УЧРЕЖДЕНИЕ
РОССИЙСКОЙ АКАДЕМИИ НАУК
Физический институт имени П.Н.Лебедева
ФИАН

ПРЕПРИНТ

Исаев Е.А., Думский Д.В., Зайцев А.Ю., Овчинников И.Л., Парунакян Д.А., Пугачев В.Д., Самодуров В. А., Беляцкий Ю.А., Китаева М.А., Герасимчук М.В.

Развитие сети науки и образования г. Пущино.

.

Москва 2010

Аннотация

В работе рассмотрена сигнальная топология городской вычислительной сети Пущинского научного центра, а также локальной вычислительной сети ПРАО АКЦ ФИАН.

Также рассмотрены различные аспекты реализации архитектуры опорной сети, ее конфигурация, а также конфигурация и функциональное назначение серверов сети. Представлена подробная информация по интеграции базовой сети в региональную и глобальную вычислительные сети и по доменам ПНЦ и ПРАО.

В справочных целях в тексте приведены наименования используемого аппаратного обеспечения.

Abstract

This paper describes the physical topology of the campus area network of the Pushchino science center of the Russian academy of science, including the LAN of the Pushchino radioastronomical observatory.

We also discuss different aspects of the backbone's implementation, its configuration, as well as configuration and functions of the networks' core servers. This paper also contains detailed information on integration of the networks in question into the rest of the Internet, traffic routes, and information on the domains belonging to the Puschino science center and the Pushchino astrophysical observatory.

For reference the text contains product names of the communication hardware used.

. Развитие сети науки и образования г.Пущино

Исаев Е.А.[(1)], Думский Д.В.[(1)], Зайцев А.Ю.[(2)], Овчинников И.Л.[(1)], Парунакян Д.А.[(2)], Пугачев В.Д.[(1)], Самодуров В. А. [(1)], Беляцкий Ю.А. [(1)], Китаева М.А. [(1)], Герасимчук М.В.[(2)]

(1) - Пущинская радиоастрономическая обсерватория астрокосмического центра ФИАН

(2) - Институт математических проблем биологии РАН

Структура ГВС ПНЦ РАН

Городская вычислительная сеть Пущинского научного центра (рис.1) построена по топологии "расширенная звезда". Центральный узел состоит из коммутатора AT-X900-24XS с дополнительным модулем AT-XEM12T, и сопряженного с ним сервера *octopus.psn.ru*, выполняющего основную работу по маршрутизации трафика. Сервер управляется операционной системой Debian GNU/Linux 5.0 - одним из наиболее надежных и удобных в конфигурации дистрибутивов.

К центральному коммутатору в ИМПБ многомодовыми оптоволоконными кабелями подключены институты ИТЭБ, ИБК, ИФХиБПП, ИБП, а также городская больница; внутри корпусов этих институтов сеть разведена силами штатного персонала. Подсети Института белка и ИМПБ подключены к центральному узлу кабелями UTP-5e, поскольку тот находится в одном здании с ними, и, как следствие, витая пара вполне подходит в качестве передающей среды для этих соединений.

В центральном вычислительном узле также расположен сервер, на котором в виртуальных машинах OpenVZ [1] работают *mail.psn.ru*, *www.psn.ru* и *ftp.psn.ru* (соответственно, почтовый, веб и файловый сервера). Такой подход позволяет достичь значительной экономии вычислительных ресурсов по сравнению с использованием нескольких физически раздельных серверов, в то же время, предоставляя лучшую безопасность и надежность работы по сравнению с эксплуатацией всех этих служб под одной операционной системой. ППНЦ, ПРАО и ИБФМ подключены к центральному коммутатору одномодовыми оптоволоконными кабелями, что позволяет достичь большей

пропускной способности на этих участках сети; структура подсети ПРАО подробнее описана в следующем разделе.

