

ПРОГРАМНО ВИЗНАЧЕНІ СИСТЕМИ ЗБЕРІГАННЯ ДАНИХ. ПЕРЕВАГИ І ОСОБЛИВОСТІ

*ДП «ЕС ЕНД ТІ УКРАЇНА», м. Київ, Україна

Анотація. З кожним роком зростання обсягів даних відбувається експоненціально, що тягне за собою збільшення як кількості, так і ємностей систем зберігання даних. Найбільшого вмісту вимагають системи зберігання даних, що використовуються для зберігання резервних та архівних даних; файлових сховищ із загальним доступом, середовищ тестування і розробок, сховищ віртуальних машин, корпоративних або публічних web-сервісів. Для реалізації подібних задач на сьогоднішній день виробники пропонують три типи систем зберігання даних: блокові та файлові сховища, що стали вже традиційними в побудові ІТ-інфраструктури, та третій тип – програмно визначені системи зберігання даних. Вони дозволяють створювати сховища даних на неспеціалізованому обладнанні, як правило, групі серверних вузлів архітектури x86-64 під управлінням операційних систем загального призначення. Основною особливістю програмно визначених систем зберігання даних є перенесення функцій зберігання з апаратної частини на програмний рівень, де ці функції визначаються не фізичними можливостями обладнання, а програмного забезпечення, обраного, виходячи з реалізації задач. На сьогодні можна виділити три основні технології, які характеризуються масштабуючою архітектурою, що дозволяє нарощувати продуктивність і об'єм дискових ресурсів шляхом додавання нових вузлів в єдиний пул: Ceph, DELL EMC VxFlex OS, HP StoreVirtual VSA. Програмно визначені системи зберігання даних мають такі переваги, як відмовостійкість, продуктивність, гнучкість і економічність. Застосування програмно визначених систем зберігання даних дозволяє підвищити ефективність використання ІТ-інфраструктури і знизити витрати на її підтримку; створювати гібридну інфраструктуру, що дозволяє використовувати як власні, так і зовнішні хмарні ресурси; підвищити ефективність роботи як сервісів, так користувачів, забезпечивши при цьому можливість надійного підключення за рахунок застосування найбільш зручних пристроїв; створювати портал як єдину точку обслуговування і контролю за ресурсами.

Ключові слова: системи зберігання даних, програмно визначені системи, інфраструктура, технології, серверні вузли, пул, кластер, дискові масиви, демони, сервіси, архітектура, масштабованість.

Abstract. Every year the amount of generated data grows exponentially which entails an increase in both the number and capacity of data storage systems. The highest capacity is required for data storage systems that are used to store backups and archives, file storages with shared access, testing and development environments, virtual machine storages, corporate or public web services. To solve such tasks, nowadays manufacturers offer three types of storage systems: block and file storages which have already become a standard used for implementing IT infrastructures, and software-defined storage systems. They allow to create data storages on non-specialized equipment, such as a group of x86-64 server nodes managed by general-purpose operating systems. The main feature of software-defined data storages is the transfer of storage functions from the hardware level to the software level where these storage functions are defined not by physical features of the hardware but by the software selected for specific tasks solving. Today there are three main singled out technologies characterized by scalable architecture that allow to increase efficiency and storage volume through adding new nodes to a single pool: Ceph, DELL EMC VxFlex OS, HP StoreVirtual VSA. Software-defined data storages have the following advantages: fault tolerance, efficiency, flexibility and economy. Utilization of software-defined storages allows to increase efficiency of IT infrastructure and reduce its maintenance costs; to build a hybrid infrastructure that would allow to use internal and external cloud resources; to increase efficiency of both services and users by providing reliable connection by using the most convenient devices; to build a portal as a single point of services and resources control.

Keywords: data storage systems, software-defined systems, infrastructure, technologies, server nodes, pool, cluster, disk arrays, domains, services, architectures, scalability.

