ДОДАТОК А
ГРАФІЧНИЙ МАТЕРІАЛ КВАЛІФІКАЦІЙНОЇ РОБОТИ

Харківський національний університет радіоелектроніки
Кафедра ЕОМ

МЕТОД МОНІТОРИНГУ МЕРЕЖНОГО ТРАФІКУ НА ОСНОВІ ТЕХНОЛОГІЙ BIG DATA

Кваліфікаційна робота

Другий (магістерський) рівень

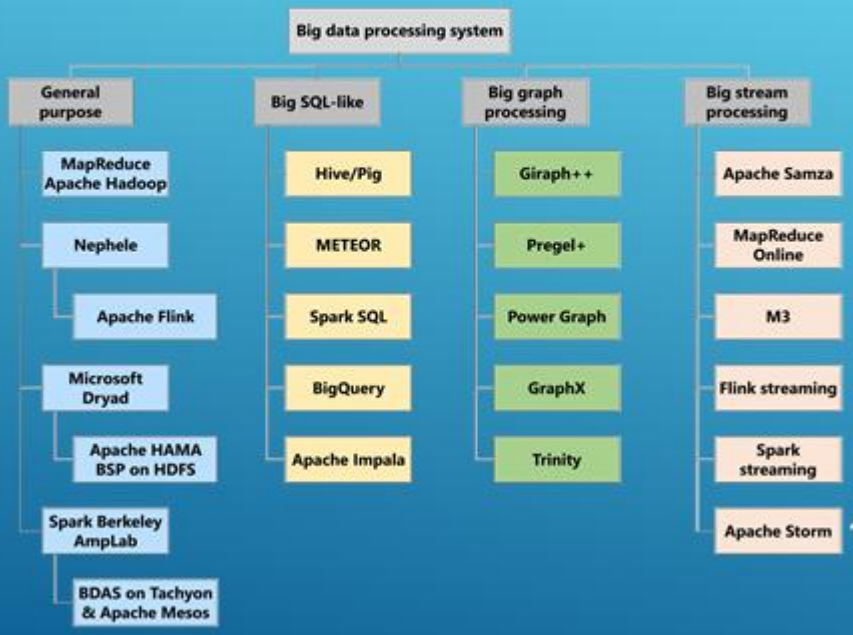
Автор:
Гнип А.К.,
студ. гр. КСМм-20-1

Керівник:
Торба А.А.,
проф. каф. ЕОМ

МЕТА І ЗАДАЧІ РОБОТИ

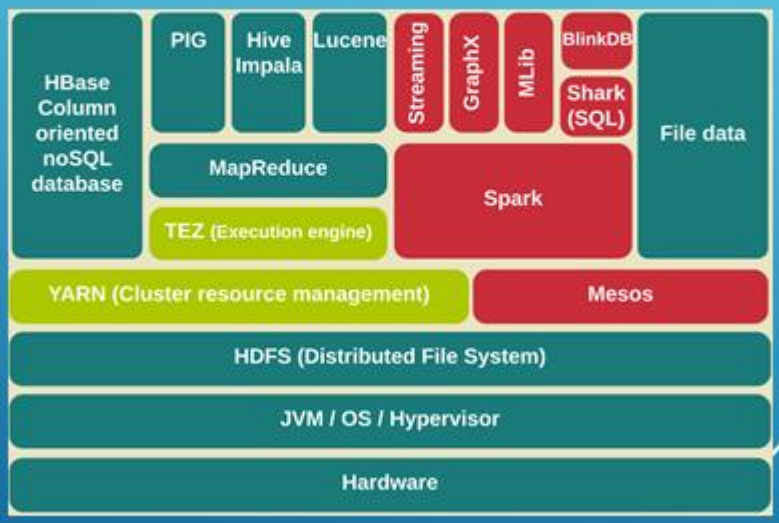
- Мета: розробка методу моніторингу трафіку комп'ютерної мережі на основі технологій великих даних.
- Задачі:
- проаналізувати принципи реалізації сучасних систем NTMA;
- дослідити інтеграцію засобів NTMA з технологіями Big Data;
- дослідити використання методів аналізу великих даних для аналізу в реальному часі;
- отримати експериментальну оцінку.

