

Висновки

За останні 30 років мобільний зв'язок перетворився зі звичайного примітивного засобу зв'язку у такий, що базується на множині високошвидкісних технологій, які здатні значно полегшити використання інтернету в житті людей. Нові покоління стандартів стільникового зв'язку з'являються доволі часто: 1G (NMT) 1979 рік, 2G (GSM) 1991 рік, 3G (W-CDMA/FOMA) 2001 рік, та 4G (LTE) 2008 рік. Початок впровадження технології 5G в Європі відбудеться до 2020 року, а в Україні через 10 років, а технологія MIMO дозволить використовувати LTE в різних частотних діапазонах.

Список використаних джерел

1. Тихвинский В. О., Терентьев С. В., Юрчук А. Б. Сети мобильной связи LTE. Технологии и архитектура – 2010.
2. Бабков В. Ю. Вознюк М. А. Михайлов П. А. Сети мобильной связи. Частотно-территориальное планирование – 2006.
3. Гельгор А. Л., Попов Е. А. Технология LTE мобильной передачи данных – 2011.
4. Кааранен Х., Ахтиайнен А. Сети UMTS. Архитектура, мобильность, сервисы – 2007.
5. Рашич А. В. Сети беспроводного доступа WiMAX – 2011.
6. Карташевский В. Г. Сети подвижной связи – 2010.
7. Гепко И. А., Олейник В. Ф., Чайка Ю. Д., Бондаренко А. В. Современные беспроводные сети: перспективы развития – 2009.
8. Суваткин В. С. WiMAX-технология беспроводной связи теоретические основы, стандарты, применение – 2005.

УДК 004.042

Т. А. Ліхоузова, П. В. Чумаченко

ОГЛЯД КОНЦЕПЦІЇ МЕРЕЖ ДОСТАВКИ КОНТЕНТУ

Анотація: Мережі доставки контенту (CDN) спрямовані на подолання обмежень, що властиві Інтернету. Дано короткий огляд основних особливостей механізму роботи CDN, наведено його типову архітектуру, охарактеризовано основні типи контенту та сервісів, з якими працює ця технологія. Також розглянуто основні підходи, що використовуються механізмом маршрутизації запитів CDN, і характеристики, на основі яких обирається найоптимальніший шлях.

Ключові слова: Мережі доставки контенту, CDN, маршрутизація запитів.

Вступ

Комерційний успіх Інтернет- та електронних послуг, разом зі швидко зростаючим використанням комплексного медіа-контенту он-лайн, проклав шлях для народження та розповсюдження мереж доставки контенту (CDN). Інтернет-трафік часто стикається з труднощами продуктивності. Нагальна потреба його користувачів у гарантіях якості викликала необхідність вивчати і розвивати нові мережеві архітектури і технології, щоб покращити якість сприйняття користувача, при цьому знижуючи затрати постачальників.

Постановка задачі

У наш час CDN підтримують доставку будь-якого типу динамічного контенту, включаючи різні форми інтерактивного медіа-стрімінгу. Провайдери CDN – це компанії, що надають свої сервери для розміщення на них контенту третіх осіб, створюючи дзеркала або дублюючи цей контент на кілька серверів по всьому світі і перенаправляючи запити користувачів на «найкращі репліки» (наприклад, найближча репліка, або та, чия затримка буде мінімальною). Тому при створенні повноцінної CDN-системи необхідно вирішити ряд технічних проблем: який тип контенту буде зберігатись на певному сервері CDN, як саме контент буде підтримуватись в актуальному стані, яка саме копія стане «найкращою реплікою» для кожного окремого користувача, які механізми необхідно використати, аби перенаправляти користувачів до цієї репліки. Правильне розміщення дублюючих серверів скорочує шлях від серверів до клієнтів, таким чином знижуючи ризик виникнення проблеми забивання каналу в неорганізованому середовищі мережі Інтернет. Механізм перенаправлення запитів встановлюється на рівні маршрутизаторів для забезпечення пошуку найбільш підходящої репліки

© Т. А. Ліхоузова, П. В. Чумаченко

на кожен запит. Обслуговуючи запит, архітектура CDN покладається на вимірювання, які виконуються маршрутизаторами спільного доступу, щоб оцінити умови трафіку, обчислювальні здатності і доступність кожної репліки.

