

УДК 004.04

Копійка О. В., к.т.н.

(Державний університет телекомунікацій. +380 (44) 249 29 23). [okopiuka@gmail.com](mailto:okopiuka@gmail.com)

## ПРОЕКТУВАННЯ СЕРВІСІВ ЗБЕРІГАННЯ ДАНИХ

**Копійка О. В. Проектування сервісів зберігання даних.** У статті розглядається задача проектування сервісів зберігання даних в Дата-центрах, які визначаються відповідно до архітектури управління даними. Сервіси призначені для зберігання значущою, перевіреною, узгодженою, несуперечливою і хронологічно цілісною інформацією, яку з досить високим ступенем впевненості можна вважати достовірною. Основним об'єктом сервісу зберігання даних є сховище даних. Мета сховища – забезпечити цілісність і підтримувати хронологію всіляких корпоративних даних, і з цієї точки зору воно нейтрально по відношенню до додатків.

**Ключові слова:** Дата-центр, сервіси зберігання даних, сховище даних

**Копейка О. В. Проектирование сервисов хранения данных.** В статье рассматривается задача проектирования сервисов хранения данных в Дата-центрах, которые определяются в соответствии с архитектурой управления данными. Сервисы предназначены для хранения значимой, проверенной, согласованной, непротиворечивой и хронологически целостной информации, которую с достаточно высокой степенью уверенности можно считать достоверной. Основным объектом сервиса хранения данных является хранилище данных. Цель хранилища – обеспечить целостность и поддерживать хронологию всевозможных корпоративных данных, и с этой точки зрения оно нейтрально по отношению к приложениям.

**Ключевые слова:** Дата-центр, сервисы хранения данных, хранилище данных

**Kopiuka O. V. Projecting of data storage services.** The article deals with the problem of projection storage services in the data centers, which are determined in accordance with the architecture of data management. Services intended for data storage which are important, verified, agreed, consistent and chronologically integral, that can be considered authentic. The purpose of storage - to ensure the integrity and support the chronology of all kinds corporate data, and from this point of view it is neutral against to applications.

**Keywords:** Data center, data storage services, data Warehouse

**Постановка проблеми.** В Україні вже 85 % підприємств малого та середнього бізнесу використовують той чи інший «хмаровий сервіс». І ця кількість буде зростати: сьогодні «хмарові технології» пропонують практично безлімітні ресурси бізнесу будь-якого розміру при вартості, яка непорівнянна з капітальними інвестиціями у власну інфраструктуру.

При реалізації концепції «хмарових сервісів» виникає завдання побудови Дата-центрів з сучасною архітектурою і наявністю повного обсягу сервісів [1...13]. У даній статті будемо розглядати розвиток сервісів зберігання даних системної архітектури ІТ інфраструктури Корпорації, яка розвивається за рахунок створення власного Дата-центру.

Метою даних досліджень є вироблення стратегії розвитку Системної Архітектури ІТ інфраструктури Корпорації на основі застосування передових методологій і концепцій провідних виробників апаратного і програмного забезпечення (HP, SUN, EMC, CISCO, Microsoft, ORACLE, Veritas) [1].

Зазвичай, з точки зору логічної структури в Корпорації створюється 2-х рівнева організація центрів обробки даних (ЦОД): центральний і регіональні ЦОД [6]. Центральний ЦОД є віддаленим резервним ЦОД для регіональних ЦОД. Інтернет Центри Обробки Даних Корпорації також є дворівневими.

**Еталонний регіональний ЦОД.** Регіональний ЦОД з точки зору забезпечення захисту від загальноміських надзвичайних обставин складається з двох офісів, розташованих на відстані до 10 км один від одного: основного і резервного (Рис. 1). У резервному офісі можуть розміщуватися служби другої категорії, розробники, служби підтримки тощо.

При повній відмові основного офісу резервний офіс бере на себе функціонування критично важливих додатків за рахунок використання кластерного програмного забезпечення, тобто є так званим офісом "гарячої" заміни.

