

Десятая независимая научно-практическая
конференция «Разработка ПО 2014»

23 - 25 октября, Москва



Улучшение производительности и повышение отказоустойчивости Data Grid системы iRODS

Андрей Неволин

EMC²

EMC Skolkovo



Андрей Пахомов

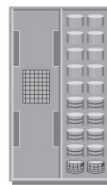
Андрей Запарий

Андрей Абрамов

Андрей Неволин

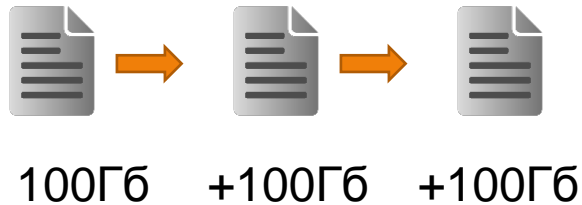
План

- Big Data в геномике
- iRODS Data Grid system
- Интеграция СХД в iRODS
- Результаты



Петабайты в геномике

- Полное секвенирование генома – \$1000
(против \$1 000 000 в 2008)
- 100Гб – 250Гб на геном
- 100,000 Genomes Project



The logo for iRODS is centered at the top of the slide. It features the word "iRODS" in a bold, sans-serif font. The letter "i" is teal, while "RODS" is dark grey. The background of the top half of the slide is a close-up, grayscale image of a dog's face, showing its eyes and fur.

iRODS

TAKE CONTROL OF YOUR DATA

Open source Data Grid middleware

- Безопасный совместный доступ к данным
- Автоматизация операций над данными
- Поиск по метаданным
- Гибкая интеграция с файловыми системами

- Совместный доступ к данным
- Поиск по метаданным (?)



||



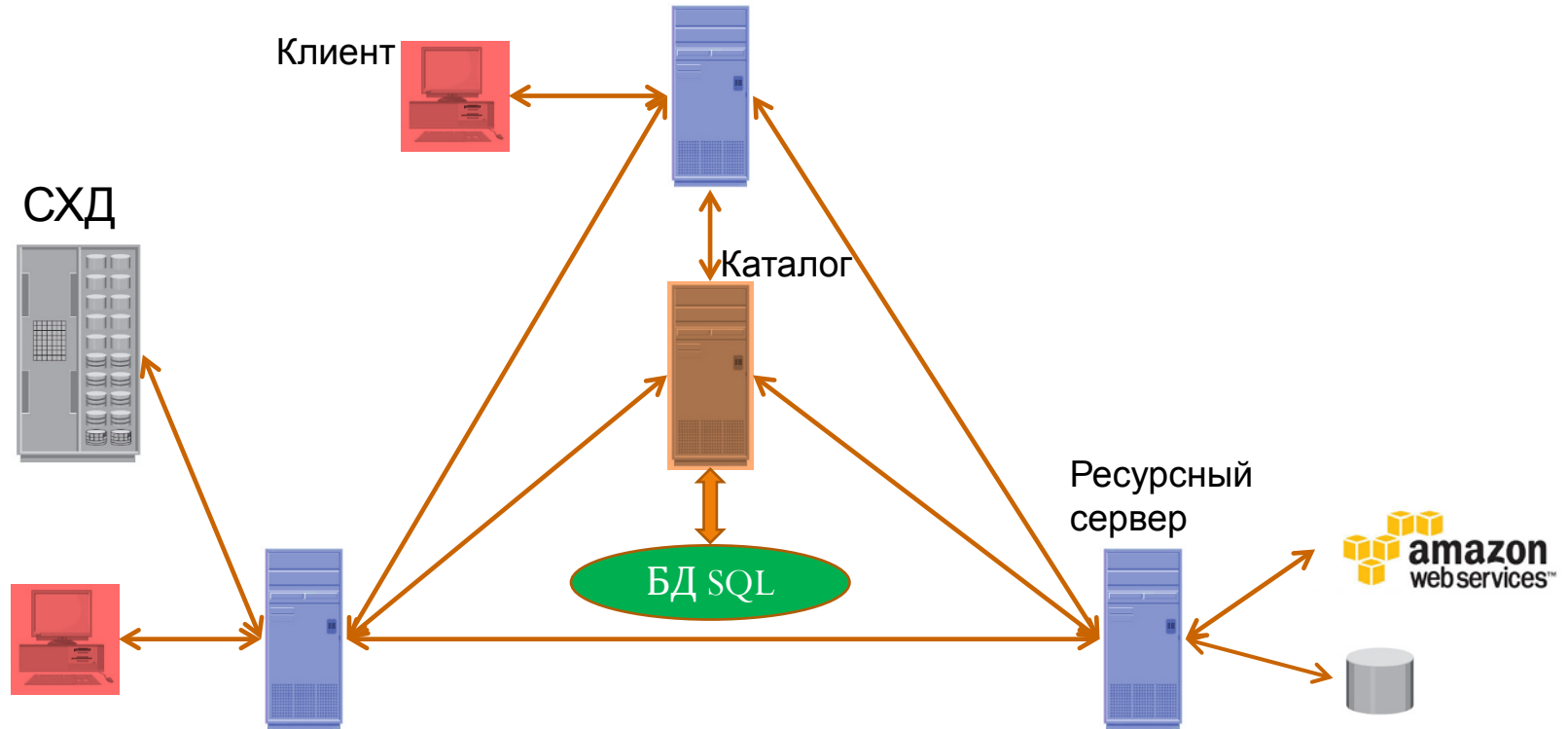
- Платный
- Закрытый код
- Нет автоматических политик управления данными
- Ориентирован на



