10. Технологии одноранговых сетей

10.1 Основы технологии одноранговых сетей

Технология одноранговых сетей (или P2P-сетей от анг. peer-to-peer – равный-к-равному) обеспечивает формирование РВС на базе принципа децентрализации [67], где разделение вычислительных ресурсов и сервисов производится напрямую посредством прямого взаимодействия между участниками сети друг с другом, без участия центрального сервера [52]. Одноранговые вычислительные сети, в какой-то степени, являются противоположностью клиент-серверным архитектурам (таким как CORBA, RMI). Можно сказать, что основным принципом P2P-сетей является воплощение идеи коммунизма – «Каждый – по способностям, каждому – по потребностям!».

10.1.1 Сравнение P2P и клиент-серверной технологий

В отличие от традиционной клиент-серверной архитектуры в P2P-сетях каждый узел, входящий в вычислительную сеть, может являться как клиентом, так и сервером, предоставляя или используя ресурсы сети. На рис. 55 представлены связи в сетях с P2P и с централизованной архитектурой [49].

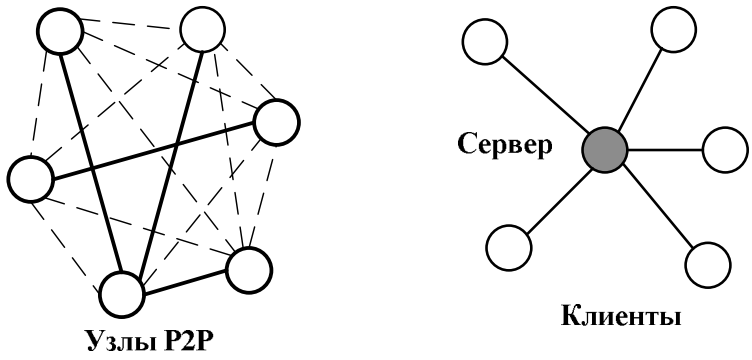

Рис. 55. Сравнение вида связей P2P и централизованной (клиент-серверной) архитектур

Можно выделить следующие проблемы клиент-серверной архитектуры, связанные с наличием централизованного сервера, обеспечивающего обработку запросов от множества клиентов:

– *Проблемы масштабируемости*. При увеличении количества клиентов растут требования к мощности сервера и пропускной способности канала. Единственным вариантом решения данной задачи является наращивание пропускной способности канала до сервера и использование более высокопроизводительных решений для аппаратной платформы сервера;

– *Зависимость*. Стабильная работа всех клиентов зависит от загруженности и функционирования одного сервера. При выходе из строя или отключении сервера, клиенты не смогут выполнять функциональные обязанности.

Этим проблемам можно противопоставить следующие преимущества P2P:

– отсутствие зависимости от централизованных сервисов и ресурсов;
– система может пережить серьезное изменение в структуре сети;
– высокая масштабируемость модели одноранговых вычислений.

На рис. 56 представлена диаграмма, позволяющая сравнить принципы этих архитектур. На основе этой диаграммы можно сделать предположение, что нет четкой границы между архитектурой P2P и клиент-серверной архитектурой. Обе модели могут быть построены с реализацией в различной степени каких-либо характеристик (например, управляемость), функциональности, структур (например, иерархии и сети) и др. Они могут выполняться на различных платформах (Интернет, Интранет и др.), и обе могут служить в качестве базы для приложений. Таким образом, понятие P2P чрезвычайно тесно переплетено с другими существующими технологиями.

