

УДК 004

**Вітер М. Б.**, к. ф.-м. н. (Міністерство фінансів України)

**Сендзюк М. А.**, к.е.н. (Київський національний економічний університет імені Вадима Гетьмана)

**Пастух О. А.**, д.т.н. (Тернопільський національний технічний університет імені Івана Пулюя)

## МЕРЕЖЕВІ ТЕХНОЛОГІЇ ЗБЕРЕЖЕННЯ ВЕЛИКИХ ОБ'ЄМІВ ДАНИХ

**Вітер М. Б., Сендзюк М. А., Пастух О. А. Мережеві технології збереження великих об'ємів даних.** Здійснено аналіз базових технологій організації мережевого збереження великих об'ємів даних. Сформульовано умови їх застосування в державних органах.

**Ключові слова:** сховища даних, мережеві засоби, система збереження даних, протокол передачі

**Витер М. Б., Сендзюк М. А., Пастух О. А. Сетевые технологии хранения больших объемов данных.** Осуществлен анализ базовых технологий организации сетевого хранения больших объемов данных. Сформулированы условия их применения в государственных органах.

**Ключевые слова:** хранилища данных, сетевые средства, система хранения данных, протокол передачи

**Viter M. B., Sendzyuk M. A., Pastukh O. A. Network technology storing large data rates.** The article provides analysis of the core technologies of network storage of large volumes of data. The conditions of their use in government are formulated.

**Keywords:** data storage, network equipment, storage systems, transfer protocol

**Вступ.** Стрімкий розвиток сучасних технологій мережевого зв'язку та обміну даними, процеси децентралізації та інформаційної інтеграції, що відбуваються в органах державної влади, перехід до міжвідомчого електронного документообігу породжують потребу в надійному зберіганні великих об'ємів інформації та забезпеченні належного доступу до них.

На заміну централізованим сховищам приходять системи організації розподіленого збереження даних, які використовують мережеві пристрої чи системи зберігання даних.

Необхідність розробки технологій оптимізації мережевого збереження даних зумовлює актуальність досліджень у цій сфері.

**Огляд останніх публікацій. Проблеми.** Сучасні дослідження у сфері організації сховищ даних здебільшого зосереджуються на питаннях структури і функціонування окремого корпоративного сховища [1...4]. Тому актуальною є задача розробки ефективних технологій для побудови мережевих сховищ інформації.

У даній статті проведено аналіз сучасних мережевих технологій зберігання даних і сформульовано рекомендації щодо їх ефективного застосування в державних органах.

**Виклад основного матеріалу.** Серед сучасних засобів, які забезпечують надійне збереження електронної інформації найпоширенішими є мережеві системи зберігання даних (СЗД).

Виділена система зберігання даних розвантажує основну мережу обчислювальних серверів і клієнтських робочих станцій, звільняючи її від потоків введення/виведення даних.

Як правило, система зберігання даних містить наступні підсистеми та компоненти:

- пристрої зберігання (дисккові масиви, стрічкові бібліотеки);
- інфраструктуру доступу до пристроїв зберігання;
- підсистему резервного копіювання та архівування даних;
- програмне забезпечення управління зберіганням;
- систему управління і моніторингу.

Сучасні СЗД покликані забезпечувати надійну і відмовостійку роботу, доступність даних, масштабованість, належну продуктивність тощо.

Для малих компаній, як правило, вистачає єдиного сховища інформації. Великі ж структури (державні органи, великі підприємства) часто об'єднують у своїх системах різноманітні підсистеми структурних підрозділів. Тому для забезпечення ефективної роботи для них є важливим адекватний розвиток відповідної мережевої інфраструктури збереження даних.

Серед сучасних мережевих топологій збереження даних можна виділити такі: DAS (Direct Attached Storage – система зберігання з прямим підключенням), NAS (Network Attached Storage – система зберігання з підключенням мережею) і SAN (Storage Area Network – мережева система зберігання).

