

## ПРОБЛЕМИ МАСШТАБОВАНOSTI В РОЗПОДІЛЕНІЙ ОБРОБЦІ ДАНИХ ДЛЯ ТУМАННИХ ІТ-ІНФРАСТРУКТУР

Цвіркун Л.І.<sup>1</sup> Соболевський І.О.<sup>2</sup>

<sup>1</sup> к.т.н., доцент, Національний технічний університет

«Дніпровська політехніка», Україна

<sup>2</sup> аспірант, Національний технічний університет

«Дніпровська політехніка», Україна

**Анотація.** У цій роботі розглядаються питання масштабованості в розподіленій обробці даних для туманних ІТ-інфраструктур. Аналіз визначив великі обсяги даних, перевантаження мережі, неоднорідність пристроїв і нестабільність мережі як ключові проблеми. Розглядаються можливі наступні кроки, такі як впровадження ефективних стратегій і використання новітніх технологій для вирішення цих проблем. Вносяться пропозиції щодо вдосконалення туманної ІТ-інфраструктури шляхом впровадження інноваційних рішень та досліджень у цій галузі. Наголошується необхідність інтегрованого підходу та співпраці між науковцями та практиками для подолання проблем масштабованості в туманних ІТ-інфраструктурах. Крім того, звертається увага на важливість безпеки даних у контексті туманної ІТ-інфраструктури та необхідність розробки відповідних заходів для її забезпечення.

**Ключові слова:** розподілена обробка даних, туманні обчислення, проблеми масштабованості, перевантаження мережі, управління ресурсами, гетерогенні пристрої, нестабільність мережі, проблеми збору даних, інноваційні рішення, технологічні досягнення, спільні дослідження, оптимізація ІТ-інфраструктури.

### Вступ

Проблема масштабованості в розподіленій обробці даних для туманних ІТ-інфраструктур виникає у зв'язку зі зростанням обсягу даних, які обробляються та передаються у таких середовищах. Зі збільшенням кількості пристроїв, що забезпечують збір та обробку даних, стає необхідним розробляти ефективні стратегії та механізми керування ними.

Однією з ключових проблем є великий обсяг даних, який може призвести до перевантаження мережі та обчислювальних ресурсів. При великій кількості даних може виникати проблема недостатньої пропускнуої здатності мережі, а також високої завантаженості серверів, що може вплинути на їхню

продуктивність. Локалізована обробка даних на кожному пристрої в туманних мережах може призвести до нерівномірного розподілу навантаження та недостатньої ефективності обробки. При цьому можуть виникати проблеми з неоднорідністю в обробці даних на різних пристроях, що може призвести до втрати консистентності та недостовірності результатів. Управління ресурсами в таких інфраструктурах може бути складним завданням через велику кількість пристроїв та їхню гетерогенність. Для ефективного використання ресурсів необхідно розробляти розумні алгоритми та стратегії, які б враховували специфіку туманного середовища та потреби додатків. Збільшення кількості точок доступу та передачі даних у туманних інфраструктурах підвищує ризик компрометації конфіденційності та цілісності інформації. Нестабільність мережі та можливі збої у спробах розподіленої обробки даних можуть призвести до втрати даних та порушення сервісу [1, 2].

Розв'язання цих проблем потребує ретельного аналізу та розробки ефективних стратегій керування даними, вдосконалення мережевих технологій та засобів безпеки, а також розробки надійних механізмів моніторингу та управління для забезпечення стабільності та надійності розподіленої обробки даних у туманних ІТ-інфраструктурах.

Для вирішення проблеми масштабованості в розподіленій обробці даних для туманних ІТ-інфраструктур потрібно розглянути різноманітні підходи та стратегії. Один із можливих напрямів – це використання розподілених систем керування базами даних, які дозволяють ефективно керувати великим обсягом даних та розподіленістю обробки. Такі системи можуть автоматично розподіляти навантаження між серверами, оптимізувати запити та забезпечувати високу доступність даних. Додатковим рішенням може бути використання технологій контейнеризації та оркестрації, таких як Docker та Kubernetes. Вони дозволяють легко масштабувати і розгортати додатки в різних середовищах, а також автоматизувати процеси управління та моніторингу. Крім того, важливо розглянути можливості оптимізації мережевого трафіку та використання кешування даних для зменшення навантаження на мережу та сервери. Техніки стиснення даних та розподіленої обробки можуть допомогти

знизити обсяг передаваних даних та збільшити швидкість їхньої передачі. Важливим елементом є також забезпечення високого рівня безпеки даних. Для цього можна використовувати шифрування даних, аутентифікацію та авторизацію, а також моніторинг та аудит доступу до інформації [1, 3].

Великий обсяг даних та перевантаження мережі

Однією з головних проблем, що виникають у контексті масштабування розподіленої обробки даних для туманних IT-інфраструктур, є великий обсяг даних та його наслідки на роботу мережі. Зі зростанням кількості підключених пристроїв та обсягу збираємих даних збільшується навантаження на мережеву інфраструктуру.

