CDN-сети и peer-to-peer

«Мы все смогли еще раз убедиться, что в условиях инновационной экономики рынок телекоммуникаций приветствует появление компаний, технологические решения которых реализованы на самом высоком уровне. Основным поставщиком видео трансляций в интернете на Олимпийских играх 2014 года в Сочи стала небольшая и молодая компания из России. Выйдя на рынок сетей доставки контента (content delivery/distribution network, CDN) осенью 2013 года, она стремительно заняла значительную долю и беззастенчиво потеснила остальных игроков, поставив их на грань выживания. По итогам 2014 года, с оборотом $1,890 млрд., эта компания контролирует уже 27% глобального рынка CDN-сетей, и у нас есть все основания считать, что в 2015 году результаты будут еще лучше.

Агрессивная политика продаж, удачное использование в системе продвижения компании фактора «Сочи-2014», когда взгляды всего мира были прикованы к событиям в России, безусловно, все это способствовало появлению новой «звезды». Тем не менее, мы сейчас прекрасно видим, что главным «коньком» русских стала техническая реализация. Поразительно, что компания, каждый месяц «перемалывая» не менее 11 Эксабайт данных и транслируя одновременно десятки миллионов видео потоков, обладает самой скромной собственной инфраструктурой. Сравнение сети новичка с сетью Aka Technologies, «раскинувшей» свои узлы по всему миру и создавшей мощную «наложенную» сеть компании №1 в 2013 году и уступившей сейчас пальму первенства, даст Вам оценку затрат обеих компаний, и сравнение будет далеко не в пользу последней.

Мы можем только рукоплескать таким замечательным показателям, и нам обязательно надо проанализировать и понять, почему так произошло…» - онлайн-газета Daily Teleglobe Arctica, март 2015.

На самом деле газеты Daily Teleglobe Arctica не существует, как не существует и Aka Technologies, да и до марта 2015 года надо ждать еще более чем 1500 дней. Фантастический «оборот» в начале статьи выполняет роль «приманки», привлекая читателей к «скучной и нудной» теме о распространении контента в сети интернет…

По данным J’son & Partners Consulting, в 2009 году через CDN-сети ежемесячно передавалось 6,9 Эксабайт, это 46,94% от общемирового трафика объемом 14,7 Эксабайт. Лидерами рынка по итогам года стали Akamai Technologies с доходом $860 млн., Limelight Networks – $131,66 млн., Level 3 Communications – $75 млн., CDNetworks – около $60 млн.

[CDN-сети и peer-to-peer]

Прогноз дохода от услуг CDN в мире, $ млн. Источник…

Как функционирует CDN-сеть?! «Пограничные серверы кэширования» (edge cache servers) размещаются в стратегически рассредоточенных ключевых областях интернет (IX, датацентры), которые обеспечивают хорошую связность с различными сегментами сети, и на них реплицируются данные от сервера источника. При запросе пользователя определяется сервер, который будет наиболее близко расположен к нему, исходя из разных параметров (кол-ва хопов, времени задержки, балансировки нагрузки и т.д.), и далее контент будет отдаваться уже с этого сервера. Прямая аналогия со старыми и добрыми прокси-серверами, кэширующими информацию из часто запрашиваемых внешних ресурсов. И там, и там, основная задача – приблизить информацию к пользователю. В CDN-сетях это решается на немного другой технологической базе и с большим размахом.

К примеру, CDN-сеть компании CDNetworks, партнером которой в России является/являлась Синтерра, включает в себя более 80 узлов в 52 регионах мира и обеспечивает суммарную пропускную способность 1,100 Gbps (по данным самой компании на август 2010 года).

[CDN-сети и peer-to-peer]

Узлы глобальной CDN-сети компании CDNetworks

Из списка ТОП-4 компаний, предоставляющих услуги CDN-сетей в мире, три являются стартапами и были созданы в последних годах прошлого – первых годах нынешнего века. Именно в это время появились первые операторы коммерческих услуг CDN-сетей и рынок получил свое развитие. В последние несколько лет, крупные операторы связи, владеющие магистральными ресурсами, стремятся самостоятельно развивать этот бизнес.

- «Вхождение» операторов связи на рынок CDN услуг началось, наверное, еще с 2005 года, когда Level 3 Communications в рамках слияния с WiTel досталось и подразделение Vyvx, занимающееся распространением контента. Уже в декабре 2007 года Level 3 Communications

объявил о своих планах покупки за $135 млн. CDN-сети компании Savvis, состоящей на тот момент из дюжины «кэширующих» узлов в Северной Америке, Европе и Азии.