1. Вступ

У сучасному світі кількість інформації, яка використовується і накопичується, зростає дуже динамічно. З кожним роком об'єми даних збільшуються експоненціально, що, у свою чергу, сприяє збільшенню як кількості систем зберігання даних (СЗД), так і вмісту самих СЗД. Найбільшого вмісту, як правило, вимагають СЗД, що використовуються для зберігання резервних та архівних даних; файлових сховищ із загальним доступом, середовищ тестування і розробок, сховищ віртуальних машин, корпоративних або публічних web-сервісів [1].

Для реалізації подібних задач на сьогоднішній день виробники пропонують три типи систем зберігання даних (рис. 1): блокові та файлові сховища, що стали вже традиційними в побудові ІТ-інфраструктури, та третій тип, що все більше набуває популярності, програмно визначені сховища СЗД (Software-defined storage, SDS). Вони дозволяють створювати сховища даних на неспеціалізованому обладнанні, як правило, групі серверних вузлів архітектури x86-64 під управлінням операційних систем (ОС) загального призначення (Linux, Windows, FreeBSD) [2].

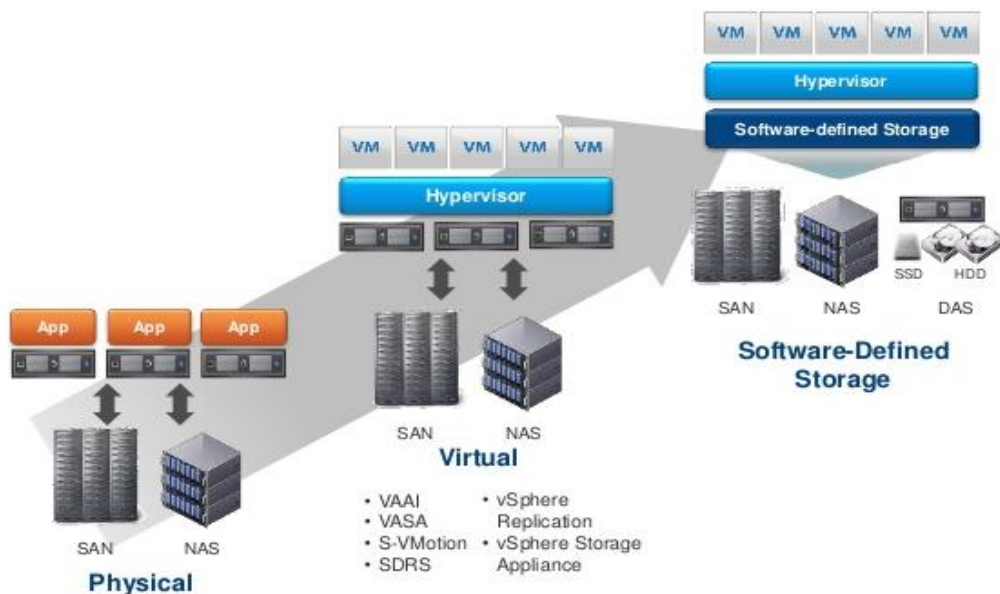


Рисунок 1 – Типи СЗД

Мета даної статті – розглянути основні технології створення програмно визначених СЗД, а також їх переваги та особливості.

2. Основні технології створення програмно визначених СЗД

На сьогодні можна виділити три основні технології, які характеризуються масштабуючою (scale-out) архітектурою, яка дозволяє нарощувати продуктивність і об'єм дискових ресурсів шляхом додавання нових вузлів в єдиний пул.