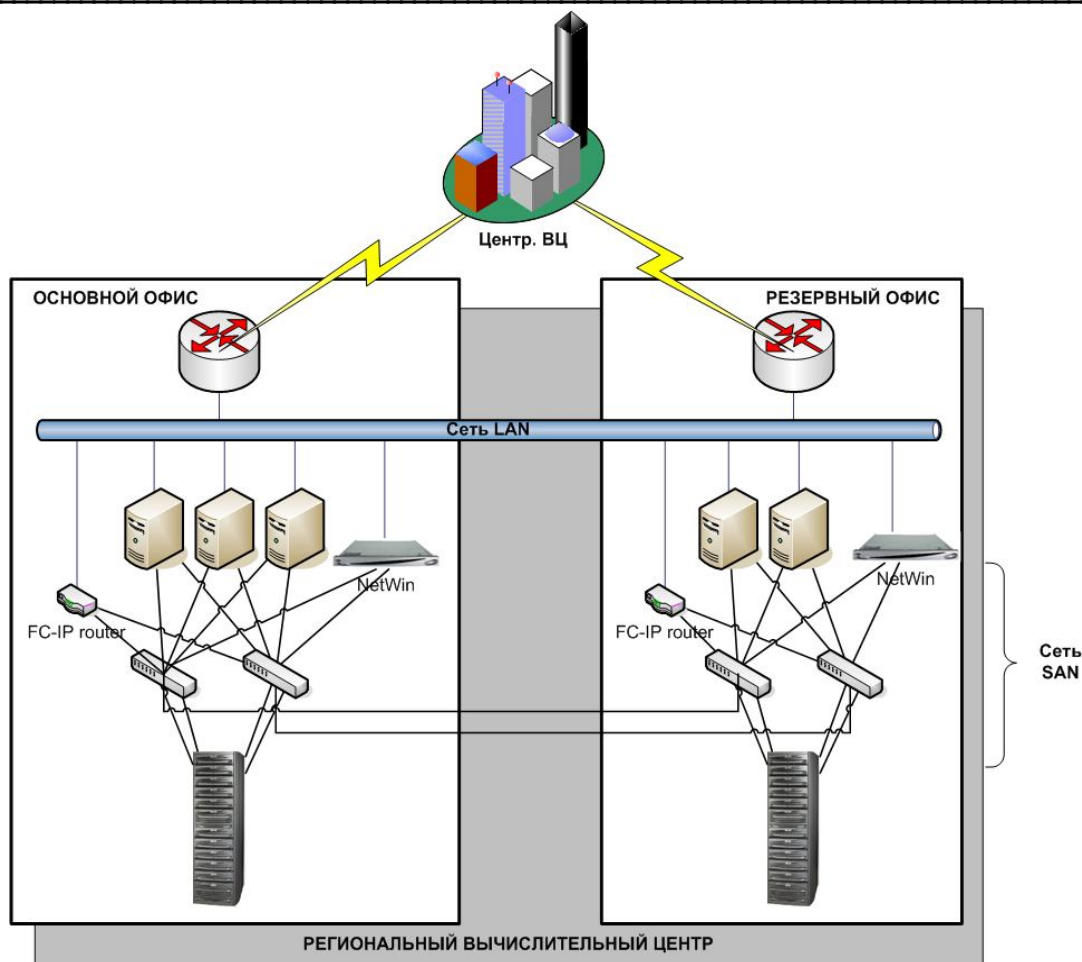


Рис. 1. Мережа зберігання даних в регіональному ЦОД

Система зберігання даних кожного локального офісу є системою без єдиної точки відмови:

- Сервери додатків вищої першої категорії підключаються до централізованого зовнішньому дискового масиву за допомогою Fibre Channel SAN і використовують для зберігання своїх даних високопродуктивні диски Fibre Channel.
- Сервери додатків другої категорії підключаються до централізованого зовнішньому дискового масиву за допомогою NAS пристрої NetWin і, по можливості, використовують для зберігання своїх даних відносно дешеві диски, наприклад CLARiiON ATA.
- Пристрій NetWin є сервером NetWorker локального резервного копіювання на відносно дешеві диски CLARiiON ATA.
- Дані додатків першої категорії синхронно реплікуються між офісами за допомогою вбудованого в дисковий масив програмного забезпечення MirrorView /S.
- Дані додатків першої категорії асинхронно реплікуються в центральний ЦОД за допомогою вбудованого в дисковий масив програмного забезпечення MirrorView / A і FC- to -IP маршрутизатора.
- Дані додатків другої категорії забезпечення безперервності бізнес операцій асинхронно реплікуються для цілей подальшої обробки (аналіз, централізоване резервне копіювання) в центральний ЦОД за допомогою програмного забезпечення RepliStor встановленого на пристроях NetWin.
- По можливості для Intel - based серверів використовуються технології віртуалізації VMware, в тому числі і технології переміщення працюючих додатків між різними фізичними серверами VMotion.
- Для централізованого управління ресурсами мереж SAN і сховищ даних використовуються програмні продукти VisualSAN і VisualSRM, відповідно.