Примерно в то время, когда получили свое развитие первые CDN-сети, появились и «пожиратели» пропускной способности сетей передачи данных – пиринговые (peer-to-peer, P2P) сети для обмена файлами. Громкое имя будущим P2P-сетям сделала Napster. Эта сеть для поиска музыкальных композиций действовала с июня 1999 года по июль 2001 года. На пике своей популярности она имела 80 млн. пользователей и оперировала более чем 25 млн. песен.

Основа P2P-сетей – децентрализация на принципах «клиент – клиент»: технология обеспечивает построение распределенной сети, в рамках которой каждый узел может одновременно выступать как потребителем ресурсов (клиентом), так и поставщиком (сервером). При этом, топология сети зависит от принципа расположения ее управляющих центров / серверов:

- с одним управляющим центром, когда центральным сервер контролирует и обеспечивает функционирование всей сети (Napster);

- большое множество управляющих центров / серверов (eDonkey);

- полностью децентрализованная сеть (без «выделенных» серверов), когда для подключения нового узла необходимо указать IP-адрес любого существующего клиента, который также является и сервером (Gnutella).

[CDN-сети и peer-to-peer]

Источник: Книга «СКАЧИВАЕМ фильмы, музыку и программы из интернета. Пиринговые сети: eDONKEY, BitTorrent, KaZaA, DirectConnect». Наука и Техника, 2006

Расплодившиеся пиринговые сети, с их взрывным ростом, надолго стали драйвером развития вендоров, вынуждая операторов связи при введении безлимитных тарифов обращать особое внимание на свою инфраструктуру, которая должна была выдержать возросшие запросы пользователей. По оценке компании CacheLogic, в 2006 году, в общем объеме трафика интернет, доля P2P-сетей занимала порядка 70%, из которой 65% приходилось на видео файлы, 11% аудио и 24% на файлы других типов.

К 2009 году ситуация меняется кардинально. Несмотря на определенные допущения, связанные с тем, что большинство P2P-сетей используют случайные порты, а 40% шифруют трафик, доля пирингового обмена данными уменьшается. Именно так следует из отчета, представленного Arbor Networks на конференции NANOG47 на основании анализа 110 ISP разного типа в 17 странах, почти 3000 маршрутизаторов и в общей сложности 264 Эксабайт данных.

[CDN-сети и peer-to-peer]

Сравнение оценок доли и динамики роста трафика P2P в 2006 и 2009 годах

Technology.Review, 14 октября 2009 года: Hendrik Schulze, главный инженер ipoque, отмечает, что снижение доли пирингового трафика не означает, что его стало меньше, он просто не растет с такой скоростью, как трафик других приложений сети интернет. Doug Knopper, соучредитель рекламного агентства FreeWheel говорит, что результаты отчета соответствуют тенденциям, наблюдаемые его компанией…

[CDN-сети и peer-to-peer]Cisco Visual Network Index Forecast: К 2014 году доля трафика file sharing (peer-to-peer and web-based file sharing) составит 26,96% с объемом 11 340 Петабайт ежемесячно, против доли в 45,81% в 2009 году с объемом 4 091 Петабайт. Трафик видео и ТВ трансляций в интернете, составляющий в 2009 году объем 2 883 Петабайта в месяц (доля 32,28%) уже в 2014 году будет составлять 23 543 Петабайта (доля 55,96%).

Действительно, набравшая в последние годы тенденция просмотра видео/кино/теле в онлайн по всей видимости получит и далее свое развитие. Только в первой половине этого года в России были широко и с помпой анонсированы два проекта, появившиеся на свет с разницей чуть более месяца: ivi.ru (26 февраля 2010) и TVzavr.ru (9 апреля 2010). На данном фоне естественно нельзя не отметить и развивающиеся решения на базе приставок (STB) для подключения к телевизору (или интегрированные в телевизор) в парадигме WebTV (ранее MSN TV), которые продвигают, к примеру, компании Apple и Google.

Повышение доли «тяжелого», для сетей, потокового видео-трафика и его коммерциализация будет играть на руку операторам CDN-сетей. Важно не только обеспечить обработку и хранение контента – надо еще и доставить его потребителю с соответствующим качеством. Именно CDN-сети, перемещая контент по своей географически распределенной сети «ближе» к пользователям, могут лучше обеспечить требуемую скорость его доставки и параметры отклика (но не всегда и не везде). Ведь видео, которое постоянно прерывается из-за буферизации, смотреть практически невозможно. С

другой стороны, рост потоковой передачи мультимедиа контента будет сложной задачей и для CDN-сетей:

- распределенная географически архитектура уменьшает число транзитных узлов на пути к потребителю, помогает уменьшить задержку, оптимизирует полосу пропускания и снижает вероятность пиковой нагрузки. Однако, это не означает что локальное соединение будет свободно от перегрузок. Система все еще страдает от проблемы отказа/аварии в единственной точке (как и схема с одним сервером источником), хотя и для только определенного/локального сегмента сети интернет;

