

Integration of the Access Service to Satellite Data in the Cloud Platform C2

Andrey A. Proshin¹, Vitaliy A. Antonenko²

¹Space Research Institute, Russian Academy of Sciences, Moscow, Russia
kashnitskiy@mail.ru

²Lomonosov Moscow State University, Moscow, Russia
vantonenko@arccn.ru

Abstract

The paper describes the current model of satellite data access system integrated into the C2 cloud platform, developed jointly by the specialists of IKI RAS and CMC MSU faculty. The main goal of the joint work is the development of technologies that allow the easy deployment of remote sensing data access systems on different universities infrastructure. Such information system mainly can be used to teach students the basics of working with satellite data and in perspective also for solving various research problems. A key advantage of technologies being developed is the ability to rapidly deploy satellite data access complexes and dynamically scale them, depending on current needs and available resources.

Keywords: Earth remote sensing, information systems, satellite data access systems, cloud platform, Network Functions Virtualization (NFV), OpenStack platform

ИНТЕГРАЦИЯ СЕРВИСА ДОСТУПА К СПУТНИКОВЫМ ДАННЫМ В ОБЛАЧНУЮ ПЛАТФОРМУ С2

А.А.Прошин¹, В.А. Антоненко²

¹Институт космических исследований РАН, Москва, Россия
andry@iki.rssi.ru

²Московский государственный университет имени М.В. Ломоносова, Москва, Россия
vantonenko@arccn.ru

Статья посвящена описанию действующего макета системы доступа к спутниковым данным, интегрированной в облачную платформу С2. Макет был реализован совместными усилиями специалистов ИКИ РАН и факультета ВМК МГУ. Основной целью его создания является выработка технологии, позволяющей на площадках различных университетов развертывать комплексы доступа к данным ДЗЗ, предназначенные для обучения студентов основам работы со спутниковыми данными, а в перспективе также и для решения различных исследовательских и прикладных задач. Ключевыми преимуществами разрабатываемой технологии являются возможность быстрого развертывания комплексов доступа к спутниковым данным, а также динамическое масштабирование сервисов доступа к спутниковым данным в зависимости от текущих потребностей пользователей и имеющихся ресурсов.

Ключевые слова: дистанционное зондирование Земли, информационные системы, системы доступа к спутниковым данным, облачная платформа, технология виртуализации сетевых функций (NFV), платформа OpenStack.

Введение

Данные дистанционного зондирования Земли в настоящее время широко используются для решения самых различных исследовательских и прикладных задач, связанных с мониторингом окружающей среды и антропогенных факторов. При этом в последние десятилетия наблюдается стремительный рост как количественных, так и качественных характеристик систем ДЗЗ, что существенно расширило сферу эффективного применения таких данных. Важно заметить, что развитие информационных технологий в последние годы позволило на качественно новом уровне реализовать сервисы доступа к спутниковым данным, которые теперь не только обеспечивают доступ к имеющимся в архивах данным, но и предоставляют пользователям широкий спектр инструментов для их обработки и анализа в режиме, приближенном к реальному времени [1]. Функционирование таких сервисов базируется на использовании сложных

распределенных информационных систем, обеспечивающих прием, обработку и предоставление спутниковых данных конечным потребителям. При этом для работы таких систем требуется масштабная и дорогостоящая инфраструктура, как для хранения данных, так и для их обработки.

На наш взгляд, наиболее перспективным является совместное использование инфраструктуры систем доступа к спутниковым данным в интересах большого числа различных проектов и конечных потребителей. Также целесообразным является использование ресурсов, предоставляемых различными дата-центрами и другими площадками, располагающими требуемой технической инфраструктурой. Эффективнее всего доступ к таким ресурсам может быть реализован на основе современных облачных технологий [2], позволяющих выделять те или иные вычислительные и другие ресурсы по мере того, как они становятся необходимы и высвобождать их, когда в них нет необходимости. Такой подход позволяет принципиально снизить начальные капиталовложения и время создания новых информационных систем, существенно упростить их масштабирование и повысить устойчивость их работы.

Настоящая статья посвящена описанию действующего макета системы доступа к спутниковым данным, интегрированной в облачную платформу С2. Макет был реализован совместными усилиями специалистов ИКИ РАН и факультета ВМК МГУ. Основной целью его создания является выработка технологии, позволяющей на площадках различных университетов развертывать комплексы доступа к данным ДЗЗ, предназначенные для обучения студентов основам работы со спутниковыми данными, а в перспективе также и для решения различных исследовательских и прикладных задач. Ключевыми преимуществами разрабатываемой технологии являются возможность быстрого развертывания комплексов доступа к спутниковым данным, а также динамическое масштабирование сервисов доступа к спутниковым данным в зависимости от текущих потребностей пользователей и имеющихся ресурсов.

