

Н.М. Куссуль, С.В. Скакун, А.Ю. Шелестов

Інститут космічних досліджень НАНУ–НКАУ, Київ

ГЕОІНФОРМАЦІЙНА ІНФРАСТРУКТУРА МОНІТОРИНГУ НАВКОЛИШНЬОГО СЕРЕДОВИЩА ТА НАДЗВИЧАЙНИХ СИТУАЦІЙ



Розглянута загальна архітектура розподіленої системи для обробки геопросторових даних і представлення результатів, яка призначена для моніторингу довкілля і надзвичайних ситуацій. Зокрема, в системі реалізована підтримка автоматизованої обробки радіолокаційних даних супутника ENVISAT/ASAR, а також інформаційні технології оцінки площ затоплених територій. Результати отримані при виконанні науково-технічного проекту НАН України за підтримки міжнародної програми UN-SPIDER і космічного агентства NASA.

Ключові слова: інформаційні технології, обробка геопросторових даних, розподілена система, моніторинг надзвичайних ситуацій.

Значне збільшення кількості природних надзвичайних ситуацій за останні роки продемонструвало людству важливість моніторингу та попередження катастроф. Зміни клімату призвели до збільшення злив, ураганів, повеней, посух та інших погодних аномалій. У звіті [1], підготовленому за сприяння Організації Об'єднаних Націй (ООН), зазначено помітне збільшення кількості надзвичайних ситуацій за останні роки. Так, наприклад, за 2007 р. сталося 414 природних надзвичайних ситуацій, які призвели до загибелі 16 847 людей та збитків на суму більше 70 млрд. дол. США. У звіті також відзначається, що більше половини природних надзвичайних ситуацій складають гідрометеорологічні явища — повені та зсуви ґрунтів, обумовлені опадами. Причому за останнє десятиріччя кількість таких лих збільшується в середньому на 7,4 % щорічно.

Україна також вразлива до повеней, особливо в Західному регіоні, де паводки спостерігаються майже кожного року. Наприклад, під час

повені в 2008 р. в західних регіонах України загинуло 34 людини, близько 25 000 людей були евакуйовані. Завдані повенями збитки оцінюються в 3–4 млрд. грн. Іншим стихійним лихом, яке завдає значних втрат економіці України, є посухи. Весною та влітку 2007 р. південні регіони України постраждали від значної засухи. За даними Міністерства надзвичайних ситуацій України внаслідок посухи було повністю знищено близько 1400 тис. га сходів (тобто 16 % від загальної площі посівів) та пошкоджено посіви сільськогосподарських культур загальною площею близько 8500 тис. га. Сума завданих збитків становила понад 450 млн. грн. (надзвичайна ситуація державного рівня).

З огляду на вищесказане надзвичайно важливого значення набули розробки інформаційних інфраструктур для моніторингу довкілля і надзвичайних ситуацій. За оцінками спеціалістів США, залучення даних дистанційного зондування Землі дає можливість зменшити затрати на моніторинг зрошуваних земель більш ніж на 80 % порівняно з наземними методами [2].

На сьогоднішній день попередження та моніторинг надзвичайних ситуацій неможливі без залучення даних спостереження Землі з космосу. Існує ряд міжнародних ініціатив, які спрямовані на використання супутникових даних для попередження та ліквідації надзвичайних ситуацій та екстреного реагування. До них слід віднести міжнародну «систему систем» GEOSS, Міжнародну хартію по космосу та великих катастрофах, Партнерство по комплексній стратегії глобальних спостережень, Глобальний моніторинг в інтересах охорони навколишнього середовища та безпеки (GMES), Програму попередження та зменшення наслідків стихійних лих Всесвітньої метеорологічної організації, Платформу ООН UN-SPIDER, Міжнародну ініціативу International Charter «Space and Major Disasters». Крім того, такі ініціативи, як RESPOND-GMES та UNOSAT, надають комерційні інформаційні продукти для прийняття рішень (заходів) при надзвичайних ситуаціях.