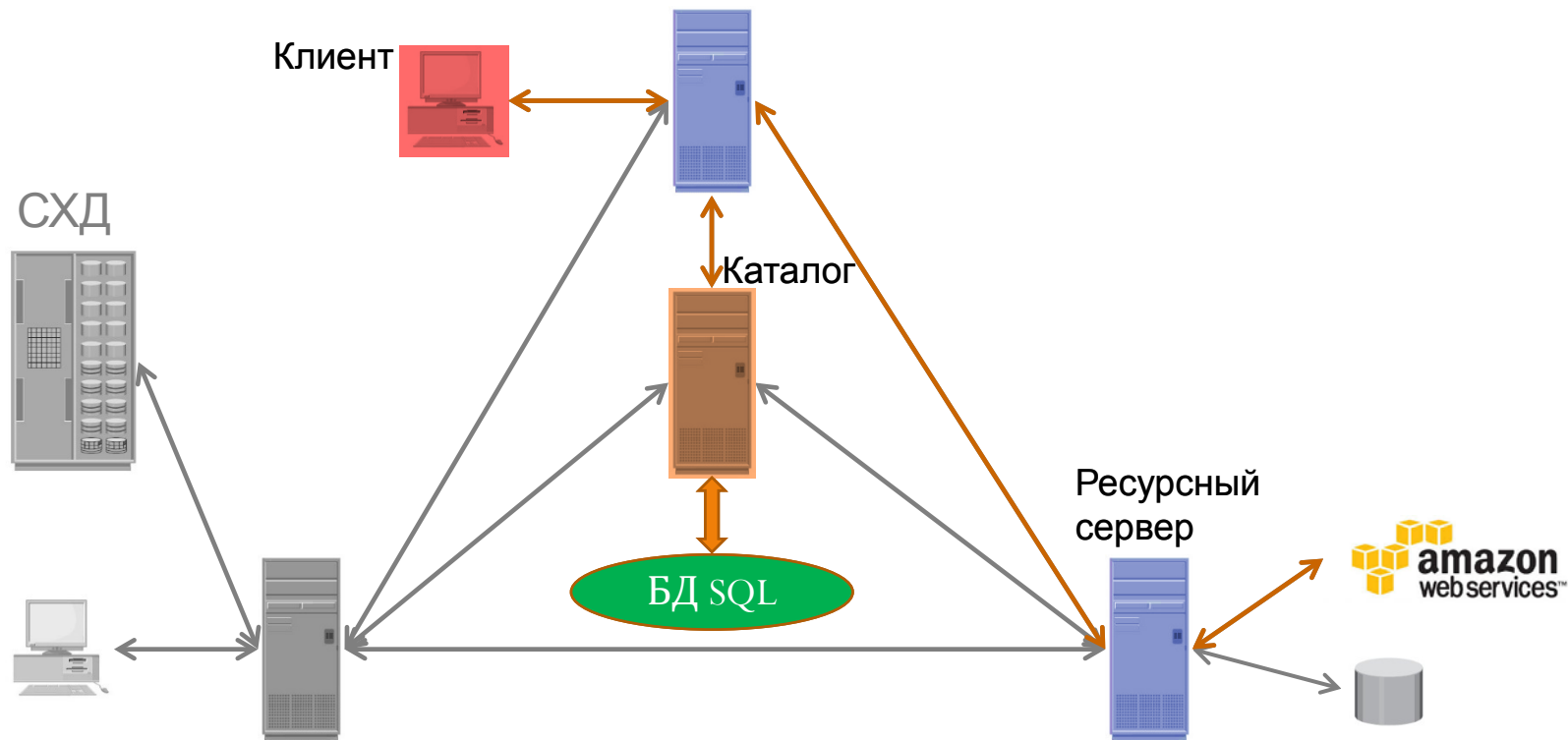
Пользователи iRODS

- Beijing Genomics Institute (BGI)
- The Wellcome Trust Sanger Institute (WTSI) ← 2Пб
- The Broad Institute
- The International Neuroinformatics Coordinating Facility (INCF)
- “Life Science Industrial Users”

