

В.М. Котов (Объединенный институт ядерных исследований)

В 1963 г. окончил Харьковский авиационный институт по специальности «радиоинженер». В настоящее время — ведущий научный сотрудник Объединенного института ядерных исследований.

Н.А. Русакович (Объединенный институт ядерных исследований)

В 1975 г. окончил Белорусский государственный университет по специальности «физика». В настоящее время — главный научный секретарь Объединенного института ядерных исследований.

Система удаленного мониторинга экспериментов Большого адронного коллайдера

Сбор, предварительная обработка и анализ данных, полученных при дистанционном зондировании Земли (ДЗЗ) с помощью космических радиолокаторов с синтезированной апертурой (РСА), связаны с накоплением и обработкой информации, объемы которой в десятки тысяч раз превышают средний уровень потребностей и достигают десятков терабайт в год.

С 90-х гг. прошлого столетия пользователи имеют возможность получать данные космических аппаратов, имеющих РСА (Envisat, RADARSAT, TerraSAR, COSMO-SkyMed), уже собран большой архив таких данных [1].

В настоящее время и в России в рамках Федеральной космической программы ведется разработка космических аппаратов («Метеор-М №3», «Аркон-2М»), оснащенных многофункциональной РСА с активной фазированной антенной решеткой, характеристики которой соответствуют характеристикам современных РСА [2].

Традиционные подходы к разработке инфраструктуры и обработки таких объемов данных (получивших название «большие данные») не годятся. Необходимы новые подходы к обработке и анализу данных для современных РСА, в том числе и реализация интерактивного взаимодействия в режиме удаленного доступа для многоуровневой, распределенной географической вычислительной системе обработки данных.

Аналогичная проблема стояла и в области экспериментальной физики высоких энергий еще в 90-е гг. прошлого столетия при формировании программы исследований крупнейшего проекта современности в

области фундаментальной науки: создании Большого адронного коллайдера (БАК) в Европейском центре ядерных исследований (ЦЕРН), г. Женева, Швейцария.

В результате многолетней работы большого коллектива физиков, инженеров и программистов в ЦЕРНе была разработана, отлажена, запущена и успешно работает второй год в режиме реальных измерений на пучках ускорителя БАК система сбора и обработки данных эксперимента ATLAS БАК, объединяющая в единую систему все компоненты обработки, управления и передачи «больших данных» (это передача и фильтрация в реальном времени 100 Gb/s и офлайн-обработка и анализ экспериментальных данных объемом в десятки петабайт в год) [3].

Объединенный институт ядерных исследований (ОИЯИ) является участником программы исследований на БАК. Важной составляющей такого участия в обработке и анализе данных, получаемых в экспериментах на БАК, является создание в ОИЯИ системы удаленного доступа реального времени (СУДРВ) и интеграция ее в глобальную сервис-ориентированную архитектуру ГРИД-системы сбора и обработки данных экспериментов на БАК [4].

На сегодня не существует технологии, обеспечивающей необходимую функциональность и эффективность обработки таких потоков и объемов экспериментальных данных в одной системе. Поэтому система сбора и обработки данных эксперимента ATLAS БАК была разделена на два уровня: онлайн — сбор и предварительная обработка — и офлайн — полная

Generic Data Acquisition and Computing system

No technology today provides the functionality and performance required by the complexity of the interconnections and the overall throughput