Важливим аспектом при використанні супутникових даних в умовах виникнення надзвичайних ситуацій є оперативність отримання і обробки даних та надання відповідних інформаційних продуктів кінцевим користувачам. Наприклад, під час повені в 2008 р. в Західній Україні виявилось неможливим забезпечити оперативний прийом даних від зарубіжних супутників, які б дозволили швидко оцінити масштаби катастрофи та скоординувати зусилля для її ліквідації.

Тому Інститут космічних досліджень (ІКД) НАНУ–НКАУ ініціював виконання науково-технічного проекту зі створення сучасної геоінформаційної інфраструктури екологічного моніторингу довкілля та надзвичайних ситуацій. Метою науково-технічного проекту стало створення геоінформаційної інфраструктури для забезпечення швидкого та ефективного доступу, обробки та надання геопросторової інформації з метою моніторингу навколишнього середовища та надзвичайних ситуацій. В рамках інформаційної інфраструктури впроваджені сервіси оцінки площ затоплених територій та стану рослин-

ності за даними супутникового спостереження. Зазначені сервіси впроваджені в межах системи GEO-Ukraine, платформи ООН UN-SPIDER та розроблюваного в межах системи GEOSS пілотного проєкту Caribbean Flood Pilot.

ЗАГАЛЬНА АРХІТЕКТУРА СИСТЕМИ АВТОМАТИЧНОЇ ОБРОБКИ ДАНИХ СУПУТНИКА ENVISAT

В межах геоінформаційної інфраструктури забезпечено автоматизований цикл отримання, обробки та візуалізації радіолокаційних супутникових даних ENVISAT/ASAR з метою швидкого картографування затоплених повенями територій. Для доступу до супутникових даних Європейського космічного агентства (ЄКА) створено графічний інтерфейс користувача (портал) та програмні інтерфейси. В архівах ЄКА дані зберігаються протягом 1–2 год з моменту їх отримання. Реалізація відповідних інтерфейсів дає можливість оперативно (1–3 год) отримувати необхідні дані, які використовуються для створення інформаційних продуктів щодо оцінки площ затоплених територій. Як вхідна інформація задаються координати території, де відбувається (відбудеться) надзвичайна ситуація (повені). В результаті знаходиться перелік відповідних зображень. Отримані дані зберігаються в межах Grid-інфраструктури, яка створена в ІКД НАНУ–НКАУ та ІК НАНУ. Для підвищення ефективності обробки даних автоматизований цикл реалізовано в межах Grid-середовища. В результаті виконання автоматизованого циклу надаються інформаційні продукти для прийняття рішень (заходів) при надзвичайних ситуаціях. Для цього використовуються міжнародні стандарти OGC та INSPIRE. Інформація щодо затоплень надається у вигляді карт через мережу Internet та систему Google Earth (файли KML).

Розглянемо сервіси, реалізовані в межах геоінформаційної інфраструктури для виконання автоматизованого циклу обробки супутникових даних (рис. 1).