- Для оптимізованого зберігання даних баз даних Oracle використовується DatabaseXtender .
- Для оптимізованого зберігання даних Exchange використовується EmailXtender .
- Для оптимізованого зберігання даних файлових систем використовується DiskXtender .

**Еталонний центральний ЦОД.** Центральний ЦОД (Рис. 2) відрізняється від регіонального ЦОД тільки додаванням функцій і, відповідно, апаратури, програмного забезпечення і технологій, які відповідають за централізовані резервне копіювання на магнітні стрічки і роботу з архівами.

Для організації централізованого архіву з on-line доступом в центральному ЦОД разом з впровадженням Documentum Enterprise Management Platform встановлюється CAS пристрій Centera, яке захищається від катастроф масштабу міста/регіону, реплікацією, за допомогою технології CentraStar, на резервний офіс центрального ЦОД.

Пристрій NetWin, який отримує інформацію з регіональних ЦОД, а також має доступ до локальних реплік дискового масиву CLARiiON, є сервером резервного копіювання NetWorker і використовується для організації LAN-free резервного копіювання на стрічкову бібліотеку, підключену за допомогою інтерфейсу Fibre Channel до SAN (виробництва ADIC , StorageTek або HP). Стрічкова бібліотека встановлюється в резервному офісі центрального ЦОД.

