

Нові рішення Oracle та Hitachi Vantara

перевизначають межі корпоративних датацентрів

З розвитком технологій стають можливими речі, що раніше здавалися недосяжними. Наприклад, зберігання геномів і архівних даних про населення цілої країни.

У світі, де дані перетворилися на найцінніший актив, а вимоги до швидкості, доступності та безпеки сягають критичних відміток, інфраструктурні рішення перестали бути просто «залізом». Компанія SNT Ukraine вже більш ніж 30 років тримає руку на пульсі новітніх технологій й допомагає замовникам обирати найперспективніші варіанти впровадження IT-рішень, зазвичай на основі продуктів глобальних вендорів. Два такі гравці, Oracle зі своєю платформою Exadata Database Machine та Hitachi Vantara з VSP One Block, представили нові ітерації своїх флагманських продуктів, які не просто оновлюють технічні специфікації, а пропонують принципово новий погляд на архітектуру корпоративного сховища та обробку даних.

Інфраструктура БД по-новому з Oracle Exadata

Oracle Exadata давно є синонімом неперевершеної продуктивності для баз даних Oracle, однак нові моделі та програмні оновлення переходять від логіки «апаратного прискорення» до концепції повноцінної «автономної інфраструктури».

1. Exadata на базі процесорів Intel Xeon Scalable 4-го покоління (Sapphire Rapids) та новітніх NVMe-накопичувачів. Це дає зростання продуктивності процесорних ядер до 40%, а пропускної здатності — до 2,4 раза. Але суть не в числах, а в тому, як ця потужність керується.

2. Autonomous Health Framework та AI-оптимізація. Система тепер набагато глибше інтегрує штучний інтелект для проактивного виявлення, діагностики та навіть усунення понад 99% потенційних проблем до того, як вони вплинуть на бізнес-процеси.

Це стосується не лише апаратної частини, а й роботи СУБД: автоматична індексація, налаштування запитів, розподіл ресурсів у реальному часі.

3. Exadata Cloud@Customer як стандарт. Oracle остаточно прибирає межі між публічною хмарою та локальним датацентром. Замовник отримує повний стек Exadata як сервіс у власному ЦОД, з повною сумісністю API, керуванням через OCI Console і єдиною моделлю оплати. Це відповідь на вимоги щодо суверенітету даних, низької затримки та міграції без змін у застосунках.

4. Революція у резервуванні: Always-On Duplexing. Нова програмна архітектура одночасно записує дані на двох незалежних серверах зберігання через RDMA-мережу. Це забезпечує нульові втрати даних (RPO=0) і відновлення за лічені секунди навіть при повній відмові одного з серверів і без впливу на продуктивність. Раніше подібний рівень надійності вимагав значно складніших і дорожчих рішень.

Компанії, прив'язані до Oracle Database, отримують не просто швидше апаратне рішення, а самокеровану еластичну платформу, яка функціонує як продовження хмари. Це кардинально знижує операційні витрати, звільняє адміністраторів БД для стратегічних завдань та забезпечує безпрецедентний рівень SLA для найкритичніших навантажень — фінтеху, телекому, держсектору.

Hitachi VSP One Block: міст між поколіннями

Hitachi Vantara, легенда ринку систем зберігання, зробила неочікуваний і блискучий хід. Замість того, щоб створити ще один масив для

flash-накопичувачів, вони переосмислили свою флагманську платформу Virtual Storage Platform (VSP), створивши VSP One Block. І це не оновлення, а нова архітектурна філософія.

1. Єдина платформа для всіх блокових робочих навантажень. Система розроблена як єдиний масив для будь-яких блокових даних: від мейнфреймів IBM до відкритих систем (x86, UNIX) та, що найважливіше, сучасних контейнеризованих середовищ. Це кладе край ері фрагментованої інфраструктури, що тривала десятиріччями.

2. Повна програмна визначеність та підтримка хмарних практик. VSP One Block повністю побудована на програмно-визначеній архітектурі, що працює на стандартних серверах. Це дає неймовірну гнучкість: ресурси можна динамічно перерозподіляти між різними типами навантажень.

