

Оценка времени обработки файлов метрик инстансов облачной инфраструктуры различными способами

Пестов Игорь Евгеньевич

старший преподаватель Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича, г. Санкт-Петербург, Россия, ip@sut.ru

Федоров Павел Олегович

студент Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича, г. Санкт-Петербург, Россия, pavel_lenin@mail.ru

Федорова Екатерина Сергеевна

студент Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича, г. Санкт-Петербург, Россия, ek.chukina@yandex.ru

Смулов Илья Александрович

студент Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича, г. Санкт-Петербург, Россия, ismurov@mail.ru

АННОТАЦИЯ

Введение: В работе рассматриваются варианты оценки времени обработки файлов метрик инстансов облачной инфраструктуры. Приводится описание методов сбора данных о метриках инстансов облачной инфраструктуры. **Постановка задачи:** В работе необходимо было оценить общее время обработки файлов метрик инстансов облачной инфраструктуры разными методами. **Методы:** В исследовании рассматриваются агентный и безагентный метод сбора метрик. **Результаты:** В рамках исследования был описан авторский метод сбора и первичной обработке метрик инстансов облачной инфраструктуры и получение с помощью него файла с данными. Рассматривалось представление этих данных в двух форматах файла XML и JSON, а также были выведены формулы по расчету времени генерации файлов для обоих форматов. **Практическая значимость:** данная оценка позволяет повысить эффективность моделирования безопасности облачной инфраструктуры. **Обсуждение:** в качестве дальнейшего исследования требуется детальный анализ времени обработки файлов метрик для выбора оптимального решения при построении модели безопасности облачной инфраструктуры.

КЛЮЧЕВЫЕ СЛОВА: информационная безопасность; облачная инфраструктура; метрика; инстанс.

Введение

В соответствии с современными тенденциями возрастает потребность в облачных вычислениях. Как следствие этого увеличивается нагрузка на систему и количество используемого трафика. В связи с этим появляется необходимость производить мониторинг в ЦОД для обеспечения корректной работы всей инфраструктуры, так как оперативное реагирование на угрозы и неполадки в системе положительно влияет на стабильность и безопасность системы [1-20].

Мониторинг, или отслеживание состояния компонентов системы и управление облачной инфраструктурой, дает возможность контролировать качество услуг, которые предоставляет облачный сервис, распределять нагрузку и количество затрачиваемых ресурсов, обеспечивать должный уровень информационной безопасности и надёжности, необходимый как клиентам, так и самим провайдером, предоставляющим облачные услуги.

Мониторинг облачной среды представляет из себя комплексный процесс и состоит из ряда подтипов, которые выполняют различные функции. Одним из подтипов является мониторинг инфраструктуры, в ходе которого собирается информация об использовании различных вычислительных ресурсов и хранилищ: процессорного времени, оперативной памяти, состояния контейнеров и дисковых накопителей.

Сбор данных о метриках инстансов облачной инфраструктуры и оценка.

Сбор данных о метриках инстансов облачной инфраструктуры может происходить как агентным так и безагентным способом. Безагентный подход, означает отсутствие установленных программ агентов на виртуальных машинах или контейнерах, развернутых в рамках облачной инфраструктуры, агентный метод же в свою очередь использует специальную программу сборщик. Данные полученные при использовании безагентного метода на этапе сбора представляются в виде временного ряда, что включает в себе научную новизну данного метода. Полученные при безагентном методе сбора данные формируются в файл формата JSON, файл получаемый при агентном методе сбора имеет формат XML. Файл метрик инстансов облачной инфраструктуры при использовании безагентного метода обрабатывается более эффективно и с наименьшей затратой времени по сравнению с классическим агентными методами обработки.

Под метриками понимаются числовые значения загрузки основных ресурсов инстанса. Метрики загрузки, содержат информацию о загрузке центрального процессора - CPU, оперативной памяти - RAM, загрузки системы хранения информации - Storage, а также загруженности сетевого интерфейса - Network, виртуальной машины или контейнера. Пример файла содержащий метрик, представлен на рисунке 1.