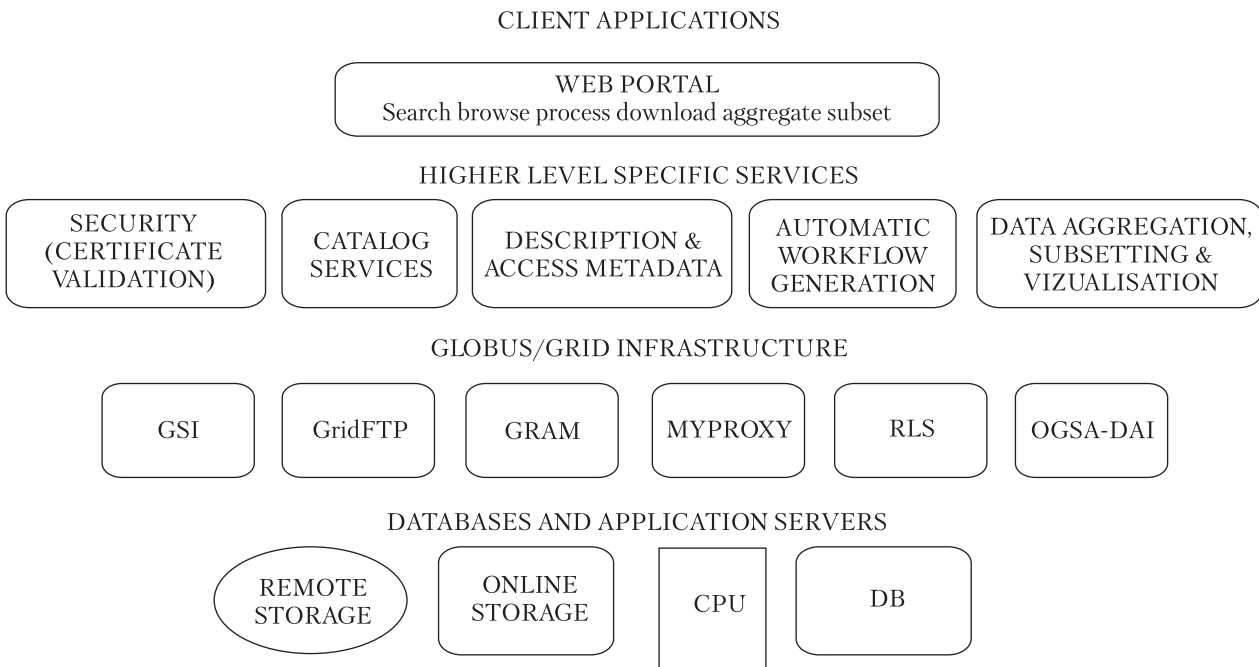


Рис. 1. Перелік сервісів в межах розробленої геоінформаційної інфраструктури

Точкою входу у систему є портал (рис. 2). Портал забезпечує прозору для користувача взаємодію з ресурсами різного рівня, даючи можливість отримати за створеним власноруч запитом оброблені супутникові дані, що покривають потрібну територію. Портал безпосередньо пов'язаний з сервісами вищого рівня. До них відносяться: підсистема безпеки; каталог даних; метадані (опис, доступність); сервіс автоматичного генерування workflow; сервіс агрегації, вибору та візуалізації даних.

Високорівневі сервіси, що потребують використання значних обчислювальних ресурсів, взаємодіють із сервісами нижчого рівня, які реалізовані в межах Grid-інфраструктури [3–4]. Ці сервіси призначені для планування та розподілу задач, що виконуються, на обчислювальних ресурсах.

До складу Grid-інфраструктури входять такі компоненти:

- ✦ інфраструктура безпеки Grid (GSI);
- ✦ підтримка протоколу передачі даних Grid (GridFTP);

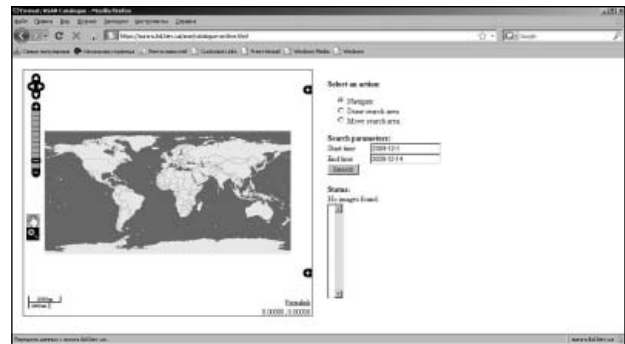


Рис. 2. Портал системи

- ✦ управління та призначення Grid-ресурсів (GRAM);
- ✦ система збереження сертифікатів та управління повноваженнями (MyProxy);
- ✦ система визначення місцезнаходження реплік (RLS);
- ✦ система організації спільного доступу до даних (OGSA-DAI).

Найнижчим рівнем у ієрархії є сервіси спеціалізованих серверів та бази даних. До них відносяться: віддалене сховище даних; опера-

ційне сховище даних; система обчислення; сервери баз даних.

Стандартизованість та сумісність різнорівневих компонентів дозволяє в разі необхідності масштабувати систему, не заважаючи її керованості. Складність взаємодії окремих підсистем прихована від кінцевого користувача.

ВИСОКОРІВНЕВІ СЕРВІСИ

Каталог даних. Супутникові дані зберігаються на ресурсах Grid-інфраструктури. Для пошуку та отримання доступу до даних створено відповідні сервіси.