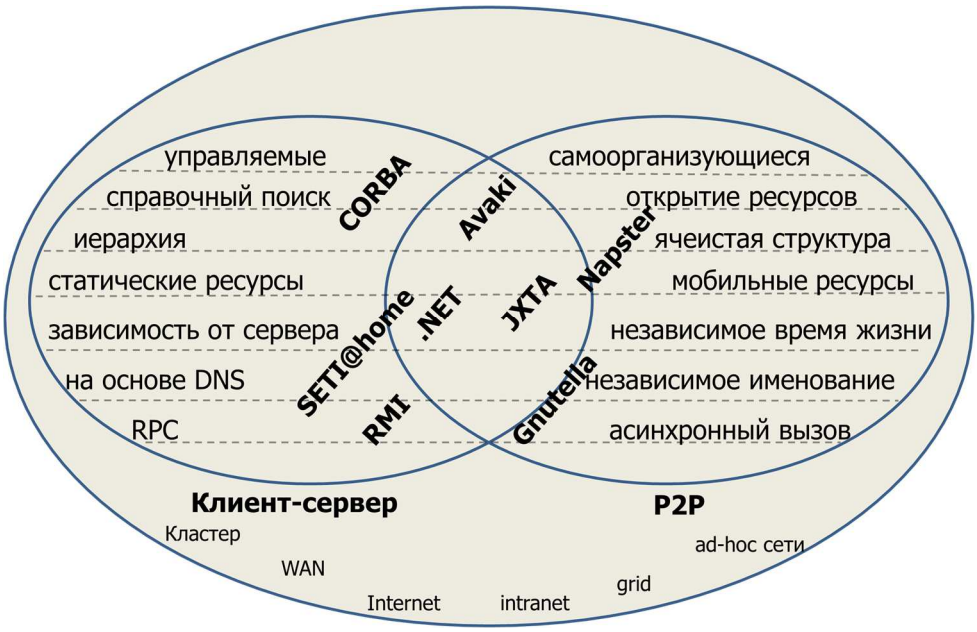

Рис. 56. Сравнение P2P и централизованной (клиент-серверной) архитектур

10.1.2 Задачи P2P сетей

Можно выделить следующие основные задачи, которые с легкостью решают P2P сети:

1. *Уменьшение/распределение затрат*. Серверы централизованных систем, которые обслуживают большое количество клиентов, обычно несут на себе основной объем затрат ресурсов (денежных, вычислительных и др.) на

поддержание вычислительной системы. P2P архитектура может помочь распределить эти затраты между узлами сети. Так как узлы, как правило, автономны, важно, чтобы затраты были распределены справедливо.

2. *Объединение ресурсов.* Каждый узел в P2P-системе обладает определенными ресурсами (вычислительные мощности, объем памяти). Приложения, которым необходимо большое количество ресурсов, например ресурсозатратные задачи моделирования или распределенные файловые системы, используют возможность объединения ресурсов всей сети для решения своей задачи. При этом важны как объем дискового пространства для хранения данных, так и пропускная способность сети.
3. *Повышенная масштабируемость.* Поскольку в сетях P2P отсутствует сильный центральный механизм, важной задачей является повышение масштабируемости и надежности системы. Масштабируемость определяет количество систем, которые могут быть достигнуты из одного узла, сколько систем могут функционировать одновременно, сколько пользователей может пользоваться сетью, сколько памяти может быть использовано. Надежность сети определяется такими параметрами как количество сбоев в работе сети, отношение времени простоя к общему времени работы, доступностью ресурсов и т.д. Таким образом, основной проблемой становится разработка новых алгоритмов обнаружения ресурсов, на которых базируются новые P2P платформы.
4. *Анонимность.* Бывает, пользователь не желает, чтобы другие пользователи или поставщики услуг знали о его нахождении в сети. При использовании центрального сервера трудно обеспечить анонимность, так как серверу, как правило, необходимо идентифицировать клиента, по крайней мере через интернет адрес. При использования P2P-сети пользователи могут избежать предоставления любой информацию о себе. FreeNet является ярким примером того, как механизмы анонимности могут быть встроены в P2P-приложения. В системе FreeNet используется схема переадресации сообщений при которой невозможно отследить первого отправителя. Также степень анонимности увеличивается за счет использования вероятностных алгоритмов.

10.1.3 Основные элементы P2P сетей

Пир (Peer) – это фундаментальный составляющий блок любой одноранговой сети:

– каждый пир имеет уникальный идентификатор;

– каждый пир принадлежит одной или нескольким группам;
– каждый пир может взаимодействовать с другими пирами, как в своей так и в других группах.

Можно выделить следующие виды пиров:

– *Простой пир:* обеспечивает работу конечного пользователя, предоставляя ему сервисы других пиров и обеспечивая предоставление ресурсов пользовательского компьютера другим участникам сети;
– *Роутер:* обеспечивает механизм взаимодействия между пирами, отделенными от сети брандмауэрами или NAT-системами.

Группа пиров – это набор пиров, сформированный для решения общей задачи или достижения общей цели. Группы пиров могут предоставлять членам своей группы такие наборы сервисов, которые недоступны пирам, входящим в другие группы.