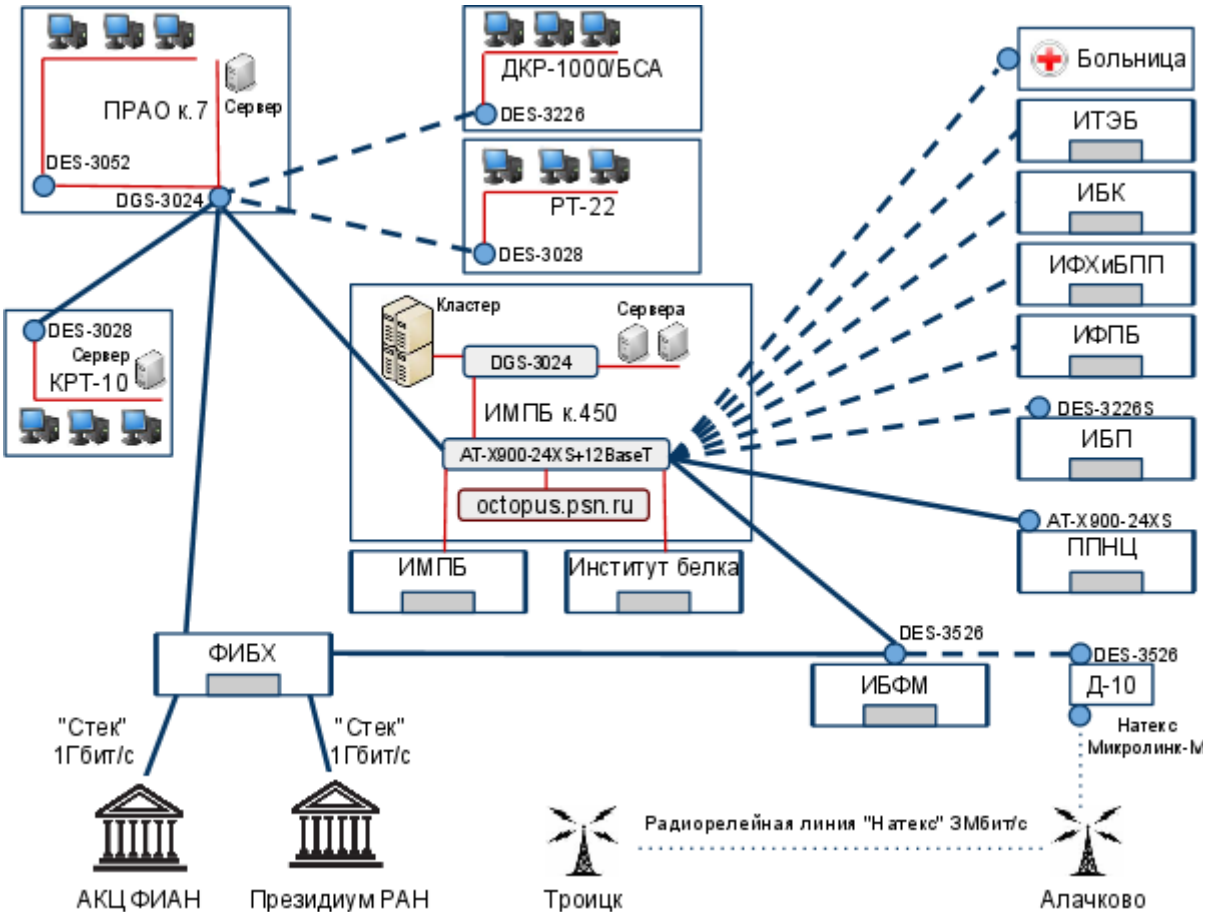

Рис.1. Физическая топология городской вычислительной сети Пущинского Научного Центра РАН, включая ЛВС ПРАО АКЦ ФИАН им.Лебедева (блок IP: 194.149.64.0/21, автономная система №9056). Тонкими сплошными линиями обозначены соединения по витой паре 5-й категории (UTP-5/5e), толстыми сплошными линиями - соединения по одномодовому оптоволоконному кабелю, толстыми штриховыми - многомодовому; пунктиром обозначены радиорелейные линии связи. Расшифровка сокращенных названий институтов приведена в конце статьи.