Метадані. В межах порталу користувачеві надаються можливості пошуку та вибору необхідних супутникових даних. Для цього використовуються метадані, які зберігаються у відповідному каталозі. До важливої інформації, яка використовується для пошуку, належить ім'я файлу, дата зйомки, географічне покриття, режим зйомки (Wide Swath Mode – WSM, Image Mode – IM, Alternating Polarization Mode – APM). При зверненні до цієї підсистеми формується відповідь, що є основою для відображення на порталі необхідної інформації.

Сервіс автоматичного генерування потоку виконання задач. Різні компоненти, які відповідають за кроки попередньої обробки, тематичної обробки та візуалізації геопросторових даних, об'єднані в межах єдиного автоматизованого циклу. Для підвищення ефективності всього процесу цей цикл реалізовано в межах Grid-середовища. Крім того, створено паралельну версію алгоритму тематичної обробки супутникових зображень. Паралелізацію виконано таким чином: вихідне супутникове зображення автоматично розбивається на рівні частини, які обробляються на різних обчислювальних вузлах з використанням інтерфейсу OpenMP Application Program Interface (www.openmp.org). Для керування циклом обробки використано стандартні засоби для роботи з потоком виконання задач в Grid-системах (мова Karajan). Використання паралельної версії дало можливість зменшити час, не-

обхідний для виконання обчислень з 30 хв на одному обчислювальному вузлі до менше 1 хв в межах Grid-системи.

Сервіс агрегації, вибору та візуалізації даних. Результати обробки даних необхідно надати у вигляді, доступному для користування у широко розповсюдженому програмному забезпеченні. Для цього використовуються міжнародні стандарти візуалізації та представлення інформації (KML, WMS, ESRI shape). Сервіс агрегації, вибору та візуалізації даних призначений для перетворення внутрішнього представлення результатів обробки даних у загальнодоступний формат.

СЕРВІСИ GRID-ІНФРАСТРУКТУРИ

Інфраструктура безпеки Grid (GSI). Підсистема безпеки призначена для встановлення ідентифікації клієнта (аутентифікації) та керування правами доступу (авторизації). Основою для функціонування підсистеми безпеки є система користувацьких сертифікатів. Сучасні криптографічні алгоритми дозволяють перевірити аутентичність користувача з невеликими затратами. Єдиною умовою для надійності процесу є довіра до органу (сертифікаційного центру), що видає сертифікати. Сертифікаційний центр є публічно відомим суб'єктом постійного контролю репутації.

Інфраструктура безпеки Grid забезпечує підтримку політик доступу до даних, специфічну для Grid-систем. Кожна взаємодія в межах інфраструктури потребує двосторонньої аутентифікації: клієнт має переконатися, що сервер є тим, за кого себе видає, та навпаки. Для цього також використовуються сертифікати (користувацькі та серверні), підписані довіреним сертифікаційним центром.

Специфічною для Grid-систем є концепція делегування прав від користувача до сервера. Для проведення обчислень сервер повинен мати можливість виконувати дії від імені користувача. Реалізується це за допомогою короткострокових проксі-сертифікатів. Їх використання дає змогу досягнути гнучкості, зберігаючи безпеку.

Підтримка протоколу передачі даних Grid (GridFTP). Обмін даними між вузлами Grid-інфраструктури здійснюється за певним протоколом, який є розширенням стандартних протоколів передачі даних. Необхідність розширення стандартних протоколів зумовлена вимогами, по-перше, до швидкої передачі даних великих об'ємів, і, по-друге, до безпеки системи в цілому. Надійність та доступність збережених даних забезпечується за рахунок використання сховищ великого об'єму.

Управління та призначення Grid-ресурсів (GRAM). В умовах високої завантаженості важливо оптимально використовувати наявні ресурси. За розподілення задач між обчислювальними ресурсами відповідає окрема підсистема. Вона зберігає та аналізує інформацію про поточне завантаження ресурсів і приймає рішення щодо умов виконання задач, які надходять. Компонентам, які використовують GRAM, надається прозорий інтерфейс для запуску задач і отримання результатів їх виконання.