Сервисы – это функциональные возможности, которые может привлекать отдельный пир для полноценной работы с удаленными пирами. В качестве примера сервисов, которые может предоставлять отдельный пир можно указать сервисы передачи файлов, предоставления информации о статусе, проведения вычислений и др. *Сервисы пира* – это такие сервисы, которые может предоставить конкретный узел P2P. Каждый узел в сети P2P предоставляет определенные функциональные возможности, которыми могут воспользоваться другие узлы. Эти возможности зависят от конкретного узла и доступны только тогда, когда узел подключен к сети. Как только узел отключается, его сервисы становятся недоступны. *Сервисы группы* – это функциональные возможности, предоставляемые группой входящим в нее узлам. Возможности могут предоставляться несколькими узлами в группе, для обеспечения избыточного доступа к этим возможностям. Как только к группе подключается узел, обеспечивающий необходимый сервис, он становиться доступной для всей группы.

P2P — это не только сети, но еще и *сетевой протокол*, обеспечивающий возможность создания и функционирования сети равноправных узлов, их взаимодействия. Множество узлов, объединенных в единую систему и взаимодействующих в соответствии с протоколом P2P, образуют пиринговую сеть. P2P относятся к прикладному уровню сетевых протоколов и являются *наложенной сетью*, использующей существующие транспортные протоколы стека TCP/IP – TCP или UDP. *Протоколы* сети P2P обеспечивают:

– поиск узлов в сети;
– получение списка служб отдельного узла;

– получение информации о статусе узла;
– использование службы на отдельном узле;
– создание, объединение и выход из групп;
– создание соединений с узлами;
– маршрутизацию сообщений другим узлам.

Одну из удачных попыток стандартизации протоколов P2P предприняла компания Sun Microsystems в рамках проекта JXTA. *Платформа JXTA* позиционируется как базовая платформа для организации P2P сетей на основе гетерогенных вычислительных сетей.

10.2 Алгоритмы работы P2P сетей

10.2.1 Структура P2P сети

Структура P2P сети определяет принципы поиска новых узлов и замены узлов вышедших из состава сети новыми. Можно выделить два основных типа P2P сетей: централизованные и децентрализованные [46].

Централизованная структура P2P сети подразумевает наличие выделенного индексного сервера (*трекера*) собирающего информацию об узлах, входящих в P2P-сеть и обеспечивающего поиск и предоставление необходимых сервисов одних узлов другим. Первой P2P сетью с централизованной структурой была сеть Napster, центральный узел которой отвечал за хранение идентификаторов всех узлов в сети и списков файлов, доступных на каждом из узлов. Еще одним примером сети с централизованной структурой является сеть BitTorrent. Центральным узлом данной сети является *трекер* – сервер, содержащий информацию о списке узлов, подключенных к сети, и сервисах, предоставляемых каждым узлом (например, список файлов, доступных для загрузки с данного узла). Для получения необходимого файла, узел посылает трекеру запрос, содержащий уникальный идентификатор необходимого файла. На данный запрос трекер возвращает список узлов, на которых доступен требуемый файл. Естественно, степень централизованности системы BitTorrent значительно меньше, чем была у системы Napster, т.к. BitTorrent позволяет работать сразу с большим количеством трекеров, в то время как Napster предполагал наличие только одного центрального сервера.

Децентрализованная структура P2P сети предполагает отсутствие выделенного сервера. Поиск и предоставление сервисов производится путем процедуры пошагового поиска, в которой могут участвовать все узлы, входящие в сеть. Типичным примером одноранговой сети с децентрализованной структу-

рой является система Gnutella. В данной сети, обнаружение и подключение к узлам сети происходит посредством процедуры случайного обхода. Каждый узел содержит таблицу соседей, содержащую IP адрес и порт известного узла Gnutella. При запуске новый узел Gnutella переходит в режим начальной загрузки, в котором посредством одного из доступных источников (список узлов на одном из узлов интернет; внутренний предустановленный список узлов и др.) формирует начальный список соседей. После чего, соседям высылается сообщение обнаружения, пересылаемое далее по цепочке систем. Таким образом обеспечивается обнаружение ресурсов, предоставляемых всеми узлами подключенными к сети.

