

## ■ Тенденции развития высокопроизводительных вычислений в нефтегазовой отрасли: технологические и экономические

Е.А. Курин, Е.Л. Музыченко, А.П. Бурцев  
ООО "ГЕОЛАБ"

*ekurin@geo-lab.ru*



## ■ HPC в нефтегазовой отрасли - источники информации

- Rice Oil & Gas HPC Workshop ([www.og-hpc.org](http://www.og-hpc.org))
- SEG Annual Meeting ([www.seg.org](http://www.seg.org))
- EAGE Conference & Exhibition ([www.eage.org](http://www.eage.org))
- Конференция «Суперкомпьютерные технологии в нефтегазовой отрасли» ([www.hpc-oilgas.com](http://www.hpc-oilgas.com))\*

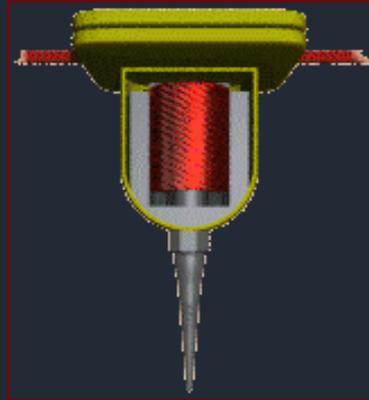
\* с 2018 года – часть конференции «ГеоЕвразия» ([www.gece.moscow](http://www.gece.moscow))