Система збереження сертифікатів та управління повноваженнями (MyProxy). Незважаючи на принципову загальнодоступність сервісу існують обмеження, пов'язані з необхідністю наявності у користувача його особистого сертифікату при здійсненні запиту до системи. Для того щоб полегшити використання однією особою або компонентом системи ресурсів з різних місць, введено окремий сервіс, який керує повноваженнями. Клієнт, переконавшись у аутентичності сервісу, може передати йому повноваження, використовуючи свій сертифікат. Ця делегація повноважень обмежена у часі, але є достатньо довгою для зручної роботи (зазвичай один–два тижні). Протягом цього часу користувач може безпечно використовувати Grid-сервіси від свого імені, не маючи безпосереднього доступу до власного сертифікату, за тієї умови, що використовувані сервіси довіряють серверу MyProxy.

Система визначення місцезнаходження реплік (RLS). Сервіс RLS працює у зв'язці з GridFTP та створює більш високий рівень аб-

стракції для роботи по збереженню даних. Сервіс дозволяє отримувати доступ до даних за системою імен незалежно від фізичного місцезнаходження даних. Користувач згідно з цією схемою ізолюваний від деталей розподілення даних по ресурсах. Сервіс дає можливість отримати меншу затримку при доступі до даних у разі використання реплікацій за рахунок автоматичного вибору найближчої репліки.

Система організації спільного доступу до даних (OGSA-DAI). Сервіс OGSA-DAI призначений для уніфікації доступу до даних, що зберігаються у багатьох місцях та у різних форматах. Цей сервіс дозволяє трансформувати дані із однієї форми представлення у іншу, інтегрувати дані з різних джерел у єдину базу та надавати дані кінцевим користувачам за допомогою загальноновживаних механізмів передачі даних.

Спеціалізовані сервери та бази даних. Віддалене сховище дає можливість ефективно зберігати дані, які є важливими, але не потребують швидкого операційного доступу. Дані зберігаються на спеціалізованому обладнанні, налаштування якого передбачає автоматичне створення резервних копій та підтримку надлишковості збережуваних даних. Для зберігання поточних та проміжних результатів необхідно *операційне* сховище, яке буде забезпечувати мінімальну затримку при збереженні та доступі до даних. Таке сховище є спільним для декількох сервісів. Воно фізично розташоване окремо і доступне через мережу, але робота з ним ідентична роботі з локальною файловою системою.

Система високопродуктивних обчислень. Для обчислень використовуються вузли високопродуктивного кластера. Кластер налаштований таким чином, щоб завдання, які надходять для обчислень, не конкурували при доступі до спільних ресурсів. Локальний планувальник кластера відповідає за організацію роботи з чергою завдань, обліком результатів та веденням статистики.