10.2.2 Алгоритмы работы P2P сетей

Поскольку сегодня в перенасыщенном информацией мире задача полноты поиска отводится на второй план, то главная задача поиска в пиринговых сетях сводится к быстрому и эффективному нахождению наиболее релевантных откликов на запрос, передаваемый от узла всей сети. В частности, актуальна задача — уменьшение сетевого трафика, порождаемого запросом (например, пересылки запроса по многочисленным узлам), и в то же время получение наилучших характеристик выдаваемых документов, т.е. наиболее качественного результата. Для решения данной задачи применяют несколько подходов.

Централизованный индекс: узлы публикуют информацию о своих сервисах в центральном индексе. При подключении к сети, пир формирует список документов и сервисов, доступных на узле. Данный список передается на центральный сервер, хранящий полную информацию о текущем состоянии вычислительной сети. Когда любой узел P2P-сети хочет получить информацию о том, какие ресурсы доступны, он отправляет поисковый запрос на центральный сервер. В ответ на этот запрос, центральный сервер выдает список доступных ресурсов и адреса узлов, на которых они располагаются. Недостатком является малая масштабируемость сети (в связи с возрастающей нагрузкой на центральный сервер) и высокая зависимость от центрального сервера. Данные недостатки решаются посредством масштабирования центрального сервера (например, использование нескольких независимых центральных серверов). *Пример*: Napster, BitTorrent, eDonkey.

Широковещательные запросы (Breadth Search): процесс поиска информации производится посредством отправки поискового запроса всем подключенным узлам, известным отправителю запроса. Данный запрос транслируется всем дальнейшим узлам, пока не получен ответ или не достигнут предел коли-

чества пересылок. Недостатками данного метода является большая нагрузка на сеть, генерируемая поисковыми запросами, а также негарантируемая достижимость результата. Однако есть метод, позволяющий избежать перегрузки всей сети сообщениями. Он заключается в приписывании каждому запросу параметра времени жизни (time-to-live, TTL). Параметр TTL определяет максимальное число переходов, по которому можно пересылать запрос. *Пример*: Gnutella

Маршрутизация документов. Данный метод обеспечивает поиск документов без участия центрального индекса, средствами самой вычислительной сети. В основе данного метода лежит принцип присвоения уникальных идентификаторов каждому узлу вычислительной сети, а также каждому ресурсу (документу, сервису и др.), который данная сеть может предоставлять.

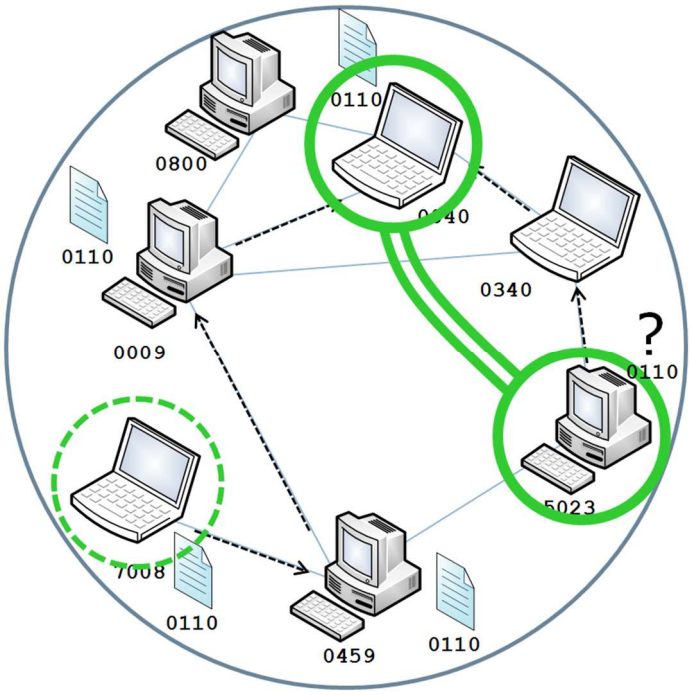

Рис. 57. Трассировка и поиск ресурса посредством алгоритма маршрутизации документов

На рисунке 57 представлен пример работы алгоритма маршрутизации документа. Когда какой-либо пир (на рисунке – узел 7008) производит публикацию ресурса в сети, каким-либо образом вычисляется идентификатор данного ресурса (на рисунке – документ 0110). При этом, идентификаторы узлов сети и ресурсов имеют единую область возможных значений. Далее, копия данного ресурса (или ссылка на него) трассируется к узлу, который имеет наиболее похожий идентификатор (на рисунке показана трасса документа 0110: узлы 7008 – 0459 – 0009 – 0040). Данная процедура производится следующим образом:

1. Производится сравнение идентификатора ресурса с идентификаторами всех соседних узлов.

