ФАЗЗИНГ СЕТЕВЫХ ПРОТОКОЛОВ С ИСПОЛЬЗОВАНИЕМ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ

ПЕЧЁНКИН А.И.

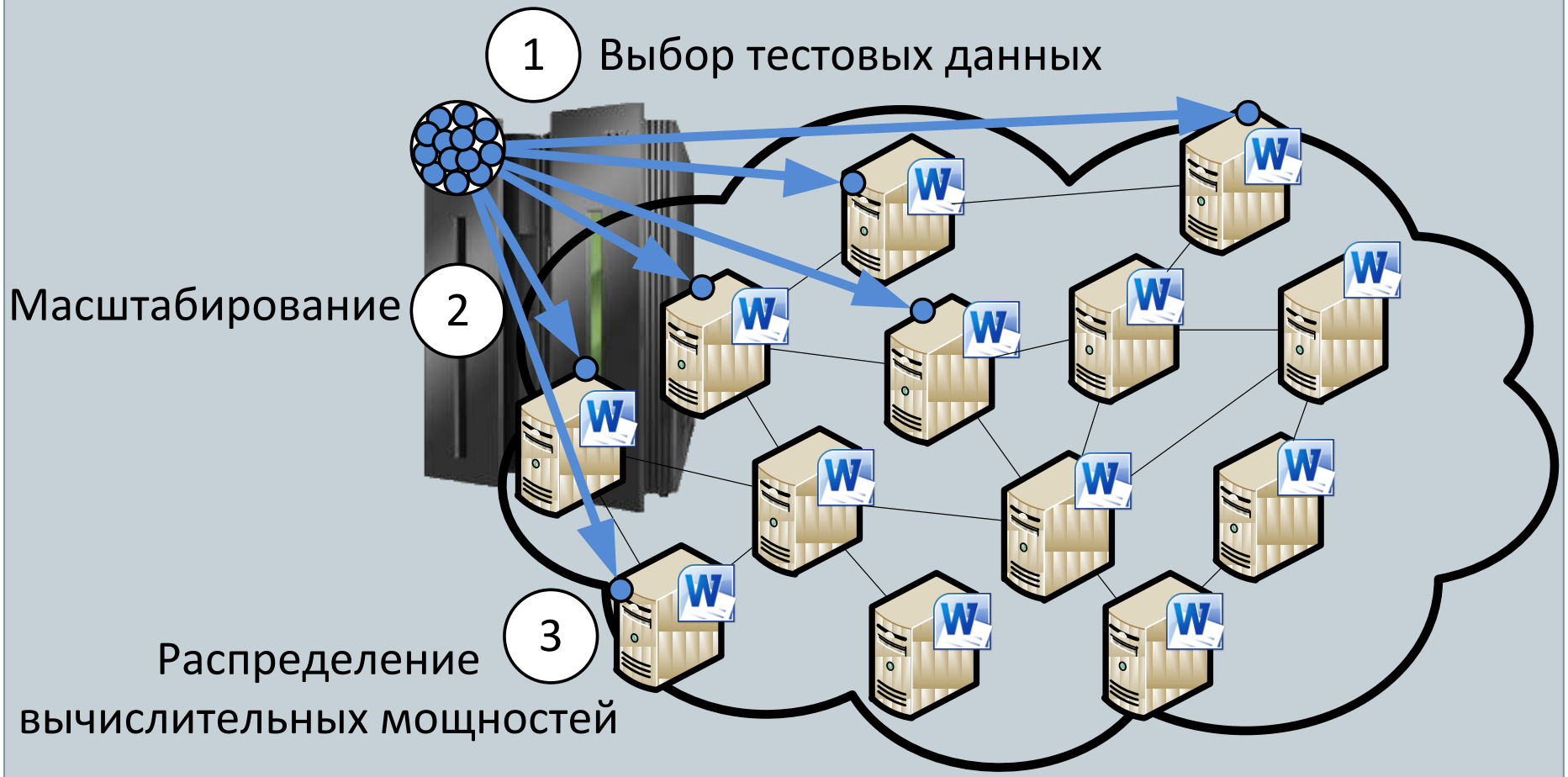
КАФЕДРА «ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ КОМПЬЮТЕРНЫХ СИСТЕМ»