2.1. Технологія Serf

Це технологія створення програмно визначених СЗД із відкритим вихідним кодом, розроблена як система зберігання з високою стійкістю до відмов за рахунок горизонтальної масштабованості всіх компонентів [3]. Завдяки такому підходу досягається висока продуктивність у порівнянні із традиційними масивами, які зазвичай повинні направляти всі операції введення/виведення через пару своїх контролерів. Технологія Serf дозволяє відновити ві-

дмову окремого диска великого об'єму протягом декількох годин, що значно швидше, ніж у традиційних системах із RAID, де подібна операція може зайняти значно триваліший проміжок часу – від доби до декількох тижнів. Сьогодні Ceph є найбільш популярною технологією для середовища OpenStack, так як вони обидві з відкритим вихідним кодом і вже мають достатньо масштабну реалізацію. Для корпоративних замовників, що вимагають високого рівня технічної підтримки цих технологій, компанія RedHat пропонує свої технології Red Hat Ceph Storage та Red Hat OpenStack Platform [4].

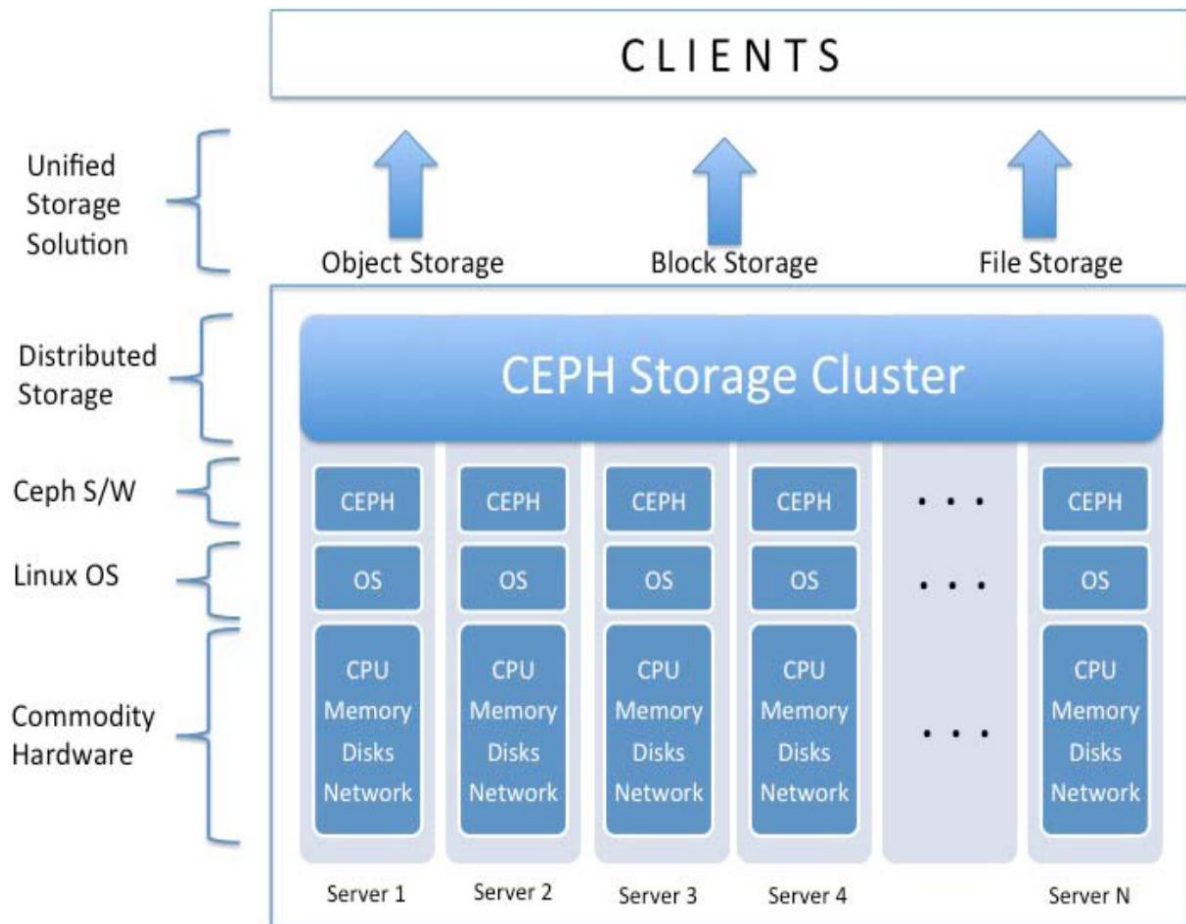


Рисунок 2 – Архітектура програмно визначених СЗД на технології Ceph

Програмно визначені СЗД на базі технології Ceph можливо реалізувати на загально-доступному обладнанні з ОС LINUX і розгорнути сервіси (демони Ceph), які виконують різні ролі кластера:

- Mon – демон монітора. Виконує роль координатора і містить інформацію про «здоров'я» та стан кластера, обмінюючись даними з іншими моніторами. Цей демон встановлюється і налаштовується першим, і як тільки запускається один робочий монітор, ми отримуємо Ceph-кластер. Клієнти звертаються до моніторів, щоб дізнатися, на які OSD писати/читати дані.

- OSD – демон сховища. Другий демон по черзі розгортання в Ceph-кластері. Виконує роль сховища для запису даних користувачами. Після закінчення розгортання цього демона вже отримуємо кластер із доступними ресурсами і готовий до роботи.

- MDS – сервер метаданих. Встановлюється, якщо планується розгортання файлової системи CephFS.

Крім демонів, у Ceph реалізовано три основні типи зберігання даних:

- RADOS Gateway – об'єктні. Дозволяє додаткам встановлювати з'єднання із сховищем об'єктів Ceph. Забезпечує додатки API-інтерфейсом, сумісним із RESTful S3/Swift для зберігання даних у форматі додатків.

- RADOS Block Device – абстракція блочного пристрою, що надає користувачам можливість створювати і використовувати віртуальні блокові пристрої будь-якого розміру. Інтерфейс RADOS Block Device дозволяє працювати з цими пристроями в режимі читання/запису і виконувати такі операції: зміна розміру, клонування, створення і повернення до знімка стану і т.д.

- CephFS – файлові системи.

Першопочатковий кластер Ceph можна створити з декількох машин, поєднуючи на них ролі кластера. Потім, зі збільшенням кластера і додаванням нових серверів, ролі можна дублювати на інших машинах або повністю виносити на окремі сервери (рис. 2).

2.2. Технологія DELL EMC VxFlex OS

Представляє собою технологію корпоративного рівня, що дозволяє масштабувати СЗД до рівня Центру обробки даних (ЦОД). У даному рішенні застосовуються принципи серверної візуалізації до локальних дисків серверів x86, перетворюючи їх у високопродуктивний і загальнодоступний блоковий пул. Додаток більше не буде обмежено ресурсами локального сервера, так як ресурси об'єднані у кластер і використовуються спільно. Такий підхід дозволяє підтримувати оптимальний баланс між усіма ресурсами кластера, задовольняючи вимоги і потреби в ресурсах кожного з додатків (рис. 3).

