



Рабочие процессы системы HiTViSC (High-Throughput Virtual Screening as a Service)

▼ Наталия Никитина, Евгений Ивашко

Институт прикладных математических исследований Карельского научного центра РАН, г. Петрозаводск

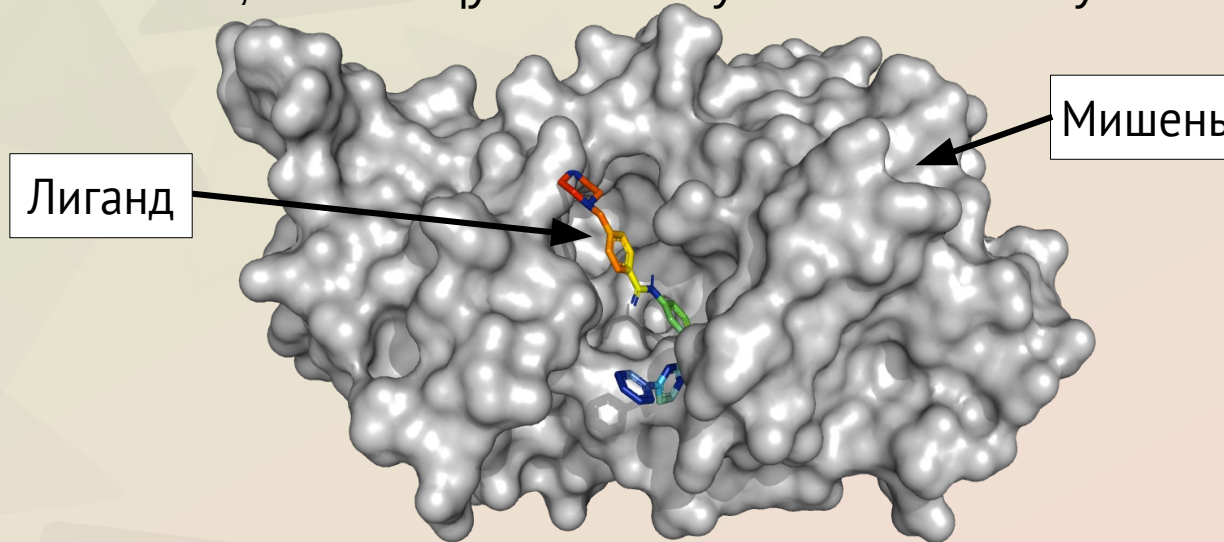
1. Разработка лекарства



- Разработка лекарства – это времяземкий и ресурсоемкий процесс.
- Для разработки лекарства требуется 10-15 лет и ~1 млрд \$.
- Методы *in silico* позволяют снизить время и стоимость начальных этапов разработки.

1. Разработка лекарства

Цель: найти малую молекулу, способную связываться с мишенью, ассоциированной с развитием заболевания, и имеющую желаемую биохимическую активность.



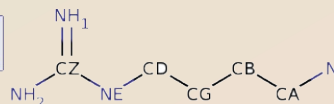
- Молекулярный докинг** - это компьютерное моделирование “стыковки” лиганда с мишенью.
- предсказание ориентации молекулы лиганда относительно мишени;
 - оценка силы связывания при помощи оценочной функции.
 - Виртуальный скрининг (идентификация и оптимизация соединений, регулирующих активность мишени, ассоциированной с развитием заболевания);
 - Репозиционирование лекарств (молекулярный докинг существующих лекарств с мишенью);
 - Поиск новых мишеней, предсказание побочных эффектов лекарств (молекулярный докинг лиганда с известными мишенями);
 - ...и др. приложения.

1. Разработка лекарства

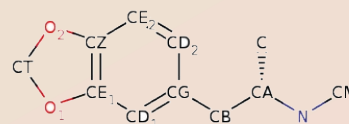
Виртуальный скрининг – это компьютеризованная процедура поиска **хитов** в библиотеках малых молекул. **Хиты** – это химические соединения, для которых предсказано наличие желаемой биохимической активности.



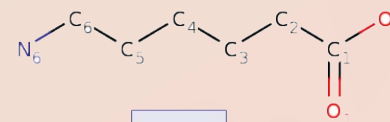
1.



2.



3.