Макет системы доступа к спутниковым данным был реализован на основе технологий и программного обеспечения, разработанных в ИКИ РАН в отделе «Технологии» спутникового мониторинга. При его создании были также использованы технические и технологические возможности ЦКП «ИКИ-Мониторинг» [3].

Используемая в настоящей работе облачная платформа С2 [4] была разработана специалистами ВМК МГУ. Она, в свою очередь, базируется на использовании технологий программно-конфигурируемых сетей (SDN) и виртуализации сетевых функций (NFV) и позволяет создавать виртуализированную инфраструктуру для каждого отдельного научно-исследовательского эксперимента.

Ниже кратко рассматривается архитектура реализованной системы доступа к спутниковым данным, а затем приводятся сведения об использовании в действующем макете возможностей облачной платформы С2. В заключении перечисляются основные преимущества разработанной технологии и описываются перспективы ее развития и внедрения.

Архитектура действующего макета

Рассмотрим архитектуру реализованного макета как совокупность взаимодействующих между собой информационных сервисов, так как такое представление наиболее соответствует задачам интеграции сложной информационной системы в облачную платформу. Упрощенная схема взаимодействия таких сервисов приведена на рис.1. Красным цветом выделены элементы, соответствующие динамически создаваемым сервисам, а именно, сервисам, динамически развертываемым по мере необходимости. На схеме пунктирными линиями изображены запросы на получение данных или метаданных, сплошными линиями – метаданные, а полыми – спутниковые изображения.



Рис. 1. Упрощенная схема взаимодействия сервисов

Сервис «WEB-интерфейс» представляет собой специализированный картографический интерфейс, позволяющий пользователю выбрать интересующие его данные, отобразить их на картографической подложке, а также воспользоваться реализованными инструментами для анализа и обработки данных. Запрос информации об имеющихся в архивах данных реализован на основе сервиса доступа к данным и метаданным в архиве. На основании полученных метаданных пользователю предоставляется список доступных ему информационных продуктов. Существенно, что большая их часть представляет из себя «виртуальные» продукты, т.е. продукты, динамически формируемые на основе имеющихся в архивах данных с использованием определенных в системе правил их построения. Такой подход позволяет избежать хранения в архивах многочисленных производных продуктов и существенно повышает гибкость системы доступа. Более сложная тематическая обработка спутниковых данных, используемая в реализованных инструментах анализа и обработки данных, реализована в асинхронном режиме на основе запросов к сервису тематической обработки, который возвращает результаты по мере их готовности.

Сервис доступа к данным и метаданным в архиве, как следует из его названия, отвечает за все запросы к метаданным и соответствующим данным в архивах. Существенно, что для получения требуемых информационных продуктов недостаточно получить только информацию о его наличии, необходимо также иметь всю связанную с выбранным экземпляром продукта справочную информацию, такую как характеристики спутникового прибора и правила формирования результирующего продукта.

Сервис формирования изображений по данным архива запрашивает требуемые для его работы расширенные метаданные и сами данные у сервиса доступа к данным и метаданным в архиве. Этот сервис является наиболее «узким местом» при возрастании одновременного числа пользователей картографического интерфейса, поэтому, на первом этапе, именно его было решено масштабировать при возрастании нагрузки.

Сервис тематической обработки данных предназначен для проведения асинхронной обработки спутниковых данных. Он исполняет роль диспетчера, а также отвечает за формирование итогового продукта на основе соответствующим образом подготовленных компонент спутниковых данных, получаемых от сервиса формирования изображений по данным в архиве.

При реализации описываемого макета системы доступа к спутниковым данным с целью упрощения было решено разместить все требуемые сервисы на трех различных виртуальных машинах: сервер доступа, сервер тематической обработки и сервер формирования изображений. На сервере доступа был развернут картографический WEB интерфейс, а также сервис доступа к данным и метаданным в архивах. На нем же в рамках реализованного макета реализовано хранение спутниковых данных и ведение соответствующих архивов. Предусмотрена поддержка работы, как с данными, размещенными локально, так и с данными, получаемыми из внешних архивов. Сервер тематической обработки спутниковых данных отвечает за работу соответствующего сервиса. Число активных серверов формирования изображений зависит от текущей нагрузки: увеличивается при возрастании числа запросов и уменьшается при сокращении их числа.

Для каждой из трех виртуальных машин был подготовлен образ, поддерживающий механизмы конфигурации на использование выдаваемого ему при инициализации IP адреса. Запросы к сервису формирования изображений направляются на один специально выделяемый для этой цели адрес, на котором размещается прокси-сервер, адресующий их на один из активных сервисов формирования изображений. Количество активных сервисов формирования изображений автоматически возрастает по мере увеличения числа запросов со стороны конечных пользователей, в роли которых, в частности, могут выступать студенты, обучающиеся основам работы со спутниковыми данными ДЗЗ.