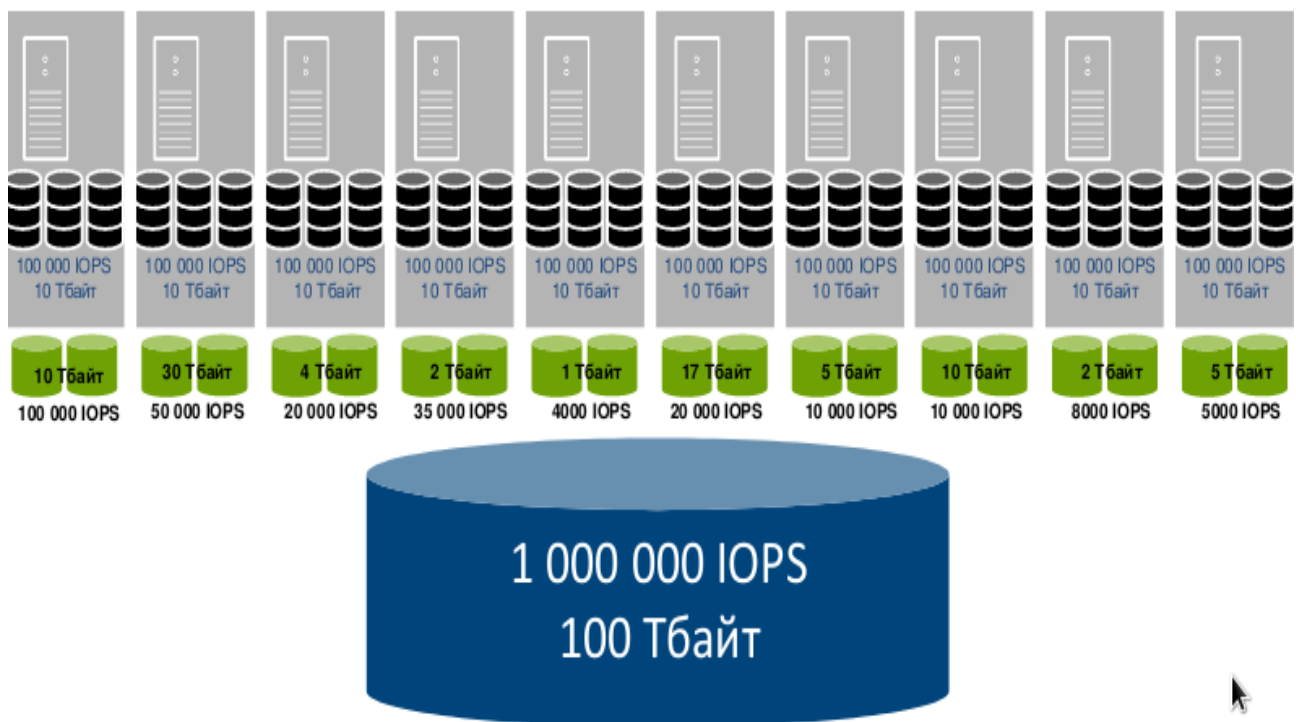


Рисунок 3 – Об'єднання серверних ресурсів в єдиний пул

VxFlex OS безперервно здійснює балансування доступних ресурсів, що, у свою чергу, дозволяє додавати нові ресурси зберігання або обчислення до вже працюючого пулу без його зупинки та впливу на інші працюючі додатки, які використовують ресурси цього пулу [5]. Компоненти програмного забезпечення (ПЗ) VxFlex OS:

- Storage Data Client (SDC) – ПЗ клієнт для вузлів, які будуть використовувати томи зберігання даних. Даний клієнт здійснює одночасні операції читання і запису даних для

всіх точок підключення, що у випадку збою дозволяє підтримувати операції введення і виведення в обхід збійних шляхів, так як дані і копії томів рівномірно розподіляються між вузлами і накопичувачами.

- Storage Data Server (SDS) – ПЗ для вузлів, що надають свої жорсткі диски та/або Flash-накопичувачі як ресурси зберігання для всієї системи, з яких надалі створюються пули зберігання даних.

- Metadata Manager (MDM) – ПЗ для вузлів, що утворюють кластер, який контролює всі операції як системи в цілому, так і окремих її частин. Встановлюється на трьох або п'яти різних вузлах, може бути розгорнутий на вузлах з SDC та/або SDS. При розгортанні на окремих вузлах може розміщуватися за межами шляху передачі даних. Забезпечує контроль налаштувань, моніторинг, балансування, відновлення пулів, поле збою.

Рішення на базі VxFlex OS рекомендується підприємствам і організаціям, які вимагають від СЗД високих показників IOPS і гарантовано низького часу відгуку.

2.3. Технологія HP StoreVirtual VSA

Одна з перших технологій створення програмно визначених СЗД для будь-якого рівня масштабування системи зберігання, починаючи від малих систем, які об'єднують 3 сервери, і закінчуючи крупними ЦОД великих підприємств і організацій. Таким чином, основним застосуванням технології є вирішення задач віртуалізації ресурсів.