СИСТЕМИ ОБРОБКИ ВЕЛИКИХ ДАНИХ



3

СТЕК HADOOP ЗІ SPARK



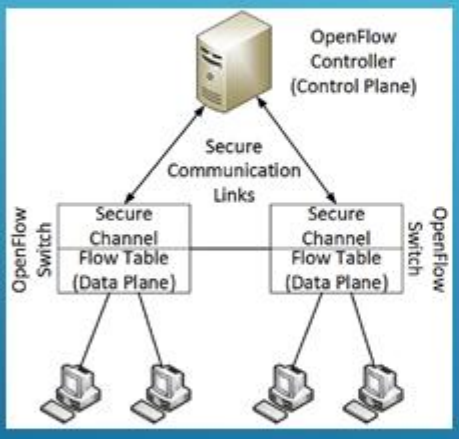
4

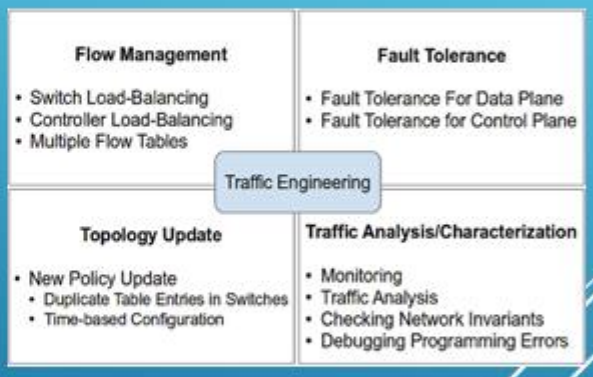
КАТЕГОРІЇ ЗАСТОСУНКІВ NTMA

Категорія	Опис
Прогноз трафіку	Обслуговування та планування мереж. Історично за допомогою прогнозування часових рядів.
Класифікація трафіку	Класифікація та розпізнавання потоків трафіку для різних цілей, наприклад, QoE, безпеки тощо.
Управління несправностями	Прогнозування та ізоляція несправностей і небажаної поведінки в мережах.
Безпека мережі	Захист мережі, реагування на зловмисні дії та загальні атаки (та їх запобігання).