3. Безкомпромісна продуктивність і нульовий вплив на основні операції. Навіть будучи програмно-визначеною, система використовує власні ASIC-чіпи для апаратного прискорення операцій введення-виведення та шифрування. Затримки залишаються мінімальними та передбачуваними, що критично для мейнфрейм-навантажень. Оновлення мікропрограми, масштабування та технічне обслуговування тепер виконуються без простою (Non-Disruptive Operations).

4. Глибока аналітика на базі AI (Hitachi Ops Center Analyzer). Платформа не просто зберігає дані, а аналізує патерни доступу, прогнозує гарячі точки, автоматично переміщує дані між різними типами накопичувачів (включно з QLC Flash) для оптимізації вартості. Для мейнфрейм-середовищ це є особливо цінним, оскільки дає безпрецедентну видимість у роботі традиційно «закритих» систем.



Рис. 1. Приклад метричної книги з архіву ЦДІАК



Рис. 2. Екстракція (процес виділення з клітин) ДНК

Компанії з гетерогенними інфраструктурами, що мають багато «успадкованих» компонентів, отримують шанс на консолідацію та модернізацію без ризику. Можна поступово переносити навантаження з мейнфреймів у контейнери, використовуючи ту саму надійну платформу зберігання. Це «міст між епохами», що захищає мільйонні інвестиції минулого та відкриває шлях до майбутніх технологій.

Завдання: сховище геномів

Найкраще уявлення про перелічені вище можливості можна отримати, обравши для прикладу масштабний та складний проект і спробувавши реалізувати його за допомогою новітніх технологій. Уявімо систему «Цифровий Ковчег», в якій знаходяться дані про 10 мільйонів повногеномних ДНК людей та оцифровані архіви України — і де ви за лічені хвилини можете знайти свого предка 1750 року та побачити його генетичний зв'язок з вами через ідеальну збірку генома T2T (Telomere-to-Telomere).

Історія роду людського захована у двох видах кодів: перший — це генетична інформація, записана у нашій ДНК, другий — рукописні нотатки в архівних документах (рис. 1). Стародавні рукописні паперові артефакти можуть будь-якої миті бути втрачені навки, але сучасні технології дають надію, що ми стоїмо на порозі технологічної революції і у галузі збереження історичної спадщини.

В державних архівах України зберігаються мільйони справ. Якісне сканування одного аркуша рукописного тексту генерує дані розміром приблизно 50 МБ. Загальний обсяг робіт навіть уявити складно, але припускаю, що тільки в архівах України доведеться оцифрувати не менш ніж 200 мільйонів аркушів. Масштабуючи на весь потенційний проект, отримуємо число $200\,000\,000 \times 50\text{ МБ} = 10$ петабайт. Варто зазначити, що процес оцифрування українських архівів уже насправді розпочато.

Повногеномне секвенування (Whole Genome Sequencing) стало доступним звичайній людині лише кілька років тому й генерує колосальні обсяги даних. Сучасний повногеномний ДНК-аналіз (рис. 2) в форматі T2T-CHM13 тільки для однієї людини генерує приблизно 150 ГБ. Масштабування на 10 мільйонів людей вже дає приблизно 1500 петабайтів (1,5 ексабайта).

Парадокс: уся писемна історія України займе в оцифрованому вигляді у 150 разів менше місця, ніж оцифрована ДНК лише третини поточного населення країни. 1,5 ексабайта даних — це не просто «багато», це більше, ніж сумарний обсяг даних більшості соціальних мереж десятирічної давнини. Зберегти — це півроботи, головне завдання — проаналізувати ці дані, знайти в отриманому океані інформації реальні взаємозв'язки та структурувати їх, згенерувавши метадані.

Потрібно задіювати дедуплікацію та компресію і сучасні методи обробки та аналізу інформації. Розрахуємо для нашого уявного проекту потенційний об'єм «сирих» даних (табл. 1).

Зберегти такий неоптимізований обсяг майже неможливо навіть на рівні держави. Чи можна щось зробити у подібній ситуації? Насправді можна.

Вдосконалюємо зберігання

Як інженер і генеалог, я знаю: ДНК — це не хаотичний набір байтів, а потенційно структуровані дані, які ідеально піддаються стисненню. Людські геноми ідентичні на 99.9%. Тому пропонуємо вашій увазі приклад розрахунку того, як створити елегантний та швидкий датацентр, використовуючи синергію Oracle Exadata (для розумних даних) та Hitachi VSP One (для масиву файлів).