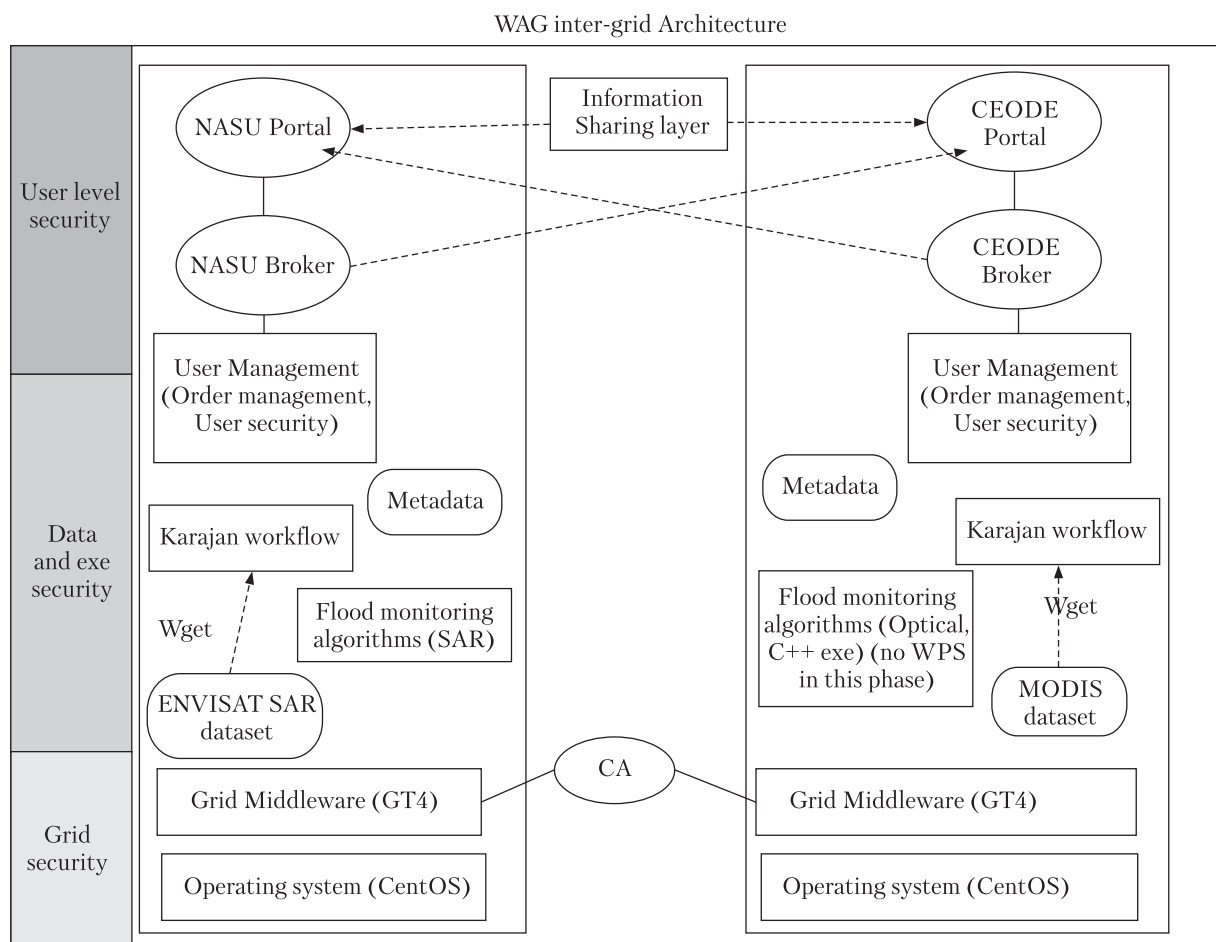


Рис. 3. Архітектура WAG

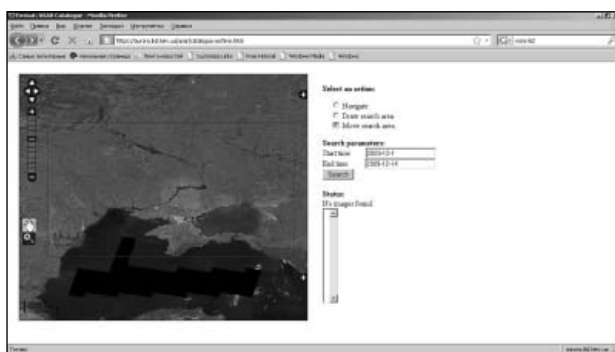


Рис. 4. Інтерфейс пошуку радіолокаційних супутникових даних

Сервери баз даних. Через значні вимоги до об'єму дані зберігаються на окремих ресурсах. Для зберігання використовуються спеціальні

структури даних та відповідним чином побудовані індекси, що дають змогу швидко знайти необхідну вибірку з даних великого об'єму.

WIDE AREA GRID – INTERGRID-ІНФРАСТРУКТУРА ДЛЯ МОНИТОРИНГУ ПОВЕНЕЙ

Метою впровадження архітектури WAG (Wide Area Grid) є об'єднання взаємодоповнюючих ресурсів ІКД НАНУ–НКАУ та центру Center for Earth Observation and Digital Earth (CEODE) Китайської академії наук. Українська сторона, зокрема, надає інформаційні продукти щодо затоплених територій з використанням радіолокаційних даних супутника ENVISAT/ASAR, а китайська сторона надає інформаційні продукти з використанням оп-

тичних спостережень супутників Terra та Aqua (прилад MODIS). Загальна архітектура інфраструктури наведена на рис. 3.

