

УДК 004.272.3

канд. техн. наук, доц Луцків А. М, Голубовський М. П.

(Тернопільський національний технічний університет ім. Івана Пулюя, Україна)

ПРОБЛЕМИ, ЯКІ ВИНИКАЮТЬ ПРИ РОЗГОРТАННІ ІНФРАСТРУКТУР ДЛЯ ОПРАЦЮВАННЯ ВЕЛИКИХ ДАНИХ

UDC 004.272.3

Ph.D., Assoc. Prof Lutskiv A, Holubovskyi M.



PROBLEMS THAT ARISE DURING DEPLOYMENT OF BIG DATA PROCESSING INFRASTRUCTURES

Використання традиційних, ручних способів розгортання, налаштування та підтримки інфраструктури для опрацювання великих даних вимагає великих затрат часу та значної кількості фахівців. Спочатку здійснюється конфігурація апаратного забезпечення – обчислювальних ресурсів та мережі для їх взаємодії. Наступними кроками є встановлення та налаштування операційних систем, програм та бібліотек від яких залежить робота додатків. Тільки після виконання усіх етапів підготовки спеціалісти виконують встановлення компонентів комплексу для опрацювання великих даних. Оскільки, до виконання таких завдань використовується ручний підхід, це спричиняє виникнення певних проблем та наслідків.

Першою проблемою є швидкість створення та підготовки до роботи такої інфраструктури. Також проблемою є непостійність при конфігуруванні ресурсів, виникнення розбіжностей при ручному конфігуруванні, зокрема, при роботі команди спеціалістів. Традиційний підхід не надає можливості здійснювати тестування перед внесенням змін у конфігурацію, перегляду історії змін та зручної командної роботи.

Для вирішення описаних проблем ефективним рішенням є використання підходу опису інфраструктури у вигляді коду. Такий спосіб створення та керування обчислювальними ресурсами передбачає їх опис у вигляді програмного коду, на відміну від ручного редагування конфігураційних файлів чи використання інтерактивних інструментів. Така інфраструктура може охоплювати фізичні сервери, віртуальні машини, а також пов'язані з ними ресурси і програмні компоненти. Таким чином, при використанні підходу “інфраструктура як код” можемо використовувати практики притаманні типовій розробці ПЗ. До переваг, які надає використання даного підходу відноситься швидкість створення, послідовність конфігурування, підзвітність описаної системи. Швидкість роботи досягається завдяки можливості створити всю потрібну інфраструктуру, використовуючи лише одну команду над конфігураційним файлом. Такий файл може використовуватися для створення ідентичних ізольованих систем та надає змогу легко проводити тестування перед внесенням змін. Вже описані елементи такої інфраструктури можуть бути повторно використані при роботі над іншими проектами. Підхід дозволяє уникнути непослідовності, помилок пов'язаних з людським фактором, оскільки, єдиним джерелом для внесення змін є конфігураційні файли, виключається необхідність внесення ручних змін. Перевагою також є наявність звітності про те, які зміни, коли і ким були внесені. Таку конфігурацію можна зберігати у системі контролю версій. Використання підходу документує стан інфраструктури, дає змогу мати задокументований стан того, як вона налаштована у поточний момент часу.

Література.

1. Morris K. Infrastructure as Code: Managing Servers in the Cloud / Kief Morris. – Sebastopol, CA: O'Reilly Media, 2016. – 362 с.