Для кожного типу даних ми застосуємо свою «особливу можливість» (табл. 2):

- геноми — використаємо Reference-based compression + апаратну дедуплікацію Hitachi;
- архіви — апаратна дедуплікація повторюваних фонів паперу та бланків (Hitachi ADR);
- метадані — гібридне колонкове стиснення (Oracle HCC).

Завдяки тому, що ми зберігаємо референсний геном один раз, а далі — лише унікальні мутації, 1500 ПБ геномних даних можуть перетворитися

Таблиця 1. Розрахунок «сирих» даних (Raw Capacity)

Тип даних	Об'єм одиниці	Кількість	Загальний «сирий» обсяг
Геном (T2T WGS)	~150 ГБ (CRAM/BAM)	10 000 000 осіб	1500 ПБ (1,5 Ексабайта)
Архівний скан	~50 МБ (TIFF)	200 000 000 сторінок	10 ПБ
Метадані/Індекси	-	-	50 ПБ
Разом			~1,6 Ексабайта

Таблиця 2. Технології та обладнання і потенційний результат їх використання

Тип даних	«Сирий» обсяг	Технологія та обладнання	Коефіцієнт (Ratio)	Фінальний об'єм на дисках
Геном (T2T)	1500 ПБ	Bio-Format (CRAM) + Hitachi VSP Dedupe Ми не зберігаємо 3 млрд однакових літер для кожного. Однакові фрагменти зберігаємо лише один раз на фізичному рівні, використовуючи дедуплікацію та стиснення даних	~30:1	50 ПБ
Архіви (Скани)	10 ПБ	Hitachi Adaptive Data Reduction (ADR) Тисячі сторінок метричних книг мають однакове пожовкле тло і типографські рамки	4:1	2,5 ПБ
Метадані	50 ПБ	Oracle Exadata HCC (Query High) Таблиці з іменами, датами та індексами стискаються в колонки. Exadata тут беззаперечний лідер	12:1	~4,2 ПБ
ВСЬОГО	~1,6 Ексабайта	Синергія BioIT + Enterprise Storage	~21:1	~76,2 ПБ

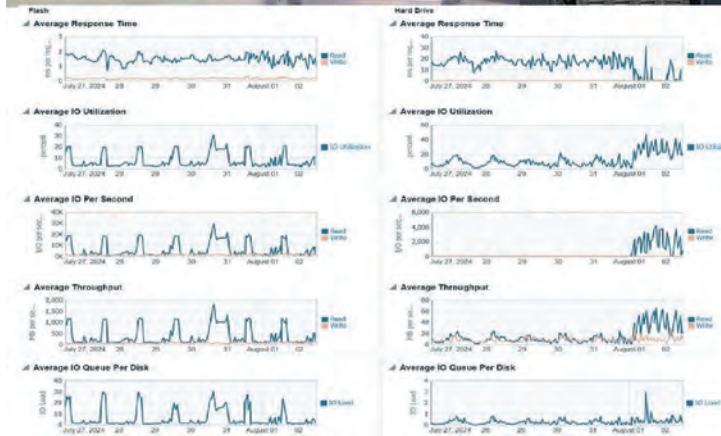
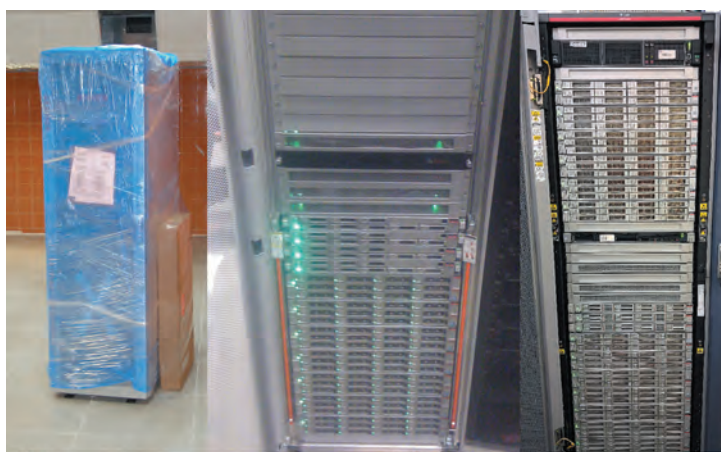


Рис. 3. Проекти з Oracle Exadata різних версій

на керовані 50 ПБ високодоступної інформації. Для такого проекту звичайні сервери та сховища даних не підійдуть.