Связь с Интернетом осуществляется через расположенный в ФИБХ узел связи, который подключен к Президиуму РАН по принадлежащему ЗАО "Стек" оптоволоконному кабелю; для этих целей арендована ширина полосы 1Гбит/с. Канал связи с Президиумом проведен по маршруту ФИБХ-ИБФМ-ИМПБ, и из

центрального узла связи в ИМПБ трафик маршрутизируется по всей остальной части ГВС.

Связь ПРАО с АКЦ ФИАН в Москве осуществляется по оптоволоконному каналу пропускной способностью 1Гбит/с, также арендованному у ЗАО "Стек".

Связь с Троицким научным центром осуществляется с помощью, установленной на доме Д-10 радиорелейной системы "Натекс Микролинк-М" через ретранслятор в Алачково. К остальной ГВС узел связи Д-10 подключен многомодовым оптоволоконным кабелем, связывающим его с коммутатором в ИБФМ.

В состав ГВС входит кластер для высокопроизводительных вычислений, расположенный в ИМПБ, а также сервера и системы хранения данных. В сети ПНЦ широко используется технология виртуальных локальных сетей по стандарту IEEE 802.1Q [2].

Для маршрутизации через протокол BGP на сервер *octopus.psn.ru* используется ПО bgpd и zebra; системные почтовые сообщения с этого сервера и других машин доставляются администратору почтовым сервером (MTA) exim4.

Сервер *octopus.psn.ru* служит пунктом сбора и визуализации статистической информации по техническому состоянию сети; каждые 5мин опрашиваются входящие в состав сети коммутаторы на предмет функционирования. В конце статьи приведен такой график на состояние октября 2009г; разрешение шкалы - 15 часов.

Структура ЛВС ПРАО АКЦ ФИАН

Отдельно остановимся на устройстве локальной вычислительной сети (ЛВС) Пущинской радиоастрономической обсерватории Астрокосмического центра ФИАН им.Лебедева (ПРАО АКЦ ФИАН).

ПРАО АКЦ ФИАН представляет собой территориально распределенную организацию, где на территории общей площадью ~2 км^2 располагаются четыре радиотелескопа, лабораторный корпус и ряд других объектов, включенные в общую вычислительную сеть. Однако основное количество компьютеров сосредоточено в лабораторном корпусе и лишь небольшое число разбросано по нескольким удаленным друг от друга на 100-1000 м. корпусам (корпуса РТ-22, корпус ДКР-1000/БСА, служба времени и др.).

Локальная сеть ПРАО организована по сигнальной топологии "расширенная звезда", т.к. эта топология обеспечивает надежность сети в случае

отказа какой-либо сетевой станции или участка структурированной кабельной системы (СКС), что существенно удешевляет поиск и ликвидацию неисправностей, а также позволяет достичь высоких значений пропускной способности во многих типовых вариантах использования ЛВС. Отказоустойчивость сети особенно актуальна в свете функционирования в составе ПРАО автоматизированных наблюдательных комплексов [3].

На рис.1 видно, что ЛВС ПРАО разбита на три крупных узла, два из которых соединены с центральным коммутатором одномодовыми оптоволоконными кабелями. Причины выбора в пользу оптоволокна как среды для передачи сигнала между узлами очевидны: спецификация IEEE 802.3 [4], описывающая широко распространенный (и используемый в ПРАО) стандарт Ethernet на витой паре UTP-5e, гарантирует передачу данных на частоте 350МГц на расстоянии, не превышающем 100м; при превышении этой длины кабеля, соединяющего сетевые устройства, возникает ряд негативных эффектов, которые приводят к резкому падению пропускной способности канала по мере увеличения его длины; в то же время на оптоволоконной передающей среде Ethernet позволяет поддерживать высокую пропускную способность (в зависимости от версии и установленных устройств связи - от 100Мбит/с до 10Гбит/с) на расстояниях до 100 км.