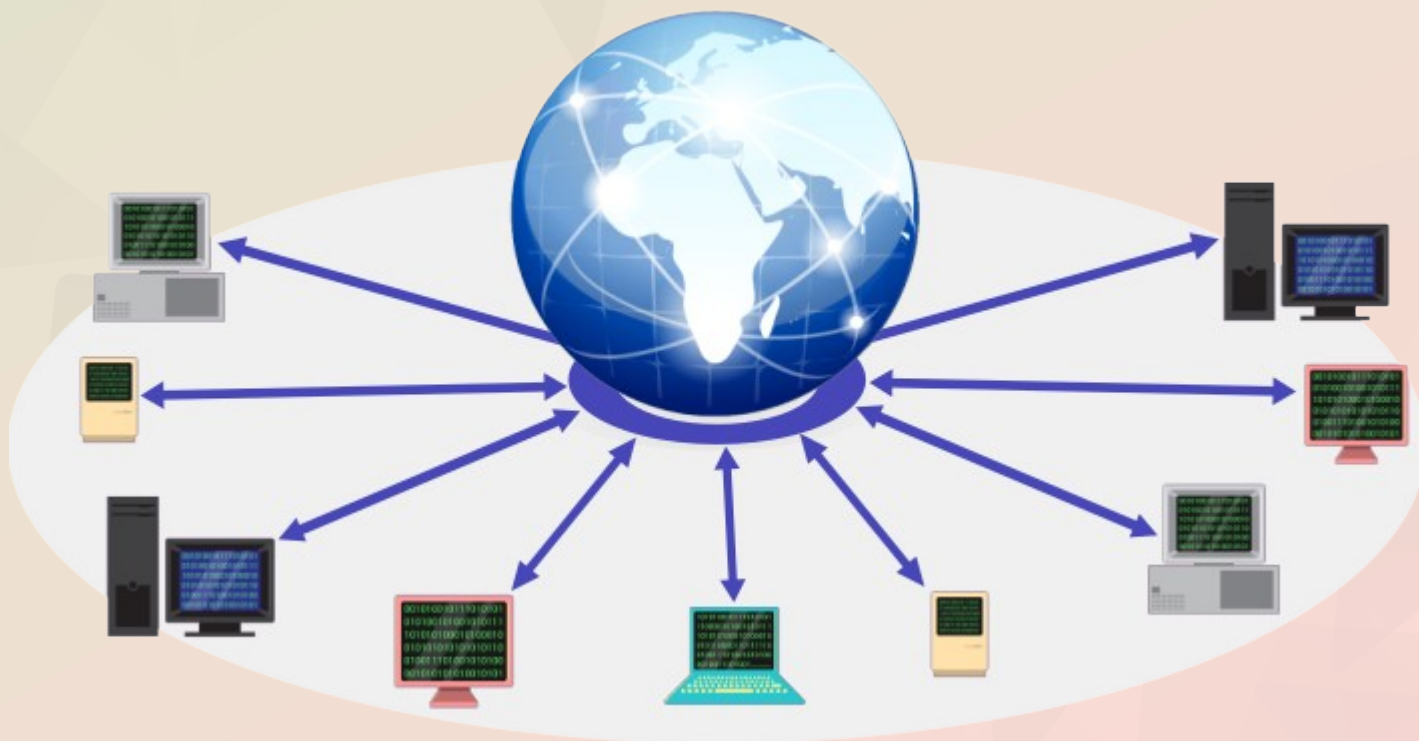
4. ...

Открытые библиотеки лигандов: ZINC (1 млрд), PubChem (116 млн), GDB-17 (166 млрд)...

=> Виртуальный скрининг требует высокопроизводительных вычислений (HPC/HTC).

2. Desktop Grid

Desktop Grid – это вычислительная система, объединяющая территориально распределенные неспециализированные вычислительные узлы (ПК, смартфоны, игровые консоли и др.), связанные с управляющим узлом локальной сетью передачи данных или сетью Интернет.



2. Desktop Grid

Преимущества:

- Высокая потенциальная пиковая производительность (2 млрд ПК, 4 млрд смартфонов);
- Высокая масштабируемость;
- Простая установка и настройка;
- Низкая стоимость;
- Простая адаптация существующего программного обеспечения.

Ограничения:

- Подходят для приложений типа «Bag-of-Tasks»;
- Относительно низкая производительность отдельных узлов;
- Неоднородность аппаратного и программного обеспечения узлов;
- Неизвестное время доступности узлов;
- Низкая надежность узлов.