Облачная платформа C2

Платформа C2 [5] базируется на технологии виртуализации сетевых функций (Network Functions Virtualization, NFV), позволяющей отделить логику сетевых функций и оборудования, на котором они выполняются, с помощью техники виртуализации физических ресурсов. Благодаря этому достигается ускорение процедуры развертывания новых сервисов и удаленная их конфигурация, что позволяет пользователям получать требуемые сервисы по запросу. Разработанные модули управления и оркестрации виртуальными сетевыми сервисами и реализованные алгоритмы планирования аппаратных ресурсов центров обработки данных позволяют достичь высокой эффективности работы платформы и снижения накладных расходов на поддержку масштабируемости и доступности сетевых сервисов.

Облачная платформа C2 является расширением проекта Self-Organized Cloud (SOC) [5] и других научных проектов [6], посвященных задачам размещения ресурсов в центрах обработки данных. В рамках этого подхода основная функциональность менеджера виртуальной инфраструктуры реализована на базе платформы OpenStack.

Жизненный цикл виртуальной сетевой функции охватывает все этапы существования функции от создания до ее удаления. К жизненному циклу относятся настройка функции, мониторинг ее основных показателей функционирования, и изменение конфигурации ПО (входящего в состав функции) в зависимости от изменения состояния конкретной функции (пример состояния: недоступность функции, или перегруженность функции). Управление полным жизненным циклом виртуальных сетевых сервисов в центрах обработки данных, по сути, является основной задачей облачной платформы C2.

Конфигурация, размещение, мониторинг экземпляров виртуальных функций и предоставление доступа к ним пользователям осуществляется на основе специально разработанного полнофункционального графического интерфейса.

Представленная технология позволяет эффективно создавать виртуальную инфраструктуру, как для обучения студентов, так и для проведения различных научных исследований. Основными преимуществами использования такой инфраструктуры являются: существенное повышение эффективности использования оборудования, содействие сотрудничеству между исследовательскими организациями, а также возможность воспроизведения результатов научных исследований, полученных различными группами ученых.

Заключение

Реализация действующего макета системы доступа к спутниковым данным, интегрированной в облачную платформу C2, позволила нам выработать подходы, которые в дальнейшем могут быть использованы как для обучения студентов основам работы со спутниковыми данными ДЗЗ, так и для других прикладных и исследовательских задач. Благодаря использованию ресурсов и возможностей ЦКП «ИКИ-Мониторинг» пользователи построенных по такому принципу систем смогут получить доступ к сверхбольшим архивам различных спутниковых данных и сервисам для их обработки и анализа.

В перспективе планируется внедрение представленной технологии на площадках различных университетов и, в частности, на вычислительных ресурсах 4-5 естественных факультетов МГУ им. Ломоносова.

При создании макета использовались ресурсы Центра коллективного пользования «ИКИ-Мониторинг», работа которого осуществляется при поддержке темы «Мониторинг», госрегистрация №01.20.0.2.00164.

References

- [1] Loupian E.A., Kashnitskiy A.V., Bourtsev M.A., Proshin A.A., Balashov I.V., Tolpin V.A. Development of Earth remote sensing data analysis tools for very large distributed data archives, *Geoinformatics Research Papers*, 2017, Vol. 5, BS1002, pp. 159. DOI: 10.2205/CODATA2017.
- [2] Lizhe Wang, Yan Ma, Jining Yan, Victor Chang, Albert Y. Zomaya, pipsCloud: High performance cloud computing for remote sensing big data management and processing, *Future Generation Computer Systems*, Volume 78, Part 1, 2018, Pages 353-368, ISSN 0167-739X, DOI: <https://doi.org/10.1016/j.future.2016.06.009>.
- [3] Loupian E.A., Proshin A.A., Burtsev M.A., Balashov I.V., Bartalev S.A., Efremov V.Yu., Kashnitskiy A.V., Mazurov A.A., Matveev A.M., Sudneva O.A., Sychugov I.G., Tolpin V.A., Uvarov I.A. IKI center for collective use of satellite data archiving, processing and analysis systems aimed at solving the problems of environmental study and monitoring, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2015, Vol. 12, No 5, pp. 263-284.
- [4] Antonenko V.A., Smelyanskiy R.L., Ermilov A.A. C2 Virtual Network Functions Management Platform, *Information Technology*, 2018, Vol. 24, No 3, pp. 165-173 UDC 004.738.

- [5] Kostenko, V., Plakunov, A., Nikolaev, A., Tabolin, V., Smeliansky, R., and Shakhova, M. Selforganizing cloud platform // In Proceedings of the 2014 international science and technology conference "Modern Networking Technologies (MoNeTec)", Moscow, Russia, 2014 <https://doi.org/10.1109/MoNeTeC.2014.6995588>
- [6] Vdovin A. V., Zotov I. A., Kostenko V. A., Plakunov A. V., and Smelyansky R. L., Comparing various approaches to resource allocating in data centers, *Journal of Computer and Systems Sciences International*, vol. 53, no. 5, pp. 689–701, 2014. <https://doi.org/10.1134/S1064>.