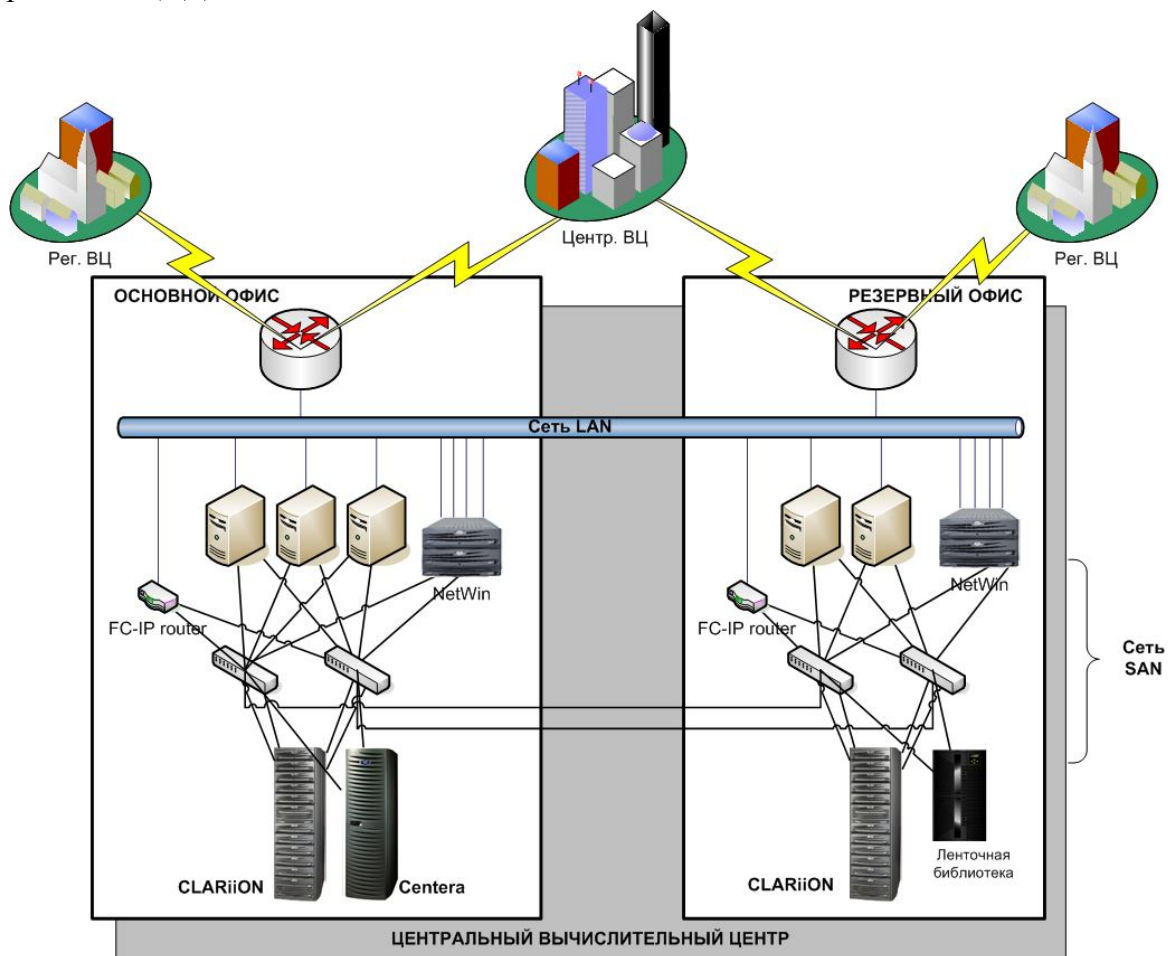


Рис. 2. Мережа зберігання даних в центральному ЦОД

**Зберігання даних.** Другий шар роботи з даними призначений безпосередньо для зберігання значущою, перевіреною, узгодженою, несуперечливою і хронологічно цілісною інформацією, яку з досить високим ступенем впевненості можна вважати достовірною.

Власне сховище даних не орієнтовано на рішення якої-небудь певної функціональної аналітичної задачі. Мета сховища – забезпечити цілісність і підтримувати хронологію всіляких корпоративних даних, і з цієї точки зору воно нейтрально по відношенню до додатків. У зв'язку з цим у більшості випадків для виконання певного комплексу функціонально замкнутих аналітичних завдань раціонально створювати вітрини даних, в основі яких може бути як багатовимірна, так і реляційна модель даних. По суті вітрина являє собою відносно невелике, але що найважливіше, функціонально-орієнтоване сховище, в якому інформація зберігається спеціальним чином, оптимізовано з точки зору вирішення конкретних аналітичних завдань деякого підрозділу чи групи аналітиків.

Зазвичай інформація потрапляє у вітрини зі сховища і в цьому випадку вітрини називаються залежними. Можлива також ситуація, коли джерелом інформації для поповнення вітрин служать безпосередньо оперативні та зовнішні транзакційні системи. Такі вітрини, що отримали назву незалежних, як правило, розглядаються як тимчасове рішення, що дозволяє досить швидко і з невеликими витратами вирішити найбільш важливі завдання, оцінити переваги нового підходу, сформулювати деякі рекомендації для більш масштабного проекту розробки загального сховища.

Сховище реалізується у вигляді реляційної бази даних, що працює під управлінням досить потужною реляційною СУБД (Рис. 3). Така СУБД повинна підтримувати ефективну роботу з терабайтними обсягами інформації, мати розвинені засоби обмеження доступу, забезпечувати підвищений рівень надійності і секретності, відповідати необхідним вимогам щодо відновлення та архівації тощо.