2. Desktop Grid

Публичные проекты виртуального скрининга на основе Desktop Grid:

- **World Community Grid** (47000 активных компьютеров, 100 TeraFLOPS). Виртуальный скрининг лекарств от малярии, туберкулеза, рака, СПИДа, вируса Зика, вируса Эбола, шистосомоза, лейшманиоза.
- **IberCIVIS: COVID-Phym** (7000 активных компьютеров, 14 TeraFLOPS). Виртуальный скрининг существующих лекарств с целью блокировки механизма репликации коронавируса; исследование *in silico* причин того, что клинические исследования показали лишь ограниченную эффективность ремдесивира и тенофовира; разработка лекарств на основе тенофовира.
- **BOINC@TACC**. Виртуальный скрининг с целью разработки молекулярной структуры ингибитора протеазы против коронавируса.
- **DreamLab: Corona-AI** (140000 мобильных устройств). Виртуальный скрининг для репозиционирования лекарств и разработки диеты при лечении коронавируса.
- **SiDock@home** (4800 активных компьютеров, 40 TeraFLOPS). Виртуальный скрининг лекарств против коронавируса.

3. High-Throughput Virtual Screening as a Service

Требования при реализации проекта виртуального скрининга:

- экспертные знания для полного цикла виртуального скрининга;
- доступ к вычислительной платформе HPC/HTC;
- технические знания и навыки для выбора, установки и использования вычислительной платформы;
- управление множеством специализированных программ и алгоритмов обработки данных.

=> для упрощения процесса требуются готовые решения и сервисы.

Существующие решения для виртуального скрининга:

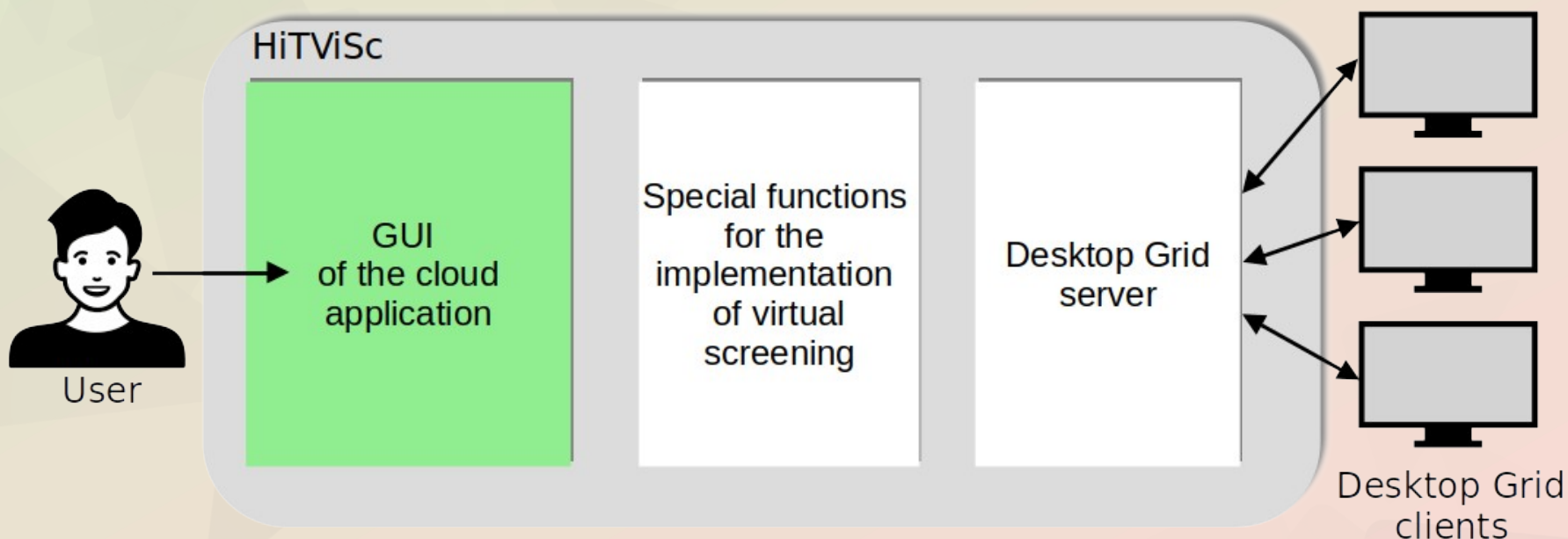
- Локальные системы (MGLTools, VirtualFlow, GriDock)
- Сервисы на основе суперкомпьютеров (EXSCALATE, GigaDocking)
- Сервисы на основе облачных вычислений (KNIME, HADDOCK, SwissDock)
- Зонтичные проекты Desktop Grid (BOINC Central)

3. High-Throughput Virtual Screening as a Service

Система **High-Throughput Virtual Screening as a Service (HiTViSc)** предоставляет пользователю хорошо масштабируемые вычислительные ресурсы Desktop Grid в виде специализированного облачного сервиса для виртуального скрининга.