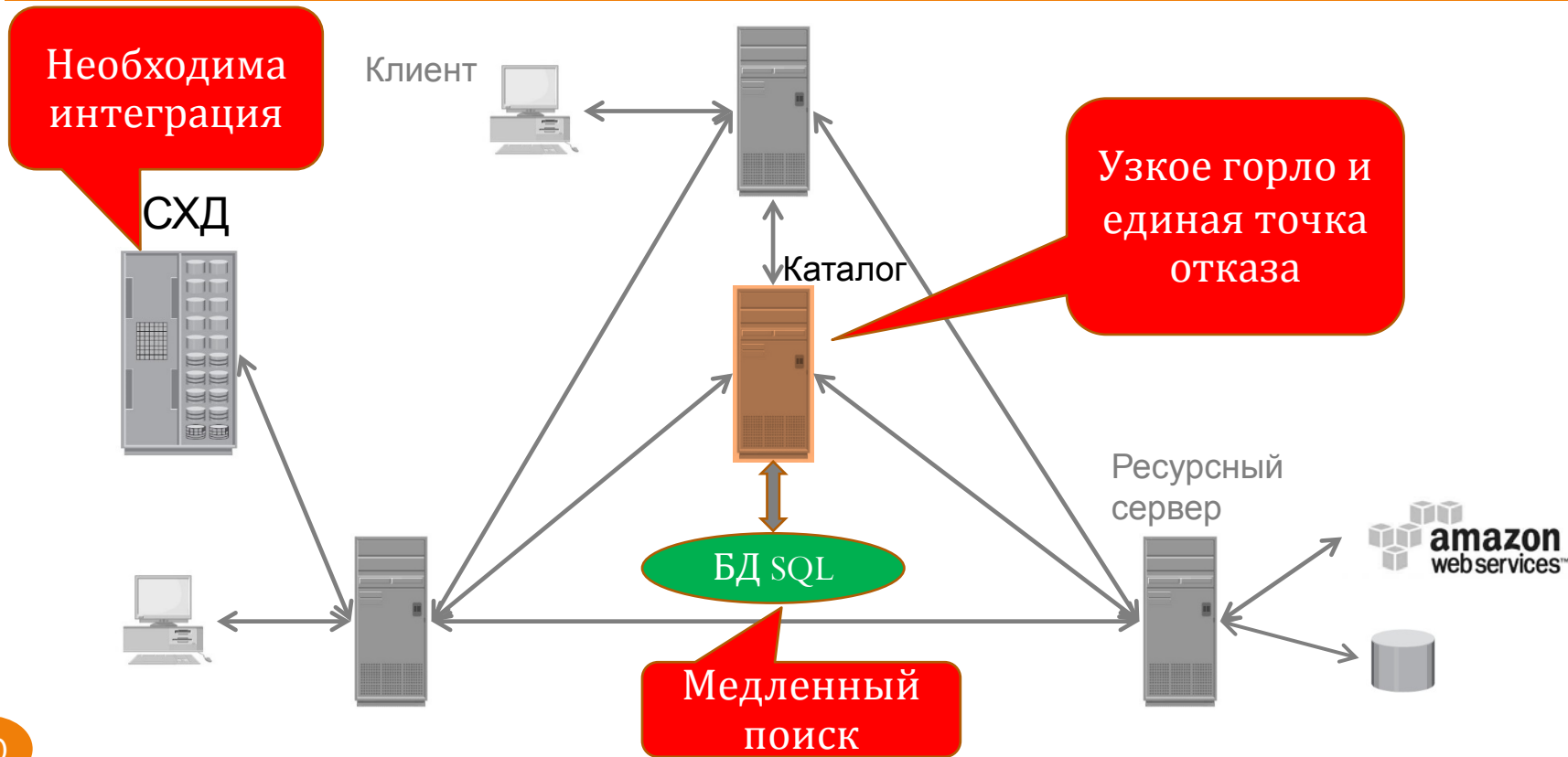
iRODS Data Grid



Передача данных в iRODS



Проблемы iRODS



Почему СХД?