Портали. Безпосередньо з сервісами спільного використання даних взаємодіють портали кожної з організацій. Портали здатні повноцінно функціонувати окремо один від одного. Надбудовані над порталами сервіси дають можливість проводити оперативну взаємодію між ними безпосередньо після отримання запиту користувача.

Брокери. Портали організацій взаємодіють з відповідними системами через *брокер*. Так, для отримання метаданих про радіолокаційні дані, генерації та виконання автоматизованого циклу обробки даних ENVISAT/ASAR портал CEODE створює запит брокера, що розташований на ресурсах ІКД НАНУ–НКАУ. Отримуючи запит, брокер виконує пошук (рис. 4) і повертає перелік знайдених даних для відображення на порталі CEODE (рис. 5). Після того як користувач вибрав дані, знову генерується запит брокера для генерації автоматизованого потоку виконання задач. Варто відзначити, що брокер транслює зовнішні запити і ініціює та викликає локальні сервіси на ресурсах ІКД НАНУ–НКАУ. Обробка радіолокаційних даних здійснюється з використанням інтелектуальної методології, описаної в [5]. Проміжні та кінцеві результати обробки зберігаються на ресурсах ІКД НАНУ–НКАУ. Кінцеві геоінформаційні продукти подаються на портал CEODE з використанням стандартів OGC WMS та KML.

Аналогічним чином працює брокер на ресурсах CEODE. Він транслює запити, отримані від порталу ІКД НАНУ–НКАУ, для ініціалізації та виклику локальних сервісів щодо отримання, обробки та візуалізації даних приладу MODIS супутників Terra та Aqua.

Підсистеми керування обліком користувачів. Для доступу до ресурсів системи користувачеві необхідно отримати відповідний сертифікат. Сервер управління сертифікатами (VOMS – Virtual Organisation Management Server) знаходиться на ресурсах ІКД



Рис. 5. Відображення результатів пошуку

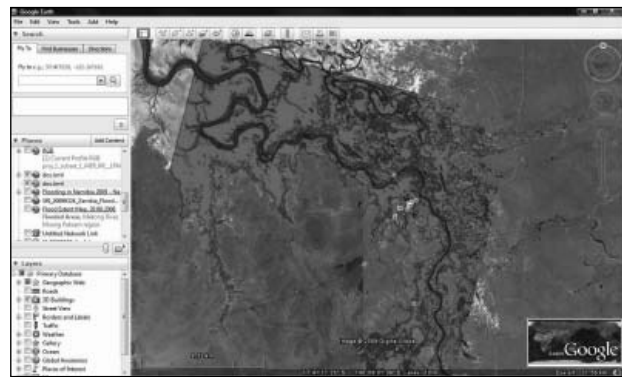


Рис. 6. Візуалізація результатів обробки у середовищі Google Earth

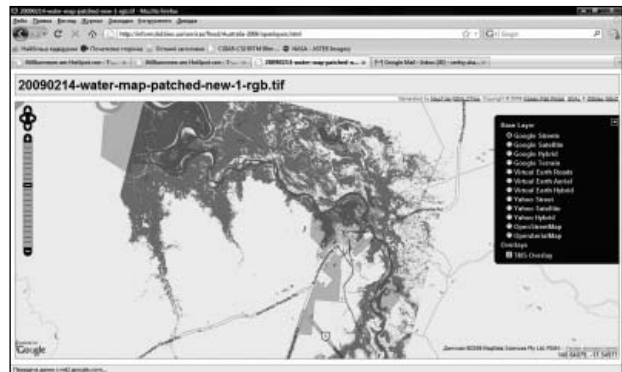


Рис. 7. Візуалізація результатів обробки за допомогою бібліотеки OpenLayers

НАНУ–НКАУ. Система моніторингу відслідковує всі запити, зроблені користувачами.

Інтеграція Grid-інфраструктур. Обчислювальні ресурси кожної з організацій поєднані з використанням безпечних механізмів аутенти-

фікації сервісів. Це можливо завдяки наявності сертифікатів, підписаних центром сертифікації, який визнається обома організаціями. Для інтеграції ресурсів української та китайської сторін розроблено брокери, що транслюють зовнішні запити для виконання локальних сервісів.