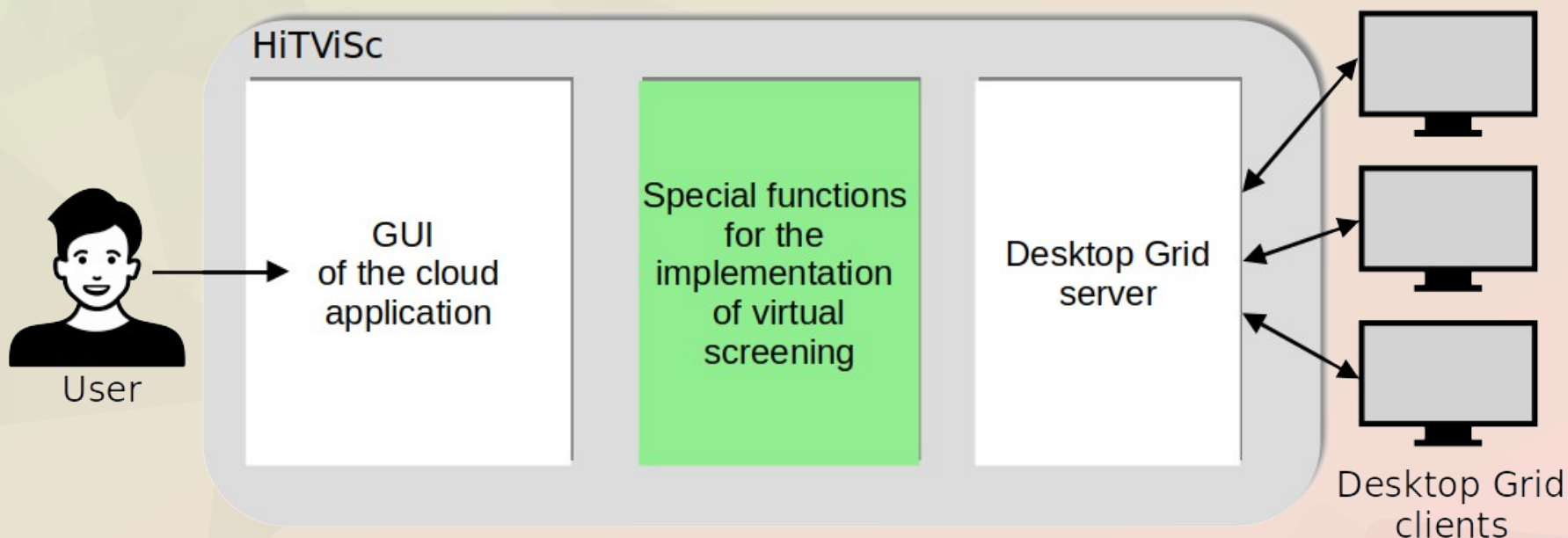
- Сервис для виртуального скрининга с удобным веб-интерфейсом;
- Система с открытым исходным кодом позволяет использовать существующее специализированное ПО и реализовывать собственные алгоритмы (**подготовка библиотеки, визуализация, анализ результатов**);
- Автоматизированная установка и использование ресурсов Desktop Grid.

