**Безпека даних у RAID і розподілених системах**

Проценко Іван

Студент ФІТ 3 курс, ДТЕУ, Україна

Науковий керівник
Юрченко Юрій Юрійович

## **Вступ**

У сучасному світі забезпечення безпеки та цілісності даних є критично важливим завданням для організацій будь-якого масштабу. Технології RAID (Redundant Array of Independent Disks) та розподілені системи зберігання є ключовими компонентами інфраструктури, які значно підвищують надійність зберігання інформації. У цій доповіді розглянемо основні аспекти безпеки даних у таких системах, методи захисту та сучасні підходи до забезпечення високого рівня збереження даних.

## **Основи технології RAID**

RAID (Redundant Array of Independent Disks) — це технологія об'єднання декількох фізичних дисків в один логічний диск для підвищення продуктивності, надійності або обох цих характеристик. Основними цілями впровадження RAID є підвищення надійності завдяки резервуванню даних, покращення продуктивності шляхом паралельного доступу до даних та збільшення ємності через об'єднання дисків у єдиний логічний том.



Існує кілька рівнів RAID, кожен з яких має свої характеристики безпеки. RAID 0 забезпечує розподіл даних без надлишковості, але не має захисту від відмов дисків. RAID 1 використовує дзеркалювання, що забезпечує високу надійність завдяки дублюванню даних. RAID 5 пропонує розподіл з розподіленим парітетом, забезпечуючи захист від відмови одного диска. RAID 6 додає додатковий рівень захисту з двома незалежними парітетами, дозволяючи витримати відмову до двох дисків. RAID 10 поєднує переваги RAID 1 і RAID 0, забезпечуючи високу продуктивність та надійність.

## **Ризики та вразливості RAID-систем**

Попри всі переваги, RAID-системи мають ряд потенційних вразливостей. Найпоширенішою проблемою є відмова дисків, особливо критична при одночасній відмові декількох дисків. Проблеми з контролерами можуть призвести до втрати доступу до всіх даних. Помилки конфігурації часто знижують рівень захисту, а деградація продуктивності проявляється при поступовому зниженні швидкості роботи через відмову компонентів. Особливо вразливим є період відновлення масиву після відмови диска.

Важливо розуміти, що RAID не є заміною резервного копіювання. Він не захищає від логічних помилок даних, вірусів та шкідливого ПЗ, помилок користувачів, фізичних пошкоджень та відмов обладнання, що виходять за рамки можливостей відмовостійкості.

## **Безпека даних у розподілених системах**

Розподілені системи зберігання даних — це інфраструктура, де дані розміщуються на декількох серверах або вузлах. Основні типи таких систем включають DAS (безпосередньо підключені сховища), NAS (мережеві сховища), SAN (мережі зберігання даних), об'єктні сховища та розподілені файлові системи, такі як HDFS, GlusterFS та Ceph.



Основними механізмами забезпечення безпеки в розподілених системах є реплікація даних, шардинг (горизонтальне розділення даних між вузлами), розподілений парітет, геораспреділення та еразурне кодування. Реплікація створює та підтримує декілька копій даних у різних вузлах. Шардинг розподіляє дані між вузлами для підвищення продуктивності та надійності. Геораспреділення розміщує дані у різних фізичних локаціях для захисту від локальних катастроф. Еразурне кодування застосовує методи зберігання даних, що забезпечують стійкість до відмов окремих частин.

## **Сучасні підходи до захисту даних**

Ефективна стратегія захисту даних включає багаторівневий захист, регулярне тестування відновлення, моніторинг стану систем та аудит безпеки. Передові технології захисту даних включають самовідновлювані системи, що автоматично виявляють та виправляють помилки, інтелектуальне прогнозування відмов на основі аналізу S.M.A.R.T. та інших метрик, миттєві знімки для створення точок відновлення, дедуплікацію та компресію для оптимізації зберігання, а також шифрування на рівні томів для захисту від несанкціонованого доступу.

## **Практичні рекомендації щодо захисту даних**

При виборі конфігурації RAID необхідно враховувати критичність даних, бюджетні обмеження, вимоги до продуктивності та масштабованість. Для критичних бізнес-даних рекомендуються RAID 10 або RAID 6, для середньокритичних даних — RAID 5 або RAID 6, а для некритичних даних з вимогами до продуктивності — RAID 0 з регулярним резервним копіюванням.

Для забезпечення безпеки в розподілених системах рекомендується дотримуватися правила 3-2-1 резервного копіювання: 3 копії даних, 2 різних типи носіїв та 1 копія офлайн або в іншій локації. Також важливо проводити регулярні перевірки цілісності даних з використанням контрольних сум, періодичного сканування на наявність помилок та автоматизованих процесів перевірки. Необхідно розробити плани відновлення після аварій з документацією процедур, регулярними навчаннями та визначенням допустимого часу та точки відновлення.

## **Приклади впровадження**

Для інфраструктури середнього бізнесу з обсягом даних до 50 ТБ, обмеженим IT-персоналом та помірними вимогами до доступності рекомендується впровадження систем NAS з RAID 6, щоденне резервне копіювання на зовнішні носії, щомісячне повне резервне копіювання з офлайн-зберіганням та моніторинг стану дисків з автоматичними сповіщеннями.

Для корпоративного дата-центру з петабайтами даних, критичними бізнес-процесами та вимогами до цілодобової доступності оптимальним є розподілена SAN-інфраструктура з реплікацією між дата-центрами, комбінація RAID 10 для високопродуктивних систем і RAID 6 для архівних даних, впровадження системи управління життєвим циклом інформації, автоматизоване тестування процедур відновлення та геореплікація даних між дата-центрами в різних регіонах.

## **Висновки**

Безпека даних у RAID і розподілених системах вимагає комплексного підходу, що поєднує технологічні рішення, процеси та методики. Хоча RAID забезпечує базовий рівень захисту від відмов дисків, справжня безпека даних досягається лише при впровадженні багаторівневої стратегії, що включає резервне копіювання, моніторинг, тестування відновлення та чіткі процедури реагування на інциденти.

Розподілені системи додають додатковий рівень складності, але також пропонують підвищену стійкість до відмов через географічну диверсифікацію та масштабованість. Поєднання технологій RAID з перевагами розподілених систем дозволяє створювати високонадійні інфраструктури для зберігання даних, що відповідають вимогам сучасного бізнесу. Головним принципом у забезпеченні безпеки даних залишається принцип "захисту в глибину", коли жоден окремий збій не здатний призвести до повної втрати даних.