- не для Big Data

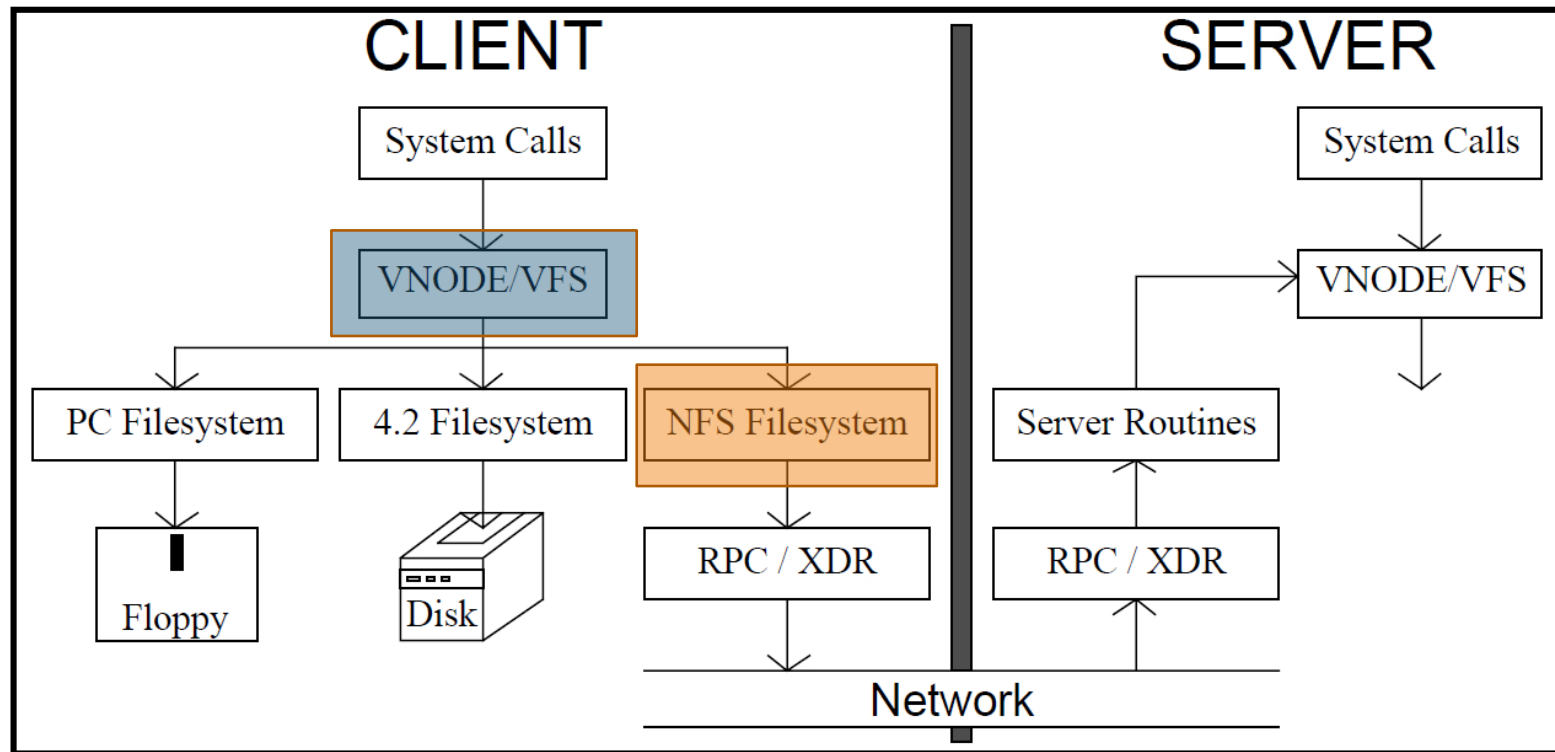


- требует дорогой полосы пропускания
- недоверие к безопасности

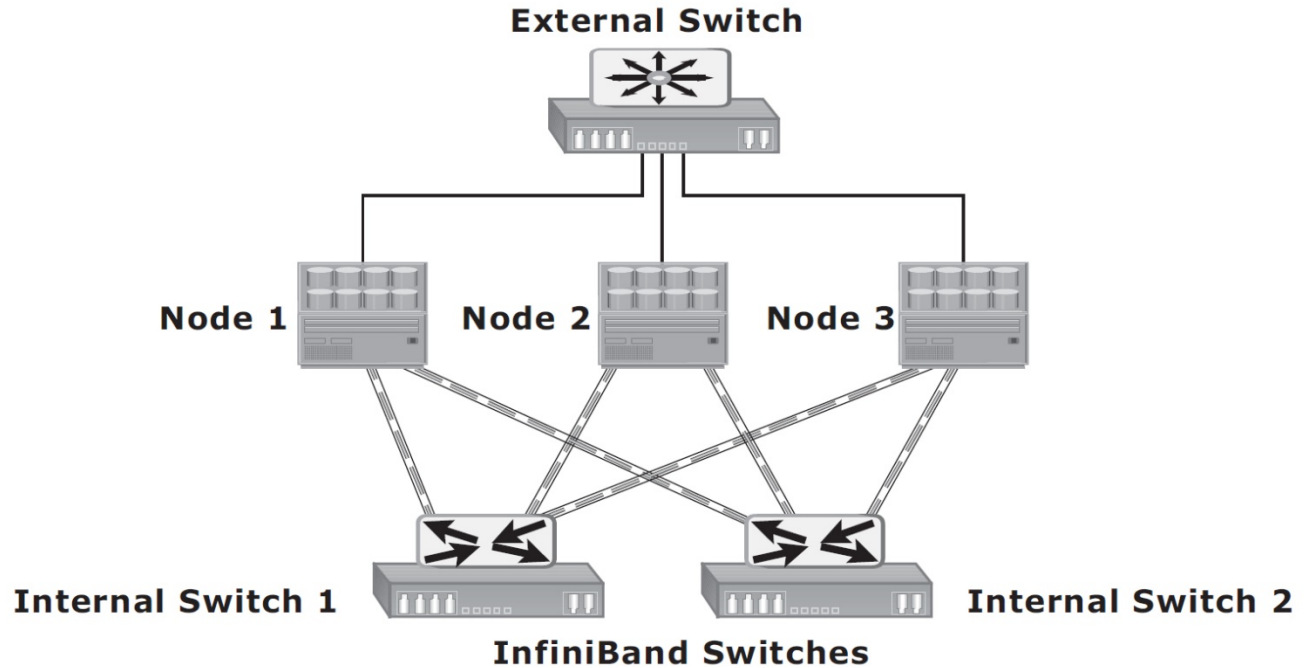


- весь контроль над данными – на стороне клиента
- интеллектуальные функции управления данными
- быстрый доступ
- экономическая выгода при больших объемах данных

EMC + iRODS сегодня



СХД горизонтального масштабирования



Задачи интеграции

- Улучшенная производительность
- Параллельный доступ к узлам СХД
- Работа в объектном стиле
- Символьные ID пользователей
- «Интеллектуальные» функции СХД
- Балансирование нагрузки на СХД

Задачи интеграции

- Улучшенная производительность
- Параллельный доступ к узлам СХД
- Работа в объектном стиле
- Символьные ID пользователей

HDFS

- «Интеллектуальные» функции СХД

REST API

- Балансирование нагрузки на СХД

REST+iRODS+plugin

HDFS



- Распределенная файловая система в составе Hadoop
- Простая
- Решает большинство проблем, для которых не подходит NFS

HDFS



- `wouter@foo:~$ time for i in {1..10}; do hadoop fs -ls / > /dev/null; done`
- `real 0m14.464s`
- `user 0m21.761s`
- `sys 0m1.148s`
- `wouter@foo:~$ time for i in {1..10}; do snakebite ls / > /dev/null; done`
- `real 0m1.639s`
- `user 0m1.072s`
- `sys 0m0.160s`



Java

Взято с <https://labs.spotify.com/2013/05/07/snakebite/>

Результаты

- Базовая интеграция завершена
- Ведется работа над расширенной интеграцией
- Готов прототип для полнотекстового поиска
- Предложения по устранению единой точки отказа и распределению нагрузки обсуждаются с консорциумом

Уроки

- Решения для BigData без доступа к реальным BigData 
нужно просить, просить и просить
- Софт для себя, который стал софтом для всех 
договариваться и договариваться об изменениях
- Open Source vs. Open Source
у каждого свой кусок vs. «у всего один папа»
(gcc, LLVM vs. iRODS)

Ссылки

Использованы материалы из:

- “[Managing Next Generation Sequencing Data with iRODS](#)”, *9th International Conference on Genomics* (слайды 4 и 6)
- “Information storage and management v2”, *EMC Education Services*, (картинки со слайдов 7-10, 12)
- “The Sun Network Filesystem: Design, Implementation and Experience”, *Sun Microsystems* (слайд 11)
- www.irods.org

Спасибо!