Вдало застосована CDN може прискорити доступ користувача до контенту, понизити рівень мережевого трафіку і знизити вимоги до машинних потужностей провайдера.

Особливості мереж доставки контенту

Проектування CDN потребує не тільки розподілення серверів-реплік у крайніх зонах мережі, але й набір сервісів для їх підтримки. Для того, щоб бути ефективним для значної кількості користувачів на широкій території, крайні сервери мають бути розгорнуті у тисячах мереж в різних географічних зонах. Оптимальна продуктивність і надійність залежить від зернистості розподілення цих серверів. Встановлення CDN потребує проектування кількох важливих особливостей.

Механізми розміщення реплік потрібні, щоб визначити місце розташування дублюючих серверів і адаптивно заповнити їх відповідним контентом ще до того, як прийдуть запити. Таким чином сервери заповнюються не після отримання запиту, як при традиційному кешуванні через проксі, а оновлюються про-активно, викликаючи одноразове розвантаження понаднормового трафіку, що не повторюється при кожному запиті до серверу оригіналу. Адаптивність в розміщенні дублікатів необхідна, щоб впоратись з постійно змінними умовами трафіку.

Механізми оновлення контенту спрямовані на автоматичну перевірку сайту хоста на наявність змін і отримання оновленого контенту для доставки на сервери. Таким чином забезпечується актуальність інформації. Стандартні механізми, що застосовуються в кешуванні через проксі, не гарантують актуальності контенту, так як він зберігається на стандартних серверах кешу і не змінюється разом з контентом-джерелом.

Механізми активного вимірювання мають бути додані до маршрутизаторів спільного доступу, щоб мати доступ до стану Інтернет-трафіку у реальному часі. Таким чином буде визначатися найшвидший шлях від користувачів, які відправляють запити, до серверів-реплік при будь-якому типі трафіку, особливо при наявності так званих «раптових натовпів», коли виникає неочікуваний сильний попит на певний сайт. Цей механізм є основою для механізму вибору репліки.

Механізми вибору реплік мають додаватись до маршрутизаторів спільного доступу, щоб точно визначити найближчий і найбільш доступний сервер, з якого користувачі можуть отримати запитуваний контент.

Архітектура CDN

Основною метою реплікації серверів в CDN є уникнення ситуації, коли великі об'єми інформації повторно запитуються з одного посилання в мережі Інтернет. На рис. 1 надано набір методів і масштабів (локальні чи на широкій території), якими можуть бути впроваджені мережі контенту. Локальними рішеннями є веб-кластери, де зазвичай розміщується один сайт, і веб-ферми, що використовуються для кількох сайтів. Рішення для широких територій включають: розподілені веб-серверні системи, що використовуються для одного або кількох сайтів; спільні мережі проксі-кешу (сервісна інфраструктура, потрібна для зниження затримки при завантаженні веб-об'єктів) і мережі з доставки контенту [1].