## ■ Что такое сейсморазведка

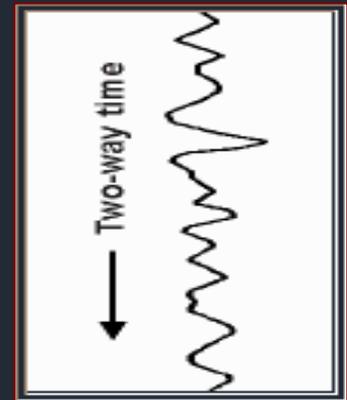
### Принцип



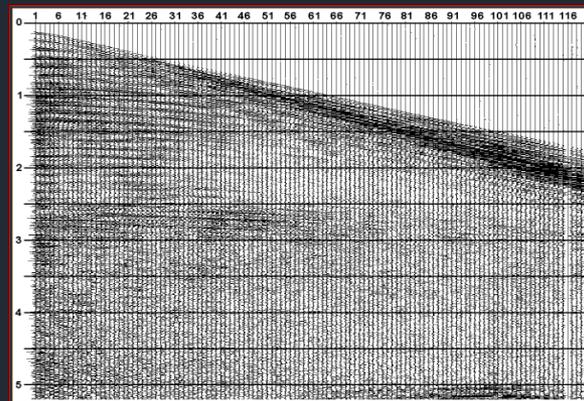
### Сейсмоприёмник



### Трасса

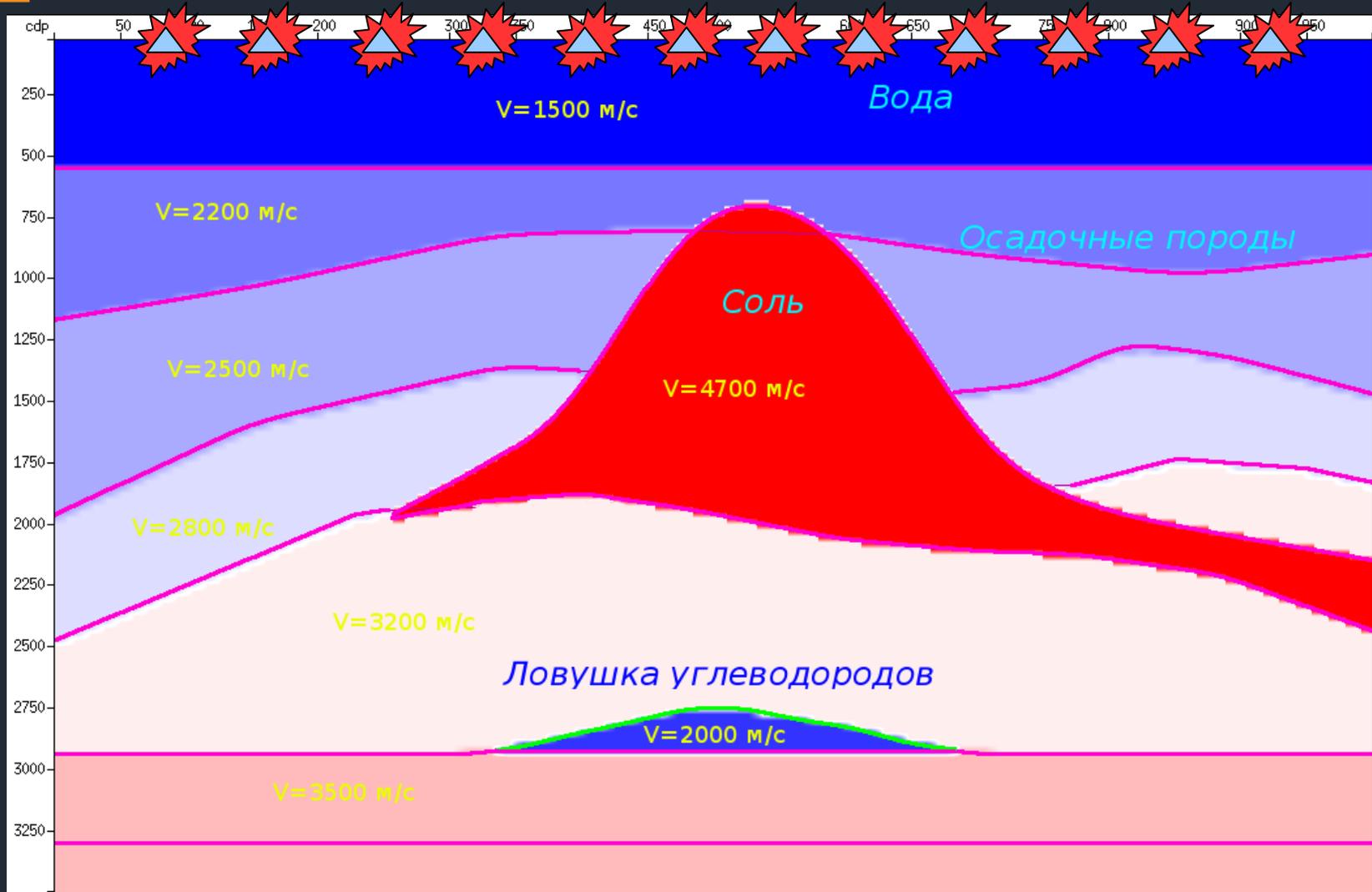


### Сейсмограмма

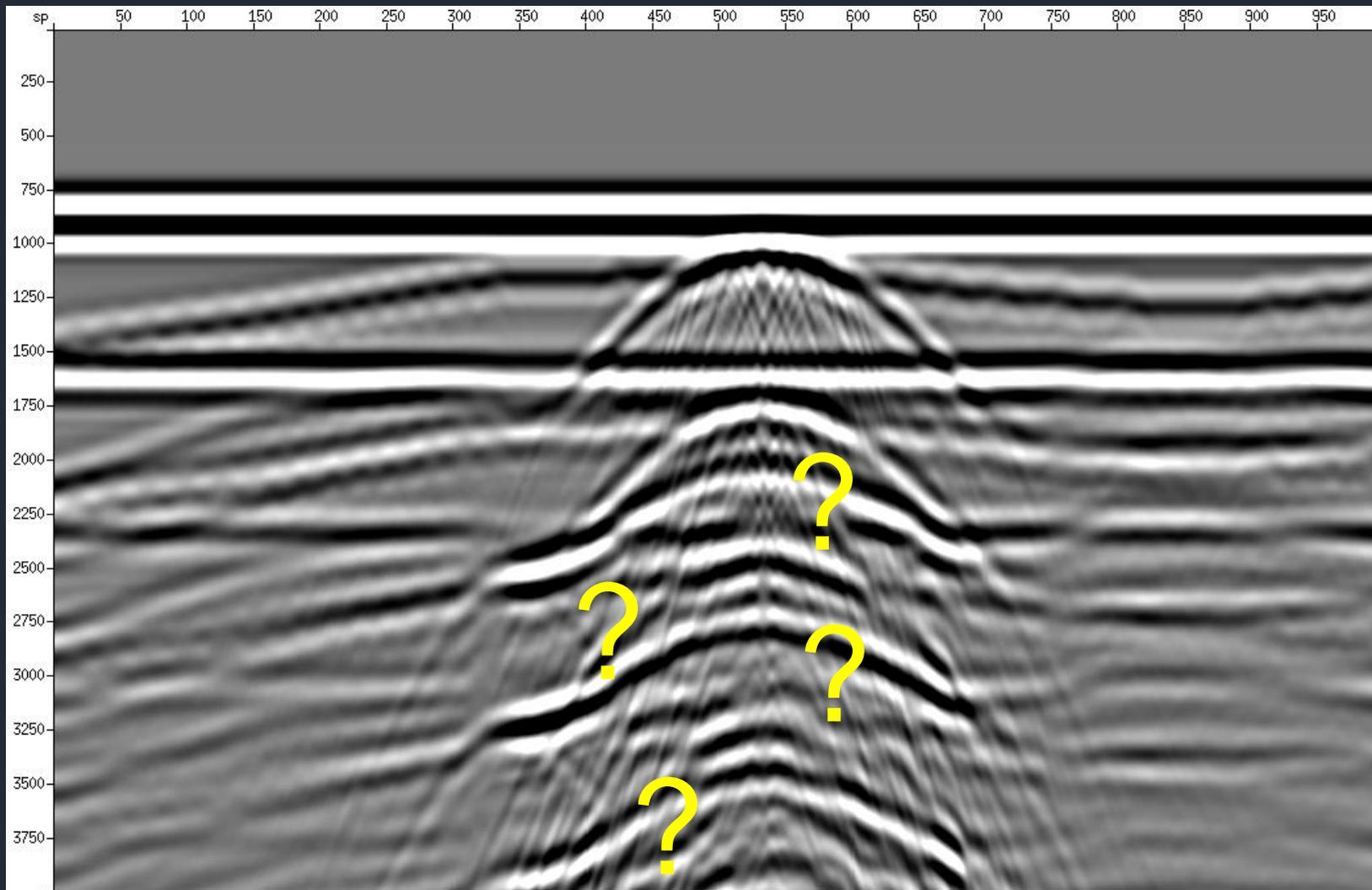


Нефть?  
Газ?