ФГБОУ ВПО «САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

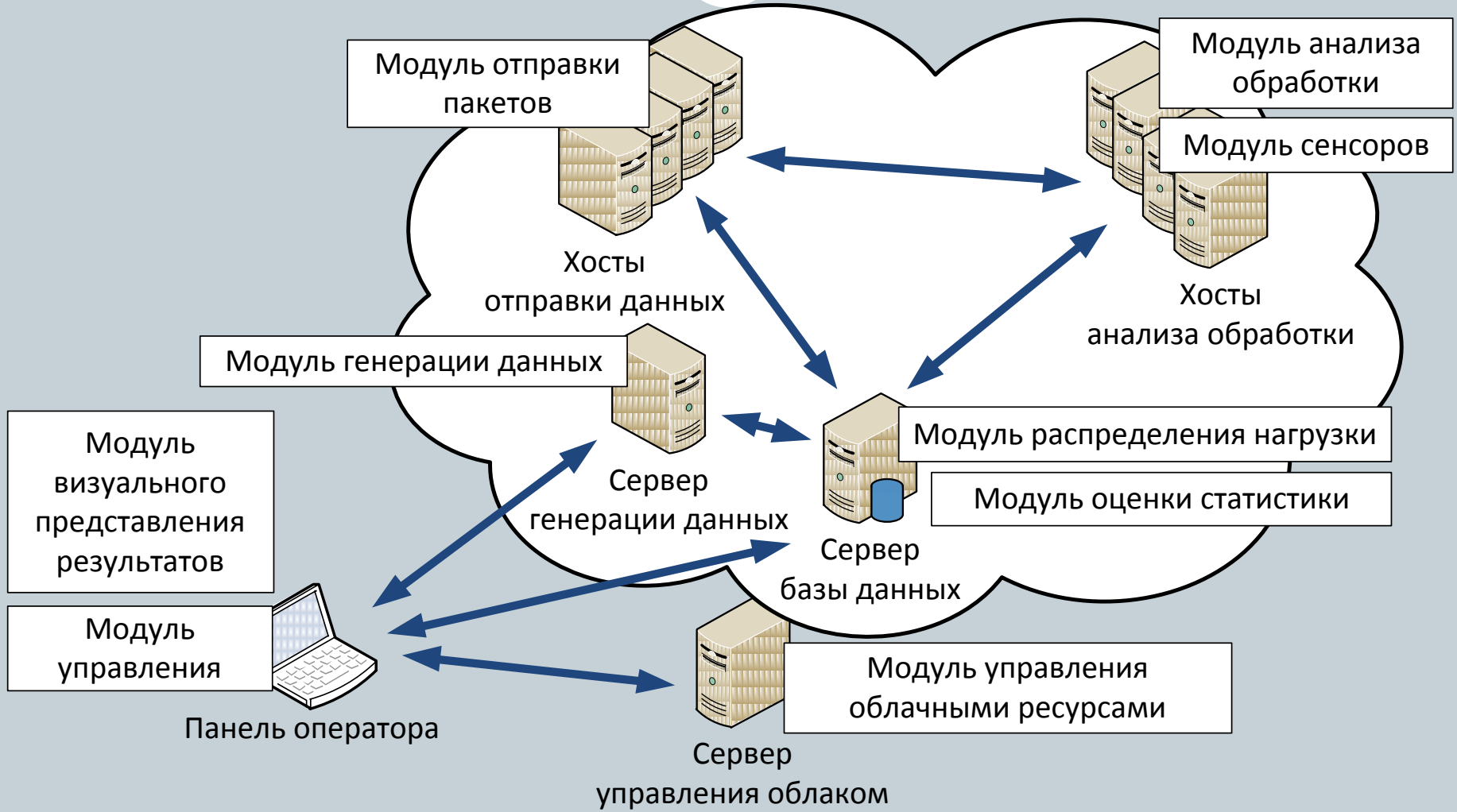
конференция
РусКрипто'2013

http://www.ruscrypto.ru/conference/

Масштабирование задачи фаззинга



Состав масштабируемой системы фаззинга сетевых протоколов



Выбор тестовых данных
Генетический алгоритм

Генетический алгоритм - метод оптимизации, основанный на концепциях естественного отбора и генетики