3. High-Throughput Virtual Screening as a Service



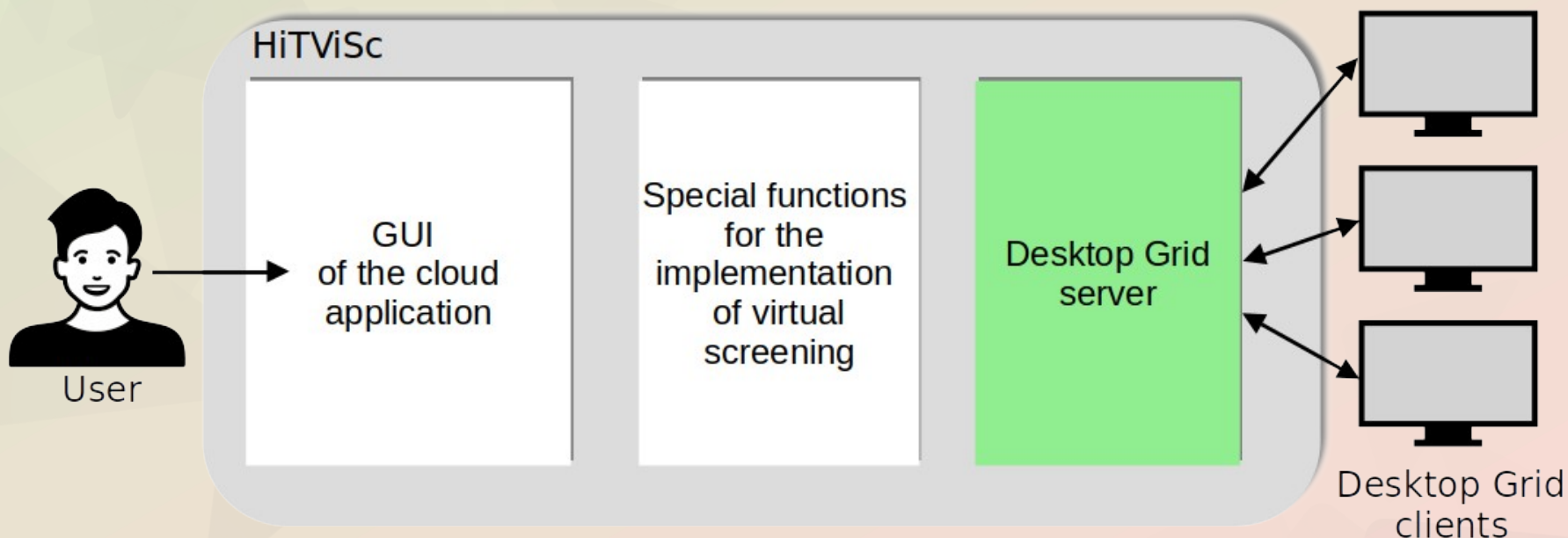
- Интерфейсы и настройки для управления виртуальным скринингом;
- визуализация прогресса;
- визуализация и анализ результатов;
- управление вычислительными ресурсами.

3. High-Throughput Virtual Screening as a Service



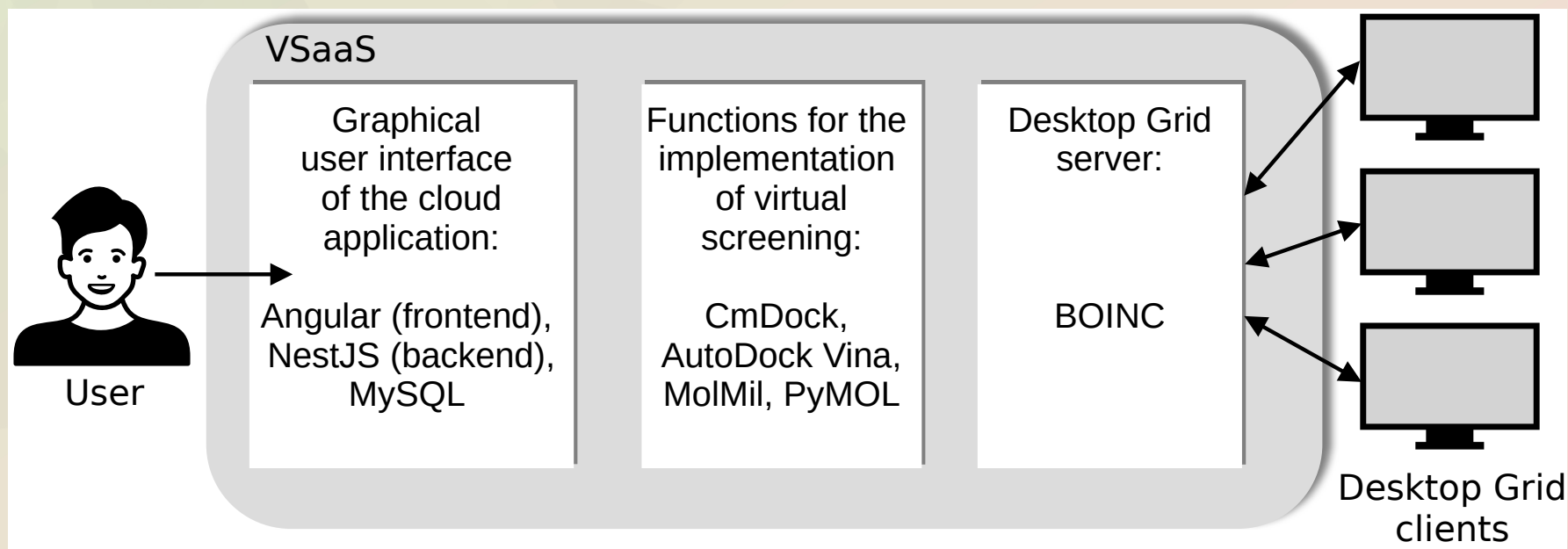
- Загрузка файла мишени и базы лигандов, выбор приложения для докинга;
- настройка параметров;
- протокол молекулярного докинга и обработки результатов;
- выбор внешних приложений для анализа и визуализации результатов.