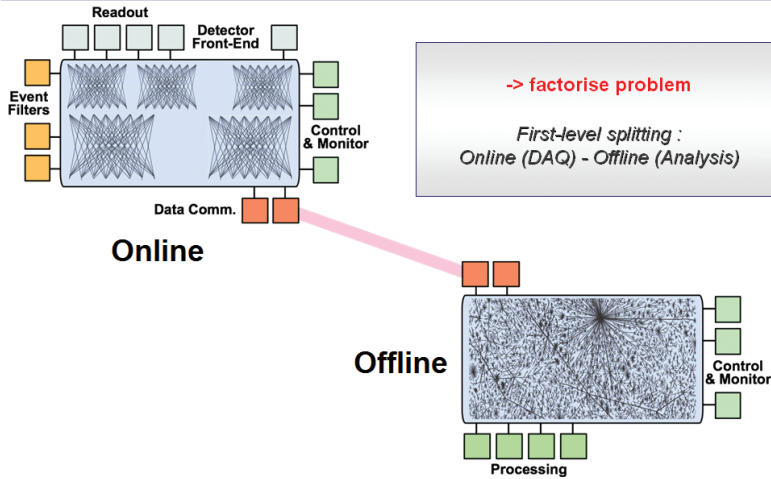


Рис. 1. Общая система онлайн и офлайн

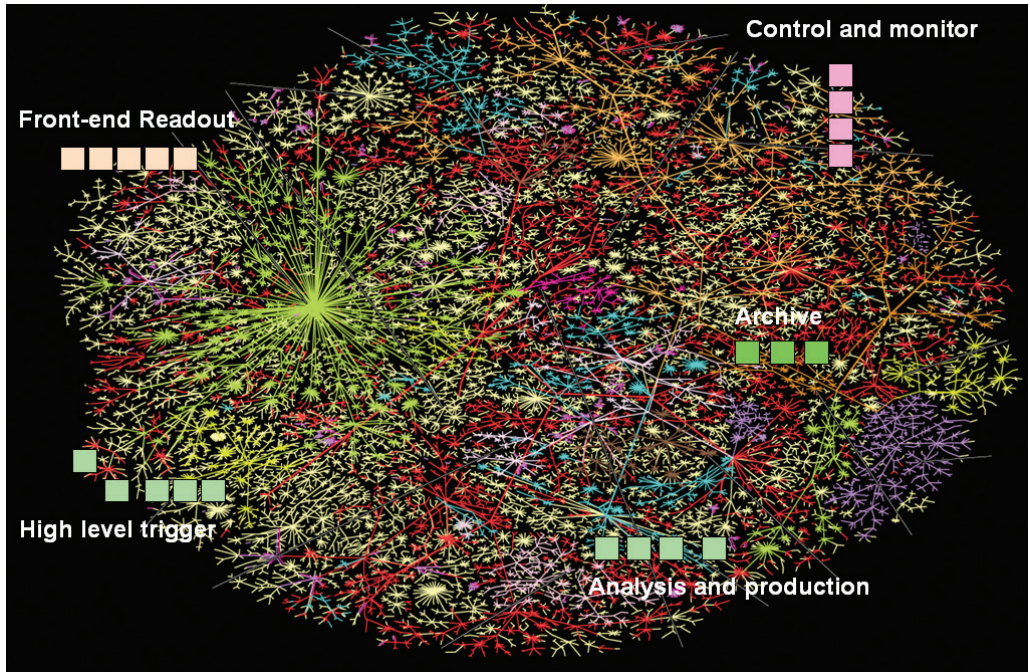


Рис. 2. Система офлайн-обработки

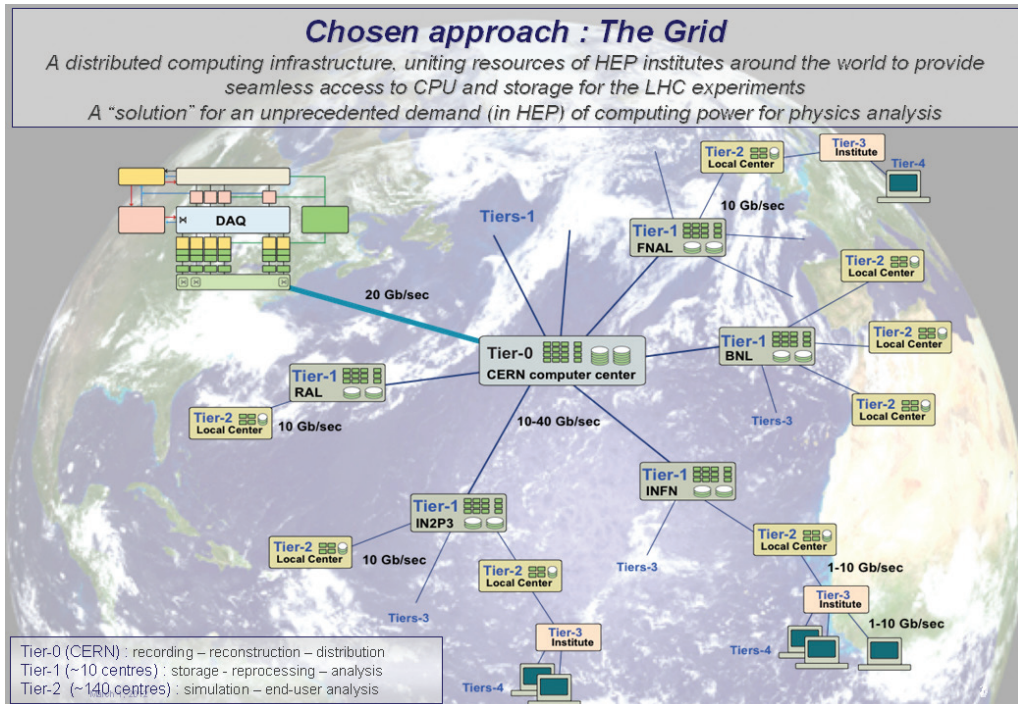


Рис. 3. ГРИД-система обработки данных экспериментов на БАК

обработка и анализ (рис. 1, 2).