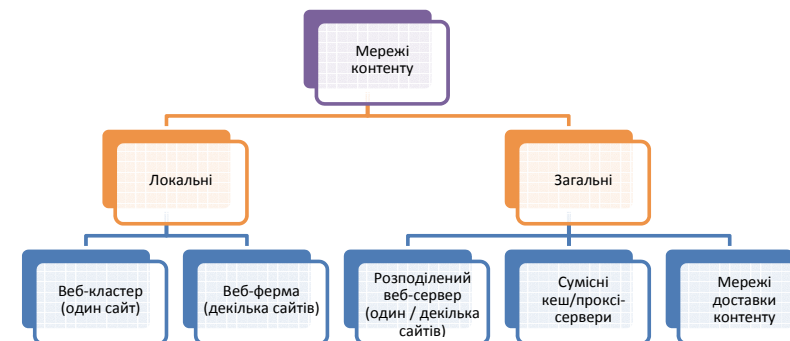


Рис. 1. Систематика мереж контенту

Типова серверна ферма є групою серверів, від двох до кількох тисяч, яка використовує так званий спільний диспетчер, що працює на рівнях OSI 4 і/або 7, щоб приховати розподілену природу системи, таким чином виглядаючи як один сайт. Веб-перемикач на рівні 4 розсилає запити між групою серверів, спираючись на таку базову мережеву інформацію цього рівня, як IP-адреса і TCP-порт. Перемикач контенту, що працює на рівні додатку, вивчає зміст запитів і розсилає їх між групою серверів. Цілі серверного кластера/ферми включають: балансування навантаження запитів між всіма серверами в групі; автоматична маршрутизація запитів від тих серверів, які надають відмову; маршрутизація всіх запитів для певної користувацької сесії до одного і того ж самого серверу, якщо необхідно постійно зберігати стан сесії.

Одним з типів мереж контенту, який використовується уже кілька років, є розгортання проксі кешування. Така мережа зазвичай впроваджується Інтернет-провайдером для підтримки користувачів з низькою пропускну здатністю. Щоб покращити продуктивність, проксі кешування розгортаються неподалік від користувачів. Цих користувачів спонукають відправляти свої веб-запити через кешовані сервери, замість оригінальних, попередньо налаштувавши їх браузері. Якщо таке налаштування виконане правильно, вся сесія користувача у браузері проходить через певний проксі-сервер кешування. Таким чином кеш проксі буде містити найбільш запитувану інформацію, яка переглядається усіма користувачами цього проксі-серверу. Провайдер, який розгортає кеш-сервери у багатьох географічних зонах, також може розгортати регіональні батьківські кеші, щоб ще сильніше поєднувати запити користувачів, таким чином створюючи архітектуру, відому як ієрархічне кешування. Це може надати додаткове покращення продуктивності і зберегти пропускну здатність. Використовуючи протоколи наслідування, допоміжні «батьки» можуть бути розгорнуті таким чином, що при виявленні будь-якої відмови у первинних «батьків» на їх місце стають допоміжні. Використовуючи подібні батьківські протоколи, запити можуть бути розподілені таким чином, що запити певного контенту відправляються до конкретних первинних «батьків». Це може допомогти максимізувати ефективність використання ресурсів проксі кешування.

При явній кращій масштабованості порівняно з одиночними серверами, як ієрархічне кешування, так і серверні ферми мають свої межі. В цих архітектурах репліки зазвичай розгортаються близько до серверів-джерел, і таким чином вони не вирішують проблеми продуктивності, що викликана перенавантаженням мережі. Проксі кешування може допомогти з цими проблемами (так як сервери розташовуються близько до кінцевих користувачів), але воно кешує лише об'єкти, на які є попит у клієнта. Реактивне кешування буде працювати незадовільно, якщо запити до певного об'єкта, яких у сукупності може бути дуже багато, розподілені між багатьма проксі кешами.

Зважаючи на ці обмеження, CDN використовують методи, засновані скоріш на про-активному, а не на реактивному кешуванні, коли контент наперед подається з серверу-джерела, а не кешується по запитам. В CDN чисельні репліки зберігають один і той самий контент. Запит з браузера до певної одиниці контенту направляється до тої репліки, яка вважається найбільш вдалою на момент цього запиту, і об'єкт постачається клієнту значно швидше, ніж це було б з серверу-джерела. Так як статична інформація про географічне розміщення і під-

ключення до мережі не є достатньою для вибору найкращої репліки, CDN зазвичай включає динамічну інформацію про стан мережі і навантаження на репліки, щоб перенаправляти запити і урівноважувати навантаження між серверами. Тому керування CDN системою є складним і затратним. Через це вони зазвичай будуються і керуються тим мережевим провайдером (або провайдером послуг), який надає послугу розподілення контенту кільком провайдерам контенту.