**Система зберігання DAS** – це, по суті, розширення дискової системи зберігання окремо взятого сервера (Рис.1). Клієнти отримують доступ до даних, звертаючись до цього сервера через мережу.

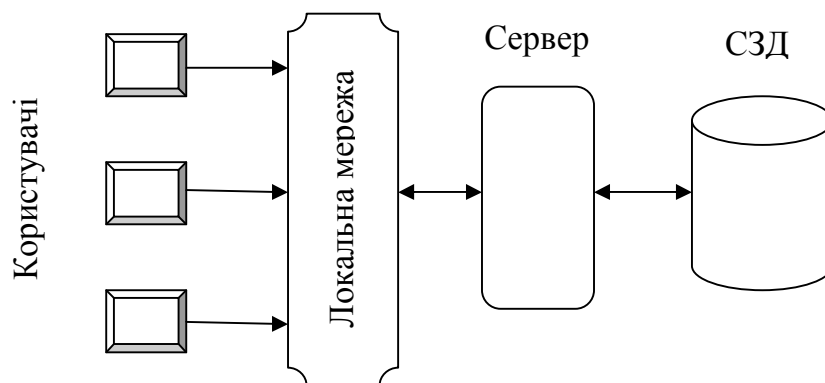


Рис. 1. Архітектура системи Direct Attached Storage

До основних переваг DAS-систем можна віднести їх невисоку вартість, простоту розгортання і адміністрування, а також високу швидкість обміну даними між системою зберігання і сервером.

Завдяки зазначеним якостям вони є найбільш ефективними для локального зберігання і використання невеликого обсягу даних. Системи зберігання DAS нерідко використовують при експлуатації Exchange Server в малих і середніх компаніях.

Одним з головних недоліків технології DAS є те, що зв'язок з дисковою системою здійснюється лише через конкретний сервер. У випадку виходу його з ладу дані в сховищі стають недоступними.

Приклади DAS-систем [5]: HP MSA 2000sa G2 фірми Hewlett-Packard, Entry Store 321 компанії Entry, Infortrend A08S-G2130-M2-KIT компанії Onix.

**Мережа зберігання NAS** – це окрема інтегрована дискова система з великим об'ємом дискового простору (Рис. 2). Часто диски в NAS об'єднують в RAID-масив. У найпростішому варіанті пристроєм NAS є звичайний мережевий сервер, що надає файлові ресурси.

Доступ до файлів в рамках єдиної NAS-системи може бути отриманий за протоколами TCP/IP, CIFS, NFS, FTP, TFTP, включаючи можливість роботи NAS, як iSCSI-target (віддаленого сховища). NAS забезпечує організацію спільного доступу до файлів за протоколом TCP/IP в мережах Ethernet [6].

Системи NAS підключаються до локальної мережі і забезпечують необхідний швидкий файловий доступ до даних для необмеженої кількості гетерогенних клієнтів (клієнтів з

різними ОС) або інших серверів. Наприклад, можлива робота в середовищі Windows Active Directory з підтримкою необхідної функціональності.

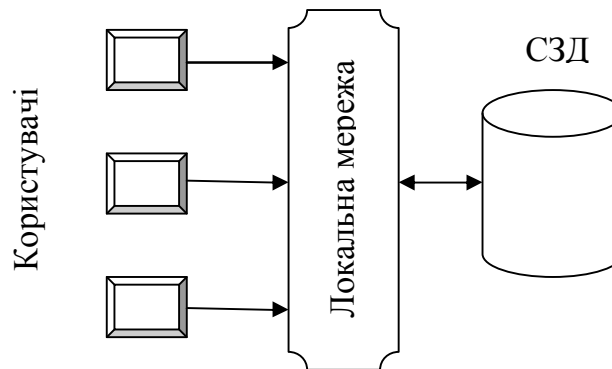


Рис. 2. Архітектура системи Network Attached Storage

Розміри NAS-систем можуть варіюватись від простих з одним портом Ethernet і двома жорсткими дисками в RAID1 до великомасштабних, що містять сотні жорстких дисків, а файловий доступ в яких забезпечується десятком спеціалізованих серверів всередині NAS-системи. Кількість зовнішніх Ethernet-портів може досягати багатьох десятків, а ємність збережених даних – кілька сотень терабайт.