## Если нефтяное месторождение устроено так:



## То сейсмические данные выглядят так:



## ■ Задачи обработки данных сейсморазведки

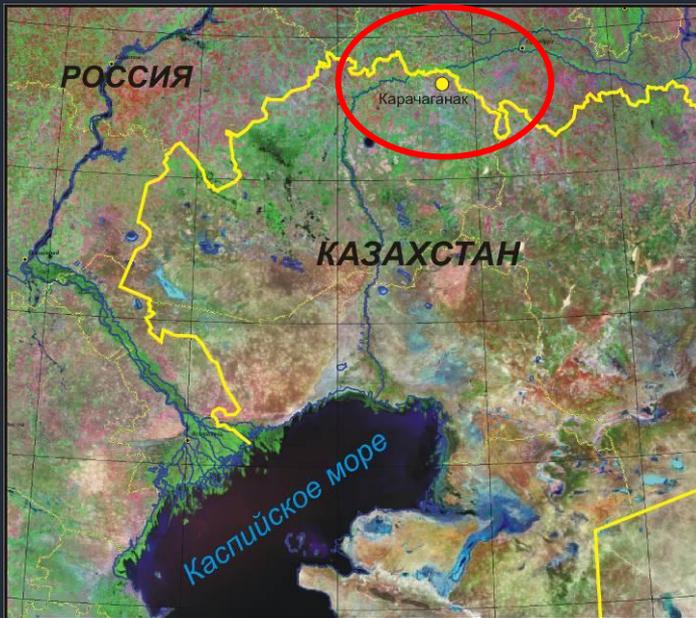
- Подавление волн-помех
- Выделение полезного сигнала
- Повышение разрешенности
- Оценка глубинно-скоростной модели среды
- Построение глубинного изображения среды
- Оценка состава и свойств горных пород

## ■ Математический аппарат

- Решение систем линейных и нелинейных алгебраических уравнений большой размерности
- Интегральные преобразования
- Расчёт и применение цифровых фильтров
- Методы оптимизации
- Решение уравнений в частных производных
- Методы теории вероятностей и математической статистики

## Задачи сейсмоки характеризует большой объём данных

Месторождение Карачаганак, 3D съёмка, 2009 г.



Площадь наблюдений ~ 900 кв.км

Количество пунктов возбуждения (ПВ) ~ 300 000

Количество пунктов приёма (ПП) ~ 100 000

Количество средних точек (ОСТ) ~ 9 000 000

Кратность сейсмограммы ОСТ 330

Количество трасс (записей) ~ 3 миллиарда

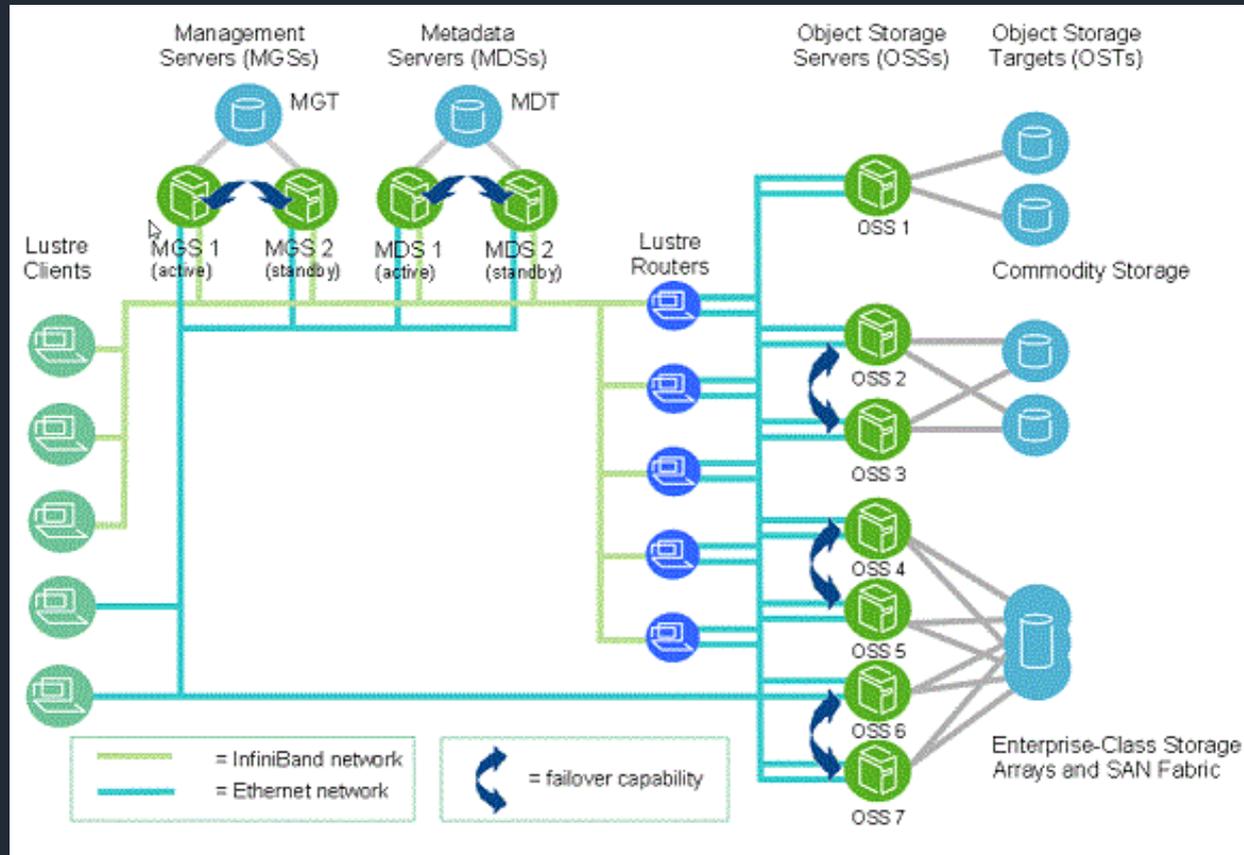
Длина записи 20 сек, 10000 отсчётов, 4 байта/отсчёт