<https://www.linkedin.com/profile/view?id=97693563>



<https://www.linkedin.com/profile/view?id=146690734>

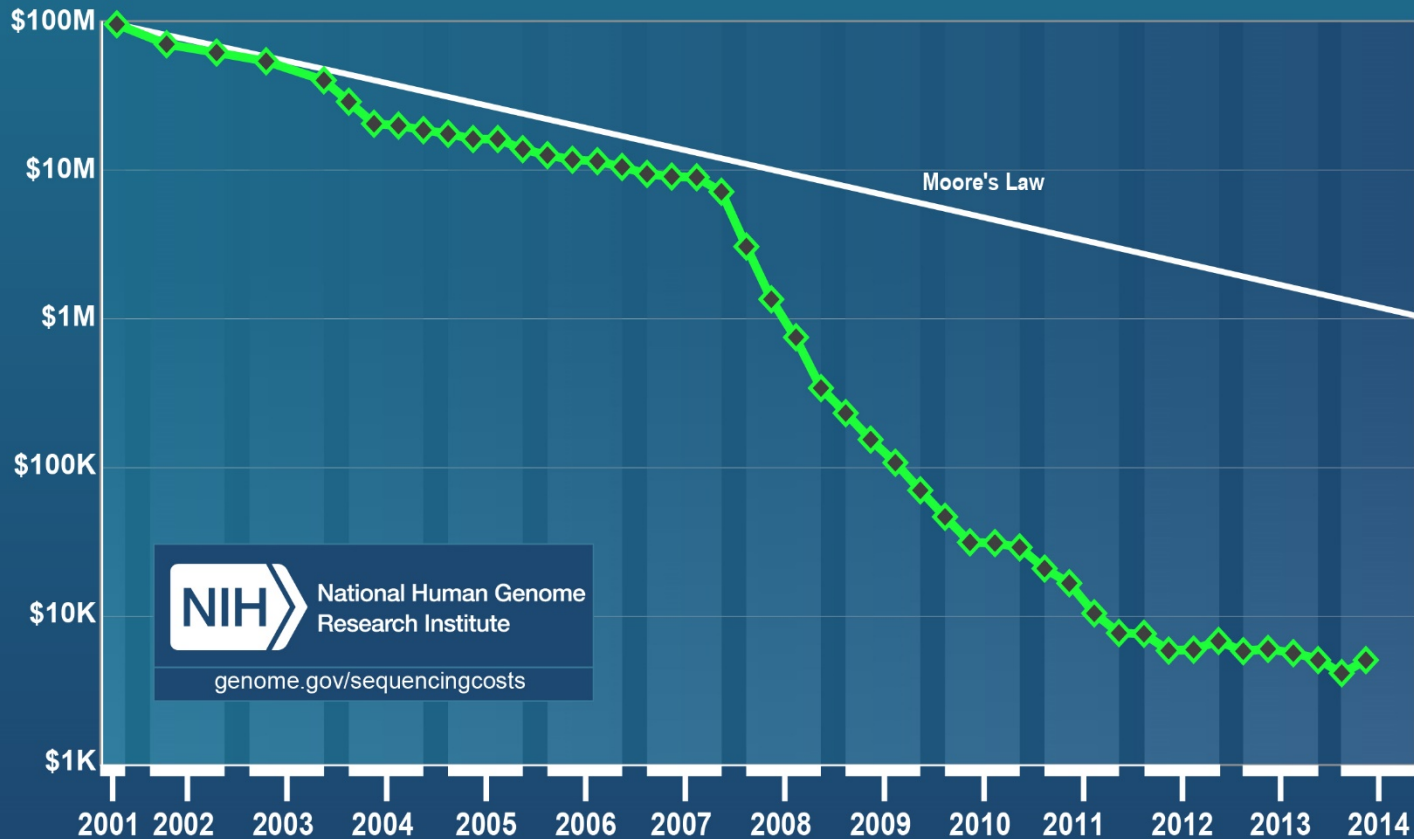


<https://www.linkedin.com/profile/view?id=368926>



Backup

Cost per Genome



Почему в геномике выбирают iRODS

- Метаданные
- Нейтральность по отношению к СХД
- Open Source
- Консорциум
- Гибкость (права доступа, cloud + local, репликация, пр.)

Баги в iRODS

Issues reported to iRODS community

Created and last modified by Andrey Nevolin on Sep 19, 2014

Dirty replicas - clarification needed #2344

Minor issue with "fileChecksum" - incomplete stack print in case of fail #2340

Memory leak in "getFile" (really critical) #2338

Resource context verification at resource creation #2336

Single-thread "get" operation doesn't check file size properly #2335

Include "runtime" package into "icat" and "resource server" packages? #2333

Registration of libraries from "runtime" package #2332

A length of resource "context" parameter is limited to 64 #2311

"oprType" should be defined in "acSetNumThreads" rule #2309

Access to 'cache'-type resources should be virtualized #2249

"irepl" cannot be executed with the specified number of threads #2241

"ibun", "iphybun" not works with non unix file system resources #2232

"ibun" documentation issue #2231

Interface between iRODS and resource plugins should be defined #2230

"irods_file_open" opens file for READ/WRITE instead of just READ #2227

Unclear logic of "unix_file_readdir_plugin" #2226

Deallocation of 'rodsDirent' objects #2218

Minor issues with "unix_file_opendir_plugin", "unix_file_closedir_plugin", "unix_file_readdir_plugin" #2214

Minor bug in "unix_file_mkdir_plugin" #2206

Error statuses cannot be used to report plugin-specific issues #2204

Access control of directories created by "..._rename_plugin" should be parameterized #2201