Технологія підтримує найпопулярніші гіпервізори, такі як VMware vSphere, Microsoft Hyper-V і KVM [6]. Для vSphere і Hyper-V виробником пропонуються допоміжні плагіни для інтеграції з консолями управління, що дозволяє виконувати задачі адміністрування і моніторингу з єдиної точки.

SV VSA встановлюється безпосередньо на вузли кластера, які взаємодіють між собою за технологією Network RAID (рис. 4). Фактично - це синхронна реплікація між вузлами кластера, а кількість копій блоків даних буде регулюватися рівнем RAID-масиву. Наприклад, кілька вузлів об'єднані у кластер, далі створений віртуальний диск із мережним RAID 10. У даному випадку кожен блок даних на цьому віртуальному диску буде дублюватися між вузлами кластера, і у разі виходу з ладу одного вузла кластер RAID 10 гарантує збереження даних і доступ до них.

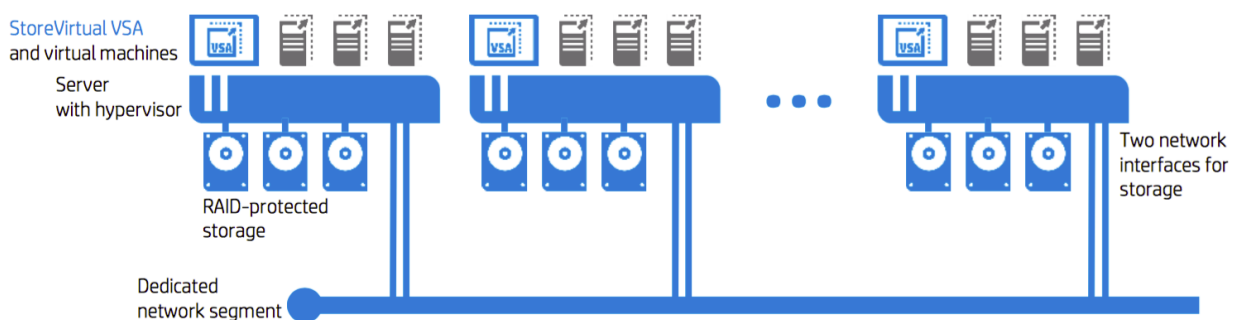


Рисунок 4 – Архітектура StoreVirtual VSA

В одному кластері можна реалізувати різні рівні Network RAID:

- Network RAID level 10 + 1. У кластері зберігаються 3 копії кожного блока даних;
- Network RAID level 10 + 2. Будується, як мінімум, на чотирьох вузлах і створює копії блоків на всіх вузлах кластера;
- Network RAID 5 і 6. На цих рівнях додаються блоки з контрольними сумами, де блоки даних і блоки з контрольних сум пишуться на різні вузли кластера. Такий підхід дозволяє знизити overhead і отримати відмовостійкість на рівні Network RAID level 10 і level

10+1, але при цьому істотно знижується продуктивність усього кластера через необхідність зчитування цих контрольних сум.

Використання StoreVirtual VSA рекомендується для підприємств і організацій із віртуальною інфраструктурою, що динамічно розвивається, одна з основних задач якої – створення відмовостійкого захищеного сховища для всього кластера.

3. Переваги та особливості програмно визначених СЗД

Ідея створення програмно визначених СЗД у цілому не нова, але отримала свій розвиток при значному підвищенні обчислювальної потужності x86 серверів, збільшенні пропускної здатності мережевої інфраструктури і активному впровадженні хмарних технологій. Програмно визначені СЗД мають такі переваги.

