять информацию о научной деятельности в базу данных информационной системы так же необходимо, как, например, писать статьи в научные журналы. А инженерный персонал, сопровождающий базу данных, должен был попытаться так построить работу по наполнению информационной системы, чтобы из разрозненных информационных элементов построить единое информационное поле, дающее полное представление об институте - ВЦ РАН.

Было решено строить систему итерационным путем и на первом этапе ограничиться достаточно простой структурой данных, т.е. построить простую информационную систему, исходя из конкретной информации ВЦ РАН. А дальше, уже в процессе эксплуатации системы, уточнить, какие новые информационные запросы и функции ИВС необходимо реализовать дополнительно.

Постепенно сложился план развертывания информационной составляющей ИВС. Кстати, именно такой план реализуем из-за ограниченности наших материальных средств. Мы не могли пойти на большие затраты (а информационные программные средства достаточно дороги), а потом убедиться в каких-то просчетах как с точки зрения функций системы, так и в плане построения базы данных.

Итак, план состоял в следующем.

1. В рамках развернутой сетевой модели (INTERNET) обеспечить для пользователей ВЦ РАН возможность работы с разнообразными информационными системами, с тем чтобы пользователь набрался опыта работы с такими системами, почувствовал вкус к новым способам получения информации в своей предметной области, задумался над вопросом представления собственной информации.

2. Изучить, как представлена и структурирована научная информация в крупных научных организациях, какие виды информационных запросов по синтаксису и семантике существуют.

3. Выбрать и начать освоение инструментальных средств, с помощью которых можно развертывать собственную базу данных.

4. Спроектировать простую структуру базы данных, продумать средства ее развертывания в рамках существующей файловой системы ИВС ВЦ РАН, определить элементарные средства защиты на базе предоставляемых средств ОС UNIX (Solaris).

5. Определить набор инструментальных средств и методику заполнения базы данных.

6. Определить круг проблем (научных, инженерных и организационных), возникающих в процессе формирования и использования базы данных.

7. Сформировать основной список информационных запросов к создаваемой ИВС.

8. Отработать методику поддержки и сопровождения созданной базы данных.

Таким образом, на первом этапе стояла задача построения пилотной модели простой ИВС без использования системы управления базой данных.

При выполнении сформулированного выше плана стало ясно, что достаточно хорошо можно реализовать ИВС на базе World Wide Web (WWW). Система WWW – широко распространенная прикладная оболочка INTERNET. Разбросанные по всему миру серверы WWW обеспечивают доступ к разнообразным информационным ресурсам. Достаточно простая по своей структуре база данных сервера WWW представляет собой гипертекст, фрагментами которого являются мультимедийные файлы (бинарные, текстовые, графические, аудио и видео) и ссылки на другие файлы. В работе [2, разд.3.5.2] мы уже излагали особенности сервера WWW.CCAS.RU.

Сервер WWW.CCAS.RU (основа существующей информационной составляющей ИВС ВЦ РАН) находится в состоянии постоянного сопровождения (модификации). Модифицируется сам сервер, его меню и база данных. Первые версии сервера носили экспериментальный характер, происходил процесс освоения, овладения техникой построения информационной структуры. На первых порах заполнение страниц базы данных сервера носило отчасти стихийный характер. Созданием информационных фрагментов занимались и системные программисты, и отдельные активные пользователи, пожелавшие "быть видимыми" в глобальной сети.

В 1997 г. была поставлена задача провести ревизию содержимого базы данных и организовать работу по ее заполнению таким образом, чтобы база данных WWW.CCAS.RU действительно отражала лицо ВЦ РАН. Предполагается, что после завершения работ первого этапа можно будет говорить о методике ведения информационной базы института. На регулярной основе будет функционировать группа специалистов, которая должна определять порядок представления материалов в базу данных, обеспечивать непосредственный ввод информации, сопровождение, осуществлять редактирование помещаемых статей, отвечать за дизайн web-сервера.

Несколько слов о наполнении базы данных сервера. Структура базы данных соответствует административной структуре ВЦ РАН. Каждый отдел имеет одну или несколько страниц, отражающих научную деятельность отдела. Первая страница – титульная, содержит основные данные об отделе, от этой страницы возможны ссылки к другим страницам отдела. Титульные страницы отделов имеют единый дизайн.

Имеется страница-каталог со ссылками на страницы отделов. Установлено, что каждая страница должна быть представлена на двух языках (русском и английском), обеспечен механизм переключения по языку.

К сожалению, в действующей версии базы данных требование двуязычия не выдерживается для всех страниц, но в перспективе этот недостаток предполагается устранить. Кроме страниц отделов имеются страницы секторов и персональные страницы отдельных сотрудников. В перспективе, когда база данных будет полностью сформирована, вся информация базы будет представлять единый гипертекст.

В базе имеется титульная страница института со ссылками на страницы отделов и страницы, соответствующие основным направлениям научной деятельности института. В 1997 г. особое внимание было уделено информации по направлениям научной деятельности. На примере работ по параллельным вычислениям была сделана попытка отработать единую методику погружения результатов исследований в информационную базу ИВС.

Защита информации в базе данных построена таким образом, что в основном только администратор может модифицировать информацию в базе. Но есть индивидуальные страницы, где сам пользователь имеет возможность их модифицировать. В этом случае должна быть определенная дисциплина взаимодействия пользователя и администратора. Администратор устанавливает ссылки на личные страницы пользователей, а пользователи не имеют право без согласования с администратором изменять имена головных файлов своих личных структур. Кроме того, хозяин индивидуальных страниц не в праве помещать произвольную информацию на свою страницу без санкции эксперта.

В роли экспертов выступают руководители подразделений. Для общеинститутской информации в роли экспертов кроме руководителей подразделений выступают директор института, его заместители и ученый секретарь. Кроме информации о научной деятельности отделов и сотрудников меню сервера содержит еще ряд интересных разделов, которые ведет и сопровождает системный администратор. Чтобы узнать о текущих изменениях в базе данных, имеется раздел "Новое", где хранятся упорядоченные по времени ссылки на вновь появившиеся страницы.

В 1998 г. в соответствии с планом работ по гранту РФФИ (96-07-89220, руководитель Г.М. Михайлов) первый этап развития информационной составляющей будет завершен.

Далее, уже на втором этапе предполагается выбрать и начать внедрение соответствующей системы управления базой данных с целью реализации более сложных информационных запросов.

ЛИТЕРАТУРА

1. Байкова И.В., Копытов М.А., Кулагин М.В., Рогов Ю.П., Михайлов Г.М., Привезенцев Ю.А. Распределенные информационно - вычислительные системы. Вып.1. Локальная сеть ВЦ РАН. М.: ВЦ РАН, 1995. 111с.
2. Байкова И.В., Копытов М.А., Кулагин М.В., Рогов Ю.П., Метелкин А.В., Михайлов Г.М., Плечов П.Ю. Распределенные информационно - вычислительные системы. Вып.2. Инфраструктура и базовые средства локальной сети ВЦ РАН. М.: ВЦ РАН, 1996. 96с.