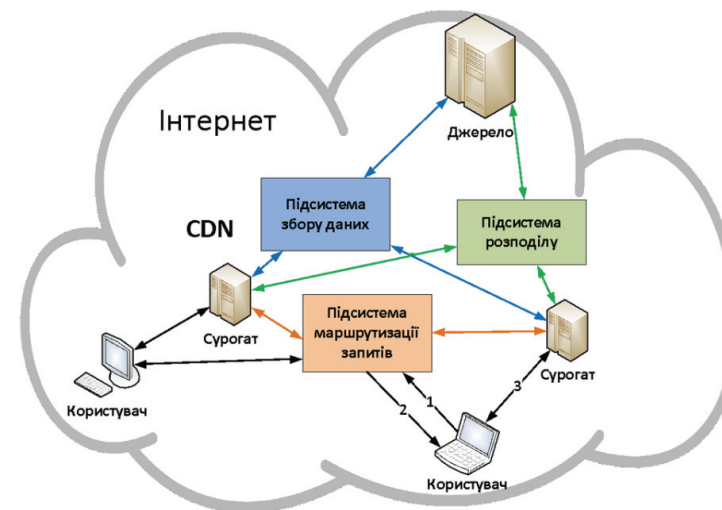


Рис. 2. Компоненти інфраструктури CDN

Архітектура доставки контенту складається з набору *серверів-реплік*, які доставляють копії контенту до користувачів, об'єднуючи різні дії.

- Інфраструктура *маршрутизації запитів* складається з механізмів для перенаправлення запитів до контенту від клієнта до прийнятної репліки.
- Інфраструктура *розподілення* складається з механізмів переміщення контенту від серверу-джерела до реплік.
- Інфраструктура *обліку* знаходить і збирає інформацію про маршрутизацію запитів, розподілення і доставку контенту всередині CDN, створюючи журнали і звіти про всі дії розподілення і доставки.

Сервер-джерело (той, що містить контент для доставки) взаємодіє з CDN двома способами:

- передає новий контент на сервери-репліки (при цьому сама репліка робить запити про оновлення контенту з серверу-джерела через інфраструктуру розподілення);

- запитує журнали та іншу облікову інформацію з CDN, або ж CDN сама надає свою інформацію серверу-джерелу через інфраструктуру обліку.

Клієнти взаємодіють з CDN через інфраструктуру маршрутизації запитів і репліки. Рис. 2 показує можливі сценарії взаємодії між клієнтами, маршрутизаторами, репліками (сурогатами) та сервером-джерелом.

Користувач відправляє (1) запит контенту через інфраструктуру маршрутизації, яка перенаправляє (2) цей запит до сурогату, до якого згодом вже клієнт відправляє запити (3) на бажаний контент.

Типи контенту і сервісів в CDN

Провайдери CDN розповсюджують контент третіх осіб для прискорення доставки будь-якого його типу, наприклад, потокове аудіо/відео, статичні сторінки html, зображення, документи або додатки. Джерелами цього контенту можуть бути медіа-компанії, великі підприємства, компанії-транслятори, провайдери веб-сервісів. Через різноманітну природу такого контенту, можуть бути прийняті різноманітні архітектури і технології при проектуванні CDN. Проаналізуємо характеристики контенту і додатків, яким скоріш за все знадобляться переваги архітектури CDN.

Статичні сервіси на веб-основі

Використовуються для доступу до статичного контенту (статичні сторінки html, зображення, документи, патчі програмного забезпечення, аудіо і/або відео файли), або до контенту, який змінюється з низькою періодичністю (тимчасові веб-сторінки). Всі провайдери CDN (Akamai Inc., Speedera Inc., AT&T inc., Globix Inc. та інші) підтримують цей тип контенту. Такий контент може легко кешуватись і його актуальність підтримується завдяки традиційним технологіям кешування контенту.