Path to "trash" should be parameterized #2200

Resource servers should be possibly ignorant of non-owned resources #2196

Referencing resource host through IP doesn't work #2195

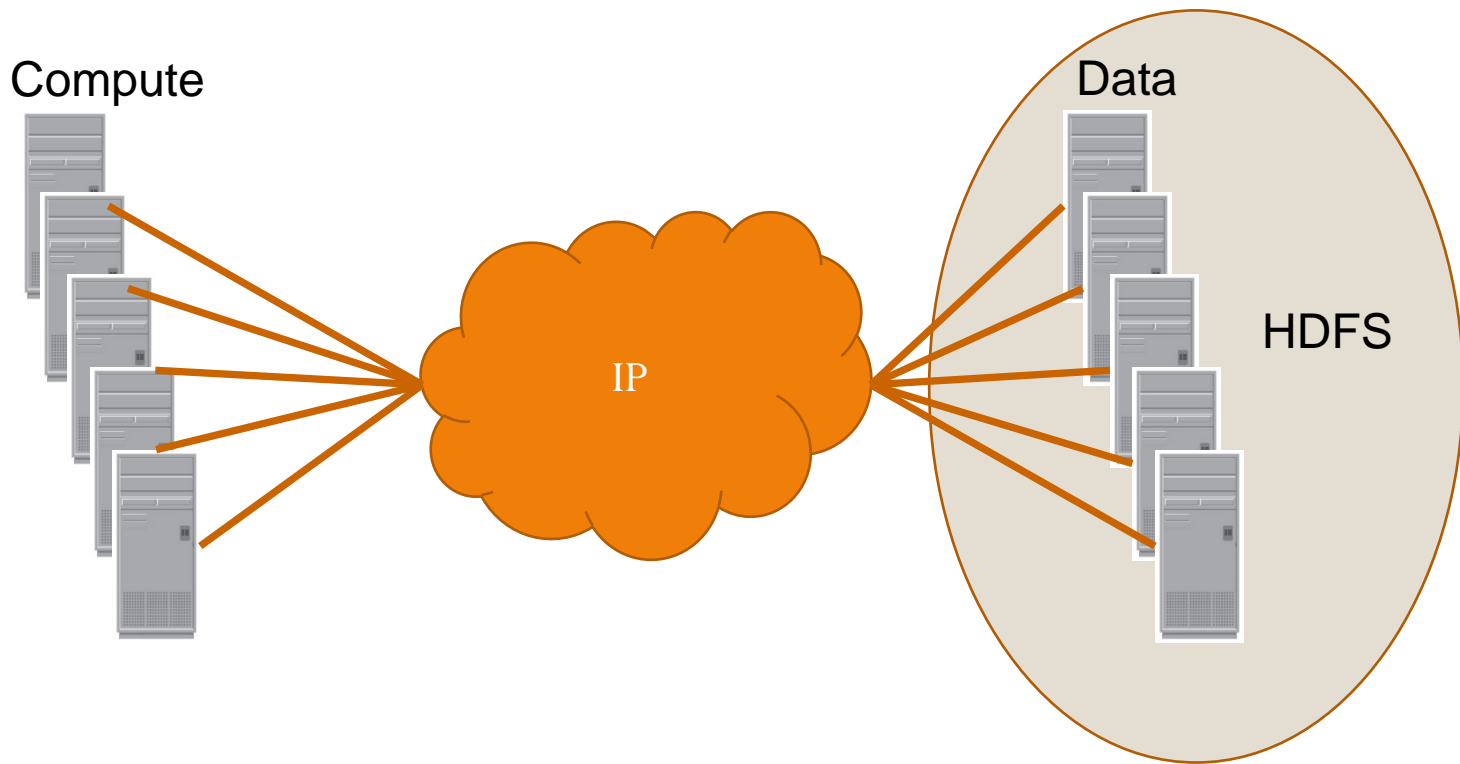
Adding file to a non-existing collection - user error is processed heavily #2194