Офлайн-обработка требует необычно больших вычислительных ресурсов (примерно 100 000 самых мощных на сегодня процессоров или 150 Kcores) и была реализована в ГРИД-системе (рис. 3).

Показанная в левом верхнем углу на рис. 3. система DAQ и есть онлайн-составляющая системы по сбору и предварительной обработке входных данных.

Опыт разработки и реализации системы сбора и обработки данных экспериментов на БАК может быть использован при создании систем обработки данных в системах ДЗЗ, и в частности для данных, полученных с помощью космических радиолокаторов с синтезированной апертурой.

Современная архитектура наземной географической распределенной системы обработки данных SAR (Synthetic Aperture Radar), предлагаемая Европейским космическим агентством (ESA), также имеет уровень предварительной обработки входных данных, необходимость которой определяется особенностью фор-

матов и структурой данных SAR для выполнения в последующем полного цикла обработки информации с SAR в пакетном режиме. Кроме того, структура радарных данных позволяет проводить эффективный анализ на уровне первичной обработки радиолокационных изображений, оперируя изображением в целом как образом для быстрого опознавания и анализа некоторого класса объектов и их характеристик в интерактивном режиме.

ESA в октябре 2007 г. заключило договор с канадской фирмой Agray Systems Computing на разработку инструментального программного обеспечения и с 2010 г. поставляет комплект программного обеспечения с открытым исходным кодом NEST (Next ESA SAR Toolbox) [5], включающий в себя функциональность всех предыдущих версий инструментального ПО BEST, BEAM и др. Состав пакета NEST, компоненты и потоки данных приведены на рис. 4.

По условиям технического задания ESA инструментарий NEST предназначен для помощи в подго-

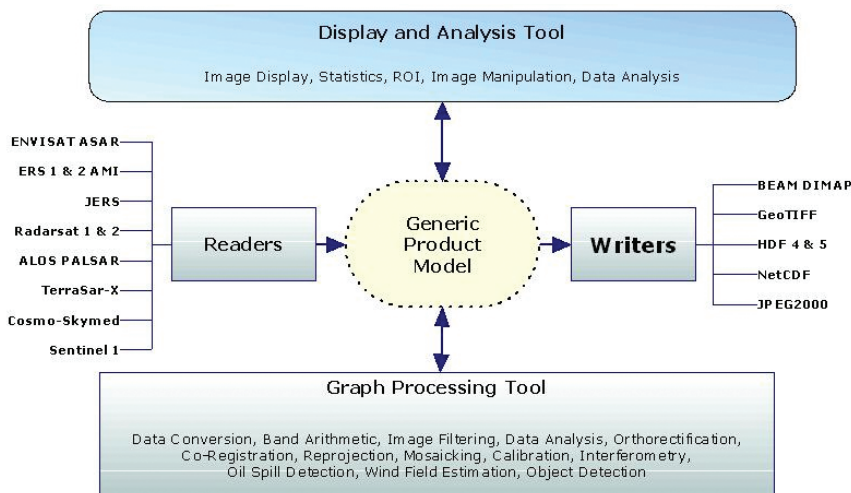


Рис. 4. Компоненты и потоки данных программного комплекса NEST

товке элементов системы обработки данных SAR, но не является SAR-процессором или готовой системой обработки SAR-данных в полном технологическом цикле.

Вместе с тем NEST как программа с открытым исходным кодом, или, как принято обозначать это движение в России, свободное программное обеспечение (СПО), имеет четыре свободы, в частности разрешает не только изучать и свободно изменять такие программы («нулевая» и «первая» свободы СПО), но и свободно распространять как копии программы («вторая свобода»), так и измененные версии с тем, чтобы принести пользу всему сообществу («третья свобода»). Может быть, поэтому функционалы известных на рынке программных продуктов для обработки радарных данных практически идентичны.

Однако современные тенденции, а тем более будущее развитие программ глобального мониторинга окружающей среды невозможны без широкого международного сотрудничества, и поэтому самым эффективным способом реализации возможностей NEST является его использование в составе PaaS (Platform as a Service), имеющей также открытый код, для разработки, тестирования, развертывания и поддержки полномасштабной системы сбора и обработки данных с космических локаторов PCA, реализующей полный технологический цикл обработки радар-

ных данных. Предполагается в составе PaaS реализовать набор инструментов для отображения и частичной обработки радарных снимков в среде ArcGIS.

В качестве основы для разработки прототипа подобной PaaS предлагается использовать систему удаленного доступа реального времени (СУДРВ) ОИЯИ. СУДРВ представляет собой композитный сервис ГРИД-системы и является фрагментом общей системы обработки данных эксперимента ATLAS БАК [6]. Согласно планам развития информационной инфраструктуры экспериментов БАК ЦЕРН и в соответствии с концепцией «облачных вычислений», где все есть Сервис (XaaS), СУДРВ ОИЯИ будет применяться как PaaS для дальнейшего развития системы обработки данных эксперимента ATLAS на БАК.

