

ГРИД и Българската Виртуална Обсерватория

Георги Петров¹, Момчил Дечев¹, Емануил Атанасов²

В средата на 2009 г. колектив от Института по астрономия на БАН предложи нов научно-изследователски проект - Българска Виртуална обсерватория.

Виртуална обсерватория (ВО), съгласно Кеймбридж Енциклопедия, е колекция от интегрирани астрономически архиви на данни и софтуерни инструменти, която използва компютърни мрежи за създаване на среда, в която могат да се провеждат научни изследвания...

ВО включва различни центрове за данни, всеки със своята уникална колекция от астрономични данни, теоретични модели, софтуер и изчислителни възможности. Астрономическите данни, получени на различни обсерватории, включват наблюдения в различни дължини на вълните – от радио- през оптиката и рентгеновия до гама диапазона. ВО позволява тези разнородни данни да бъдат интегрирани и изследвани като елемент от едно цяло. Няколко страни са започнали национални програми **виртуална обсерватория**, която ще обедини съществуващите бази данни от наземни и орбитални обсерватории.

Основните задачи на Астрономическата виртуална обсерватория са:

- Обезпечаване на възможност за сравнителен анализ на голямо количество разнородни данни от широк спектрален диапазон
- Многократно използване на данните от даден източник
- Обезпечаване на уеднаквен достъп до данните
- Да направи достъпни данните до общности с по-малки възможности
- Да служи и като инструмент за обучение

Предлаганият проект **Българска виртуална обсерватория (БГВО)** си поставя следните задачи:

- България да заеме своето достойно място сред водещите астрономични институции в света.
- Привеждането на всички наши наблюдателни данни към единна система, приета от Европейската и Международната виртуални обсерватории,
- Организиране на добре работещ сървър за данни и използването на унифицирани астрономични програми за анализ на данните

Научноизследователският проект си поставя следните специфични цели:

- Създаване на единен информационен масив с данни за българските астрономически изследвания
- Облекчаване на достъпа до всички възможни източници за студенти, докторанти, преподаватели и научни работници от България и чужбина
- Очертаване на основните направления на развитие, в които работи българската астрономична наука, чрез дигитализиране, електронно публикуване и анализ на ресурсите, които имат доказана научна стойност
- Съхраняване на оригинални ценни източници за развитието на съвременното българско общество чрез публикуването им в електронен вид.



Сред очакваните резултати от проекта са:

- Подобряване на условията за работа с архивни наблюдателни данни чрез дигитализирането им и въвеждането им в интернет
- Усвояване и тестване на методология за дигитализирането на различни по специфика източници с оглед произхода и начина на използването им
- Улесняване на възможностите за проверка и оценка на изследователските постижения
- Създаване на онлайн среда за обмен на ценни данни, теоретични и моделни разработки и др.
- Споделяне на научни теми и изследователските им достижения
- Подобряване на практическите умения на докторантите и разширяване на знанията им чрез работа по дигитализация и електронно публикуване на научни ресурси
- Внедряване и разширяване на електронно обучение в образователния процес за докторанти и студенти
- Повишаване на изследователския капацитет на Института по Астрономия и заинтересованите звена в Софийския университет, Шуменския университет, Пловдивския университет и др.
- Издаване на обобщения отчет на проекта като методическо ръководство за работа с Виртуалната обсерватория, издаване на компакт диск съдържащ дигитализирани ресурси и софтуер, поддържане на уеб-сайта на проекта и след приключване на дейностите по него

- По време на изпълнението на проекта се предвижда специфичната терминология на Виртуалната обсерватория да бъде предложена, изпитана и тиражирана между потенциалните български заинтересовани участници.
- На сървъра на БГВО ще бъдат инсталирани бази данни **POSTGRESQL** и съответните софтуерни инструменти за прехвърляне на наличните данни към стандарта на ВО – разработеният от френските колеги **SAADA tools**.

Крайна цел на проекта е да станат общодостъпни всички наблюдателни данни и документи на български учени, работили в областта на астрономията.