Объём данных > 100 терабайт

## ■ Основные направления развития систем и методов хранения данных

- Использование распределенных файловых систем, эмулирующих классические (Lustre, GPFS, BeeGFS)
- На прикладном уровне: создание новых форматов хранения данных, лояльных к распределенным файловым системам (JavaSeis, OpenCPS)
- Гибридные программно-аппаратные комплексы (сериализация случайной записи на SSD, каскадные хранилища) (ExSeisPIOL, HDF5)

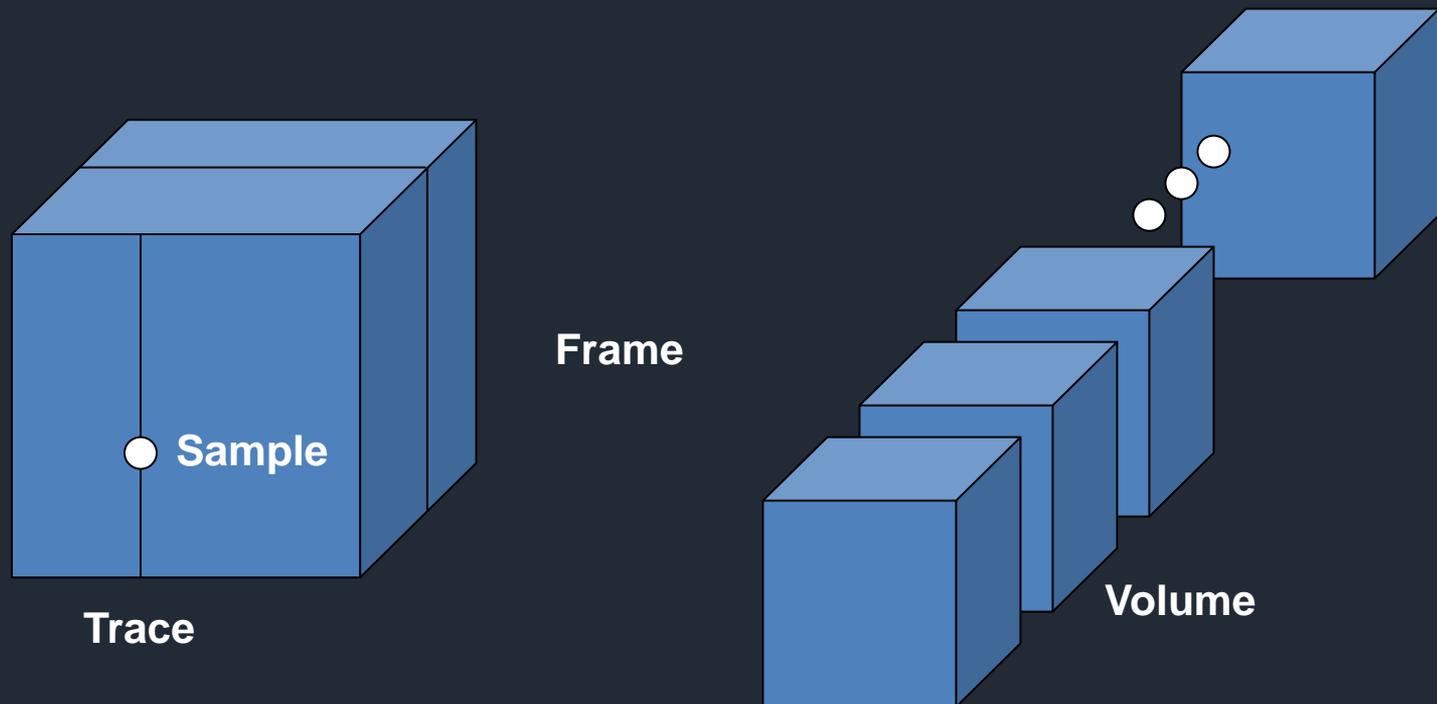
# Lustre<sup>1</sup>



<sup>1</sup><http://lustre.org/about/>

## ■ JavaSeis файловый формат<sup>1</sup>

В основе 4-х мерный массив: Volume – Frame – Trace – Sample  
Обычно представляется в виде серии 3-х мерных