Оценка общего времени обработки файлов метрик инстансов облачной инфраструктуры разными методами

Важным параметром при сборе данных о метриках инстансов облачной инфраструктуры является время обработки файла с метриками. Целью данной статьи является оценка общего времени обработки файлов метрик инстансов облачной инфраструктуры при использовании агентного и безагентного метода сбора данных. Для данного вычисления необходимо определить ряд ключевых параметров необходимых для расчета. Для оценки времени обработки файла с метриками при агентном методе сбора используется формула 1.

$$t_{\text{общ}} = t_{\text{ген}} + t_{\text{пер}} + t_{\text{обр}} \quad (1)$$

где:

$t_{ген}$ – время создания файла с метриками инстанса облачной инфраструктуры,

$t_{пер}$ – время передачи файла в кластер обработки,

$t_{обр}$ – время обработки файла с метриками.

Определить время генерации файла $t_{ген}$ для агентного метода сбора метрик можно по формуле 2.

$$t_{ген} = \frac{O_{ген}}{CPU} \tag{2}$$

где:

$O_{ген}$ – временная сложность генерации файла с метриками,

CPU - CPU – вычислительная мощность машины с установленным на ней агентом сборщиком.

Сложность обработки XML файла рассчитывается по формуле 3.

$$O = N^3 + N \tag{3}$$

где N – количество обрабатываемых элементов.

При расчете вычислительной мощности процессора используется формула 4.

$$CPU = v * \alpha * \beta * \gamma \tag{4}$$

где:

v – тактовая частота процессора компьютера с установленным на ней агентом сборщиком,

α – количество ядер процессора компьютера с установленным на нем агентом сборщиком,

β – количество потоков на одно ядро,

γ – коэффициент гипертрейдинга.

Время передачи файла в кластер обработки используемый в формуле оценки общего времени обработки файла с метриками определяется по формуле 5.

$$t_{пер} = \frac{V}{\vartheta} \tag{5}$$

где:

V - объем передаваемого файла,

ϑ - скорость передачи файла.

Для оценки будем использовать значение объема равное 1 Гбит и значение скорости передачи данных равное 1 Гбит/с.

Параметры CPU вычислительной системы при расчете теоретического времени обработки, взяты из таблицы 1, значения параметров, которые использовались являются типовыми.

Таб. 1. Параметры вычислительной системы

Параметр	Значение
Тактовая частота процессора	2 МГц
Количество ядер процессора	4 шт
Количество потоков на одно ядро	2 шт
Коэффициент гипертрейдинга	1,3

В данной таблице частота процессора соответствует количеству выполняемых операций за единицу времени, количество ядер процессора соответствует физическому количеству ядер процессора, количество потоков на одно ядро соответствует количеству обрабатываемых одновременно потоков данных, коэффициент гипертрейдинга соответствует коэффициенту последовательно параллельных выполняемых задач.

```

"VM": [{
  "vmID": 1001,
  "supportID": 1001,
  "contents": [{
    "CPU": 21,
    "RAM": 41,
    "Storage": 51,
    "Net": 1
  }]
},
{
  "vmID": 1002,
  "supportID": 1002,
  "contents": [{
    "CPU": 19,
    "RAM": 23,
    "Storeg": 42,
    "Net": 71
  }]
}
]
    
```

Рис. 1. Пример файла, содержащего метрики

В общем случае для расчета времени выполнения программы $t_{обр}$ используется формула 6.

$$t_{обр} = \frac{O}{CPU} \tag{6}$$

где:

O – сложность алгоритма,

CPU – вычислительная мощность процессора.

При расчете вычислительной мощности процессора используется формула 7.

$$CPU = v * \alpha * \beta * \gamma \tag{7}$$

где:

v – тактовая частота процессора,

α – количество ядер процессора,

β – количество потоков на одно ядро,

γ – коэффициент гипертрейдинга.

Сложность в данном случае рассчитывается, основываясь на алгоритме анализа файла. Если используется формат файла XML, необходимо учитывать уровень вложения тегов, так как от этого уровня зависит количество рекурсивных обращений внутри обрабатываемого файла, влияющее на сложность алгоритма влияет.

Сложность обработки XML файла рассчитывается по формуле 8.

$$O = N^i + N \tag{8}$$

где:

N – количество обрабатываемых элементов,

i – уровень вложенности.

Файлы типа JSON не имеют вложений, это означает что временная сложность изменяется линейно-логарифмически.

Сложность обработки JSON файла рассчитывается по формуле 9.

$$O = N^2 * LOG_2 N^2 \tag{9}$$

где N – количество обрабатываемых элементов.

Для расчета временной сложности использовался файл, содержащий N=1000 значений метрик инстансов облачной инфраструктуры.

Исходя из формулы 1 получаем значение времени генерации файла равное отношению времени обработки файла и временной сложности. В конечном итоге формулу расчета времени генерации файла для двух используемых форматов можно привести к определенному виду.