Оперирует конечным числом «особей» - популяцией

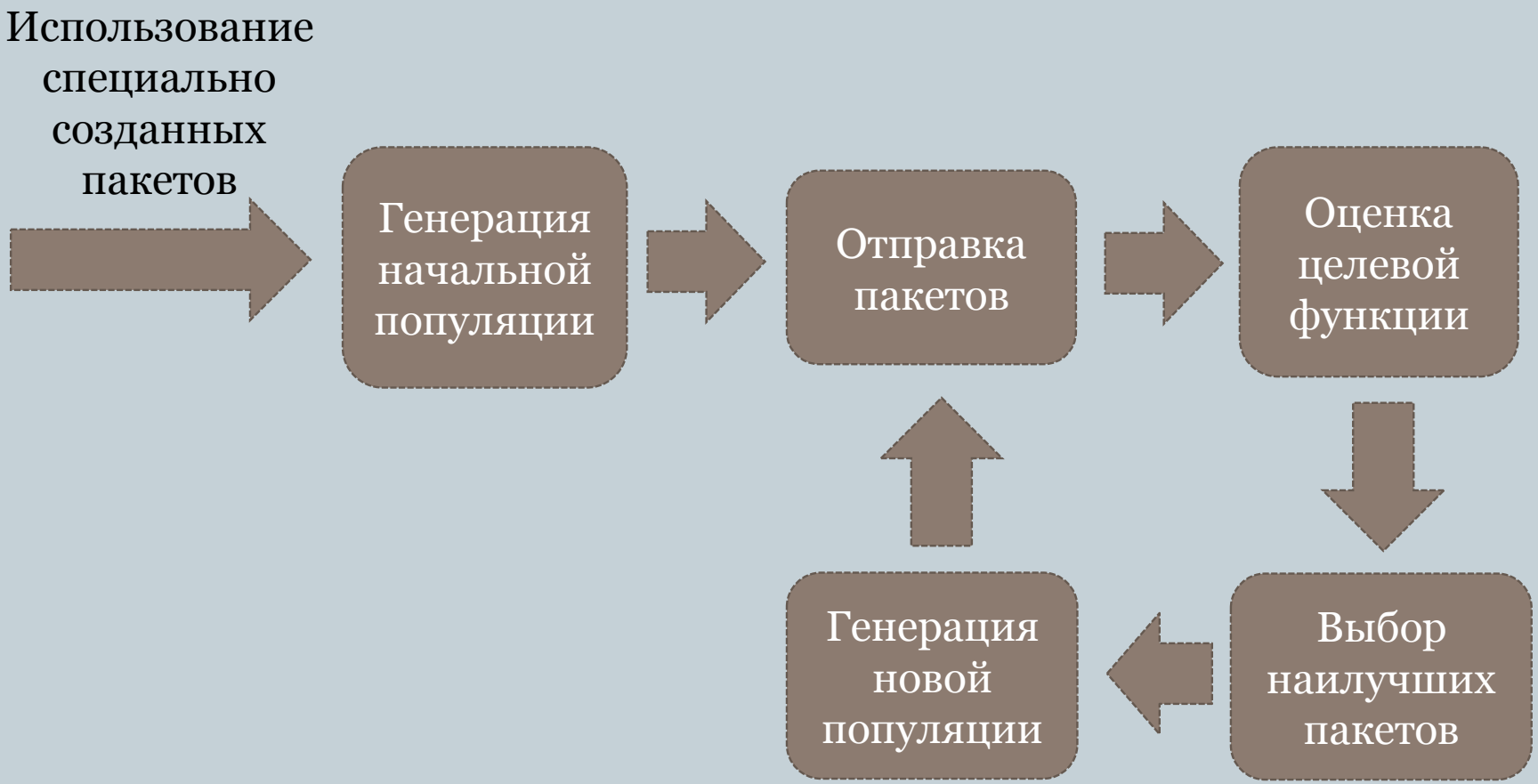
Каждая «особь» представляет собой возможное решение рассматриваемой проблемы

Критерий отбора – качество предложенного решения (значение целевой функции)

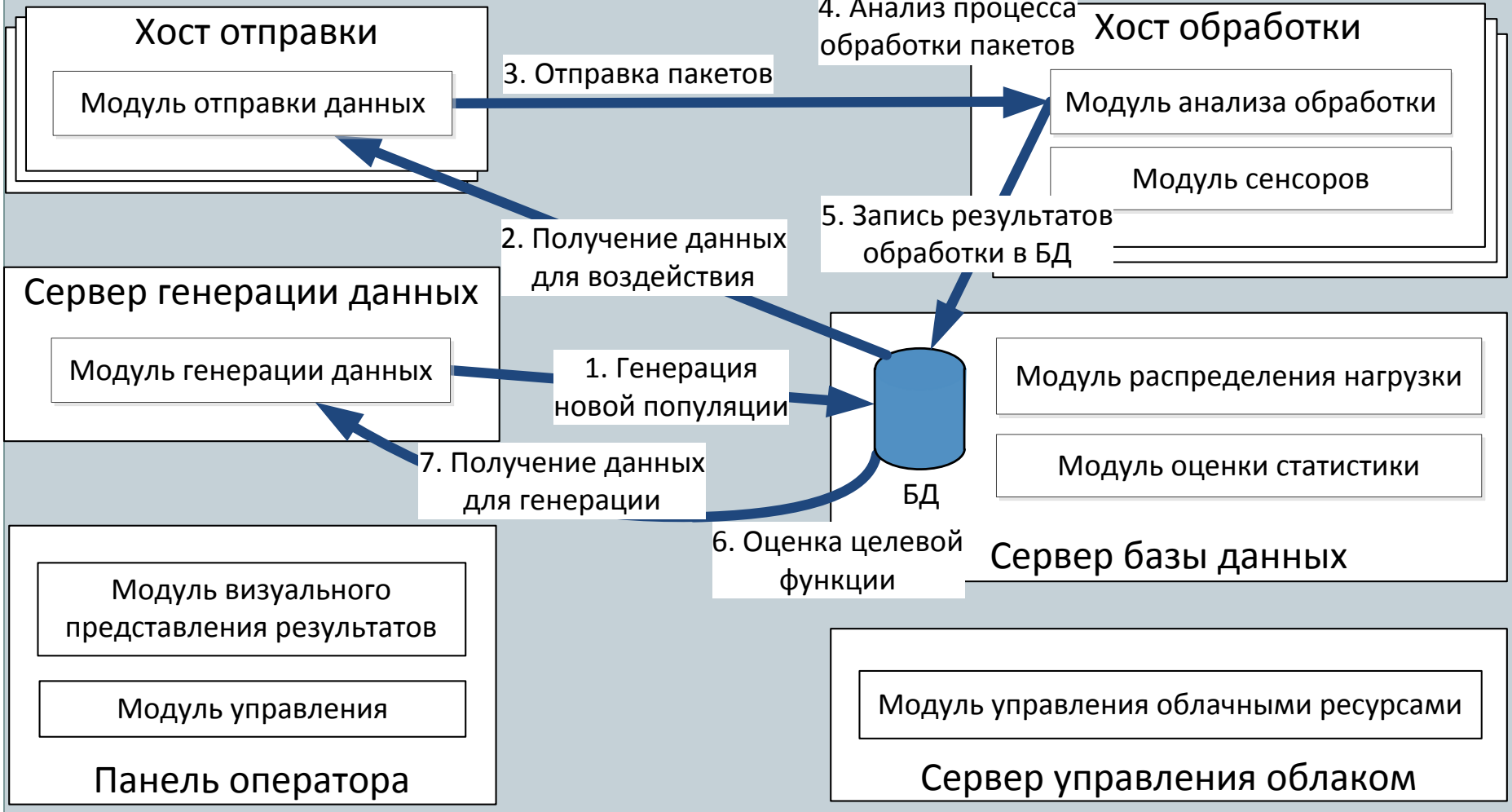
Наиболее приспособленные «особи» получают возможность воспроизводить потомство с другими «особями» популяции