Предимства на Виртуалната обсерватория

- Нова, повече, по-добра, по-бърза и по-лесна наука
- Сравнителен анализ на данните от различни инструменти
- Подобряване на подготовката, обосновката и конструирането на нови наземни и орбитални проекти
- Подобряване на предложенията за нови наблюдения
- Сравняване на реалните данни с такива от симулации и модели, което дава възможност обратно за нов поглед в същността на явленията, нови модели, а и нова физика
- Позволява нов подход на изследователите при работа в различни дължини на вълните и отваря нови възможности за открития, недостъпни до сега

Примерни научни теми, които могат да се решават успешно с развитието на Виртуалната обсерватория:

- Околосвездни дискове: от звездите преди главната последователност до звездите, приютяващи планети
- Облаци с междинни скорости
- Кои звезди ще са следващите свръхнови?
- Начална функция на масите в радиус 1 крс: от планетни до звездни маси
- Начална функция на масите за масивните звезди
- Приносът на звездите с ниска и средна маса към междузвездната среда
- Формиране и еволюция на галактиките при $z = 10$ до $z = 0.1$
- Формиране и еволюция на куповете от галактики
- Купове от купове от галактики
- Възникване на свръхмасивна черна дупка
- Корелация на фоновото космично лъчение и обзорите на галактики в радио/мм диапазони и в оптиката/близка инфрачервена област.

Българската виртуална обсерватория ще даде възможност за решаване на следните **конкретни научни задачи**: Променливост в блясъка и спектъра на галактики с активни ядра, гравитационни лещи, търсене на променливи звезди в звездни агрегати, оценка на параметри на звездни купове, търсене на двойни разсеяни купове, изследване на малки тела от Слънчевата система и др.

Софтуер на Виртуалната обсерватория

- **Откриване на данните**

Aladin: Интерактивен софтуер – небесен атлас, позволяващ на потребителя да визуализира цифровите изображения от произволна част от небето и да налага върху тях данни от различни каталози.

VO Desktop: Настолно приложение за работа с Виртуалната обсерватория. Позволява използването на различни ресурси от данни, запитвания до отдалечени каталози, автоматично провеждане на изследвания.

Datascope: Web услуга, която открива и използва данни във Виртуалната обсерватория от архиви и центрове по целия свят.

- **Спектрален анализ**

VOSpec: Инструмент за спектрален анализ по данни от различни диапазони дължини на вълните, който позволява достъп до услугите Spectral services (SSAP) и Theoretical Spectral services (TSAP).

SPLAT: Инструмент за спектрален анализ.

Specview: 1-D (едномерна) спектрална визуализация и анализ.

Euro3D: Анализирани данни, представени в Euro3D FITS формат.

- **Визуализация и манипулиране на данните**

Topcat: Интерактивен графичен метод за представяне и редактиране на таблични данни. Разпознава някои типични астрономични формати на данни, в т.ч. FITS и VOTable, а много други са в процес на добавяне.

VOPlot: Инструмент за визуализация на астрономични данни.

VisIVO: Софтуер за визуализация и анализ на астрономични данни. VisIVO може да управлява както наблюдателни, така и теоретични данни.

- **Построяване и фитиране на спектралното разпределение на енергията**

VOSED: Инструмент за построяване и фитиране на спектралното разпределение на енергията.

Yafit: Инструмент за фитиране на спектралното разпределение на енергията.

Астрономията е на кръстопът. Значителният напредък в телескопостроенето, приемниците на лъчение и компютърните технологии доведоха до това, че астрономичните обзори “произвеждат” огромно количество изображения, спектри и каталози. Тези мрежи от данни покриват цялото небе в целия диапазон от дължини на вълните – от γ - и X-rays, оптика, инфрачервен диапазон, до радиодиапазона. *Виртуалната обсерватория е международна, инициатива, позволяваща глобален електронен достъп до наличните астрономични данни – както базирани в космични станции, така и на земята.* Тя си поставя за задача да направи възможен координиран анализ на данните, което ще доведе до общоприети стандарти, ширококолентова бърза връзка и ще издигне анализа на данни до степен на изкуство.

Виртуалната обсерватория предлага «стъпка по стъпка» инструкции да решаване на конкретни задачи. Основните от тях са:

Търсене на данни за даден източник. Дава информация какъв тип данни съществуват за даден астрономичен източник и какъв достъп има до тях.