«Мозок» системи: Oracle Exadata X10M

Рішення Oracle чудово підходять для аналізу даних. Ми завантажуємо сюди метадані (ПІБ, дати, генетичну інформацію та інше). Таким чином ми отримуємо аналітичне ядро. Компанія SNT Ukraine брала участь у впровадженні та супроводі програмно-апаратних комплексів Oracle Exadata різних версій в багатьох проектах (рис. 3), від телеком-операторів до державних організацій і банків, та переконалась у надійності та масштабованості даної технології. То чому саме Exadata?

- **SmartScan:** коли я роблю запит «Знайти всіх носіїв гаплогрупи I-FTB87564 у радіусі 50 км від Києва», Exadata не тягне всі дані в процесор, а фільтрує їх на рівні дисків зберігання. Це прискорює пошук у тисячі разів.
- **Hybrid Columnar Compression (HCC):** дані дуже повторювані. HCC стискає їх у 10–15 разів, дозволяючи тримати всю базу «в оперативці» та на швидкому флеші.

➤ **RDMA (Remote Direct Memory Access):** блискавичний обмін даними між вузлами без участі процесора. Аналітика в реальному часі.

➤ **AI Vector Search (Oracle 23ai):** векторний пошук у ДНК. Геном можна представити як послідовність векторів. Exadata X10M з її 192 ядрами на сервер дозволяє виконувати AI Vector Search прямо в базі даних. Ви можете робити SQL-запити на кшталт: «Знайди найближчих генетичних родичів для клієнта X в радіусі 5 поколінь».

➤ **Oracle Text & AI OCR:** Exadata ідеально підходить для індексації результатів OCR (оптичного розпізнавання тексту). Завдяки швидкості IOPS дослідник може за секунду знайти прізвище типу «Szczeniowski» серед мільярдів записів рукописного тексту 19 століття, розпізнаного нейромережею.

➤ **GraphDatabase:** генеалогія — це простір графів, які дозволяють відстежувати в межах багатьох поколінь, хто є чийм батьком, братом, дитиною тощо. Oracle Database на Exadata має вбудовану підтримку графів. Це дає можливість будувати «дерева роду» для цілої нації в реальному часі,

знаходячи неочевидні зв'язки між родами з різних регіонів проживання.

Сейф системи: Hitachi Vantara VSP One Block

Тут зберігаються «найважчі» файли: скани метрик та повні сирі геноми (якщо треба відновити оригінал). Компанія SNT Ukraine має давні партнерські стосунки з Hitachi Vantara, а наші інженери вже кілька десятиків років переконуються на реальних проектах (рис. 4) в бездоганній надійності рішень Hitachi по зберіганню даних та їх онлайн міграції. Чому Hitachi VSP One?

➤ **Adaptive Data Reduction (ADR):** це «розумна» дедуплікація «на льоту». Вона розпізнає типи даних. Якщо ми завантажуємо мільйони типових бланків, VSP One збереже лише один примірник, а решту замінить посиланнями (дедуплікація 4:1 і вище).

➤ **Багаторівневе зберігання (Tiering):** VSP One Block автоматично визначає, які геноми аналізуються саме зараз (наприклад, для пошуку родичів), і тримає їх на надшвидкому NVMe. Ті дані, що не використовувалися певний час, прозоро переміщуються на дешевші носії (QLC flash або