Время генерации файла XML находится по формуле 10.

$$t_{генерации} = \frac{(N+N^i)*O_t}{v*\alpha*\beta*\gamma} \tag{10}$$

где:

N – количество обрабатываемых элементов,

i – уровень вложенности,

v – тактовая частота процессора,

α – количество ядер процессора,

β – количество потоков на одно ядро,

γ – коэффициент гипертрейдинга,

O_t – временная сложность.

Время генерации файла JSON определяется по формуле 11.

$$t_{генерации} = \frac{(N^2*LOG_2 N^2)*O_t}{v*\alpha*\beta*\gamma} \tag{11}$$

где:

N – количество обрабатываемых элементов,

v – тактовая частота процессора,

α – количество ядер процессора,

β – количество потоков на одно ядро,
 γ – коэффициент гипертрейдинга,
 O_t – временная сложность.

Заключение

В рамках данной статьи был описан авторский метод сбора и первичной обработке метрик инстансов облачной инфраструктуры и получение с помощью него файла с данными. Рассматривалось представление этих данных в двух форматах файла XML и JSON, а также были выведены формулы по расчету времени генерации файлов для обоих форматов.

Литература

1. Темченко В. И., Цветков А. Ю. Проектирование модели информационной безопасности в операционной системе // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2019). – 2019. – С. 740-745.
2. Красов А.В., Штеренберг С.И., Голузина Д.Р. Методика визуализации больших данных в системах защиты информации для формирования отчетов уязвимостей // Электросвязь. 2019. № 11. С. 39-47.
3. Волкогонов В. Н., Гельфанд А. М., Деревянко В. С. Актуальность автоматизированных систем управления // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2019). – 2019. – С. 262-266.
4. Гельфанд А. М., Косов Н. А., Красов А. В., Орлов Г. А., Защита для распределенных отказов в обслуживании в облачных вычислениях // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2019). – 2019. – С. 329-334.
5. Гельфанд А.М., Косов Н.А., Красов А.В., Орлов Г.А. Защита для распределенных отказов в обслуживании в облачных вычислениях // В сборнике: Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2019). сборник научных статей VIII Международной научно-технической и научно-методической конференции : в 4 т.. 2019. С. 329-334.
6. Андрианов В.И., Виткова Л.А., Сахаров Д.В. Исследование алгоритма защиты общедоступных персональных данных в информационных системах. В сборнике: Актуальные проблемы инфотелекоммуникаций в науке и образовании. сборник научных статей V международной научно-технической и научно-методической конференции. 2016. С. 227-231.
7. Красов А.В., Штеренберг С.И., Москальчук А.И. Методология создания виртуальной лаборатории для тестирования безопасности распределенных информационных систем // Вестник Брянского государственного технического университета. 2020. № 3 (88). С. 38-46.
8. Штеренберг С.И., Москальчук А.И., Красов А.В. Разработка сценариев безопасности для создания уязвимых виртуальных машин и изучения методов тестирования на проникновения // Информационные технологии и телекоммуникации. 2021. Т. 9. № 1. С. 47-58.
9. Красов А.В., Левин М.В., Фостач Е.С. Проблемы обеспечения безопасности облачных вычислений // В книге: Информационная безопасность регионов России (ИБРР-2017). Материалы конференции. 2017. С. 520-522
10. Миняев А.А., Красов А.В. Методика оценки эффективности системы защиты информации территориально-распределенных информационных систем // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия 1: Естественные и технические науки. 2020. № 3. С. 26-32.
11. Красов А.В., Сахаров Д.В., Тасюк А.А. Проектирование системы обнаружения вторжений для информационной сети с использованием больших данных // Научные исследования в космических исследованиях Земли. 2020. Т. 12. № 1. С. 70-76.
12. Красов А.В., Штеренберг С.И., Голузина Д.Р. Методика визуализации больших данных в системах защиты информации для формирования отчетов уязвимостей // Электросвязь. 2019. № 11. С. 39-47.
13. Виткова Л.А., Иванов А.И., Сергеева И.Ю. Исследование и разработка методик оценки рисков облачных ресурсов // В сборнике: Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2017). Сборник научных статей VI Международной научно-технической и научно-методической конференции. В 4-х томах. Под редакцией С.В. Бачевского. 2017. С. 152-155

14. *Виткова Л.А., Глуценко А.А., Сахаров Д.В., Чмутов М.В.* Выбор оптимального метода оценки эффективности перехода к облачной архитектуре. // В сборнике: Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2018). VII Международная научно-техническая и научно-методическая конференция. Сборник научных статей. В 4-х томах. Под редакцией С.В. Бачевского. 2018. С. 168-171.