Результат работы алгоритма – набор решений с требуемыми значениями целевой функции

Общая схема использования ГА



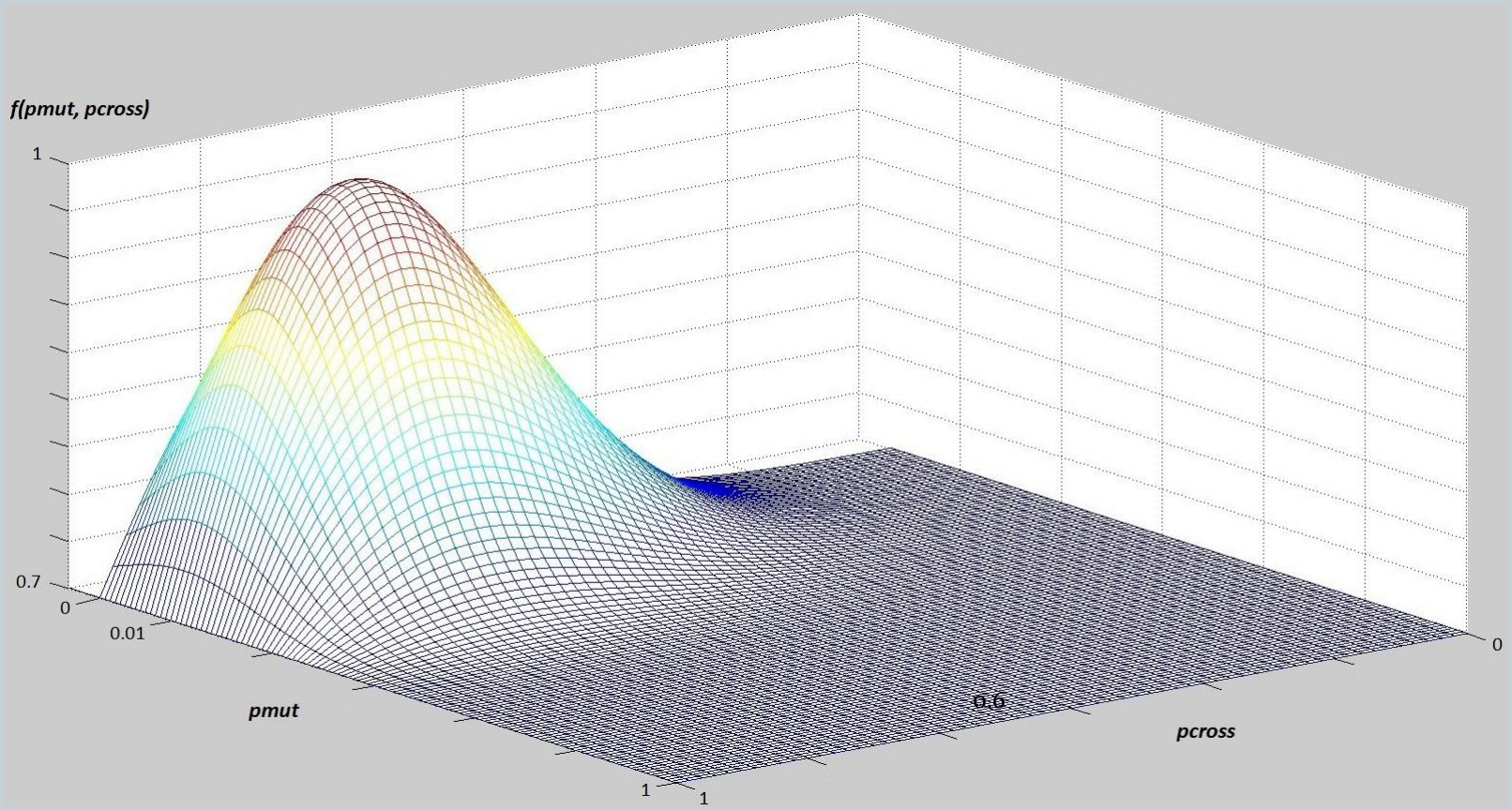
Масштабируемая система фаззинга
Выбор тестовых данных