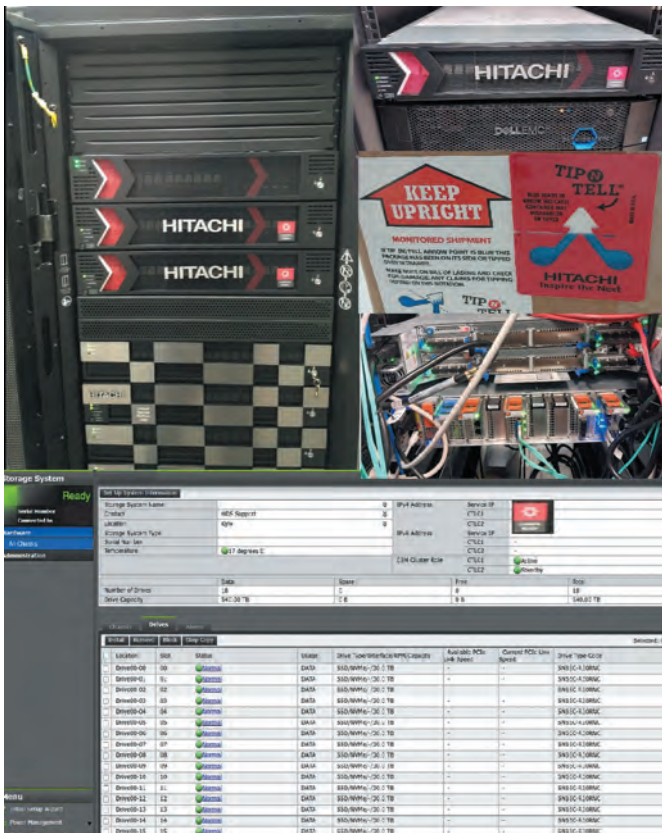


Рис. 4. SNT Ukraine впроваджує Hitachi Vantara VSP One Block



Рис. 5. Oracle Exadata X10M та Hitachi Vantara VSP One Block

об'єктне сховище), залишаючись доступними для системи.

➤ **Hitachi Ops Center Analyzer:** глибока аналітика на базі AI. Платформа не просто зберігає дані, а аналізує патерни доступу, прогнозує гарячі точки, автоматично переміщує дані між різними типами накопичувачів (включно з QLC Flash) для оптимізації вартості.

➤ **Immutable Snapshots (захист від вандалізму):** головна загроза для цифрового архіву — віруси-шифрувальники або навмисне видалення. Hitachi VSP One дозволяє робити апаратні знімки томів, які неможливо видалити чи змінити навіть з правами адміністратора протягом 50–100 років (WORM — Write Once, Read Many). Це гарантує, що метрична книга 1750 року залишиться цифровою майже навіки.

➤ **Глобальна дедуплікація:** багато архівних документів мають типові бланки або порожні поля. Hitachi VSP One ефективно оптимізує однакові блоки даних.

Синергія для сучасних інженерних задач

Чому не побудувати це на звичайних серверах («самозборі»)? Тому що на масштабі ексабайтів «дешево

залізо» стає «золотим» в обслуговуванні. Утримувати ексабайти даних — це мегавати електроенергії, нестача якої стає все більш актуальною по всьому світу. Hitachi VSP One має сертифікацію Carbon Footprint Reduction. Система вміє вимикати живлення окремих банків флеш-пам'яті, до яких немає звернень (наприклад, архів 17 століття, який рідко читають), і миттєво «будити» їх при запиті. Економія енергії — до 40%. Натомість Oracle Exadata X10M на процесорах AMD EPYC забезпечує найвищу щільність обчислень на ват енергії. Менше стійок — менше кондиціонерів і нижчі витрати на живлення (рис. 5).

Втрата скану 17-го століття чи даних про ДНК — неприпустима. Hitachi Vantara юридично гарантує, що дані не зникнуть (100% Data Availability Guarantee). Для аналізу інформації масштабних проектів також потрібна надійність,

адже згаяний час — це втрачені кошти. Oracle Exadata використовує технологію RAC (Real Application Clusters): якщо один сервер «падає», запити на пошук миттєво підхоплює інший, і користувачі навіть не помітять затримки інтерфейсу веб-порталу.

Ми живемо в унікальний час, коли синергія новітніх технологій та програмно-апаратних комплексів, таких як Oracle Exadata та Hitachi VSP One, дозволяють нам зробити те, про що раніше годі було й мріяти. І це не фантастика, це інженерне завдання, яке вирішується правильним поєднанням обчислювальної потужності Oracle та надійності зберігання Hitachi Vantara. Глибоке розуміння сучасних технологій та унікальний досвід інженерів, які десятиріччями складають ядро SNT Ukraine, дозволяють завжди обирати правильні рішення та реалізовувати проекти будь-якого масштабу.

Максим ТАТАРЕНКО, старший інженер
SNT Ukraine

+380 (44) 238-63-88
info@snt.ua, www.snt.ua