3. High-Throughput Virtual Screening as a Service



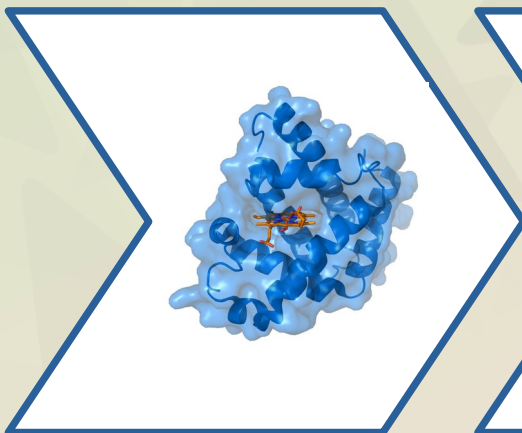
- Генерация заданий;
- коммуникация с вычислительными узлами: назначение заданий, передача входных данных и вычислительных приложений;
- сбор и обработка результатов.

3. High-Throughput Virtual Screening as a Service

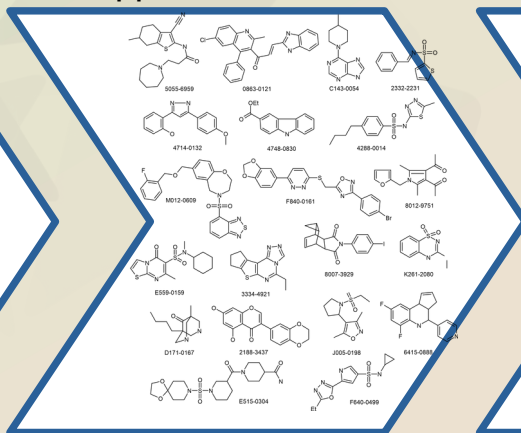


4. Рабочий процесс: вычислительный эксперимент

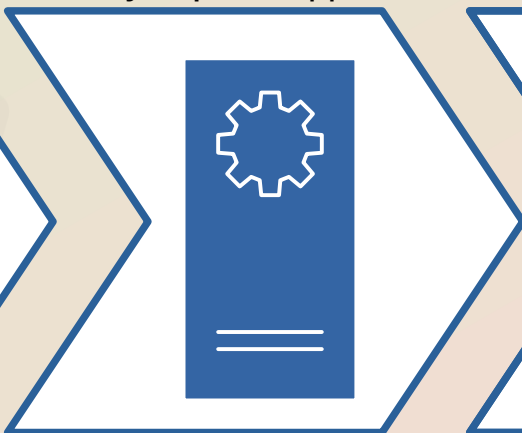
1. Выбор мишени



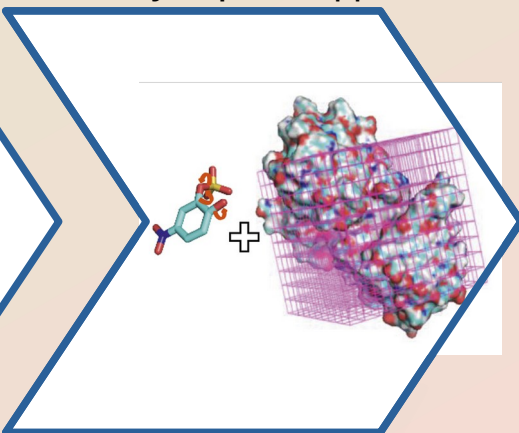
2. Выбор библиотеки лигандов



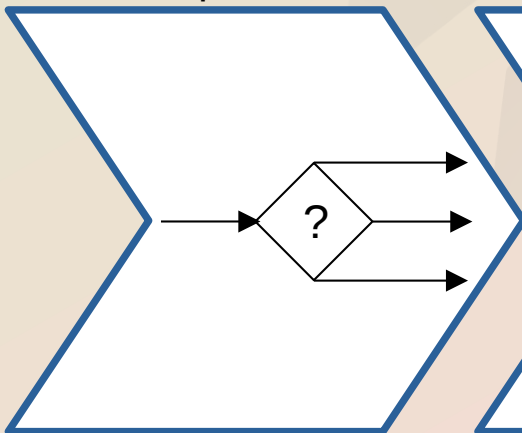
3. Выбор приложения молекулярного докинга