Великий обсяг даних призводить до збільшення мережевого трафіку, що може викликати перевантаження комунікаційних каналів та мережевих вузлів. Це може призвести до зниження швидкості передачі даних, затримок у доставці інформації та загального погіршення продуктивності мережі.

Підвищення навантаження на мережу може також вплинути на якість обслуговування (Quality of Service – QoS) для окремих додатків або послуг. Велика кількість одночасних запитів на дані може призвести до конфліктів у використанні ресурсів мережі та затримок у відповіді на запити [1, 4].

Навантаження на сервери та його вплив на продуктивність

Підвищений обсяг даних, який обробляється в туманних IT-інфраструктурах, також може призвести до перевантаження серверів. Постійне збільшення обсягу даних, що обробляються, може вимагати додаткових обчислювальних ресурсів для їх ефективної обробки.

Перевантаження серверів може викликати збільшення часу відповіді на запити, зниження загальної продуктивності системи та збільшення ризику виникнення простоїв. Для подолання цих проблем необхідно розглянути можливості оптимізації роботи серверів, використання швидкодіючих обчислювальних систем та розподіл навантаження між різними вузлами мережі [2, 3].

Нерівномірний розподіл навантаження та неоднорідність в обробці даних

Однією зі складнощів, що супроводжують масштабування розподіленої обробки даних в туманних ІТ-інфраструктурах, є нерівномірний розподіл навантаження між пристроями та неоднорідність в обробці даних.

В туманних мережах можуть бути пристрої з різними характеристиками та рівнями продуктивності. Це може призвести до ситуації, коли деякі пристрої надзвичайно завантажені, тоді як інші можуть залишатися недостатньо використаними. Нерівномірний розподіл навантаження може призвести до перевантаження деяких пристроїв, що може вплинути на їхню продуктивність та знизити ефективність обробки даних [4, 5].

Крім того, неоднорідність у характеристиках пристроїв може вплинути на консистентність та надійність обробки даних. Різні швидкості обробки на різних пристроях можуть призвести до ситуацій, коли дані обробляються в різні моменти часу, що може порушити консистентність даних та призвести до втрати надійності результатів.

Проблеми управління ресурсами та гетерогенність пристроїв

Управління ресурсами в туманних ІТ-інфраструктурах може бути складною задачею через гетерогенність пристроїв та їхніх характеристик. З урахуванням того, що в мережі можуть бути пристрої різних виробників, з різними обчислювальними можливостями та мережевими характеристиками, важливо розробляти стратегії управління ресурсами, які б враховували цю гетерогенність.

Ефективне управління ресурсами передбачає розробку алгоритмів та стратегій, які б динамічно реагували на зміни у навантаженні та характеристиках пристроїв. Наприклад, можливе розроблення алгоритмів розподілу навантаження, які б враховували поточний стан мережі та характеристики пристроїв для ефективного використання ресурсів [1, 6].

Також важливо розглядати можливості автоматизації процесів управління ресурсами за допомогою інтелектуальних систем, що базуються на аналізі даних та машинному навчанні. Це може допомогти знизити навантаження на адміністраторів систем та забезпечити більш ефективне використання ресурсів у туманних ІТ-інфраструктурах [5, 6].

### Вплив нестабільності мережі на збір даних

Нестабільність мережі є однією з ключових проблем у процесі збору даних в туманних середовищах. Ця проблема може виникати з різних причин, таких як переповнення мережі, втрати пакетів даних, а також непередбачені збої у роботі мережевих пристроїв. Внаслідок цього може страждати якість та достовірність зібраних даних.

Одним із можливих рішень для подолання цієї проблеми є розробка механізмів, що дозволяють виявляти та виправляти помилки з'єднання. Наприклад, можна використовувати протоколи, які забезпечують повторну відправку втрачених пакетів даних або автоматичне відновлення з'єднання після його втрати.

Також важливо розглядати можливості оптимізації мережевого трафіку для зменшення його впливу на збір даних. Це може включати в себе використання алгоритмів стиснення даних або розробку ефективних стратегій маршрутизації, які дозволяють уникати перевантаження мережі [2, 5, 7].

### Роль безпеки в методах збору даних для туманних середовищ

Забезпечення безпеки даних є важливою складовою процесу збору інформації в туманних середовищах. З урахуванням розподіленості та гетерогенності пристроїв, які здійснюють збір даних, стає важливою задачею застосування механізмів, що гарантують конфіденційність, цілісність та доступність інформації.

Для досягнення цих цілей можуть застосовуватися різноманітні методи, включаючи шифрування даних, механізми аутентифікації та авторизації, а також моніторинг та аудит дій з даними. Крім того, важливо розглядати можливості розробки захищених протоколів передачі даних та механізмів виявлення та відхилення від несанкціонованого доступу до інформації [1, 3, 7].

### Висновки

1. В результаті аналізу було виявлено, що проблеми масштабованості в розподіленій обробці даних для туманних IT-інфраструктур потребують комплексного підходу та розробки ефективних стратегій. Для успішного

вирішення цих питань необхідно розробляти нові технології, використовувати інтелектуальні системи та стежити за останніми тенденціями в галузі ІТ.