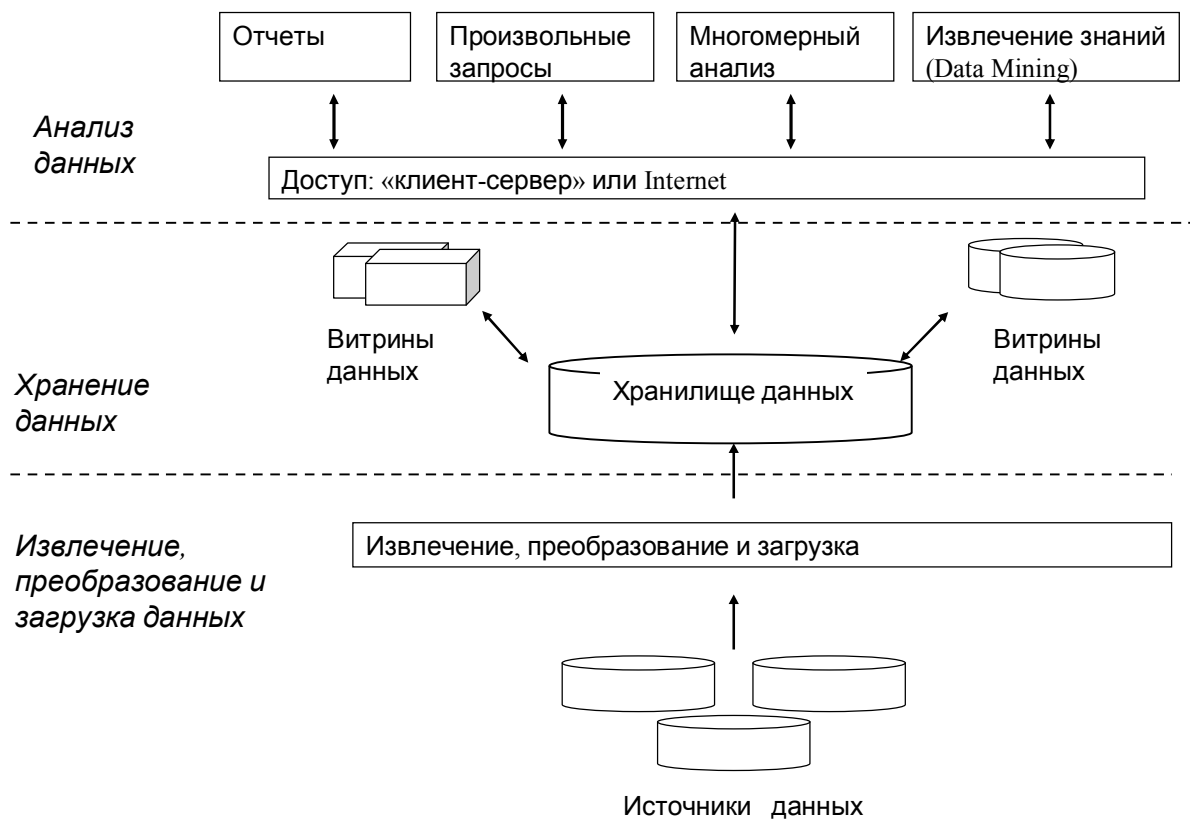


Рис. 3. Побудова моделі роботи з даними

Вітрини даних можуть будуватися на основі як реляційної, так і багатовимірної технології баз даних. Зазвичай для досить великої частини аналітичних додатків виявляється зручною та ефективною технологія інтерактивного багатовимірного аналізу і в цьому випадку вітрина являє собою багатовимірну базу даних, реалізовану в архітектурі OLAP, ROLAP або HOLAP.

Для організації доступу аналітиків до даних сховища і вітрин використовуються спеціалізовані робочі місця, підтримують необхідні технології як оперативного, так і довготривалого аналізу. Результати роботи аналітиків оформляються у вигляді звітів, графіків, рекомендацій і зберігаються як на локальному комп'ютері, так і в загальнодоступному вузлі локальної мережі.

Аналітична діяльність в рамках корпорації досить різноманітна і визначається характером розв'язуваних завдань, організаційними особливостями компанії, рівнем і ступенем підготовленості аналітиків. У зв'язку з цим сучасний підхід до інструментальних засобів аналізу не обмежується використанням якоїсь однієї технології. В даний час прийнято розрізняти чотири основних види аналітичної діяльності: стандартна звітність, нерегламентовані запити, багатовимірний аналіз (OLAP) і вилучення знань (data mining).

Кожна з цих технологій має свої особливості, певний набір типових завдань і повинна підтримуватися спеціалізованим інструментальним середовищем.

Самими продуктивними є рішення від компанії Oracle. Інструментальні засоби корпорації Oracle забезпечують повне інтегроване рішення для створення сховищ даних і ефективного використання накопиченої в ньому інформації (Рис. 4).