2. Если идентификатор *текущего* узла ближе всего (по некоторой метрике) к идентификатору документа, то процесс трассировки завершается.
3. Если один из идентификаторов соседних узлов ближе к идентификатору документа, чем идентификатор текущего узла, то ссылка на данный ресурс *копируется* на узел с более близким идентификатором и процесс повторяется с п. 1.

В результате данной процедуры ссылка на документ отправляется на вычислительный узел, идентификатор которого больше всего соответствует идентификатору документа, среди соседей начального узла (узел 0040 на рисунке).

Поиск ресурса производится по аналогичному алгоритму, но вместо копирования документа происходит трансляция запроса (в запросе содержится идентификатор запрашиваемого ресурса, например 0110 на рисунке 57) от узла, инициировавшего запрос (на рисунке – узел 5203) к узлу, идентификатор которого ближе всего соответствует идентификатору запрашиваемого документа. Соответственно, в процессе трансляции данного запроса должен появиться такой узел, который хранит информацию об интересуемом документе, если данный документ находится в соответствующей части сети.

К недостаткам такого подхода можно отнести необходимость знания точного идентификатора документа, который необходимо найти в сети (невозможность поиска по отдельным атрибутам документа), а также возможность образования «островов», затрудняющих алгоритм поиска. *Примеры*: FreeNet, протокол DHT.

Отдельно стоит отметить *алгоритм обхода брандмауэров и NAT*, применяющийся при работе с конечными узлами, установка внешнего соединения с которыми затруднена или невозможна. Для обеспечения обмена информацией, необходимо использовать возможности узла-роутера для трансляции сообщений во внешнюю сеть и получения информации из внешней сети. Узел-роутер буферизует всю информацию, предназначенную конечному узлу до момента, пока не получит от него запрос на загрузку. Как только получен запрос на загрузку, узел-роутер отвечает, выгружая всю накопившуюся информацию на конечный узел.

10.3 Применение технологий P2P

Наибольшее распространение одноранговых сетей наблюдается в системах, обрабатывающих большие объемы данных и обеспечивающих индивиду-

альный обмен информацией между пользователями. В настоящий момент, технологии P2P наиболее ярко представлены в 3-х направлениях:

- *Распределенные вычисления*: разбиение общей задачи на большое число независимых в обработке подзадач (проекты на платформе BOINC [9]);
- *Файлообменные сети*: P2P выступают альтернативой FTP-архивам, которые утрачивают перспективу ввиду значительных информационных перегрузок (однако требуются эффективные механизмы поиска) (Gnutella [33], eDonkey, BitTorrent [8]);
- *Приложения для совместной работы*: требуют обеспечения прозрачных механизмов для совместной работы. (Skype [36,59], Groove [17]).

10.3.1 Распределенные вычисления

В основном, к данному типу проектов относят системы типа проекта SETI@home (распределенный поиск внеземных цивилизаций), который продемонстрировал огромный вычислительный потенциал для распараллеливаемых задач. В настоящий момент в нем принимают участие свыше трех миллионов пользователей на бесплатной основе. Данная система основана на платформе *BOINC*.

BOINC (англ. Berkeley Open Infrastructure for Network Computing — открытая программная платформа Беркли для распределённых вычислений) — некоммерческое межплатформенное ПО для организации распределённых вычислений. Система состоит из двух основных частей:

- сервер BOINC – это набор PHP-сценариев для организации и управления проектом: регистрация участников, распределение заданий, получение результатов;
- клиент BOINC – это пользовательское приложение, позволяющее участвовать в одном или нескольких проектах. Обычно представляет собой хранитель экрана, который производит вычисления в моменты простоя компьютера.

Наиболее популярные проекты, реализованные на основе BOINC:

- SETI@home — анализ радиосигналов с радиотелескопа Аресибо для поиска инопланетных цивилизаций.
- Einstein@Home — проверка гипотезы Альберта Эйнштейна о гравитационных волнах с помощью анализа гравитационных полей пульсаров или нейтронных звёзд.