Сервіси веб-сховищ

По суті, це застосування може засновуватись на таких самих техніках, як і при доставці статичного контенту. Мають бути впроваджені деякі додаткові особливості для керування веденням журналів і для безпечної передачі файлів. Цей тип застосування може потребувати обробки на сайті-джерелі.

Сервіси передачі файлів

Міжнародне поширення програмного забезпечення, електронні навчальні матеріали підприємства для всіх його глобальних співробітників, високо деталізовані медичні зображення, які поширюються між лікарями і госпіталями та ін-

ше. Всі ці типи контенту по суті статичні, і можуть підтримуватись завдяки тим самим технологіям, які використовуються для статичних веб-сервісів.

Сервіси електронної комерції

Семантика запитів, які використовуються для пошуку в каталозі товарів не є складною, отже часті результати запитів можуть бути успішно кешовані, використовуючи традиційні техніки кешування запитів до БД [2][5]. «Корзини» можуть зберігатись і підтримуватись на репліці. Крім цього, замовлення і транзакції можуть оброблятися не в центрі: для цього потрібні довірені репліки з увімкненими транзакціями. У науковій праці [3] автори пропонують спеціальний фреймворк, що використовується для динамічного кешування контенту у сервісах електронної комерції.

Веб-додатки

Веб-транзакції, обробка інформації, доступ до баз даних, календарі, графіки роботи, усі ці сервіси зазвичай характеризуються логікою додатків, яка виконує запит клієнта, створюючи динамічну веб-сторінку. Частковим рішенням використання інфраструктури CDN при наявності динамічних сторінок є заповнення реплік контентом, який найбільш часто складає динамічно згенеровані веб-сторінки, а підтримка додатку і його дій по створенню динамічних сторінок відбувається на сервері-джерелі. Іншим підходом може бути реплікація як додатку (або його частини), так і самого контенту. Таким чином весь процес генерації контенту (логіка додатку і отримання контенту) відбувається на репліках, таким чином розвантажуючи джерело.

Сервіси каталогів

Використовуються для доступу до серверів баз даних. Наприклад, у випадку LDAP серверу, часті результати запитів або набори каталогів можуть кешуватись на репліках. Можуть також застосовуватись традиційні технології кешування запитів.

Технології вимірювання для маршрутизації запитів

Системи маршрутизації запитів використовують різноманіття метрик для визначення найкращої репліки, яка б виконувала запити клієнта. Децентралізована природа Інтернету дуже ускладнює кількісну оцінку продуктивності мережі. Збір мережевої статистики напряму на мережевих пристроях (маршрутизаторі чи сервері) може виявитись більш затратною у плані продуктивності системи. Отже, зазвичай, отримання мережевої статистики покладається на використання комбінації методів активного зондування мережі, пасивного моніторингу трафі-

ку і зворотного відгуку від реплік. Більш детальний опис, як саме впроваджуються технології вимірювання в мережу, надано у [4].

В мережах CDN можна поєднувати чисельні метрики, використовуючи як концепцію близькості, так і зворотній відгук від репліки для найкращого вибору. Вимірювання продуктивності часто є компонентом системи управління мережею і пропонує можливості спостереження та проектування продуктивності CDN. Більш того, потрібно вимірювати як внутрішню продуктивність, так і продуктивність з точки зору користувача. Типові параметри, які були б корисні для вимірювань: втрата пакетів і затримка для всіх типів контенту, особливо для потокового контенту з середньої шириною смуги пропускання. Розгортаючи стратегічне зондування всієї мережі на апаратних засобах чи програмному забезпеченні, можна співвідносити отриману інформацію з кешем і журналами серверу, щоб визначити рівень доставки і статистику QoS. Найбільш вдалим місцем для запуску зондування будуть границі мережі, тому що таким чином можна виміряти продуктивність такою, якою вона є для кінцевих користувачів через CDN. Вимірювання мережевої і географічної близькості можуть використовуватись системою маршрутизації запитів для перенаправлення користувачів до «найближчої» репліки. Більш того, репліки можуть обмінюватись вимірюваннями близькості і сутністю запитів. У багатьох випадках вимірювання близькості є односторонніми, тобто вони вимірюють лише прямий або зворотній шлях пакетів від репліки до джерела запиту. Це важливо, так як багато шляхів в Інтернеті є асиметричними. Щоб отримати набір вимірів близькості, мережа може застосувати як техніки активного зондування, так і пасивні вимірювання. Система маршрутизації запитів також може використовувати відгуки від реплік для обрання найменш завантажених вузлів. Зворотній відгук може доставлятися від кожної репліки, або ж накопичуватись відповідно до сайту або географічної зони. Далі більш детально описуються пасивні вимірювання, активне зондування і зворотній відгук.