Намиране на изображения за даден източник. Дава информация има ли в даден диапазон

от дължини на вълните изображение за даден астрономичен източник, как да бъде намерено и как да се визуализира.

Построяване на спектралното разпределение на енергията за даден източник. Дава информация съществуват ли спектри в даден спектрален диапазон за източниците в определено поле и как да построим тяхното спектрално разпределение на енергията.

Генериране и анализиране спектралното разпределение на енергията. Това е рецепта как да генерираме и анализираме спектралното разпределение на енергията, комбинирайки спектрите от различни обсерватории в различни диапазони от дължини на вълните.

Инструкция за достъп до данни. Тази “стъпка по стъпка” инструкция дава различни методи за получаване и визуализиране на астрономични изображения, спектри, данни от каталози и др. **За нуждите на радиообщността** е разработен специален инструмент с радиоинтерферометрични данни - „*Merlin Imager*“.

GRID и Виртуалната обсерватория:

Виртуалната обсерватория е силно приложение на **GRID**, включващо обединяването на много и различни източници на астрономични данни. Добра илюстрация на това е Европейската мрежа на слънчевите обсерватории – (European Grid of Solar Observations - EGSO) – начална фаза на обединяване на ресурсите със слънчеви данни, което води до създаването на Слънчева виртуална обсерватория.

От особено значение е крос-корелацията на различни мрежи от данни за намиране на нови явления. Вероятно всеки проект ще има за резултат база данни, направена достъпна чрез мрежа от Web услуги. Търсения, включващи както пространствени, така и времеви ограничения, са типична задача за Виртуалната обсерватория. Ползата от съществуването на такива Виртуални обсерватории става ясна с първите големи открития – напр. откриването на 31 недетектирани до сега свръхмасивни черни дупки в така нареченото GOODS (Great Observatories Origins Deep Survey) поле.

Предвижда се Българската Виртуална Обсерватория да използва и част от ресурсите на ГРИД, предоставяни от Института по паралелна обработка на информацията.

ГРИД е нова изчислителна инфраструктура, която осигурява достъп до изчислителни ресурси и данни, разпределени по цял свят. Тя използва Интернет като среда за пренос и се явява един вид надстройка над нея с по-тясно специализирана цел. Наименованието ГРИД (мрежа – в т.ч. електрическа, решетка) е избрано по аналогия с електрическата мрежа.

За рождена дата на тази система се приема септември 1997 г., когато се провежда семинар на тема "Изграждане на изчислителен ГРИД" в Аргонската Национална Лаборатория, Илинойс, САЩ. През 1998 г. излиза и книга със заглавие "The Grid: Blueprint for a New Computing Infrastructure" с автори Ян Фостър от Аргонската Национална Лаборатория и Карл Кеселман от Университета на Южна Калифорния.

Дефиниция на ГРИД

Първоначално Foster and Kesselman (1999) дават следната дефиниция на ГРИД:

„Компютърната мрежа ГРИД е хардуерна и софтуерна инфраструктура, която осигурява надежден и евтин достъп до висок клас изчислителни мощности“

В една от следващите си статии те прецизират дефиницията по отношение на социалния аспект, като подчертават, че предназначението на ГРИД пресмятането е *"осигуряване на съгласуван достъп до ресурси и решаване на задачи в динамични, мултиинституционални виртуални организации"* (Foster et al., 2001). В крайна сметка, една система, за да бъде определена като ГРИД, трябва да притежава три основни характеристики:

- да координира ресурси, които не са обект на централизиран контрол;
- да използва стандартни, отворени, универсални протоколи и интерфейси;

- да осигури високо качество на услугите.

Предимства на ГРИД

Числените методи са доказали своята ефективност при решаване на широк диапазон от задачи. С помощта на компютри се моделират и симулират сложни и трудоемки научни и инженерни задачи, диагностицират се заболявания, контролира се апаратурата и техническото оборудване, прогнозира се времето, управляват се капитали.

Въпреки че днешният персонален компютър (PC) е по-бърз от суперкомпютъра CRAY отпреди 10 години, той все още не може да отговори адекватно на изискванията за предопределяне на изхода от сложни операции или за избор между много възможности.