- в случае необходимости существует возможность балансировки нагрузки, переведя трафик на другой, менее загруженный «кэширующий» узел. Тем не менее, механизмы балансировки страдают определенными недостатками, так как маршрутизация на участке от узла до потребителя находится вне рамок контроля оператора CDN-сети;

- платформы на базе CDN, впрочем, как и традиционные системы (с одним сервером), спроектированные на пиковые параметры, страдают от низкой утилизации аппаратных и сетевых ресурсов в течение часов неполной загрузки. Это особенно актуально в виду ярко выраженного прайм-тайма (пик в телевизионных терминах) вечерних видео-просмотров или резко появляющегося трафика событийного уровня.

[CDN-сети и peer-to-peer]

График трафика BBC во время матча отборочного тура чемпионата Европы-2008 России – Англия, 17 октября 2007 года. Примечание: очень вероятно люди слушали онлайн Five Live

Один из вероятных путей снижения нагрузки и увеличение производительности может лежать в реализации сетей распространения контента, объединяющих в своем составе преимущества технологии peer-to-peer. В такой системе часть рабочей нагрузки, создаваемая передачей потоков данных, распределяется на рабочие станции потребителей. Возможности же системы, распределяющей ресурсы по большому множеству пользователей, великолепно продемонстрировал проект SETI@Home.

- SETI@home: An Experiment in Public-Resource Computing. Первый клиент под операционные системы Windows и Macintosh для проекта SETI@Home был выпущен в мае 1999 года. В течение недели его установили и загрузили приблизительно 200 тыс. человек. К июлю 2002 года в вычислениях проекта участвовало более чем 3,83 млн. пользователей. За 12 месяцев, с июля 2001 года по июль 2002 года средняя вычислительная мощность «распределенного компьютера» составила порядка 27,36 TerraFLOPS.

Несмотря на явно выглядящие преимущества от слияния двух схем распространения контента, CDN и peer-to-peer, - они пока являются больше альтернативой друг другу. Тем не менее, есть ряд вопросов, которые надо решить перед внедрением P2P в коммерческие сети операторов. Помимо технических проблем, кои происходят из-за разности принципов, лежащих в основе обеих моделей (надежность доставки, качество и т.д.), есть еще и другие аспекты, к примеру:

- в виду отсутствия стандартов, на рынке наличествует разнотипное клиентское программное обеспечение, реализуемое каждым оператором отдельно;

- программное обеспечение клиента можно сравнить с «вредоносным» – оно «съедает» пропускную способность канала пользователя в сеть интернет;

- широкое развитие ADSL (в США и Западной Европе) и как следствие разности в потенциалах полосы пропускания прямого и обратного канала.

Пока к объединению преимуществ архитектуры CDN и P2P огромный интерес проявляют больше академические и производственно-научные сообщества: CoralCDN, Decentralized Media Streaming Infrastructure (DeMSI), P2P-Next. Практические результаты реального понимания, как будет вести себя, к примеру, гибридная инфраструктура CDN/P2P-сети только начинает появляться: Design and Deployment of a Hybrid CDN-P2P System for Live Video Streaming: Experiences with LiveSky.

Сейчас нет ни одной реально реализованной и получившей широкое признание гибридной CDN/P2P-сети. Akamai, Limelight, да и другие крупные операторы, только присматриваются к применению P2P в своих сетях. Коммерческие операторы осторожно относятся к нововведениям. Может внимание такого крупного игрока как Microsoft даст толчок рынку, ведь P2P обладает бесспорной масштабируемостью: один PC сервер с PPLive и портом 10 Mbps с одной стороны – видео поток с битрейтом 400 kbps и 1,480 млн. зрителей с другой стороны.

[CDN-сети и peer-to-peer]

Из презентации «Locality aware P2P delivery: the way to scale Internet Video», Jin Li, Microsoft Research,

P2P Media Summit 2010

В условиях развития конкуренции на рынке услуг CDN (в том числе, и в России, со временем) - возможность переложить часть затрат на пользователя стала бы хорошим подспорьем…

PS: «Пока русские спорили, что такое программирование: наука или искусство – американцы доказали что это производство». Чем кстати не пример одной из задач, двигающей нашу экономику на путь инноваций?! Ведь все есть, производить ничего не надо – нужны только мозги, математическое образование, имеющее в России традиционно сильную школу, и программисты!

Да и формат национальной операционной системы тоже не помешает расширить, надо ввести новые функции, присущие распределенным системам по типу The Worldwide Computer. Представьте разнесенную систему хранения данных, использующую компьютеры школьников. Да и когда они, школьники, дружно маршируют на обед, процессорам тоже не стоит «простаивать» без дела…