5

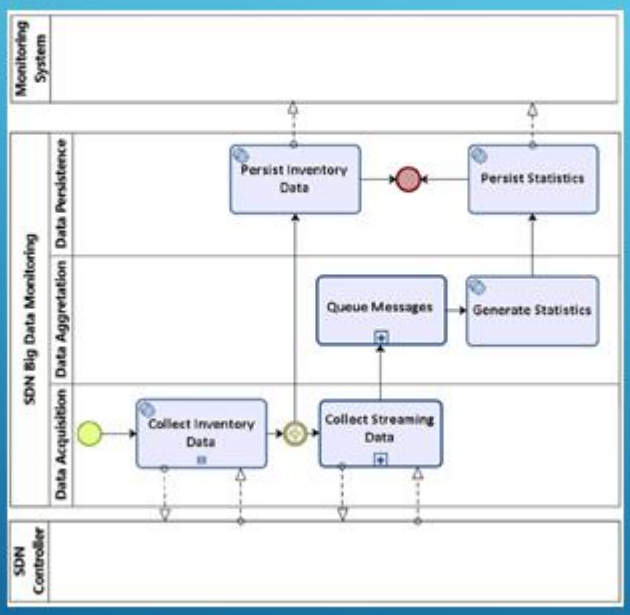
БАЗОВА АРХІТЕКТУРА OPENFLOW. ІНЖЕНЕРІЯ ТРАФІКА





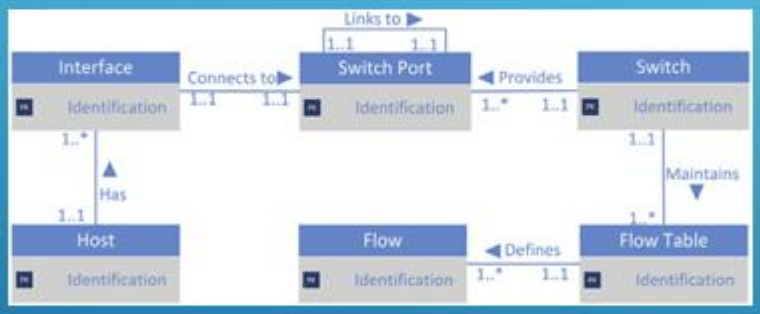
6

МОДЕЛЬ МОНІТОРИНГУ



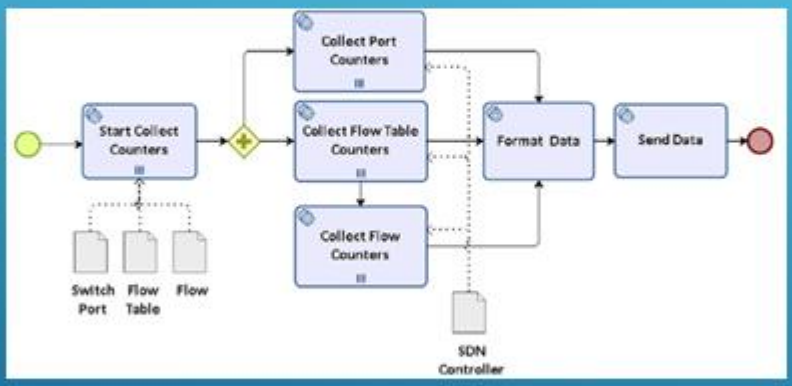
7

ER-МОДЕЛЬ КОНТРОЛЬОВАНИХ ДАНИХ



8

ЗБІР ПОТОКОВИХ ДАНИХ



9

АЛГОРИТМИ ЗБОРУ ДАНИХ, ПОЧАТКУ ЗБИРАННЯ ЛІЧИЛЬНИКІВ ТА ЗБОРУ ЛІЧИЛЬНИКІВ ТАБЛИЦІ ПОТОКІВ

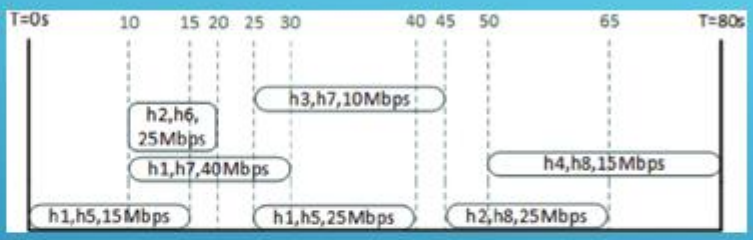
```
1  Procedure CollectData()
2      while TRUE do
3          CollectHostInventory()
4          CollectSwitchInventory()
5          CollectLinkInventory()
6          update Host, Switch, Switch Port, Flow
             Table, Flow repositories
7          sleep(n units of time)
8      end
```