Предимствата на пресмятането в ГРИД-среда са:

- Осигуряване на достъп до мащабни изчислителни ресурси, по-бърз достъп до данни, използване на незаети модули.
- Извършване на пресмятания на ново ниво по отношение на изчислителната сложност и размерността.
- Повишено сътрудничество между Виртуалните Организации (Virtual Organization – VO).

Всъщност най-силният довод в полза на използването на ГРИД се дава от самите приложения. Засега повечето от тях произлизат от науката и техниката, но в бъдеще със сигурност ще обхванат всички сфери от живота. На този етап се акцентира върху:

- Физика/Астрономия - съхраняване и обработка на данни, получени чрез различни уреди и апаратура.
- Медицина/Здравеопазване - правене на снимки, поставяне на диагноза, лечение.
- Нанотехнология - проектиране на нови материали, съразмерни с молекули.
- Инженерство - извършване на оптимизация, симулация, анализ на грешката, достъп и контрол на отдалечена апаратура.
- Природни ресурси и околна среда - прогнозиране на времето, наблюдение на Земята, моделиране и предвиждане на сложни системи: наводнения, земетресения и др.

Типове приложения върху ГРИД

Distributed Supercomputing

Приложенията от този тип използват ГРИД, за да присъединяват значителни изчислителни ресурси, с помощта на които да се атакуват задачи, чието решаване е невъзможно за единична система. В зависимост от архитектурата на ГРИД, тези нови ресурси могат да включват както повечето от суперкомпютрите в страната, така и всички работни станции само в една институция.

High-Throughput Computing

За приложенията от този тип ГРИД се използва за планиране на голям брой отделни или групирани по две задачи с цел използване на незаети процесори. В крайна сметка може да се окаже, че както при предишния тип приложения, достъпните ресурси са ангажирани само с една задача, но независимите по своята същност задачи довеждат до много различни типове проблеми и методи за тяхното решаване.

On-Demand Computing

Тези приложения използват възможностите на ГРИД-средата, за да отговорят на краткосрочни изисквания за ресурсите, които не са рентабилни или подходящо разположени. Тези ресурси могат да бъдат софтуер, хранилища за данни, специални сензори и др.

Data-Intensive Computing

При този тип приложения се акцентира върху синтезирането на нова информация от данни, които се съхраняват в различни по географско положение "хранилища", дигитални библиотеки и бази от данни. Този процес на синтез е доста интензивен по отношение на изчислителната сложност и комуникациите.

Collaborative Computing

Тези приложения са насочени основно към подпомагане на взаимодействието между потребителите при използването на изчислителни ресурси като бази от данни и симулации. Приложенията от този тип често са структурирани в термините на виртуалното пространство.

За използване на възможностите на ГРИД е създадена **Национална инициатива**, която включва:

- Институт за паралелна обработка на информацията;
- Институт за ядрени изследвания и ядрена енергетика;
- Институт по механика;
- Институт по електрохимия и енергийни системи;
- Институт по астрономия;
- Институт по математика и информатика;
- Геофизичен институт;
- Факултет по математика и информатика - Пловдивски университет;
- Факултет по математика и информатика - Софийски университет;
- Физически факултет - Софийски университет;
- Минно-геоложки университет "Св. Иван Рилски";
- Институт по органична химия с Център по фитохимия;
- Университет за национално и световно стопанство.

В рамките на тази инициатива се предвижда решаването на различни по сложност и значимост задачи. *Астрономическата колегия ще започне работата с ГРИД по следните задачи:*

Задача 1: N-body симулация на разсеян звезден куп и система от два купа.

Задача 2: Числено моделиране на нехомогенни бариогенезисни модели.

Задача 3: Монте Карло моделиране на кометни прахови структури.

Задача 4: PIXON метод за изявяване на слаби структури в галактиките.

Задача 5: Моделиране на активни явления на Слънцето.

Цялата необходима информация за ГРИД и БГВО може да бъде намерена на следните сайтове:

<http://www.grid.bas.bg> и <http://www.astro.bas.bg/bgvo/>

1) Институт по Астрономия, БАН

2) Институт за паралелна обработка на информацията, БАН