15. *Виткова Л.А., Иванов А.И.* Обзор актуальных угроз и методов защиты в сфере облачных вычислений. // В сборнике: Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2018). VII Международная научно-техническая и научно-методическая конференция. Сборник научных статей. В 4-х томах. Под редакцией С.В. Бачевского. 2018. С. 179-182.

16. *Сахаров Д.В., Левин М.В., Фостач Е.С., Виткова Л.А.* Исследование механизмов обеспечения защищённого доступа к данным, размещенным в облачной инфраструктуре. // Научные технологии в космических исследованиях Земли. 2017. Т. 9. № 2. С. 40-46.

17. *Миняев А.А., Красов А.В., Сахаров Д.В.* Метод оценки эффективности системы защиты информации территориально-распределенных информационных систем персональных данных // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия 1: Естественные и технические науки. 2020. № 1. С. 29-33

18. *Штеренберг С.И., Москальчук А.И., Коптелова В.А., Виноградова О.М.* Разработка методов обеспечения безопасности для создания уязвимых виртуальных машин и изучение методов тестирования на проникновение // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия 1: Естественные и технические науки. 2021. № 1. С. 32-38.

19. *Minyaev A.A., Krasov A.V., Saharov D.V.* The method and methodology of efficiency assessment of protection system of distributed information systems // В сборнике: 2020 12th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). 2020. С. 291-295

20. *Бирях Э.В., Виткова Л.А., Левин М.В., Чмутов М.В.* Развитие стандартов и руководств в сфере облачных технологий // В сборнике: Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2017). Сборник научных статей VI Международной научно-технической и научно-методической конференции. В 4-х томах. Под редакцией С.В. Бачевского. 2017. С. 92-95.

CALCULATION OF THE GENERATION TIME OF A FILE WITH DATA FROM METRICS OF CLOUD INFRASTRUCTURE INSTANCES

IGOR E. PESTOV,

St-Petersburg, Russia, ip@sut.ru

PAVEL O. FEDOROV,

St-Petersburg, Russia, pavel_lenin@mail.ru

EKATERINA S. FEDOROVA,

St-Petersburg, Russia, ek.chukina@yandex.ru

ILIA A. SMUROV,

St-Petersburg, Russia, ismurov@mail.ru

ABSTRACT

Introduction: This work discusses options for estimating the processing time for metric files of cloud infrastructure instances. Provides a description of methods for collecting data on metrics of cloud infrastructure instances. Statement of the problem: In this article, it was necessary to estimate the total processing time for metric files of cloud infrastructure instances by different methods. **Methods:** This work discusses the agent-based and agentless method for collecting metrics. **Results:** Within the framework of this work, the author's method for collecting and primary processing of metrics was described cloud infrastructure instances and using it to get a data file. The presentation of this data in two file formats, XML and JSON, was considered, and formulas were derived to calculate the file generation time for both formats. **Practical significance:** this assessment improves the efficiency of cloud infrastructure security modeling. **Discussion:** as a further study, a detailed analysis of the processing time of metric files is required to select the optimal solution when building a cloud infrastructure security model.

Keywords: cloud infrastructure; information security; metric; instance; graph theory; traversal algorithm.