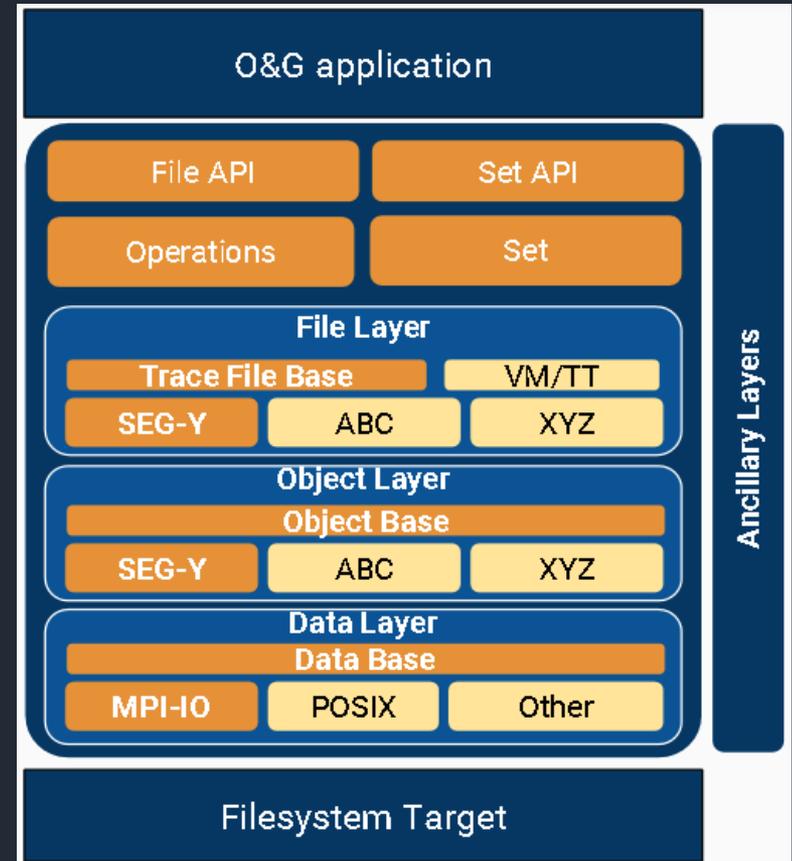
<sup>1</sup><http://javaseis.sourceforge.net/documents/ParallelArrays.ppt>

## ■ ExSeisPIOL<sup>1</sup>

Многоуровневое решение (уровни отделены)

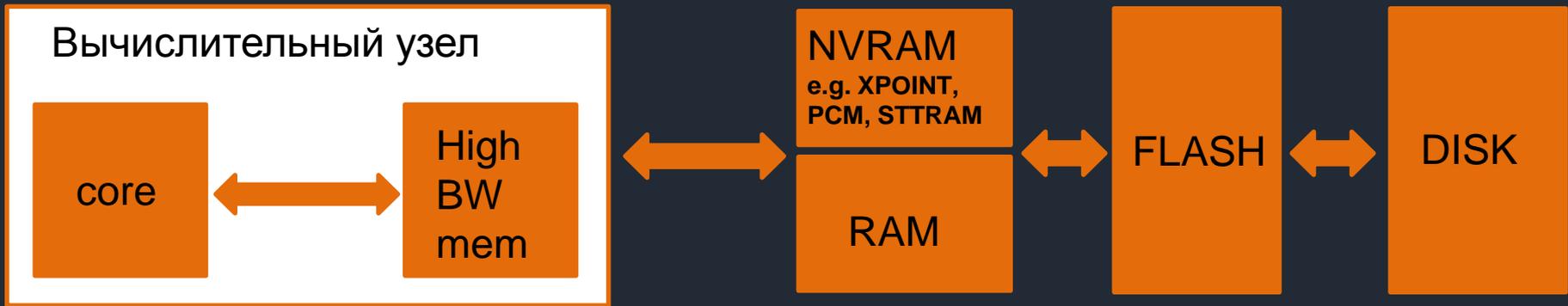
Два уровня доступных публичных API (High-level set API: Implicit I/O и Low-level file API: Explicit I/O)

Оптимизировано для обработки сейсмических данных



<sup>1</sup>Facilities, Infrastructure & Networking: Extreme-Scale Parallel I/O for Seismic Workflows, Dr. Cathal Ó Broin, RICE 2017 Oil & Gas HPC Conference

## Каскадное хранение<sup>1</sup>



Node BW	1 TB/s	100 GB/s	20 GB/s	5 GB/s
Cluster BW	1 PB/s	100 TB/s	5 TB/s	100 GB/s
Software	Language level	Language level NVL-C	HDF5 DAOS	Parallel FS Campaign Storage
Key features	transparent computation	transparent computation ultra-fast storage apps small object container	name space scientific formats FS style container	bulk data movement - many files - subtrees of MD

<sup>1</sup> The Future of Large Scale IO: Opportunities & Obstacles, Peter Braam, RICE 2017 Oil & Gas HPC Conference

## ■ Алгоритмы применяемые в сейсмике имеют высокую вычислительную сложность

### Задача построения глубинного изображения<sup>1</sup>

Одна итерация алгоритма для вычислительного ядра 16-ого порядка в трансверсально-изотропной среде ~ **6000 FLOPs**

Размер сетки: 1580 x 1580 x 1130 ~  **$2.82 \times 10^9$**

Число операций для одного источника:

$2.82 \times 10^9 \times 3000$  (время) x 2 (прямое + обратное) ~  **$10^{17}$  FLOPs**

Количество источников ~ **30000**

Итого на одну итерацию ~  **$3 \times 10^9$  TFLOPs**

Число итераций ~ **15**

**Общее число операций ~  $4.6 \times 10^{10}$  TFLOPs**

<sup>1</sup> Programming: Leveraging Symbolic Math for Rapid Development of Applications for Seismic Modeling, Navjot Kukreja, M. Louboutin, F. Luporini, P. Witte, M. Lange, F.Hermann, G. Gorman, RICE 2017 Oil & Gas HPC Conference

**■ Большая вычислительная сложность при достаточно высокой степени параллелизма позволяет решать многие задачи сейсмологии на сопроцессорах (ускорителях)**

### Решение задач волнового моделирования

- Для графических процессоров (GPU)<sup>1</sup>
- Для сопроцессоров Intel Xeon Phi (поколение KNL)<sup>2</sup>