- Climate Prediction — построение модели климата Земли для предсказания его изменений на 50 лет вперёд.
- World Community Grid — Различные проекты. Организатор — IBM.
- Malaria Control Project — Контроль распространения Малярии в Африке (AFRICA@home).
- Predictor@home — моделирование 3-хмерной структуры белка из последовательностей аминокислот.
- LHC@home — расчёты для ускорителя заряженных частиц в CERN (Centre Europeen de Recherche Nucleaire).

10.3.2 Файлообменные сети

По статистическим данным на конец 2006 года объем трафика, генерируемого файлообменными сетями на базе P2P-сетей, составил более 70% всего сетевого трафика. На сегодняшний день существует большое число P2P-сетей, ориентированных на обмен файлами между пользователями. Они могут развиваться и функционировать как в глобальном сетевом пространстве, так и в отдельных подсетях.

Самым ярким примером таких сетей, является система BitTorrent. Протокол BitTorrent был разработан в 2001 Брэмом Коэном. В соответствии с протоколом BitTorrent файлы передаются не целиком, а частями, причем каждый клиент, закачивая эти части, в это же время отдает их другим клиентам, что снижает нагрузку и зависимость от каждого клиента-источника и обеспечивает избыточность данных.

Если узел хочет опубликовать файл или набор файлов, то программа-клиент BitTorrent сети разделяет передаваемые файлы на части и создает *файл метаданных* (идентификатор раздачи), который содержит следующую информацию:

- URL трекера;
- Общая информация о файлах (имя, длина и пр.);
- Хеш-суммы SHA1 сегментов раздаваемых файлов;
- Passkey пользователя – ключ, который однозначно определяет пользователя загрузившего файл;
- Хеш-суммы файлов целиком (не обязательно);
- Альтернативные источники – адреса альтернативных трекеров, на которых можно найти информацию по данному файлу (не обязательно).

Алгоритм загрузки документа производится следующим образом:

- клиент подключается к трекеру по URL из файла метаданных;
- сообщает хеш-идентификатор требуемого файла;
- получает адреса пиров скачивающих и раздающих данный файл;
- клиенты соединяются между собой и обмениваются информацией без участия трекера.

В последнее время стала распространяться альтернативная технология поиска и загрузки документов на основе «магнитных ссылок» (magnet links) и подхода распределенных хеш-таблиц (Distributed Hash Table — DHT) по сути дела представляющих собой реализацию алгоритма маршрутизации документов, описанную ранее. Причина возникновения этой технологии – дальнейшее развитие деперсонализации и попытка торрент-трекеров защититься от юридического преследования правообладателей. Торрент-файл для такой раздачи создаётся без адреса трекера и клиенты находят друг друга через распределенные хеш-таблицы.

DHT – это система распределенного хранения данных о скачиваемых файлах. Все клиенты, подключенные к DHT-сети и сами становятся «узлами», чем-то вроде мини-трекеров. Каждый узел имеет уникальный идентификатор – «node ID». Все узлы хранят информацию об узлах, «близких к ним», кроме того узел должен хранить информацию о пирах в раздачах, чей хеш напоминает «node ID».

При этом торренты представляют собой Magnet-ссылки, которые в основном идентифицируют файлы не по их расположению или имени, а по содержанию, точнее — по хеш-коду.

Одно из преимуществ magnet-ссылок — их открытость и независимость от платформы: ссылка может быть использована для загрузки файла при помощи разнообразных приложений на практически всех операционных системах. Благодаря тому, что magnet-ссылка представляет собой короткую строку текста, пользователи могут использовать обычные операции копирования-вставки и отправить ее по электронной почте или программе мгновенного обмена сообщениями.

10.3.3 Приложения для совместной работы

Приложения для совместной работы, это такие приложения, которые обеспечивают возможность общения, совместной работы и т. п. различных географически-распределенных пользователей.

- Jabber: мгновенный обмен сообщениями

- Skype: голосовое общение, видеоконференции, приложения для совместной работы
- Groove: система для совместной работы

Самой популярной на сегодняшний день службой Интернет-телефонии является Skype, созданная в 2003 году шведом Никласом Зеннстромом и датчанином Янусом Фриисом, авторами известной пиринговой сети KaZaA. В настоящее время Skype принадлежит корпорации Microsoft, которая приобрела ее за $8,5 млрд. в мае 2011 года.