В качестве центрального коммутатора используется коммутатор 2-го уровня марки D-Link DGS-3024 с 20 портами 10/100/1000BASE-T для подключений Gigabit Ethernet по медному кабелю, и 4 комбинированными портами 1000BASE-T /SFP для подключений к сети ПНЦ РАН и к оптическим каналам опорной сети. Каждый из узлов, в свою очередь, состоит из управляемого коммутатора 2-го уровня и подключенных к нему рабочих станций и абонентских устройств других типов; по корпусам разведена сеть на неэкранированной витой паре 5е категории; пиковая пропускная способность этих коммутаторов составляет 100Мбит/с. Также к центральному коммутатору подключены каналами связи типа Gigabit Ethernet сервера обсерватории.

Из рис.1 видно, что выход ЛВС ПРАО в Интернет обеспечивается с помощью оптоволоконного соединения между ИМПБ РАН и центральным маршрутизатором ПРАО (1 Гбит/с по SFP).

Администрирование и мониторинг серверов осуществляется в штатном режиме через удаленный вход в систему по протоколу SSH v.2. Использование первой версии этого протокола не допускается в связи с обнаруженными в ней уязвимостями [5]. На случай аварийных ситуаций машинный зал оборудован

консолью (монитор и клавиатура), к которой через KVM-переключатель подсоединены все сервера.

В качестве основной операционной системы для серверов используется свободно распространяемая операционная система с открытым исходным кодом Debian GNU/Linux 5.0, известная своей надежностью, стабильностью работы, и сертифицированная по системе CGL (Carrier-Grade Linux).

Для организации удобной рабочей среды ЛВС ПРАО предоставляет ряд важных сетевых сервисов, среди них следует особо выделить следующие:

- DHCP-сервер и DNS-сервер;
- HTTP-сервер;
- сервер баз данных;
- сервер электронной почты;
- сервер точного времени NTP.

Эксплуатация большого количества служб, необходимых для эффективной работы сети ПРАО, может отрицательно сказываться на уровне загрузки центрального сервера, а также неизбежно ведет к снижению уровня безопасности; в связи с этим в ЛВС ПРАО службы, доступные извне ЛВС (к примеру, веб-сервер, DNS и электронная почта) физически и логически отделены от внутренней сети, и выделены в так называемую демилитаризованную зону. Входящих трафик из Интернета может достичь и повлиять только на службы находящиеся в ДМЗ, и не пропускается во внутреннюю сеть.

Службы физически распределены по серверам следующим образом:

- Сервер №1: маршрутизатор, брандмауэр, DHCP-сервер;
- Сервер №2 (расположен в КРТ-10): HTTP-сервер и почтовый сервер; каждый из этих серверов работает под управлением виртуальной машины openvz.
- Сервер №3: сервер баз данных;
- Сервер №4: DNS-сервер, сервер точного времени NTP.

Назначением DHCP-сервера является автоматизация процесса выдачи абонентским устройствам сети IP адресов из доступного пула по заранее заданным администратором ЛВС правилам. Это позволяет существенно сэкономить время, необходимое для подключения к сети новых персональных компьютеров, а также снизить вероятность возникновения конфликтов IP-адресов, когда под одним и тем же IP-адресом пытаются функционировать два и более компьютера одновременно.

Маршрутизатор обеспечивает доступ пользователям локальной сети ПРАО к ресурсам Интернет и сети ПНЦ РАН, а также защиту находящихся в ЛВС ПРАО машин от несанкционированного доступа и сетевых атак с помощью брандмауэра. Маршрутизация и управление трафиком осуществляется с помощью пакета программ iproute2 (*http://lartc.org/howto*); политики брандмауэра реализованы с помощью модуля ядра Linux iptables (*http://www.netfilter.org/projects/iptables*).