### SLR разложение при решении уравнения Гельмгольца

- Для manycore систем (систем с большим числом вычислительных ядер)<sup>3</sup>

<sup>1</sup> Programming: Revisiting Wave Propagation Applications on GPUs: Improving the Accuracy and Performance Tradeoff and Implementing Portability via OCCA, Thomas Cullison, Scott Morton, Sergey Terentyev, Arturo Vargas, RICE 2017 Oil & Gas HPC Conference

<sup>2</sup> Programming: Many-core Implementation of Numerical 3D Isotropic Acoustic Wave Equation and Analysis of its Performance Portability, Reid Atcheson, RICE 2017 Oil & Gas HPC Conference

<sup>3</sup> Algorithms and Performance: High Performance Low Rank Schur Complement for the Helmholtz Equation, Hatem Ltaief, RICE 2017 Oil & Gas HPC Conference

## ■ Нейронные сети и методы глубокого обучения

### Задача прямого продолжения поля<sup>1</sup>

На абстрактном уровне глубокая нейронная сеть это:

- Ациклический направленный граф
- Узлы: вычислительные сущности
- Рёбра: потоки данных

Таким образом нейронные сети в различных формах подходят для решения широкого спектра задач



<sup>1</sup> Scaling Deep Learning Applications: Theoretical and Practical Limits, Janis Keuper, RICE 2017 Oil & Gas HPC Conference

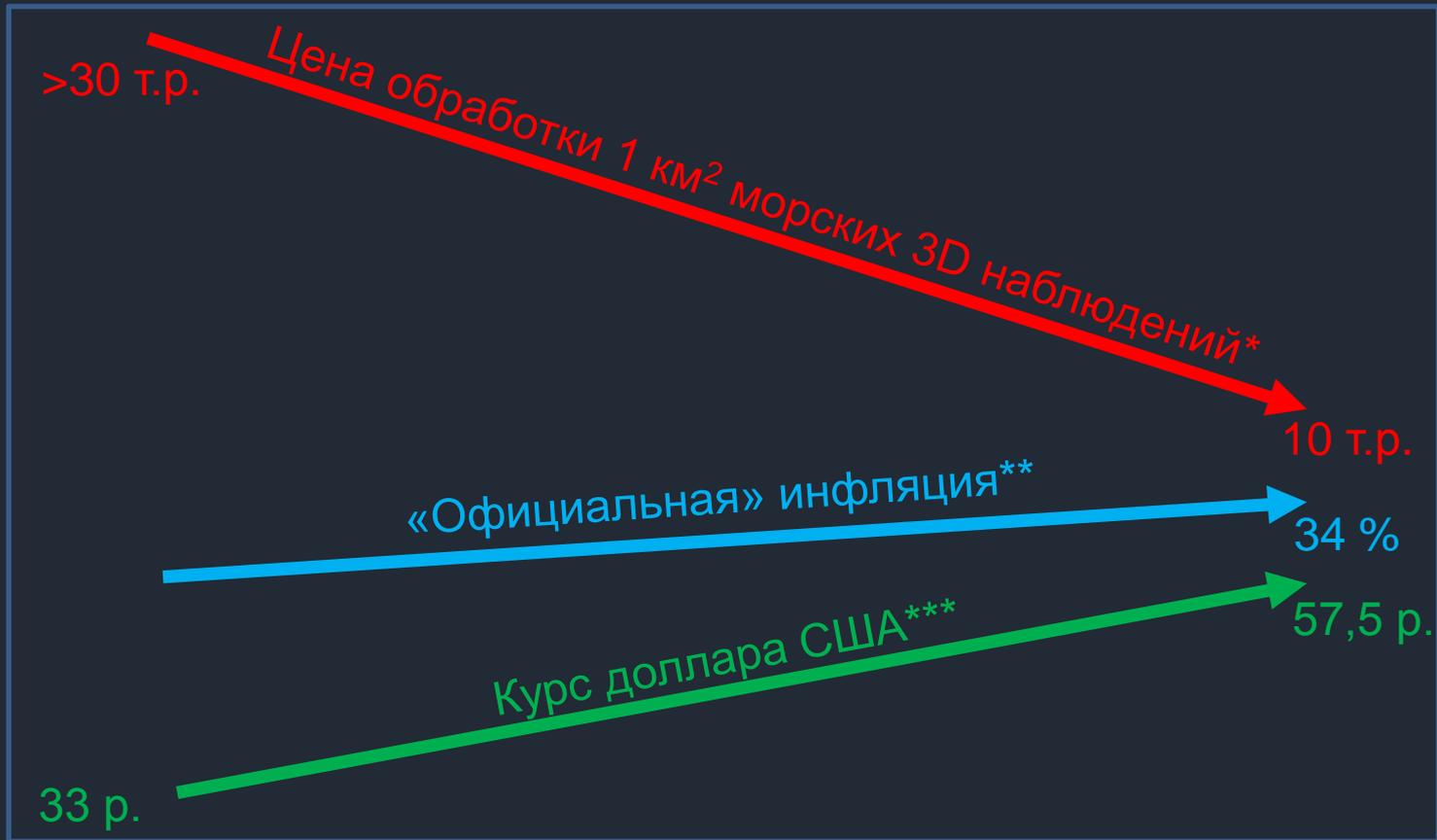
## ■ НРС-разработки алгоритмов и ПО для сейсморазведки в России\*

- ПО для решения широкого спектра геолого-геофизических задач (проект **Союзного государства Беларуси и России «СКИФ-Недра»**)
- Алгоритмы и программы для решения прямых и обратных задач сейсморазведки (**ИВМиМГ СО РАН, МФТИ**)
- ПО для обработки данных сейсморазведки (**Яндекс-Терра**)
- ПО для обработки сейсмических данных (**ГЕОЛАБ**)
- Решение отдельных ресурсоёмких задач (**МГУ, СпбГУ и др.**)

\* Список не претендует на полноту, он отражает только научные интересы автора

## ■ Экономика...

2014 г.