Структура Skype-сети состоит из обычных узлов (normal/ordinal node/host/nest), обычно обозначаемых аббревиатурой SC (Skype Client), и super-узлов (super node/host/nest), которым соответствует аббревиатура SN. Любой узел, имеющий публичный IP-адрес (т.е. тот, который маршрутизируется в Интернет) и обладающий достаточно широким каналом, автоматически становится super-узлом и пропускает через себя трафик обычных узлов, помогая им преодолеть защиты типа брандмауэров или трансляторов сетевых адресов (NAT) и равномерно распределяя нагрузку между хостами. Единственным централизованным элементом является Skype-login сервер, отвечающий за процедуру авторизации Skype-клиентов и гарантирующий уникальность «позывных» для всей распределенной сети.

Такая архитектура позволяет установить и использовать прямое соединение между любыми вычислительными узлами в одной подсети, увеличивая скорость и уменьшая накладные расходы. Связь между узлами через интернет осуществляется не напрямую, а через цепочку super-узлов. «Серверов» в общепринятом смысле этого слова (таких, например, как в сети eDonkey) в Skype-сети нет и любой узел с установленным Skype-клиентом является потенциальным сервером, которым он автоматически становится при наличии достаточных системных ресурсов (объема оперативной памяти, быстродействия процессора и пропускной способности сетевого канала, не защищенного никакими средствами защиты).

Каждый узел Skype-сети хранит перечень IP-адресов и портов известным ему super-узлов в динамически обновляемых кэш-таблицах (Host Cache Tables, HC-tables). Начиная с версии Skype 1.0, кэш-таблицы представляют собой простой XML-файл, в незашифрованном виде записанный на диске в домашней директории пользователя.

Таким образом, в состав системы входят следующие элементы:

- *Skype-login сервер* – единственный централизованный элемент Skype-сети, обеспечивающий авторизацию Skype-клиентов.
- *Обычный узел (Skype Client)* – обычный конечный узел в сети.
- *Супер-узел (Super node)* – узлы, играющие роль роутеров в сети Skype. Любой узел, обладающий публичным IP и обладающий широким каналом становится супер-узлом.
- *Выделенные узлы* для установки связи со стационарными телефонными линиями.

10.4 Достоинства и недостатки P2P

Можно выделить следующие основные преимущества одноранговых сетей:

- высокая масштабируемость, связанная с равномерным распределением вычислительной нагрузки на всех участников сети;
- стабильность работы сети, обусловленная отсутствием «узкого места» – выделенного сервера, обрабатывающего все сетевые запросы;
- возможность объединения ресурсов отдельных участников сети, и их предоставление другим участникам;
- распределение совокупных затрат на предоставление ресурсов между участниками сети.

С другой стороны, отдельно стоит упомянуть о следующих недостатках и особенностях функционирования P2P-сетей:

- в одноранговых сетях не может быть обеспечено гарантированное качество обслуживания: любой узел, предоставляющий те или иные сервисы, может быть отключен от сети в любой момент;
- индивидуальные технические характеристики узла могут не позволить полностью использовать ресурсы P2P сети (каждый из узлов обладает индивидуальными техническими характеристиками что, возможно, будет ограничивать его роль в P2P-сети и не позволят полностью использовать ее ресурсы: низкий рейтинг в torrent-сетях, LowID в eDonkey могут значительно ограничить ресурсы сети, доступные пользователю);
- при работе того или иного узла через брандмауэр может быть значительно снижена пропускная способность передачи данных в связи с необходимостью использования специальных механизмов обхода;

- участниками одноранговых сетей в основном являются индивидуальные пользователи, а не организации, в связи с чем возникают вопросы безопасности предоставления ресурсов: владельцы узлов P2P-сети, скорее всего, не знакомы друг с другом лично, предоставление ресурсов происходит без предварительной договоренности;
- при увеличении числа участников P2P сети может возникнуть ситуация значительного возрастания нагрузки на сеть (как с централизованной, так и с децентрализованной структурой);
- в случае применения сети типа P2P приходится направлять значительные усилия на поддержку стабильного уровня ее производительности, резервное копирование данных, антивирусную защиту, защиту от информационного шума и других злонамеренных действий пользователей.