Відмовостійкість. Одна з найважливіших властивостей будь-якої СЗД. На відміну від традиційних систем, де резервування відбувається на рівні апаратної частини шляхом подвійного резервування всіх модулів (блоки живлення, контролери і т.д.), програмно визначені СЗД вирішують подібні завдання на рівні ПЗ, створюючи копії даних, зберігаючи їх на фізично різному обладнанні. У випадку виходу з ладу одного з серверів дані, за допомогою інструментів реплікації, легко переміщуються на резервну систему, і робота при цьому не зупиняється.

Продуктивність. Підтримка технологій NVMe і NVMe over Fabrics у програмно визначених СЗД дозволяє істотно збільшити продуктивність операцій введення/виведення, що дає можливість організувати багаторівневе зберігання, наприклад, включивши туди пристрої NVDIMM, NVMe і HDD.

Економічність. Не вимагають використання обладнання якогось одного виробника і дозволяють використовувати стандартизоване обладнання будь-якого виробника, що, у свою чергу, знижує вартість володіння ІТ-інфраструктурою. Можливість відмовитися від однієї або декількох класичних СЗД і задіяти дисковий простір серверів, які мають більший об'єм, при цьому буде істотно дешевше.

Гнучкість. Централізоване управління дозволяє контролювати та управляти усіма ресурсами системи зберігання і при необхідності легко масштабувати її горизонтально, що дозволяє нарощувати як об'єм дискового простору, так і продуктивність усього кластера до необхідних показників.

Основною особливістю програмно визначених СЗД є перенесення функцій зберігання з апаратної частини на програмний рівень, де ці функції визначаються не фізичними можливостями обладнання, а ПЗ, обраним, виходячи з реалізації задач. Застосування програмно визначених СЗД дозволяє підвищити ефективність використання ІТ-інфраструктури і знизити витрати на її підтримку; створювати гібридну інфраструктуру, що дозволяє використовувати як власні, так і зовнішні хмарні ресурси; підвищити ефективність роботи як сервісів, так користувачів, забезпечивши при цьому можливість надійного підключення за рахунок застосування найбільш зручних пристроїв; створювати портал як єдину точку обслуговування і контролю за ресурсами.

4. Висновки

У рамках однієї статті досить складно описати весь спектр та можливості програмно визначених СЗД, так як останнім часом вони активно розвиваються, і на ринку з'являються нові компанії, які пропонують свої технології як для вузькоспеціалізованих задач, так і для широкомасштабних. Розвиток хмарних технологій спільно із програмно визначеними технологіями кардинально змінює підхід до стандартних та до апробованих ІТ-рішень. Уже сьогодні такі моделі надання сервісів, як Paas, Iaas, Saas, набагато переважніші для багатьох підприємств і організацій, так як допомагають знизити витрати на володіння ІТ-

інфраструктурою за рахунок використання недорогого обладнання, що забезпечує таку ж відмовостійкість, як і дорогі СЗД High-end класу.

СПИСОК ДЖЕРЕЛ

1. Лисецкий Ю.М. Технологии хранения и управления данными. *Актуальные проблемы информационных систем и технологий*: монография / Н. Борисова, В. Васянин и др.; под науч. ред. проф. В. Вычужанина. Одесса: НУ «ОМА». 2020. 296 с.
2. Lysetskyi Y.M., Seredovich Y.P. Software-Defined Storages as a Tool of Lowering IT Infrastructure Costs. *Управляющие системы и машины*. 2019. № 4. С. 42–48.
3. Документация по продукту CEPH. URL: <https://docs.ceph.com/docs/master/>.
4. Сборник авторских статей. URL: <http://vasilisc.com/ceph>.
5. Обзор решения VxFlex OS. URL: https://www.dellemc.com/kk-kz/collaterals/unauth/white-papers/products/converged-infrastructure/vxflex_overview_whitepaper.pdf.
6. Документация продукта StoreVirtual VSA. URL: <https://h20195.www2.hp.com/v2/default.aspx?cc=kz&lc=ru&oid=5306917>.

Стаття надійшла до редакції 25.01.2021