2017 г.

Источники

\* [rn.tektorg.ru](http://rn.tektorg.ru)  
\*\* Росстат  
\*\*\* [rbc.ru](http://rbc.ru)

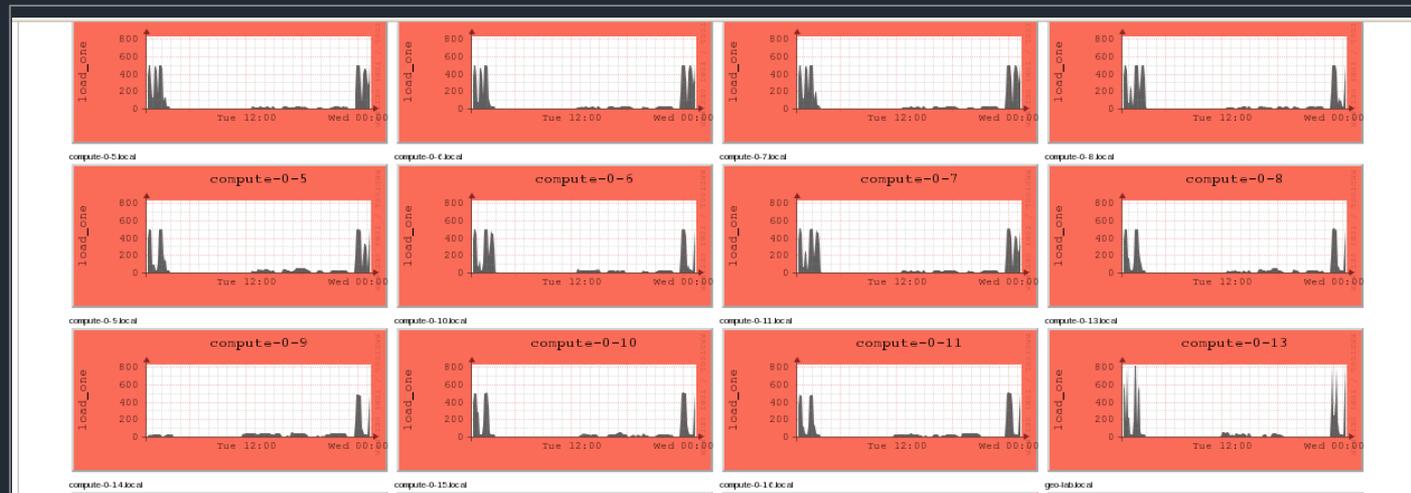
## ■ Что делаем мы?

Исследуем и оптимизируем

Алгоритмы, ПО, оборудование ← *Perpetual Beta*

Цель

«Выжать» максимум из имеющегося оборудования!



Окно программы мониторинга Ganglia при оптимальной загрузке кластера "ГЕОЛАБ"

## ■ Экспериментируем с оборудованием



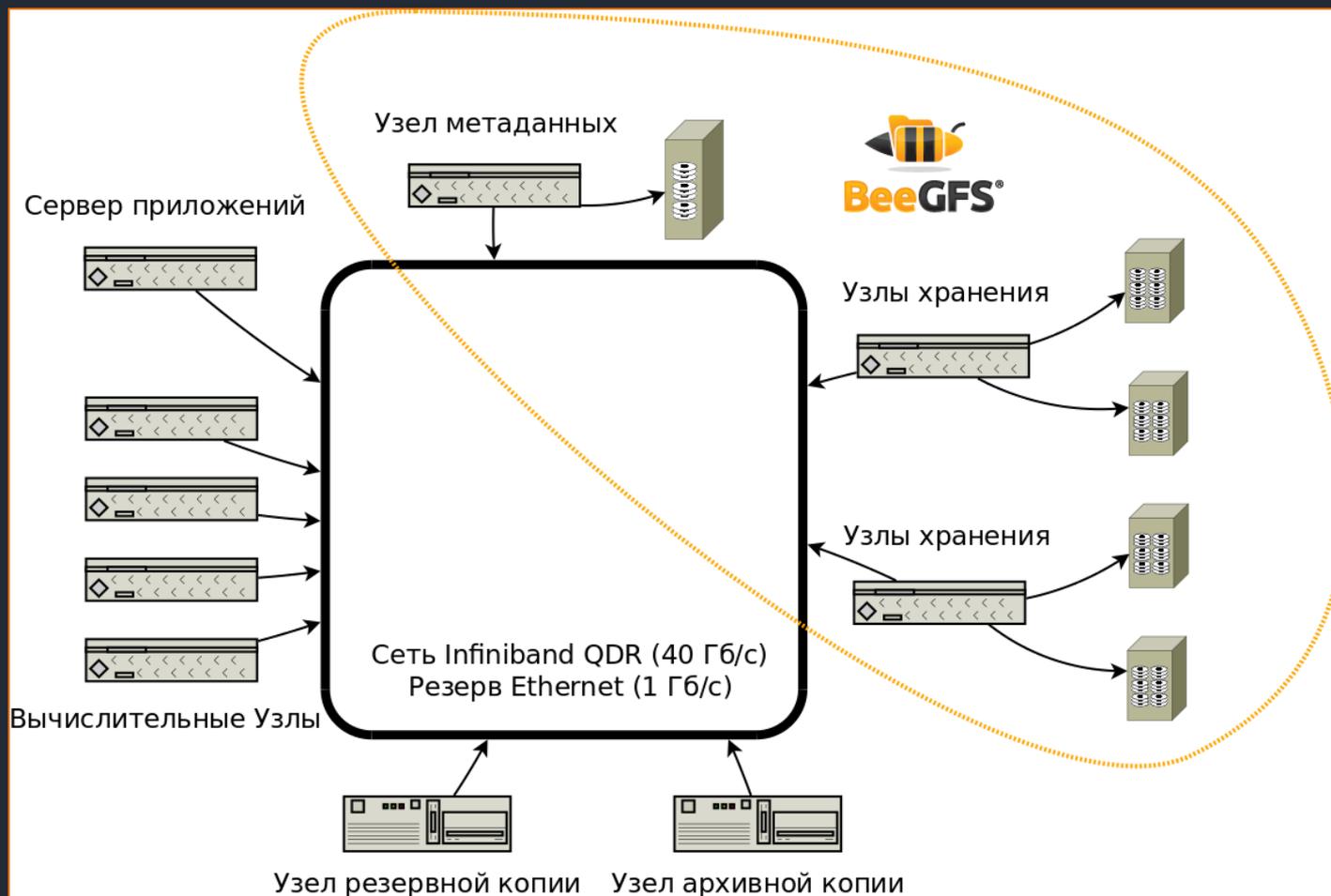
Выбор оптимальной конфигурации RAID массивов с использованием SSD-кэша