Пасивні вимірювання (Passive Measurement)

Пасивні вимірювання можуть отримуватись, коли клієнт виконує передачу інформації до/від сервера репліки. Як тільки клієнт під'єднується, вимірюється дійсна продуктивність передачі даних. Ця інформація передається на систему маршрутизації запитів через зворотній зв'язок. Прикладом пасивних вимірів може бути спостереження за втратою пакетів від користувача до репліки, або за затримкою, з якою сприймає контент користувач, спостерігаючи за поведінкою TCP [6]. В цілому, потрібен хороший механізм для забезпечення того, щоб не

кожна репліка тестувалася для кожного користувача задля отримання цієї інформації.

Активне зондування (Active Probing)

Активне зондуванням виконує сканування об'єктів, які у минулому відправляли запити, використовуючи одну або кілька технік для визначення відповідних метрик(-и), що в подальшому будуть використані для вибору оптимальної репліки. Прикладом зондувальної техніки є запит ICMP ECHO (ping), який періодично відправляється з кожної репліки до об'єкту, що потенційно може зробити запит. Через певні причини ця техніка є досить обмеженою. Вимірювання можуть здійснюватися лише періодично, бо в іншому випадку вони створюватимуть значне навантаження на канал пропускання. Також фаєрволи і NATи у мережах є перешкодою для результативного зондування. Нарешті, зондування часто викликає спрацювання тривоги у системах виявлення мережевих атак.

Зворотній відгук (Feedback information)

Ця технологія заснована на отриманні інформації за допомогою періодичних запитів до спеціальних додатків репліки, які повертають відповідні виміри. Проблемами зондування реплік є те, що досить важко отримати саме актуальну («real-time») інформацію, а неактуальна («non-real-time») часом є неточною і застарілою. Отже, інформація зворотного відгуку може бути отримана за допомогою агентів, що інтегровані до реплік. Вони у свою чергу мають змогу обмінюватись між собою повідомленнями про метрики, що використовуються на власному вузлі. Технологія використовує два методи для отримання інформації: статичний, при використанні якого обирається маршрут, що або мінімізує кількість «стрибків», або оптимізує інші статичні параметри; динамічне зондування (зондування в реальному часі), який дозволяє обчислити час подорожі в обидві сторони або інші параметри QoS в режимі реального часу.

Метрики для перенаправлення запитів

Вибір репліки виконується з метою максимізації продуктивності з точки зору кінцевого користувача. Далі буде надана класифікація метрик (табл. 1), які можуть застосовуватись для обчислення продуктивності мережі, а отже і CDN системи, що використовується для вибору, куди перенаправити запит клієнта.