NAS-пристрої можуть також бути частиною SAN-мережі і не мати власних накопичувачів, а лише надавати файловий доступ до даних, що знаходяться на блокових пристроях зберігання. У цьому випадку NAS бере на себе функцію потужного спеціалізованого сервера, а SAN – пристроїв зберігання даних.

Кілька систем NAS можуть управлятися з єдиного центру, дозволяючи економити час та кошти.

*Переваги NAS:*

- дешеві і доступні ресурси не тільки для окремих серверів, але і для будь-яких комп'ютерів мережі;
- простота колективного використання ресурсів;
- простота розгортання і адміністрування;
- універсальність для клієнтів (один сервер може обслуговувати клієнтів MS, Novell, Mac, Unix).

*Недоліки NAS:*

- доступ до інформації через протоколи "мережевих файлових систем" часто повільніший, ніж як до локального диску;
- підвищене навантаження на мережу.

При інтенсивному використанні декількох систем NAS в локальній мережі можливі проблеми. Тому необхідно правильно розрахувати розташування і кількість NAS-пристроїв.

Приклади NAS- систем [5]: Sun Storage 7000 фірми Sun Microsystems, Hitachi Data Systems Simple Modular Storage 100 (SMS100) компанії Hitachi Data Systems.

**Мережа зберігання SAN** призначена для розподілу ресурсів високопродуктивних пристроїв зберігання даних, таких як дискові (RAID) масиви, стрічкові бібліотеки, оптичні накопичувачі тощо між мережевими серверами і робочими станціями. При цьому її операційна система розпізнає їх як локальні [7].

У найпростішому випадку SAN складається з СЗД, комутаторів і серверів, об'єднаних оптичними каналами зв'язку (Рис. 3).

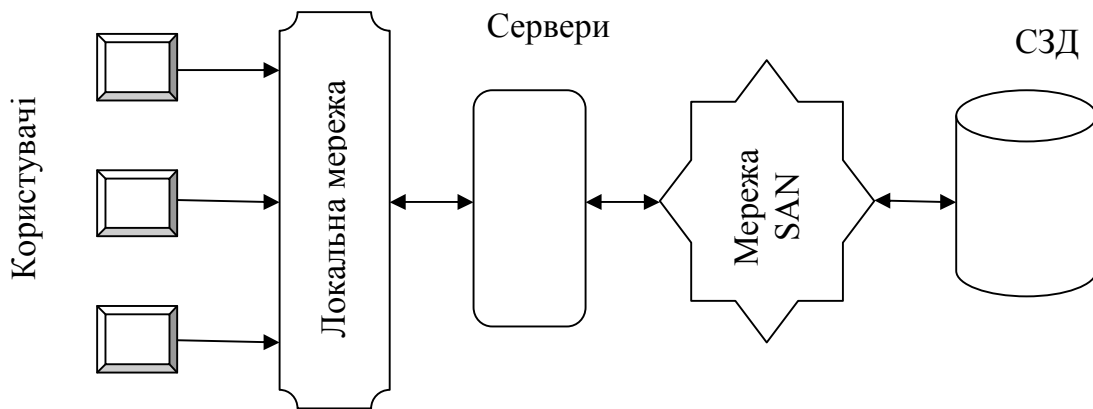


Рис. 3. Архітектура системи Storage Area Network

В основі концепції SAN лежить можливість з'єднання будь-якого з серверів з будь-яким пристроєм зберігання даних, які працюють за протоколом Fibre Channel (FC) або iSCSI (IP-SAN).

Перевагою Fibre Channel, крім швидкості, є ще й те, що він не має жорстких обмежень за довжиною лінії мережі.

Технологія iSCSI не завжди забезпечує таку ж швидкість, як Fibre Channel, але все ж завдяки своїй універсальності і нижчій ціні рішення на базі iSCSI користуються більшою популярністю.