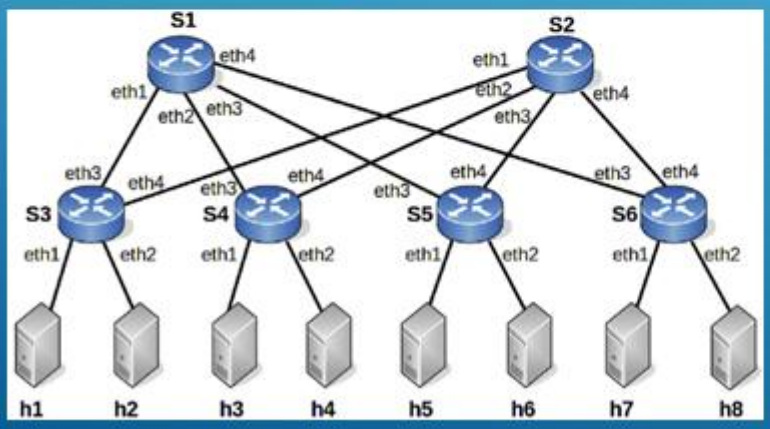
```
1  Procedure StartCollectCounters()
2      while TRUE do
3          read Switch Port repository
4          for each switch port do
5              if collect counters not started then
6                  start collect port counters task
7              end
8          end
9          read Flow Table repository
10         for each flow table do
11             if collect counters not started then
12                 start collect flow table counters
                     task
13             end
14         end
15         sleep (n units of time)
16     end
```

```
1  Procedure CollectFlowTableCounters()
2      while TRUE do
3          CollectFlowTableCounters(switch,flow
             table)
4          send counters to Format Data
5          for each flow in the flow repository do
6              if collect counters not started then
7                  start collect flow counters task
8              end
9          end
10         sleep(n units of time)
11     end
```

10

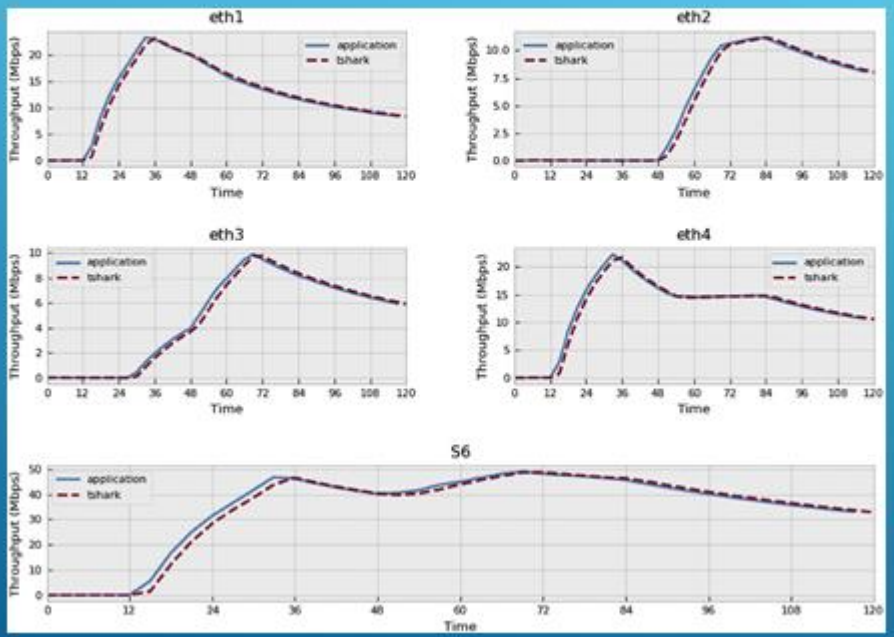
ТОПОЛОГІЯ МЕРЕЖІ ТА ДІАГРАМА ПЛАНУВАННЯ





11

ОЦІНКА ПРОПУСКНОЇ ЗДАТНОСТІ КОМУТАТОРІВ І ПОРТІВ



12

АПРОБАЦІЯ РЕЗУЛЬТАТІВ ДОСЛІДЖЕНЬ

- За темою кваліфікаційної роботи опубліковано тези доповіді в рамках дев'ятої міжнародної науково-технічної конференції «Проблеми інформатизації»

Проблеми інформатизації : дев'ята міжнародна науково-технічна конференція

МЕТОД МОНІТОРИНГУ ТРАФІКУ З ВИКОРИСТАННЯМ ТЕХНОЛОГІЙ BIG DATA ДЛЯ МЕРЕЖ МОБІЛЬНОГО ЗВ'ЯЗКУ

Гавін А.К., Торба А.А.
Харківський національний університет радіоелектроніки, Харків, Україна