Настраиваемые параметры ГА

- Данные о сетевых пакетах
  - Максимальный размер пакета – размер сетевого пакета в байтах. Параметр позволяет сузить размер перебора и повысить эффективность работы системы, поскольку многие сетевые сервисы принимают пакеты фиксированного размера
  - Входные заранее сформированные сетевые пакеты в начальную, либо последующие популяции
- Размер популяций
  - Размер популяции – максимальное количество сетевых пакетов, которые будут создаваться в рамках одной популяции
  - Размер выборки – количество пакетов, которые будут выбираться из предыдущего поколения для генерирования нового
- Характеристики рождения
  - Вероятность мутации – значение вероятности, с которой будет осуществляться мутация байтов в сетевом пакете при создании новой популяции
  - Вероятность скрещивания – значение вероятности, с которой будет осуществляться скрещивание двух сетевых пакетов

Выбор параметров ГА



$$f(pcross, pmut) = \frac{Fitness(pcross, pmut)}{Fitness\_\max}$$

$\Rightarrow$ `pmut = 0.01, pcross = 0.6`

Модуль анализа обработки

Результаты анализа

Оценка времени обработки сетевого пакета

Получение пути обработки пакета

Получение API-функций, выполняющихся при обработке

Возможные объекты анализа

Процессы

Службы

Модули ядра

Используемые целевые функции

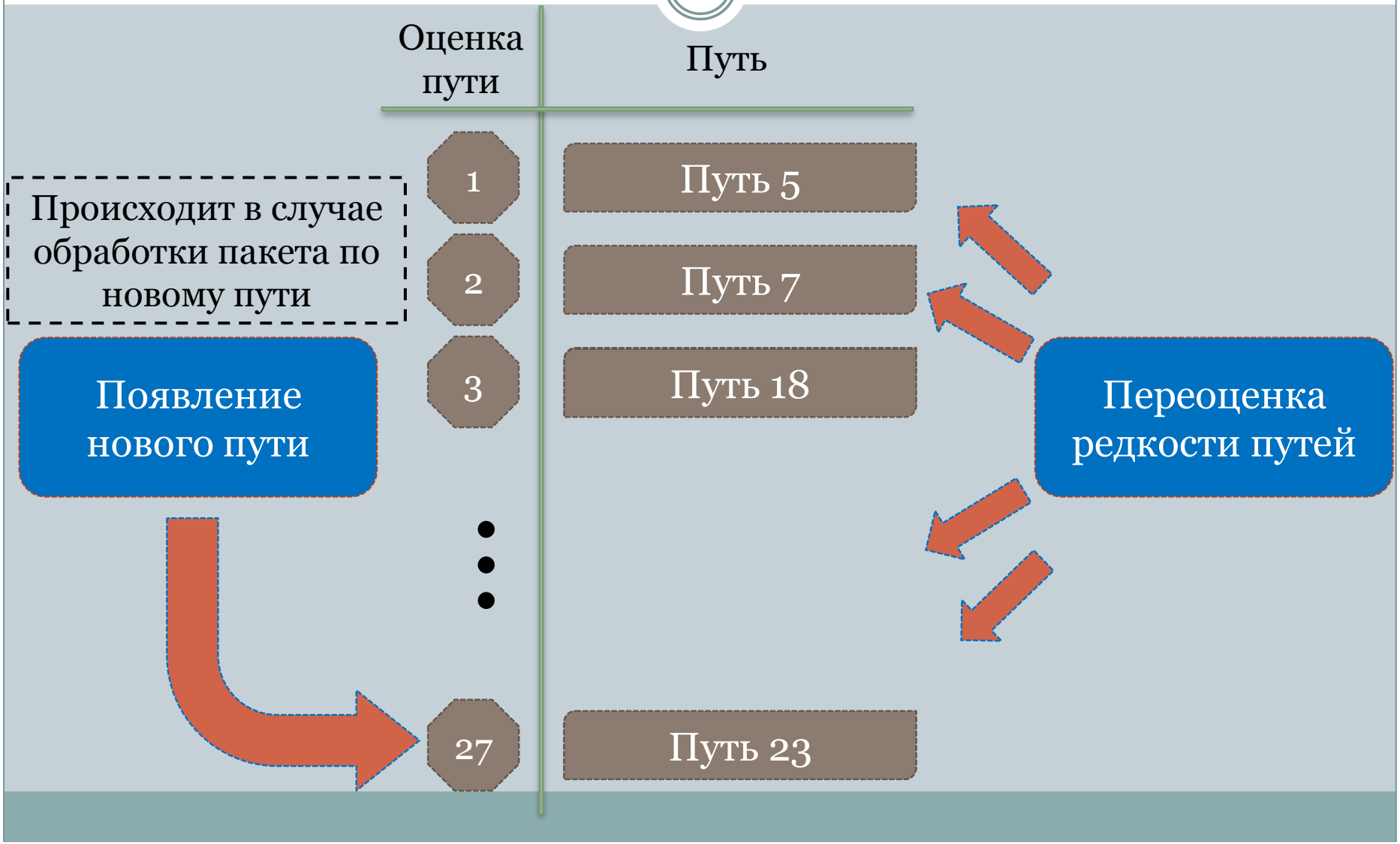
Целевые функции

- Время обработки пакета
- Количество выполненного кода
- Количество вызванных API-функций
- «Новизна» пути
- «Редкость» пути
- Близость пути к необходимому месту

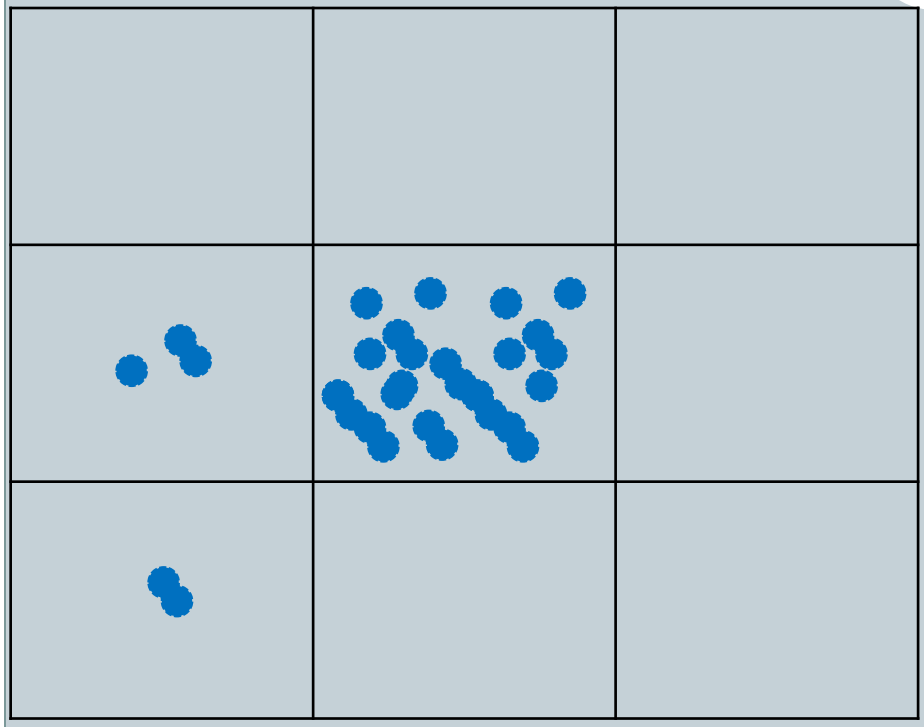
Экспериментальные исследования

- Для целевых функций «время обработки», «количество кода», «количество API-функций» в определённый момент происходит насыщение, и следующие популяции не приводят к улучшению результата

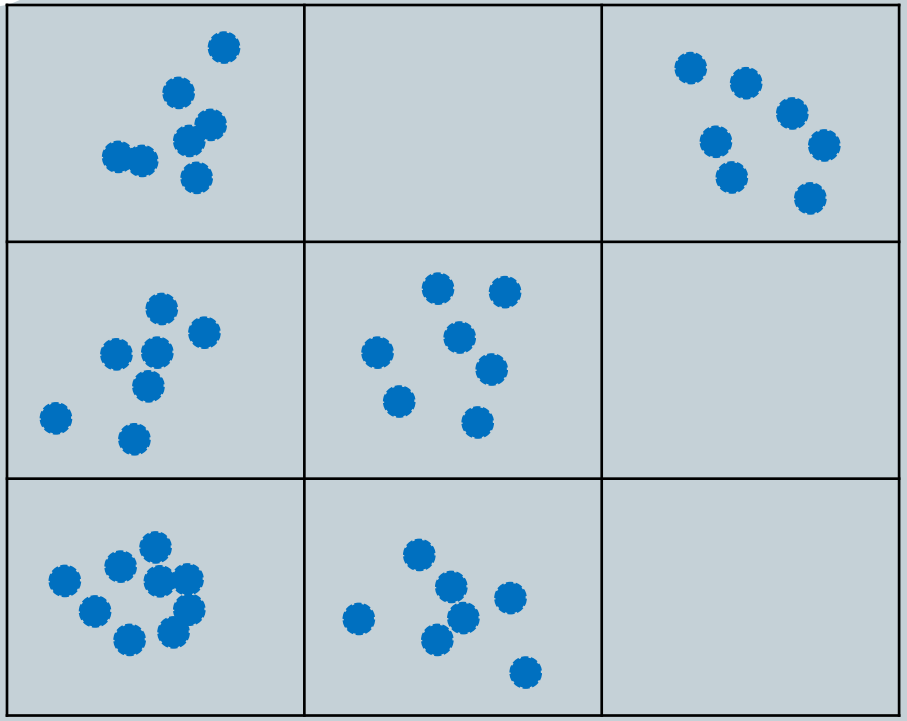
Оценка целевой функции
«Редкость» пути обработки