В качестве MTA (mail transfer agent: программа, взаимодействующая с другими почтовыми серверами в сети Интернет по протоколу SMTP для обмена электронной почтой и ее маршрутизации) в ЛВС ПРАО используется созданный в Кембриджском университете почтовый сервер Exim4; в качестве MDA (mail delivery agent - программа доставки почты) используется программа высокого уровня защищенности Dovecot, что позволяет пользователям скачивать свою почту по протоколам POP3 и IMAP. В качестве альтернативного способа работы с почтовой системой предусмотрен веб-интерфейс SquirrelMail. Для фильтрации спама и вирусов используются программы spamassasin, greylist и clamav.

Роль HTTP-сервера выполняет Apache httpd: быстрый в эксплуатации, гибкий в настройке и обладающий богатым спектром функций веб-сервер, в течение многих лет поддерживаемый Apache Software Foundation. Веб-сервер Apache поддерживает переконфигурацию "на лету", без полной перезагрузки процесса, а также обладает возможностью расширения своих функций путем подключения библиотек-модулей. На эксплуатирующийся веб-сервер возложена задача обслуживания веб-сайта ПРАО *www.prao.ru* и сайтов некоторых отделов обсерватории; также, благодаря его наличию, возможно осуществлять доступ к хранящимся в реляционной СУБД (системе управления базами данных) PostgreSQL астрономическим базам данных и результатам наблюдений через удобный веб-интерфейс. В целях дальнейшей оптимизации использования вычислительных ресурсов планируется переход на веб-сервер Lighttpd.

Для трансляции IP-адресов в доменные имена в обсерватории используется DNS-сервер, построенный на основе программного пакета djbdns, включающего в себя набор утилит для обслуживания и разрешения DNS зон.

Сервер времени используется для автоматической синхронизации времени серверов и рабочих станций по протоколу NTP (Network Time Protocol, RFC 1305 [6]). Данный протокол синхронизирует время в географически распределенных сетях по порту 123 протокола UDP (User Datagram Protocol). Протокол NTP поддерживает множественные избыточные источники данных о

времени, что обеспечивает его постоянную синхронизацию, и позволяет установить время на локальной машине с точностью до одной миллисекунды.

В главном корпусе ПРАО установлена и функционирует система цифровой коммутации (СЦК) «ЭЛКОМ», которая представляет собой цифровую автоматическую телефонную станцию (АТС). Система построена в виде отдельных модулей, связанных между собой и обеспечивает как внутреннюю связь в корпусах обсерватории, так и выход на городские телефонные линии. Телефонная связь с корпусом, обслуживающим радиотелескоп РТ-22, осуществляется по существующим оптоволоконным линиям связи с помощью, так называемых абонентских выносов. Для передачи голосового трафика внутри сети по протоколу SIP (*Session Initiation Protocol* - протокол установления сеанса)[7] используются голосовые шлюзы Nateks VC-130-1 и VC-130-2. Необходимость такого решения была вызвана частичным выходом из строя старых телефонных линий.

Интеграция ГВС ПНЦ РАН и ЛВС ПРАО АКЦ ФИАН в глобальную вычислительную сеть

Данный раздел посвящен описанию BGP-связей ПНЦ и смежных сетей, а также общей структуре сегмента Интернета, к которому они подключены. Информация представлена в виде таблиц и автоматически построенных графов (таблица 1, рис. 2-4). BGP (англ. Border Gateway Protocol, протокол граничного шлюза) — основной протокол динамической маршрутизации в Интернете. BGP, в отличие от других протоколов динамической маршрутизации, предназначен для обмена информацией о маршрутах не между отдельными маршрутизаторами, а между целыми автономными системами, и поэтому, помимо информации о маршрутах в сети, переносит также информацию о маршрутах на автономные системы. BGP не использует технические метрики, а осуществляет выбор наилучшего маршрута исходя из правил, принятых в сети. BGP поддерживает бесклассовую адресацию и использует суммирование маршрутов для уменьшения таблиц маршрутизации. С 1994 года действует четвёртая версия протокола, все предыдущие версии являются устаревшими. BGP является протоколом прикладного уровня и функционирует поверх протокола транспортного уровня TCP (порт 179). BGP, наряду с DNS, является одним из главных механизмов, обеспечивающих функционирование Интернета. Автономная система (AS) в Интернете — это система IP-сетей и