--disable-parallel mode (covered by PARA_OPT define) is broken #2130

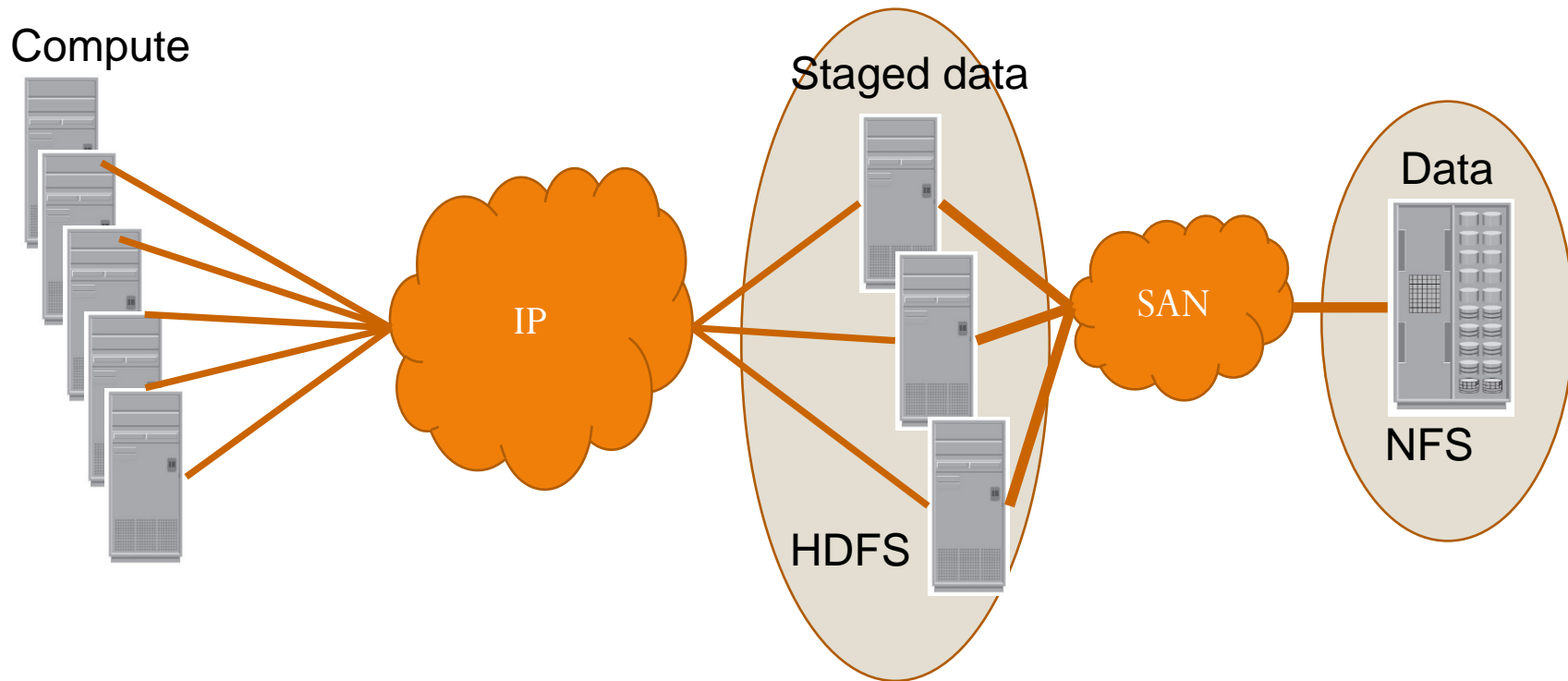
Server availability is not properly tested in Perl scripts #2108

27 багов описано
на текущий
момент

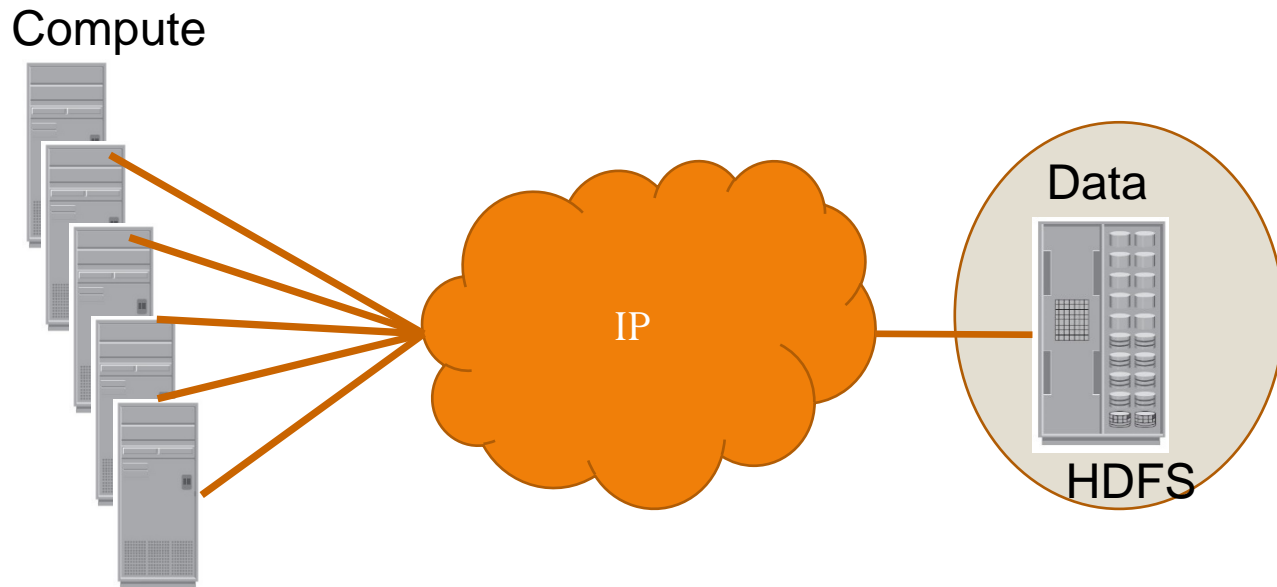
Нadoop позавчера



Нadoop вчера




Нadoop сегодня



Поиск по метаданным (проблемы)

attr_id	attr_name	attr_value
1	Autor	John
5	Grant number	123579

object_id	...
15	
21	



attr_id	object_id
1	21
5	15

Инвертированный индекс

be	1	
important	0	
is	0	
not	0	1
or	1	
questioning	0	
stop	0	
to	0	1
the	0	
thing	0	

c:\docs\einstein.txt:

1

The important thing is not to
stop questioning.

c:\docs\shakespeare.txt:

2

To be or not to be.