## ■ Экспериментируем с оборудованием



Оптимизация подсистемы быстрого ввода данных с внешних носителей

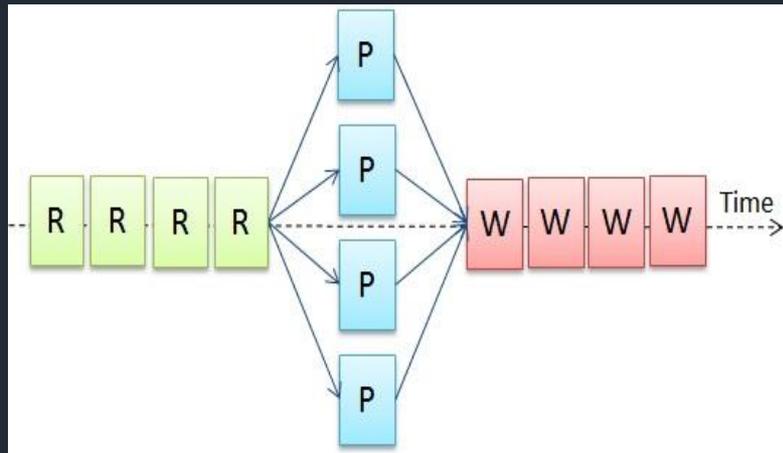
## Результат экспериментов – очередной проект ВЦ



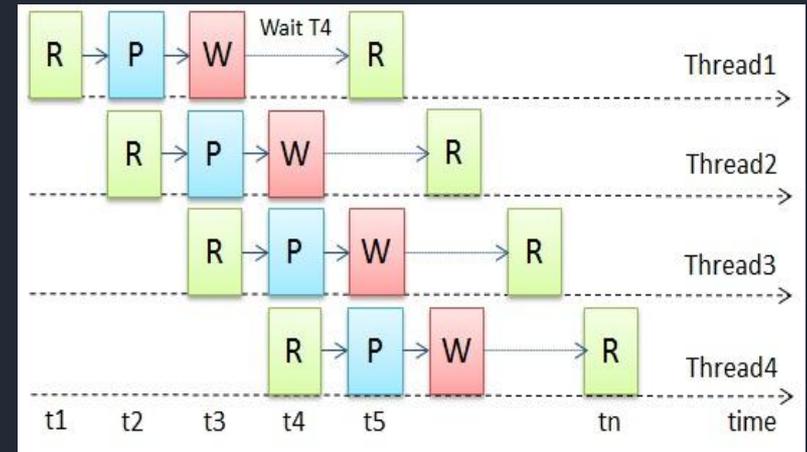
## Оптимизируем программы

Подавление кратных волн-помех при помощи алгоритма SRME - рефакторинг

До рефакторинга



После рефакторинга



Серьёзную помощь оказывают инструменты: Intel® Vtune Amplifier и Intel® Advisor

Далее – векторизация с использованием AVX2

## ■ HPC в нефтегазовой отрасли - источники информации

- Rice Oil & Gas HPC Workshop ([www.og-hpc.org](http://www.og-hpc.org))
- SEG Annual Meeting ([www.seg.org](http://www.seg.org))
- EAGE Conference & Exhibition ([www.eage.org](http://www.eage.org))
- Конференция «Суперкомпьютерные технологии в нефтегазовой отрасли» ([www.hpc-oilgas.com](http://www.hpc-oilgas.com))\*

\* с 2018 года – часть конференции «ГеоЕвразия» ([www.gece.moscow](http://www.gece.moscow))

- Конференция и выставка «ГеоЕвразия-2018», секция «Суперкомпьютерные технологии в нефтегазовой отрасли»
  - Время 5-8 февраля 2018 года
  - Место Центр международной торговли на Красной Пресне
  - Сайт [www.gece.moscow](http://www.gece.moscow)



# GEO LAB

Geophysical Innovations

[WWW.GEO-LAB.RU](http://WWW.GEO-LAB.RU)