Таблиця 1

Метрики, що використовуються при виборі репліки

Метрика	Цілі	Способи виміру
Затримка	Вибір репліки з найнижчою затримкою	Активне зондування / Пасивний збір даних
Втрати пакетів	Вибір шляху з найменшою кількістю втрат (використовується для потокового трафіку)	Активне зондування / Пасивний збір даних (інформація з заголовків ТСП)
Безпосередня близькість	Вибір найкоротшого шляху	Активне зондування
Пропускна здатність каналу	Вибір найкращого шляху для потокового трафіку	
Частота кадрів		
Географічна близькість	Перенаправлення запитів з регіону до того ж POP	Інформація з заголовків IP
Завантаженість CPU, жорстких дисків, мережних інтерфейсів; кількість активних підключень	Вибір серверу з найменшим навантаженням	Агенти зворотного відгуку / Активне зондування

Географічна близькість часто використовується для перенаправлення всіх користувачів всередині певного регіону до певного серверу.

Вимірювання *мережевої близькості* зазвичай отримується активним зондуванням вузлів маршрутизації.

Стан завантаження серверу може бути обчислений, використовуючи SNMP або посередників для зворотного відгуку на серверній стороні, спираючись на інформацію про стан завантаження серверних компонентів (CPU, жорсткого диску, пам'яті, мережних інтерфейсів) або на певну спільну характеристику продуктивності, таку як пропускна здатність чи час відгуку сервера.

Усі отримані виміри слугують основою для механізму вибору серверу. Також на цей вибір впливає ще один важливий компонент – знання про *ідентичність користувача*, за допомогою яких визначається пріоритетність у доступі до контенту чи сервісів. Як приклад, користувач, що оплачує послуги, має доступ до кращого рівня послуг, ніж той, який не платить. Ідентичність користувачів, які оплачують послуги, може бути визначена завдяки файлам cookies, отриманим з системи клієнта, або ж через процес автентифікації.

Висновки

Сьогоднішні CDN діють на традиційних мережних протоколах різних рівнів, покладаючись на динамічний і активний кеш контенту та на автоматичне розгортання і переміщення межі мережі ближче до кінцевих користувачів. Репліки контенту в CDN розподілені географічно, що дозволяє швидко і надійно доставляти його до будь-якої кінцевої точки призначення: завдяки сервісам CDN, актуальний контент отримується кінцевими користувачами скоріш локально, ніж навпаки.

Покращення роботи комп'ютерних мереж засобом накладання усвідомлення контенту створює нову архітектурну парадигму CDN. Було запропоновано багато рішень для розв'язку найголовніших проблем [5,6,7], найбільш перспективні з яких засновані на усвідомленні контенту, що постачається. Традиційні «сліпі до контенту» мережні інфраструктури [1,2,3,4] не можуть забезпечити якість обслуговування для всіх користувачів в динамічному і постійно зростаючому трафіку. Нові протоколи і інтегровані рішення мають бути як і в мережі, так і на сервері, щоб поширювати, знаходити і завантажувати контент по всій мережі Інтернет.

Список використаних джерел

1. *Verma D. C.* Content Distribution Networks: An engineering approach / Wiley Inter-Science, 2002.
2. *Florescu D.* Database techniques for the worldwide web: A survey / D. Florescu, A. Y. Levy, A. O. Mendelzon, SIGMOD Record, 27(3) : 59–74, 1998.
3. *Candan K. S.* Enabling dynamic content caching for database-driven web sites / K. S. Candan, W.-S. Li, Q. Luo, W.-P. Hsiung, D. Agrawal, In Proc. of ACM/SIGMOD Conference 2001, Santa Barbare, California, USA, May 2001.
4. *Subharthi P.* Architectures for the future networks and the next generation Internet: A survey / P. Subharthi, P. Jianli, J. Raj, Computer Communications, 34 : 2–42, 2011.
5. *Jacobson V.* Networking named content / V. Jacobson, D.K. Smetters, J. D. Thornton, M. F. Plass, N. Briggs, Communications of the ACM, 55: 117–124, 2012.
6. *Goldman R.* WSQ/DSQ: A practical approach for combined querying of databases and the web / R. Goldman, J. Widom, In SIGMOD Conference, pages 285–296, 2000.
7. *Perino D.* A Reality check for content centric networking / D. Perino, M. Varvello, ICN'11, August 19, Toronto, Ontario, Canada, 2011.