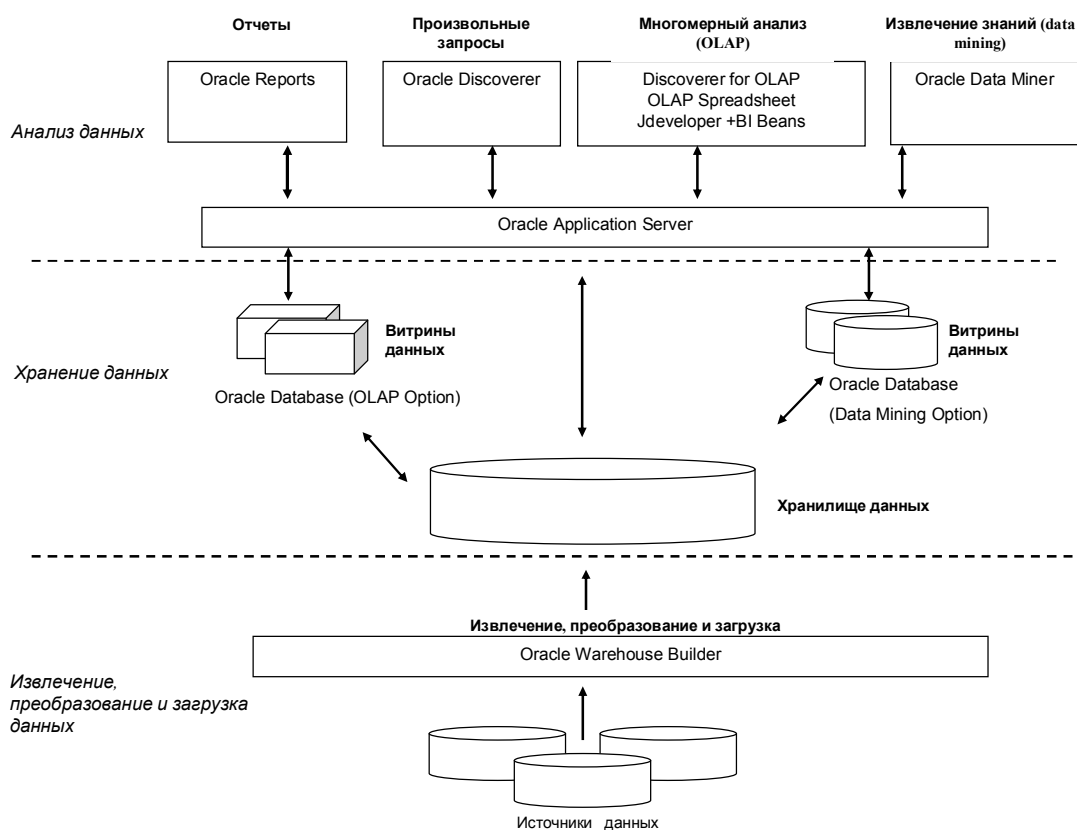


Рис . 4. Інструментальні засоби Oracle для побудови сховищ даних

Загальний перелік продуктів Oracle, необхідних для реалізації технології сховищ даних і аналітичних додатків, наводиться в Табл. 1 відповідно до виділених в попередньому розділі компонентами - вилучення та завантаження даних, зберігання, аналіз.

**Перелік продуктів Oracle для сховища даних**

**Табл. 1**

Тип засобів	Продукт	Коментарі
Витяг, перетворення та завантаження	<b>Oracle Warehouse Builder, ETL-засоби Oracle Database, Oracle Workflow</b>	Витяг, перетворення і загрузка. Підтримка процесів вилучення, перетворення і завантаження даних в сховище.
Зберігання даних	<b>Oracle Database</b>	СУБД для сховища даних і реляційних вітрин даних
	<b>Oracle OLAP Option</b>	Опція СУБД для багатовимірних вітрин даних
Аналіз даних	<b>Oracle Reports</b>	Регламентована звітність
	<b>Oracle Discoverer</b>	Довільні запити
	<b>Oracle Discoverer for OLAP, OLAP Spreadsheet, Jdeveloper + BI Beans</b>	Багатовимірний аналіз
	<b>Oracle Data Mining Option + Oracle Data Miner</b>	Опція СУБД для доступу до знань (data mining) і графічний інтерфейс
Адміністрування	<b>Oracle Enterprise Manager</b>	Адміністрування сховища

Як середовище зберігання інформації в реляційних сховищах і вітринах даних використовується сервер Oracle Database. Центральним інструментальним засобом створення сховищ і вітрин є Oracle Warehouse Builder, побудований на базі сучасної архітектури Common Warehouse Metadata. Він призначений для опису структури сховища і вітрин, проектування і створення процедур вилучення, погодження та завантаження даних, а також генерації метаданих для засобів доступу, таких як Discoverer.

Проектувати сховище можна і за допомогою стандартного інструменту Oracle Designer, а потім автоматично перенести опис проекту в репозиторій метаданих Oracle Warehouse Builder.

Засоби аналізу даних охоплюють весь спектр аналітичних завдань – для стандартної звітності використовується Reports, для генерації нерегламентованих звітів і запитів – Discoverer, для складного багатовимірного аналізу - продукти сімейства Express, а для задач "вилучення знань" – Oracle Data Mining Suite . Крім того, існують готові програми для вирішення спеціалізованих завдань - Enterprise Planning and Budgeting (бюджетування і планування), Ariba (ABC ) і Balanced Scorecard (збалансована система показників), Demand Planning , Value Based Management та інші.