Кроме того, в соответствии с концепцией «Открытой инновационной лаборатории» эксперимента ATLAS-LAB (ATLAB) на рабочем совещании ЦЕРН—ОИЯИ «Brainstorming workshop on applications from ATLAS using EU-funding for R&D-upgrades» в г. Дубне 24 октября 2010 г. по обсуждению доклада ОИЯИ «Real time remote access system for ATLAS» было поддержано предложение о возможности прикладного использования СУДРВ ОИЯИ в области космического мониторинга, проводимых совместно с ЦЕРН при поддержке ESA.

В настоящее время обсуждается возможность включения партнерства ЦЕРН—ОИЯИ в состав миссии содействия программы Global Monitoring for Environment and Security (GMES). Программа GMES, проводимая Европейским агентством окружающей среды (European Environment Agency, EEA) в партнерстве с ESA, предполагает создание единой системы, объединяющей данные наблюдений на суше, море и в атмосфере с данными космического мониторинга, от успешного использования которых в значительной степени будет зависеть успех программы в целом.

Основой космической составляющей GMES будут не только спутники Sentinel, специально создаваемые для этой программы (оператор ESA), но и другие национальные и международные спутниковые системы ДЗЗ, партнерство с которыми будет оформлено как миссии содействия GMES (GMES Contributing Missions (GCMs)).

Структура программного обеспечения СУДПВ и NEST использует объектноориентированный подход проектирования и соответствует стандарту PSS05 ESA. Объединение NEST и СУДПВ в единую платформу обеспечит интеграцию NEST в общую систему ГРИД-обработки данных экспериментов БАК, а значит, и возможность отладки в последующем и оффлайн-режима обработки данных космического мониторинга в географически распределенной вычислительной системе ГРИД-обработки и партнерство с участием ESA и ЦЕРН.

Следует отметить, что, кроме инструментария NEST, отражающего специфику обработки радарных данных, предлагаемая платформа PaaS будет обеспечивать также доступ к сервисам сбора и обработки данных, необходимых для функционирования NEST в составе СУДПВ.

В качестве базового ядра предлагаемой системы удаленного доступа для сбора и обработки космической радиолокационной информации предполагается использовать набор компонентов из системы сбора и обработки данных в реальном времени эксперимента ATLAS БАК [7].

Интеграция инструментального пакета NEST ESA и системы удаленного доступа СУДПВ ОИЯИ позволит создать развитую платформу (PaaS) для разработки прототипов (макетов) систем предварительной обработки радарных данных космических лока-

торов с синтезированной апертурой высокого разрешения и отладки их в условиях реальной работы в большой географически распределенной вычислительной ГРИД-системе обработки данных экспериментов на БАК.

СПИСОК ЛИТЕРАТУРЫ

1. Кантемиров Ю.И. (Компания «Совзонд»), Обзор современных радиолокационных данных ДЗЗ и методик их обработки с использованием программного комплекса SARSCAPE
2. Костюк Е.А., Веремчук Ю.А., Денисов П.В. Перспективные технологии обработки космической радиолокационной информации в НКПОР Оператора КС ДЗЗ // V Международная конференция «Космическая съемка — на пике высоких технологий».
3. Mapelli L. Spanning from Data Acquisition to GRID – Today and a view of tomorrow. // XXIII International Symposium on Nuclear Electronics & Computing NEC'2011.
4. Кореньков В.В., Котов В.М., Русакович Н.А., Яковлев А.В., Система удаленного доступа реального времени (СУДПВ) как композитный сервис распределенной ГРИД-системы обработки данных экспериментов на Большом Адронном Коллайдере (БАК) // Параллельные вычислительные технологии (ПаВТ, 2010): Труды международной научной конференции (Уфа, 29 марта – 2 апреля 2010 г.).
5. Software Architecture Document (SAD) for the Next ESA SAR Toolbox (NEST) (ARR-NEST-RS07-016); http://www.array.ca/nest/Software_Architecture_Document_v2.0.pdf
6. Кореньков В.В., Котов В.М., Русакович Н.А., Яковлев А.В., Модель и технология интеграции online-сервисов эксперимента ATLAS на Большом Адронном Коллайдере (БАК) и сервисов ГРИД-инфраструктуры // Параллельные вычислительные технологии (ПаВТ, 2011): труды международной научной конференции (Москва, 28 марта – 1 апреля 2011 г.).
7. ATLAS High-Level Trigger, Data Acquisition and Controls. Technical Design Report (ATLAS TDR-016); <http://atlas-proj-hlttdaqdcs-tdr.web.cern.ch/atlas-proj-hlttdaqdcs-tdr/tdr-v1-r4/PDF/TDR.pdf>