4. Настройка молекулярного докинга



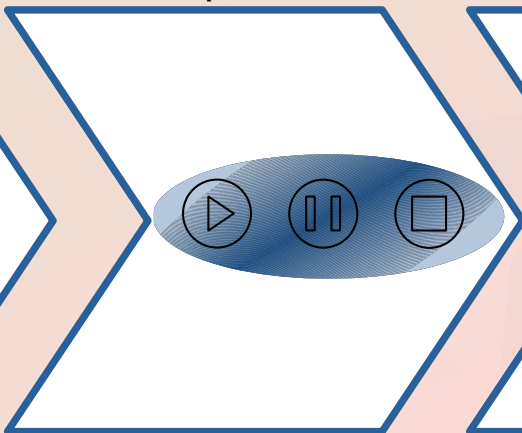
5. Настройка протокола вычислительного эксперимента



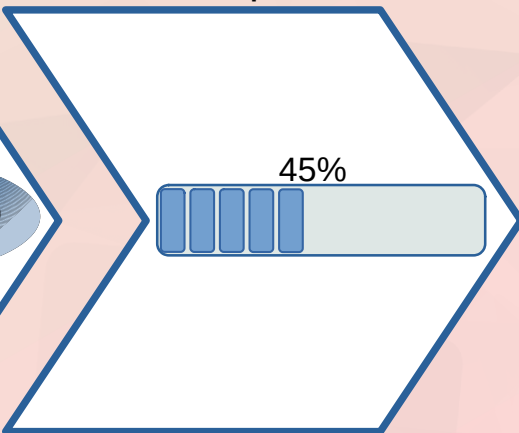
6. Выбор вычислительных ресурсов



7. Запуск/остановка вычислительного эксперимента



8. Отображение прогресса вычислительного эксперимента



Screenshot from: Jianhua Li et al. "Molecular Docking for Ligand-Receptor Binding Process Based on Heterogeneous Computing", Scientific Programming, vol. 2022, Article ID 9197606, 2022. <https://doi.org/10.1155/2022/9197606>

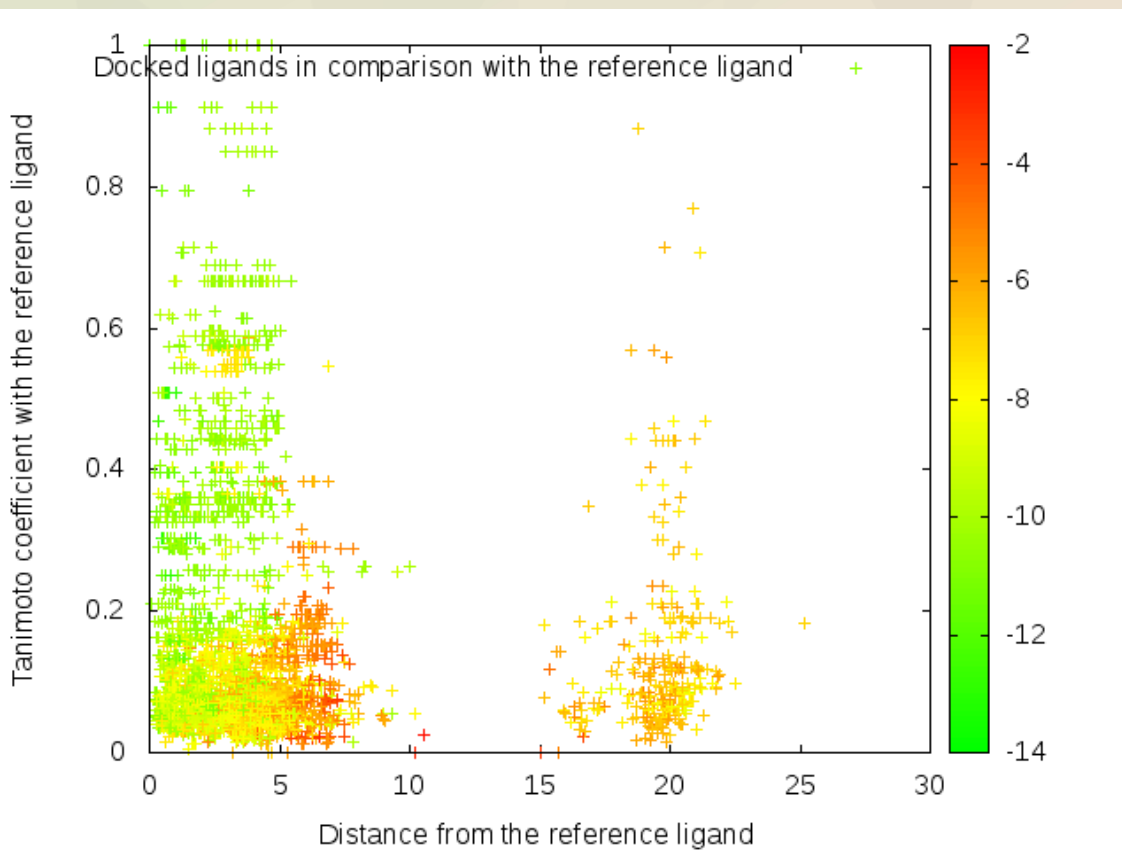
4. Рабочий процесс: управление вычислительными ресурсами