SAN масштабується і дозволяє передавати терабайти інформації з високою швидкістю великій кількості станцій. Зв'язок між згаданими об'єктами забезпечують відповідні мережеві засоби: комутатори, мости, концентратори тощо.

Дискові масиви і стрічкові бібліотеки, не обладнані інтерфейсами Fibre Channel, можна підключити до SAN використовуючи маршрутизатори Fibre Channel-SCSI.

Використання SAN дозволяє забезпечити:

- централізоване управління ресурсами серверів і систем зберігання даних;
- підключення нових дискових масивів і серверів без зупинки роботи всієї системи зберігання;
- використання раніше придбаного обладнання спільно з новими пристроями зберігання даних;
- оперативний і надійний доступ до накопичувачів даних, що знаходяться на великій відстані від серверів, без значних втрат продуктивності;
- прискорення процесу резервного копіювання і відновлення даних.

Можна виділити такі елементи SAN:

- Host Bus Adaptors (HBA) – плата адаптера SAN;
- ресурси зберігання даних;
- пристрої, що реалізують інфраструктуру SAN;
- програмне забезпечення.

Сервери для мереж SAN відрізняються від звичайних серверів додатків тільки однією деталлю: крім мережевого адаптера, для взаємодії сервера з локальною мережею вони оснащуються HBA-адаптером, що дозволяє підключати їх до SAN-мереж на основі Fibre Channel.

На відміну від DAS або NAS, які забезпечують спільний доступ до даних на файловому рівні, SAN організовує передавання інформації у вигляді великих блоків даних. Це особливо важливо для таких додатків, як бази даних, програм обробки зображень тощо. Внаслідок високого рівня організації системи, складного управління і високої вартості SAN традиційно застосовуються там, де найбільше значення мають надійність і висока продуктивність.

Різноманітні топології SAN (точка-точка, петля з арбітражною логікою, комутація) замінюють традиційні шинні з'єднання «сервер – пристрої зберігання» і надають в порівнянні з ними велику гнучкість, продуктивність і надійність.

Системи SAN мають вбудовані функції захисту від збоїв, що гарантують стабільну роботу і мінімальний час відновлення. На відміну від систем з прямим підключенням, в моделі SAN ресурси зберігання можна об'єднувати в групи, що призводить до підвищення ефективності їх використання.

Використання мереж Fibre Channel з волоконно-оптичними з'єднаннями пристроїв дозволяє значно збільшити відстань між структурами мережі до десятків кілометрів [8], забезпечуючи швидкість до 16 Гбайт/сек. У великих установах, де оперують величезними потоками інформації, такі мережі просто незамінні.

#### *Переваги SAN:*

- висока швидкодія і надійність доступу до даних, що знаходяться на зовнішніх системах зберігання;
- незалежність топології SAN від використовуваних СЗД і серверів;
- централізоване зберігання даних (надійність, безпека);
- зручне централізоване управління комутацією і даними;
- масштабованість і гнучкість логічної структури SAN;
- проста схема резервного копіювання;
- високий рівень безпеки.
- можна використовувати блочні методи доступу, зберігання «нефайлової» інформації (часто використовується для баз даних, а також для поштової бази Exchange).
- деякі програми працюють тільки з «локальними дисками» і не працюють на NAS (приклад – MS Exchange).

#### *Недоліки SAN:*

- висока вартість (зокрема, програмного забезпечення);
- складність в налаштуванні FC-систем;
- більш жорсткі вимоги до сумісності та валідації компонентів.

Приклади SAN-систем [5]: EqualLogic серії PS4000 фірми Dell, IBM SystemStorage DS3400, Sun StorageTek ST2540.

Сьогодні виробники шукають шляхи об'єднання технологій NAS і SAN в єдину мережеву інфраструктуру зберігання, яка забезпечить консолідацію даних, централізацію резервного копіювання, спростить загальне адміністрування, масштабованість і захист даних. Така єдина інфраструктура дозволяє консолідувати дані на рівні файлів і блоків у рамках загальних масивів [9].