2. Наступними кроками є реалізація запропонованих стратегій, впровадження нових технологій та продовження досліджень у цьому напрямку для постійного вдосконалення туманних ІТ-інфраструктур.

3. У підсумку, розробка інноваційних рішень та співпраця між науковцями та практиками є ключовими для подолання проблем масштабованості в розподіленій обробці даних для туманних ІТ-інфраструктур.

### **ЛІТЕРАТУРА**

1. R. Mayer, L. Graser, H. Gupta, E. Saurez and U. Ramachandran, "EmuFog: Extensible and scalable emulation of large-scale fog computing infrastructures," 2017 IEEE Fog World Congress (FWC), Santa Clara, CA, USA, 2017, pp. 1-6
2. Charles C. Byers and Patrick Wetterwald. 2015. Fog Computing Distributing Data and Intelligence for Resiliency and Scale Necessary for IoT: The Internet of Things (Ubiquity symposium). Ubiquity 2015, 12 pages.
3. F. Bonomi, R. Milito, J. Zhu, S. Addepalli. 2012. Fog computing and its role in the internet of things. In Proceedings of the first edition of the MCC workshop on Mobile cloud computing (MCC '12). Association for Computing Machinery, New York, USA, 13–16.
4. R. Mayer, H. Gupta, E. Saurez and U. Ramachandran, "FogStore: Toward a distributed data store for Fog computing," 2017 IEEE Fog World Congress (FWC), Santa Clara, CA, USA, 2017, pp. 1-6
5. Shanhe Yi, Cheng Li, and Qun Li. A Survey of Fog Computing: Concepts, Applications and Issues. In Proceedings of the 2015 Workshop on Mobile Big Data (Mobidata '15). Association for Computing Machinery, New York, USA, 2015, pp. 37–42.
6. S. Yi, Z. Hao, Z. Qin and Q. Li, "Fog Computing: Platform and Applications," 2015 Third IEEE Workshop on Hot Topics in Web Systems and Technologies (HotWeb), Washington, DC, USA, 2015, pp. 73-78,
7. I. Stojmenovic and S. Wen, "The Fog computing paradigm: Scenarios and security issues," 2014 Federated Conference on Computer Science and Information Systems, Warsaw, Poland, 2014, pp. 1-8.

### **SCALABILITY ISSUES IN DISTRIBUTED DATA PROCESSING FOR FOGGY IT INFRASTRUCTURES**

Tsvirkun Leonid, Sobolevskiy Ivan

**Abstract.** *This paper addresses scalability issues in distributed data processing for fog IT infrastructures. The analysis identified large data volumes, network congestion, device heterogeneity, and network instability as key issues. Possible next steps are considered, such as implementing effective strategies and using the latest technologies to address*

*these issues. Proposals are made to improve fog IT infrastructures through the introduction of innovative solutions and research in this area. The abstract emphasizes the need for an integrated approach and cooperation between scientists and practitioners to overcome scalability issues in fog IT infrastructures. In addition, attention is drawn to the importance of data security in the context of fog IT infrastructures and the need to develop appropriate measures to ensure it.*

**Keywords:** *distributed data processing, fog computing, scalability issues, network congestion, resource management, heterogeneous devices, network instability, data collection challenges, innovative solutions, technology advancements, collaborative research, IT infrastructure optimization.*

#### **REFERENCE**

1. R. Mayer, L. Graser, H. Gupta, E. Saurez and U. Ramachandran, "EmuFog: Extensible and scalable emulation of large-scale fog computing infrastructures," 2017 IEEE Fog World Congress (FWC), Santa Clara, CA, USA, 2017, pp. 1-6
2. Charles C. Byers and Patrick Wetterwald. 2015. Fog Computing Distributing Data and Intelligence for Resiliency and Scale Necessary for IoT: The Internet of Things (Ubiquity symposium). Ubiquity 2015, 12 pages.
3. F. Bonomi, R. Milito, J. Zhu, S. Addepalli. 2012. Fog computing and its role in the internet of things. In Proceedings of the first edition of the MCC workshop on Mobile cloud computing (MCC '12). Association for Computing Machinery, New York, USA, 13–16.
4. R. Mayer, H. Gupta, E. Saurez and U. Ramachandran, "FogStore: Toward a distributed data store for Fog computing," 2017 IEEE Fog World Congress (FWC), Santa Clara, CA, USA, 2017, pp. 1-6
5. Shanhe Yi, Cheng Li, and Qun Li. A Survey of Fog Computing: Concepts, Applications and Issues. In Proceedings of the 2015 Workshop on Mobile Big Data (Mobidata '15). Association for Computing Machinery, New York, USA, 2015, pp. 37–42.
6. S. Yi, Z. Hao, Z. Qin and Q. Li, "Fog Computing: Platform and Applications," 2015 Third IEEE Workshop on Hot Topics in Web Systems and Technologies (HotWeb), Washington, DC, USA, 2015, pp. 73-78,
7. I. Stojmenovic and S. Wen, "The Fog computing paradigm: Scenarios and security issues," 2014 Federated Conference on Computer Science and Information Systems, Warsaw, Poland, 2014, pp. 1-8.