Распределение пакетов по путям обработки

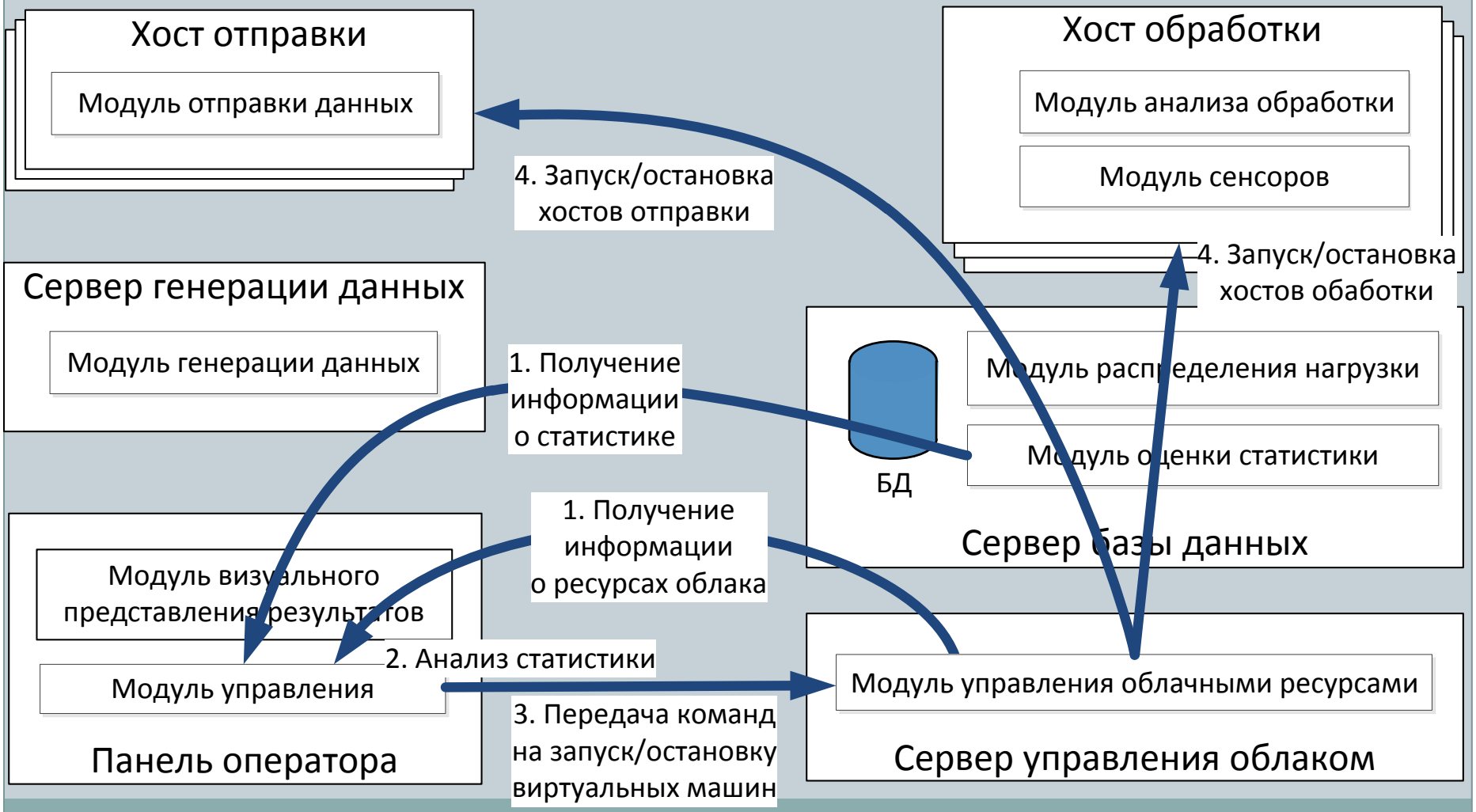


Случайные пакеты



Генерация пакетов с целевой функцией «редкость» пути

Масштабируемая система фаззинга
Масштабирование

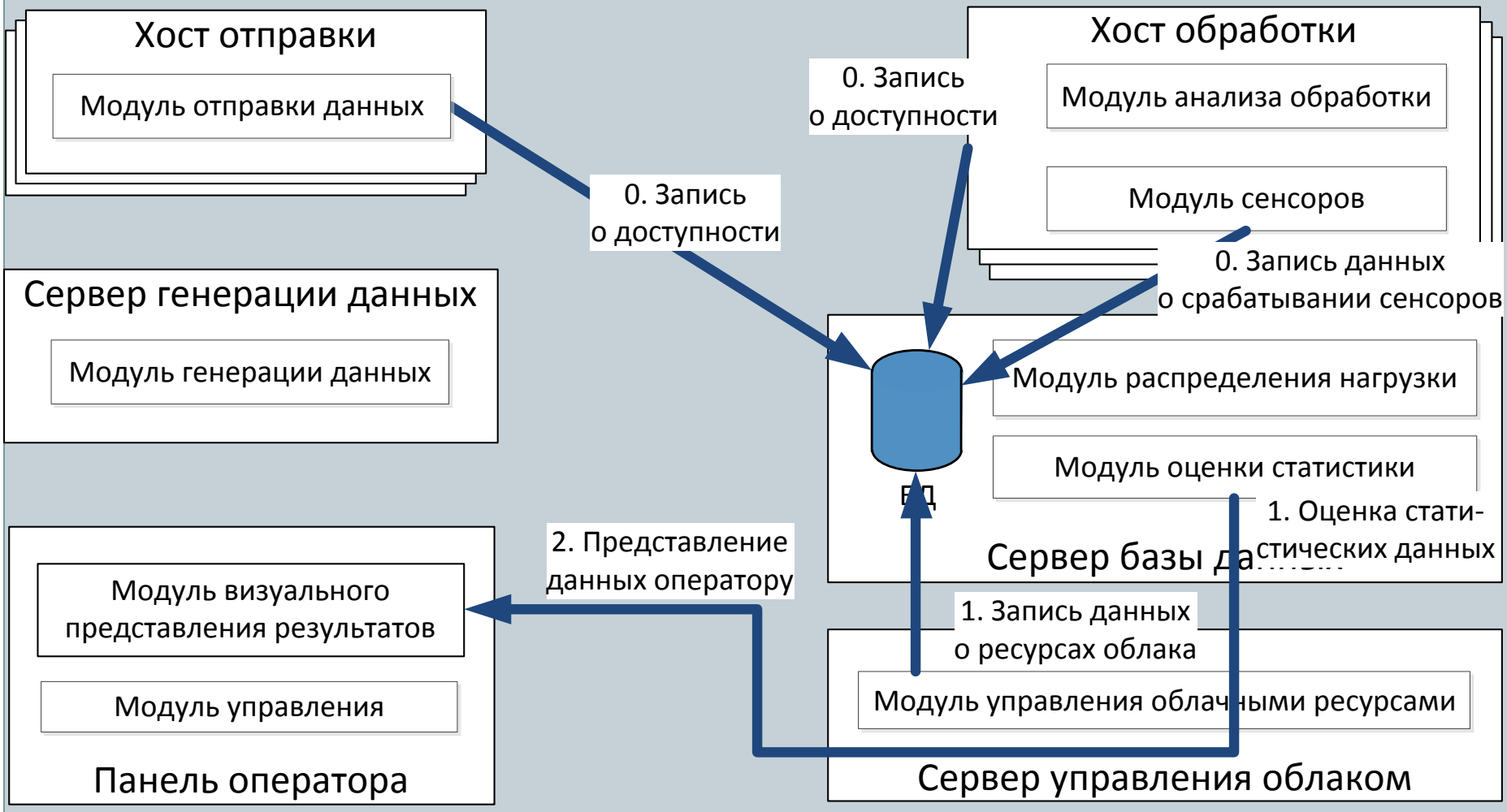


Данные о масштабируемости

- Персональный компьютер Core i7 2,8 ГГц, 8 Гб оперативной памяти
  - 1 сервер генерации, совмещенный с БД
  - 1 хост отправки данных
  - 3 хоста анализа обработки
- Многопроцессорный сервер 160 ядер (320 потоков), 600 Гб RAM
  4-ре IBM x3850 X5 по 4 процессора Intel(R) Xeon(R) CPU E7-8860 2.27GHz (10 ядер, 20 потоков), 150 Гб RAM
  - 1 сервер генерации
  - 1 сервер БД
  - ~30 хостов отправки данных
  - ~200 хостов анализа обработки

Масштабируемая система фаззинга
Распределение нагрузки

Масштабируемая система фаззинга
Оценка статистики



Статистические данные

Параметр	Значение
Хосты-жертвы	100 шт.
Хосты-отправители	10 шт.
Общее время функционирования	~ 2 недели
Популяции пакетов	~1 млн.
Обработанные пакеты	~500 млн.
Среднее скорость обработки пакетов	~500 пак./с
Протокол	NTP
% покрытия кода (w32time.dll)	~45%

Возможности системы фаззинга

- Выбор тестовых воздействий
  - Использование ГА для генерации пакетов
  - Возможность выбора параметров ГА
- Модульная архитектура
  - Возможность добавления собственных методов генерации данных
  - Возможность добавления новых целевых функций
  - Возможность добавления модулей распределения нагрузки
- Эффективное использование аппаратных ресурсов
  - Автоматическое масштабирование
  - Автоматическое распределение нагрузки
  - Автоматическая обработка сбоев и возобновление работы (работа 24/7)