маршрутизаторов, управляемых одним или несколькими операторами, имеющими единую политику маршрутизации с Интернетом. См. RFC 1930 для дополнительной информации по данному определению. Уникальный номер AS (или ASN) присваивается каждой AS для использования в BGP-маршрутизации. Именно ASN однозначно идентифицирует каждую сеть в Интернете.

Пиринг	**Ипорт**	**Экспорт**
AS2643 IHEP-SU AS	AS2643	AS9056 AS13161
AS2683 RADIO-MSU	ANY	AS9056 AS2643
AS3058 RAS-AS Russian Academy of Sciences	ANY ANY	AS9056 AS9056
AS5568 RBNet Russian Backbone Network	AS-RBNET	AS9056
AS21453 FLEX-AS Autonomus System for Wireless Network in Moscow region Russia	ANY	AS9056
AS28700 INFORMTELECOM-AS Informtelecom XXI LTD.		AS-INFORMTELECOM

Таблица 1. Пиринговые связи системы AS9056

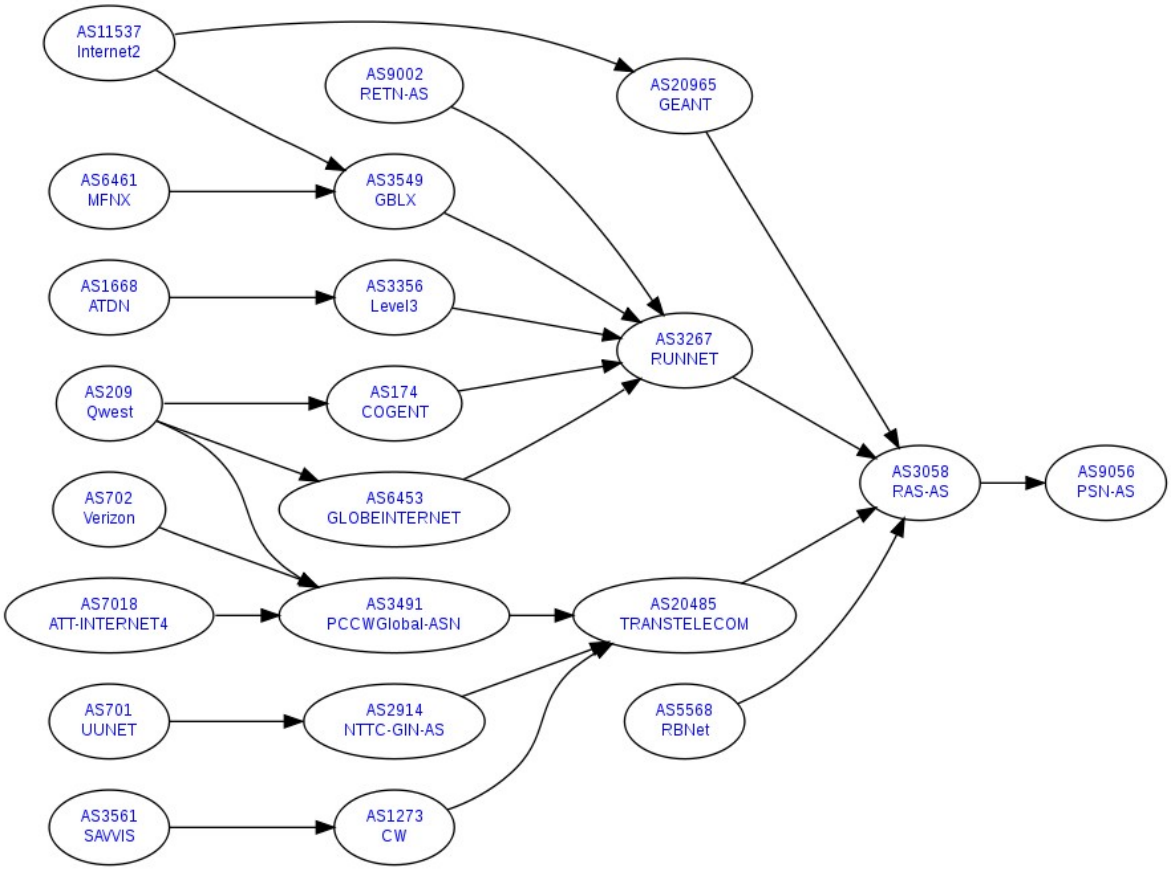

Рис.2 Визуализация графа маршрутизации внешнего трафика AS9056 - автономной системы ПНЦ РАН (подсеть 194.149.64.0/21).