REFERENCES

1. Temchenko V.I., Tsvetkov A.Y. Designing an information security model in the operating system // Actual problems of infotelecommunications in science and education (APINO 2019). - 2019. - S. 740-745.
2. Krasov A.V., Shterenberg S.I., Goluzina D.R. Big data visualization technique in information security systems for vulnerability reporting // Telecommunications. 2019. No. 11. S. 39-47.
3. Volkogonov V. N., Gelfand A. M., Derevyanko V. S. Relevance of automated control systems // Actual problems of infotelecommunications in science and education (APINO 2019). - 2019. - S. 262-266.
4. Gelfand A. M., Kosov N. A., Krasov A. V., Orlov G. A., Protection for distributed denial of service in cloud computing // Actual problems of infotelecommunications in science and education (APINO 2019). - 2019. - S. 329-334.
5. Gelfand A.M., Kosov N.A., Krasov A.V., Orlov G.A. Protection for distributed denial of service in cloud computing // In the collection: Actual problems of infotelecommunications in science and education (APINO 2019). Collection of scientific articles of the VIII International scientific-technical and scientific-methodological conference: in 4 volumes. 2019. P. 329-334.
6. Andrianov V.I., Vitkova L.A., Sakharov D.V. Study of the algorithm for protecting publicly available personal data in information systems In the collection: Actual problems of info-telecommunications in science and education. Collection of scientific articles of the V international scientific-technical and scientific-methodical conference. 2016. S. 227-231.
7. Krasov A.V., Shterenberg S.I., Moskalchuk A.I. Methodology for creating a virtual laboratory for testing the security of distributed information systems // Bulletin of the Bryansk State Technical University. 2020. No. 3 (88). pp. 38-46.
8. Shterenberg S.I., Moskalchuk A.I., Krasov A.V. Development of security scenarios for the creation of vulnerable virtual machines and the study of penetration testing methods // Information technologies and telecommunications. 2021. V. 9. No. 1. S. 47-58.
9. Krasov A.V., Levin M.V., Fostach E.S. Problems of ensuring the security of cloud computing // In the book: Information Security of Russian Regions (ISRR-2017). Conference materials. 2017. S. 520-522
10. Minyaev A.A., Krasov A.V. Methodology for evaluating the effectiveness of the information protection system of geographically distributed information systems // Bulletin of the St. Petersburg State University of Technology and Design. Series 1: Natural and technical sciences. 2020. No. 3. S. 26-32.
11. Krasov A.V., Sakharov D.V., Tasyuk A.A. Designing an intrusion detection system for an information network using big data // Science-intensive technologies in space research of the Earth. 2020. V. 12. No. 1. S. 70-76.
12. Krasov A.V., Shterenberg S.I., Goluzina D.R. Big Data Visualization Technique in Information Security Systems for Vulnerability Reporting // Elektrosvyaz. 2019. No. 11. S. 39-47.
13. Vitkova L.A., Ivanov A.I., Sergeeva I.Yu. Research and development of methods for assessing the risks of cloud resources // In the collection: Actual problems of infotelecommunications in science and education (APINO 2017). Collection of scientific articles of the VI International scientific-technical and scientific-methodical conference. In 4 volumes. Edited by S.V. Bachevsky. 2017. S. 152-155.
14. Vitkova L.A., Glushchenko A.A., Sakharov D.V., Chmutov M.V. Choosing the optimal method for evaluating the effectiveness of the transition to cloud architecture. // In the collection: Actual problems of infotelecommunications in science and education (APINO 2018). VII International scientific-technical and scientific-methodical conference. Collection of scientific articles. In 4 volumes. Edited by S.V. Bachevsky. 2018. S. 168-171.
15. Vitkova L.A., Ivanov A.I. Overview of current threats and protection methods in the field of cloud computing. // In the collection: Actual problems of infotelecommunications in science and education (APINO 2018). VII International scientific-technical and scientific-methodical conference. Collection of scientific articles. In 4 volumes. Edited by S.V. Bachevsky. 2018. S. 179-182.
16. Sakharov D.V., Levin M.V., Fostach E.S., Vitkova L.A. Study of mechanisms for providing secure access to data hosted in the cloud infrastructure. // Science-intensive technologies in space research of the Earth. 2017. V. 9. No. 2. S. 40-46.

17. Minyaev A.A., Krasov A.V., Sakharov D.V. Method for evaluating the effectiveness of the information security system of geographically distributed information systems of personal data // Bulletin of the St. Petersburg State University of Technology and Design. Series 1: Natural and technical sciences. 2020. No. 1. S. 29-33
18. Shterenberg S.I., Moskalchuk A.I., Koptelova V.A., Vinogradova O.M. Development of security methods for creating vulnerable virtual machines and the study of penetration testing methods // Bulletin of the St. Petersburg State University of Technology and Design. Series 1: Natural and technical sciences. 2021. No. 1. S. 32-38.
19. Minyaev A.A., Krasov A.V., Saharov D.V. The method and methodology of efficiency assessment of protection system of distributed information systems // In the collection: 2020 12th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). 2020. C. 291-295
20. Birikh E.V., Vitkova L.A., Levin M.V., Chmutov M.V. Development of standards and guidelines in the field of cloud technologies // In the collection: Actual problems of infotelecommunications in science and education (APINO 2017). Collection of scientific articles of the VI International scientific-technical and scientific-methodical conference. In 4 volumes. Edited by S.V. Bachevsky. 2017. S. 92-95.