Найважливішою рисою аналітичних інструментальних засобів і додатків Oracle є їх готовність до роботи в середовищі Internet . Менеджери та аналітики, де б вони не знаходилися, можуть отримувати інформацію з Сховищ і Вітрин Даних в захищеній Інтранет-архітектурі за допомогою сервера додатків Oracle Application Server.

**Висновки.** При проведенні даних досліджень ставилась задача вироблення стратегії розвитку Системної Архітектури ІТ інфраструктури Корпорації на основі застосування передових методологій і концепцій провідних виробників апаратного і програмного забезпечення (HP , SUN , EMC , CISCO , Microsoft , ORACLE , Veritas ). При цьому, багато уваги приділялось розробці стандартних сервісів ІТ-інфраструктури. У статті розглянуто принципи проектування сервісів зберігання даних в Дата-центрах, які визначаються відповідно до архітектури управління даними. Сервіси призначені для зберігання значущою, перевіреної, узгодженої, несуперечливої і хронологічно цілісної інформації, яку можна вважати достовірною. У якості основного об'єкту сервісу зберігання даних ми виділяємо

сховище даних. Мета сховища – забезпечити цілісність і підтримувати хронологію всіляких корпоративних даних, і з цієї точки зору воно нейтрально по відношенню до додатків.

### **Література**

1. Еталонні архітектури MSA. – К.: Майкрософт Україна; К.: Видавнича група ВНН, 2005. – 352 с.
2. Копейка О. В. Архитектура системы управления ИТ-инфраструктурой в современных Дата-центрах / О. В.Копейка // Наукові записки Українського науково-дослідного інституту зв'язку. – 2014. – № 1(29). – С.29-37.
3. Oleg Koriyuka Telecommunication Systems Architectures Structural Synthesis with Modern Servises Providing / Oleg Koriyuka, Alexander Drobyk, Iurii Kovalchuk// XII International Conference Modern Problems of Radio Engineering TCSET 2014. – 2014. – S.527-528.
4. Копейка О. В. Архитектура системы безопасности ИТ-инфраструктуры в дата-центрах / О. В. Копейка // Сучасний захист інформації. – 2014. – №1. – С.48-57.
5. Копейка О. В. Сетевые службы и службы сетевых устройств в дата-центрах / О. В. Копейка // Системи управління, навігації та зв'язку: наукове періодичне видання. – 2013. – №4(28). – С.98-104.
6. Довгий С. О. Засади регіональної інформатизації / С. О. Довгий, О. В. Копійка, Ю. Т. Черепін. – К.:ВПЦ «ТИРАЖ», 2004. – 304 с.
7. Довгий С. А. Новые технологии в телекоммуникации: выбор технологической архитектуры. Современные тенденции развития / С. А. Довгий, О. В. Копейка, С. П. Поленок. – К.: Укртелеком, 2001. – 281 с.
8. Сучасні телекомунікації: мережі, технології, економіка, управління, регулювання / [С. О. Довгий, О. Я. Савченко, О. В. Копійка та ін..] ; за ред. С. О. Довгого. – К.: Український видавничий центр, 2002. –502 с.
9. О. Копейка, І. Тарасенко, А. Kisselevskiy, А. Karichenskiy, Т. Valiulin. Softline applies TMF standards as a guide when building Resource Inventory solution for nation-wide carrier Ukraine Telecom // TM Forum Case Study Handbook, Volume 3, May 2007 – P. 27
10. Jew, Jonathan. BICSI Data Center Standard: A Resource for Today's Data Center Operators and Designers // BICSI News Magazine, May/June 2010. – P. 28.
11. Niles, Susan. Standardization and Modularity in Data Center Physical Infrastructure // 2011, Schneider Electric / – P. 4.
12. Telecommunications Infrastructure Standard for Data Centers//TIA STANDARD TIA-942. Telecommunications Industry Association / – April 2005. – 135 p.
13. ANSI/BICSI 002-2011 Data Center Design and Implementation Best Practices // Committee Approval. – January 2011. First Published : March 2011. – 367 p.

Дата надходження в редакцію: 16.04.2014 р.

Рецензент: д.т.н., проф. Савченко Ю. Г.