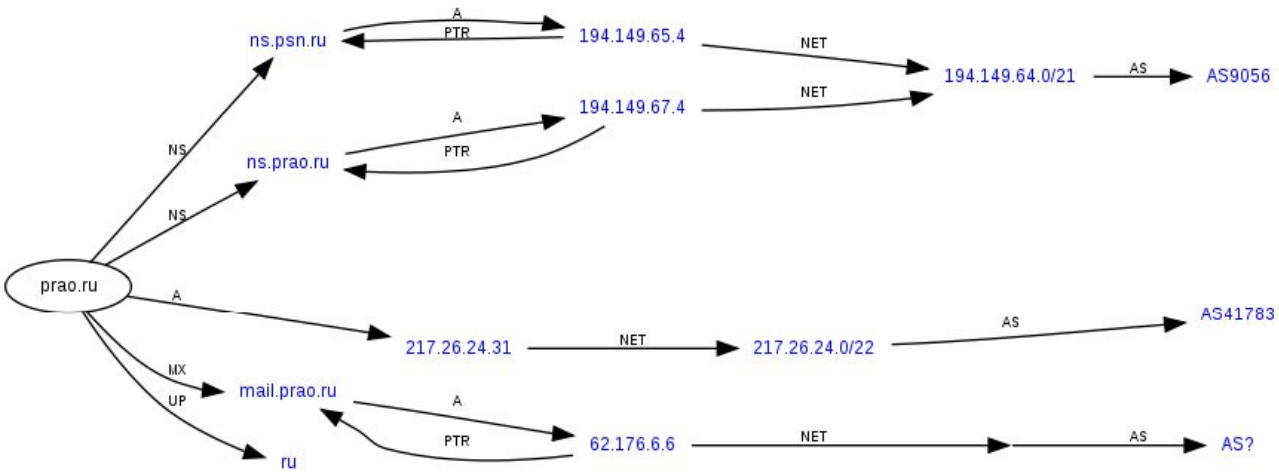

Рис.3 Схема DNS-поиска доменных имен домена prao.ru

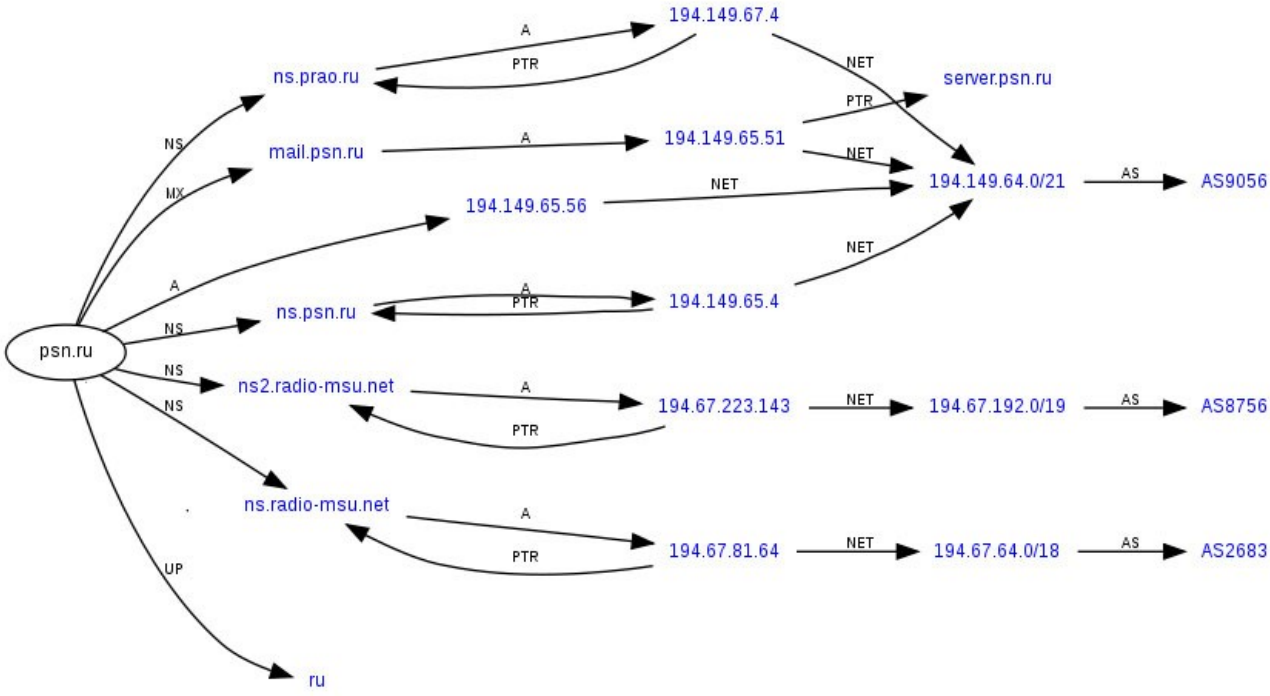

Рис.4 Схема DNS-поиска доменных имен домена psn.ru

Список литературы

[1] K Kolyshkin. Virtualization in Linux. *http://mirrors.unbornmedia.com/openvz/doc/openvz-intro.pdf*, 2006.

[2] IEEE Std. 802.1Q-2005, Virtual Bridged Local Area Networks, ISBN 0-7381-3662-X.

[3] В.В. Китаев. Распределенная система обработки и сбора данных ПРАО АКЦ ФИАН. II Базовая локальная вычислительная сеть; препринт ФИАН №52, Москва, 1997.
[4] IEEE 802.3 LAN/MAN CSMA/CD (Ethernet) Access Method, *http://standards.ieee.org/getieee802/802.3.html*
[5] CORE SDI S.A. SSH Insertion attack.1998. US CERT Vulnerability Note VU#13877
[6] David L. Mills, University of Delaware. Network Time Protocol (Version 3). Specification, Implementation and Analysis. IETF RFC1305. March 1992.
[7] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler. SIP: Session Initiation Protocol. IETF RFC 3261. June 2002.

Список аббревиатур

ПРАО: Пущинская радиоастрономическая обсерватория
АКЦ: астрокосмический центр
ФИАН: физический институт академии наук
ПНЦ: Пущинский научный центр ИТЭБ: институт теоретической и экспериментальной биофизики
ИБК: институт биофизики клетки
ИФХиБПП: институт физико-химическеских и биологических проблем почвоведения
ИФПБ: институт фундаментальных проблем биологии
ИБП: институт биологического приборостроения
ИМПБ: институт математических проблем биологии
ФИБХ: филиал института биоорганической химии
ИБ: институт белка
ИБФМ: институт биохимии и физиологии микроорганизмов
ППНЦ: президиум Пущинского научного центра
ЛВС: локальная вычислительная сеть
ГВС: городская вычислительная сеть