Моніторинг та аналіз мережевого трафіку мають теоретичне та практичне значення для оптимізації мережевих ресурсів та поліпшення роботи користувачів. Однак існуючі рішення, які зазвичай покладаються на високопродуктивний сервер із великою смністю, не є масштабованими для детального аналізу великого обсягу даних про трафік [1]. Щоб задовольнити зростаючу вимогу до пропускної здатності, мобільні оператори розгортають більше мережевих ліній на рівні гігабіт, таких як 10G та 40G. У такому високошвидкісному середовищі, навіть з розвантажувальними мережевими картами та оптимізованим кодом ядра, програмні системи моніторингу трафіку все ще недостатні для моніторингу в режимі реального часу. Єдиним варіантом для мереж з кількома лініями зв'язку 10 Гбіт/с є індивідуальна колекторна система, заснована на апаратному забезпеченні. Отже, доцільно створити гнучку систему моніторингу, засновану на масштабованій архітектурі апаратно-програмного забезпечення [2], щоб бути придатною до модифікації та доповнення вимог до моніторингу, а також до майбутніх збільшень швидкості.

Метою доповіді є обґрунтування методу, що полягає у використанні для аналізу отримуваних великих даних набору програм, що входять до модуля програмування MapReduce. Пропонується виконувати аналіз даних про трафік у чотирьох аспектах. Перш за все, це статистика мережевого трафіку, яка обчислюється відповідно до різних індексів групування, таких як IP-адреса, протокол транспортного рівня та п'ять кортежів TCP/IP. По-друге, це аналіз рівня застосунків: створювані записи потоку та записи сеансу на рівні програми дають можливість досліджувати деякі алгоритми видобутку даних для видобуту інформації рівня застосунку та вивчення характеристики трафіку з точки зору програми. По-третє, аналіз постачальника веб-послуг: веб-трафік досліджується через стільникові мережі передачі даних з боку постачальника послуг, що відіграє важливу роль у мобільному Інтернеті. По-четверте, аналіз поведінки користувачів, для чого трафік характеризується та моделюється з боку мобільного клієнта та кінцевого користувача.

Список літератури

1. W. Xu, Y. Xu, C. H. Lee, Z. Feng, P. Zhang, and J. Lin. 2018. Data-Cognition-Empowered Intelligent Wireless Networks: Data, Utilities, Cognition Brain, and Architecture. *IEEE Wireless Communications 25*, 1 (February 2018), 56-63.

2. H. N. Dai, R. Wang, H. Wang, Z. Zheng, A. Vasilakos. (2019). Big Data Analytics for Large Scale Wireless Networks: Challenges and Opportunities.

ВИСНОВКИ

- Представлено новий підхід до вимірювання трафіку в мережах SDN. Запропоновано метод моніторингу на основі збору і обробки великих даних для формування статистики трафіку за допомогою значень лічильників.
- Оскільки запропонований метод зберігає всю згенеровану статистику в базі даних NoSQL, стає можливим використовувати аналітику великих даних для аналізу мережевого трафіку. Це може допомогти при розв'язанні задач ТЕ, таких як балансування навантаження комутатора та відмовостійкість площини даних.
- Для підтвердження ефективності методу виконана реалізація з використанням спеціальних інструментів Big Data. Досліджено використання методів аналізу великих даних для аналізу в реальному часі великої кількості збережених даних.

ВИСНОВКИ

- Експериментальні результати показали, що запропонований метод надає уявлення про те, як мережевий трафік впливає на пристрої, маршрути, що з'єднують будь-яку пару хостів OD, і зв'язки між будь-якою парою пристроїв.
- Дослідження може бути розвинуте у напрямку зменшення інтервалу між показаннями лічильника. Проте зменшуючи часовий інтервал між показаннями лічильника, збільшиться кількість запитів, що надсилаються потоками до контролера за одиницю часу.

15