Три типа вычислительных ресурсов, которые можно использовать при виртуальном скрининге в качестве вычислительных узлов Desktop Grid:

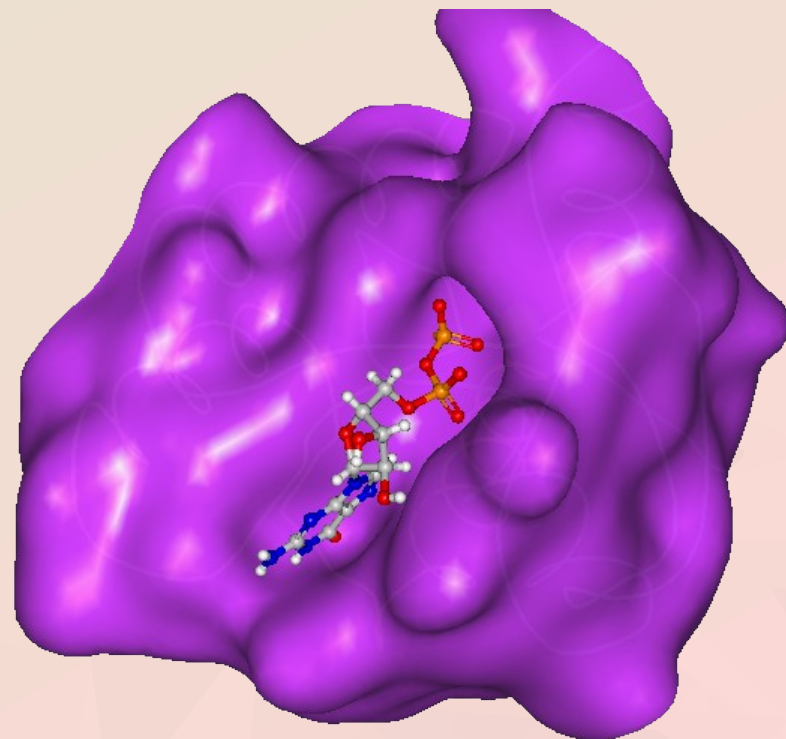
- ▼ Тестовые узлы. Ограниченное количество высокодоступных и надежных вычислительных ресурсов, используемых для проверки параметров при небольшом объеме вычислений.
- ▼ Вычислительные ресурсы, предоставленные в общее пользование добровольцами и другими пользователями HiTViSc и используемые для проведения экспериментов.
- ▼ Собственные и партнерские вычислительные ресурсы, подключенные к системе, которые могут быть использованы для собственных экспериментов, партнерских проектов или предоставлены в общее пользование.

4. Рабочий процесс: экспертный анализ результатов

- ▼ Снижение размерности и кластеризация;
- ▼ визуализация;
- ▼ загрузка результатов для обработки сторонним ПО



(a) 2D-представление множества хитов



(b) 3D-представление лиганда

5. Многопользовательский доступ

Возможности, доступные пользователю:

- ▼ Доступ к аккаунту для проведения вычислительных экспериментов, хранения и анализа результатов;
- ▼ использование вычислительных ресурсов;
- ▼ обмен с другими пользователями результатами вычислительных экспериментов по одному из трех вариантов доступа: результаты доступны всем пользователям системы, определенной группе пользователей или только автору.

Все сгенерированные данные (подготовленные мишени и библиотеки лигандов, найденные хиты, оценки энергии связывания и т.д.) могут быть использованы широким кругом пользователей для дальнейшего экспертного анализа с сохранением авторства.

Заключение

High-Throughput Virtual Screening as a Service (HiTViSc)

— это система с открытым исходным кодом, предоставляющая ресурсы Desktop Grid в виде специализированного облачного сервиса для виртуального скрининга.

Планы на будущее:

- Проектирование и реализация системы;
- разработка и внедрение новых алгоритмов планирования заданий, специализированных для виртуального скрининга;
- апробация в проекте виртуального скрининга SiDock@home.

Спасибо за внимание!

E-mail: nikitina@krc.karelia.ru