В процесі автоматичної обробки даних користувач може перевірити статус виконання кожного кроку, натиснувши кнопку. Після виконання всіх етапів обробки з'являється посилання на файли візуалізації результатів обробки (рис. 6, 7).

ВИСНОВКИ

В результаті виконання проекту створено геоінформаційну інфраструктуру, яка забезпечує швидкий та ефективний доступ, обробку та представлення геопросторової інформації з метою моніторингу навколишнього середовища та надзвичайних ситуацій, зокрема повеней. В межах виконання даного науково-технічного проекту розроблено автоматизований цикл обробки радіолокаційних даних супутника ENVISAT/ASAR та впроваджено інформаційні продукти щодо оцінки площ затоплених територій.

Даний проект виконувався у тісній співпраці з міжнародними організаціями UN-SPIDER та NASA, які виступали партнерами проекту, надаючи доступ до супутникових даних при виникненні надзвичайних ситуацій. За результатами виконання проекту в Державному департаменті інтелектуальної власності зареєстровано авторське право на комп'ютерну програму автоматичного картографування повеней на основі радіолокаційних супутникових даних ENVISAT.

ЛІТЕРАТУРА

1. Scheuren J.-M., le Polain de Waroux O., Below R., et al. Annual Disaster Statistical Review – The Number and Trends 2007 // Center for Research of the Epidemiology of Disasters (CRED), Jacoffsset Printers, Melin, Belgium. – 2008.
2. Morse A., Kramber WJ., Allen R.G. Cost Comparison for Monitoring Irrigation Water Use: Landsat Thermal Data Versus Power consumption Data // Proceedings

of the 17th William T. Pecora Memorial Remote Sensing Symposium. – 2008. – Denver, CO, USA.

3. Kussul N., Shelestov A., Skakun S. Grid System for Flood Extent Extraction from Satellite Images // Earth Science Informatics. – 2008. – Vol. 1, No. 3. – P. 105–117.
4. Kussul N., Shelestov A., Skakun S. Intelligent Computations for Flood Monitoring // International Book Series «Advanced Research in Artificial Intelligence» (ed. Markov K., Ivanova K., Mitov I.). – 2008. – 2. – P. 48–54.
5. Куссуль Н.Н., Лулян Е.А., Шелестов А.Ю. и др. Нейросетевой метод мониторинга затопленных территорий с использованием радиолокационных спутниковых данных // Исследование Земли из космоса. – 2008. – 4. – С. 29–35.

Н.Н. Куссуль, С.В. Скакун, А.Ю. Шелестов

ГЕОИНФОРМАЦИОННАЯ ИНФРАСТРУКТУРА МОНИТОРИНГА ОКРУЖАЮЩЕЙ СРЕДЫ И ЧРЕЗВЫЧАЙНЫХ СИТУАЦИЙ

Рассмотрена общая архитектура распределенной системы для обработки геопространственных данных и представления результатов, которая предназначена для мониторинга окружающей среды и чрезвычайных ситуаций. В частности в системе реализована поддержка автоматизированной обработки радиолокационных данных спутника ENVISAT/ASAR, а также информационные технологии оценки площадей затопленных территорий. Результаты получены при выполнении научно-технического проекта НАН Украины при поддержке международной программы UN-SPIDER и космического агентства NASA.

Ключевые слова: информационные технологии, обработка геопространственных данных, распределенная система, мониторинг чрезвычайных ситуаций.

N. Kussul, S. Skakun, A. Shelestov

GEO-INFORMATIONAL INFRASTRUCTURE FOR ENVIRONMENT AND DISASTERS MONITORING

Common architecture of the distributed system for geospatial data processing and visualisation designed for environment and disaster monitoring is considered. In particular, automated processing support of the radar data ENVISAT/ASAR and informational technologies for flooded area estimation have been realised in the infrastructure. The results were obtained within the scientific-technical project of National Academy of Sciences of Ukraine, which was supported by the international program UN-SPIDER, and NASA Space Agency.

Key words: information technology, geospatial data processing, distributed system, disaster management.

Надійшла до редакції 23.03.10