Серед відомих світових виробників СЗД слід назвати Dell, EMC, HP, IBM, Fujitsu, Hitachi Data Systems, Sun Microsystems.

**Висновки.** При організації систем зберігання даних в органах державної влади необхідно визначитись, які задачі вони повинні розв'язувати, яка структура центру збереження даних створюється – єдина локальна чи розподілена. При цьому повинна максимально повно враховуватись організаційно-технічна структура організації.

Наприклад, якщо не планується спільне використання ресурсів, то, очевидно, що найефективнішою системою збереження буде DAS. Правда, при цьому важливим фактором є відстань між користувачами і системою, так як DAS функціонує на невеликій відстані.

Важливим фактором є обсяг і тип даних (файловий чи блочний рівень), які необхідно зберігати. У випадку файлового доступу і спільного використання інформації для додатків на різномірних серверних платформах в локальній мережі доцільно використовувати систему NAS. Такі системи можуть застосовуватись в регіональних структурах (обласних і районних) державних органів.

Мережі SAN є актуальними у випадку, коли потрібен високопродуктивний блоковий доступ до баз даних, консолідація, висока надійність і ефективність збереження даних. Такі вимоги до СЗД є в центральних органах державної влади, що здійснюють оперативні обробку великих масивів інформації ( Міністерство фінансів України, Міністерство доходів і зборів України, Державна казначейська служба України).

Крім того, архітектура повинна враховувати число одночасно працюючих користувачів. Якщо для рішень середнього класу головними цілями можна вважати мінімізацію часу доступу до даних і підтримку як можна більшого числа користувачів, то для рішень high-end (елітний рівень) ключові характеристики будуть дещо іншими – це надійність, масштабованість і широкі можливості інтеграції з корпоративною мережею.

### Література

1. Додонов О. Г. Інформаційні потоки в глобальних комп'ютерних мережах: монографія / О. Г. Додонов, Д. В. Ланде, В. Г. Путятін (НАН України, Інститут проблем реєстрації інформації). – К. : Наукова думка, 2009. – 295 с.
2. Матов О. Я. Сучасні технології Інтеграції інформаційних ресурсів / О. Я. Матов, І. О. Храмова // Реєстрація, зберігання і обробка даних. – 2009. – Т. 11, № 1. – С. 33-42.
3. Хаджинов В. В. Модель інформаційного сховища в Інтернет-середовищі / В. В. Хаджинов, І. О. Храмова // Реєстрація, зберігання і обробка даних. – 2002. – Т. 4, № 4. – С. 104-110.
4. Кадошук И. Т., Липчинский Е. А. Обзор технологий хранилищ данных [Электронный ресурс] // – Режим доступа: <http://www.sever.eduhmao.ru/info/1/3611/22381/> (29.10.2013).
5. Кучеренко С. СХД для «среднячков» [Электронный ресурс] // – Режим доступа: [http://sib.com.ua/arhiv\\_2009/2009\\_3/shd/3\\_2\\_2009.htm](http://sib.com.ua/arhiv_2009/2009_3/shd/3_2_2009.htm) (29.10.2013).
6. Сюртуков И. Современные системы хранения данных [Электронный ресурс] // – Режим доступа: <http://citforum.ru/hardware/data/overview/> (29.10.2013).
7. Дюран Алаби NAS, DAS или SAN: выбор технологии хранения [Электронный ресурс] // – Режим доступа: <http://rus.625-net.ru/625/2005/02/nas.htm> (29.10.2013).
8. Голубев Д. Сети хранения данных (SAN) / Д. Голубев, А. Лобанов // Jet Info: Информационный бюлетень. – 2002. – № 9 (112).
9. Борзенко А. Конвергенция SAN и NAS // Intelligent Enterprise. – 2003.– №18 (83) [Электронный ресурс] // – Режим доступа: <http://www.iemag.ru/analytics/detail.php?ID=18014> (29.10.2013).