



Managed OpenSearch:

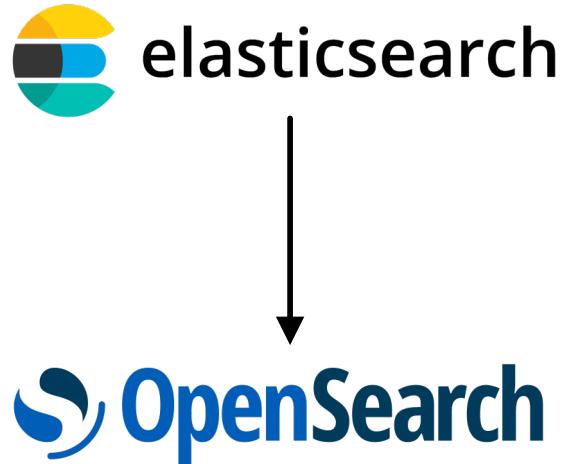
Yandex Cloud

Владислав Таболин,
руководитель группы
Специальные СУБД Yandex Cloud



OpenSearch

Современная платформа
для анализа данных



- Опенсорс-платформа для поиска, анализа и визуализации данных
- Форк Elasticsearch + Kibana (OpenSearch + Dashboards)
- Поддерживается Amazon

Что такое OpenSearch: сценарии

- **Логи и Observability**

Система хранения и анализа логов,
observability-платформа

- **Поисковый бэкенд**

Поиск по большим объёмам
данных

- **Векторный поиск**

RAG-системы

Почему OpenSearch?

Комплексное решение

Хранение данных, удобный веб-интерфейс, автоматизация анализа, расширенная система контроля доступа



Масштабируемость

Индексация больших объёмов данных, горизонтальное масштабирование



Гибкость

Возможность построения сложных систем с неоднородной конфигурацией хостов



Трудности, с которыми сталкиваются пользователи

- Сложность настройки и поддержки коробочного решения
- Высокая стоимость хранения больших объёмов данных
- Несовершенная логика балансировки данных при горизонтальном масштабировании

Решение: Managed OpenSearch

Простая настройка, большие возможности

Поддержка самые
различные конфигурации
кластеров

Hot-cold

Поддержка
внешних IDP



Опыт эксплуатации больших кластеров

Внутри Яндекс используют
кластеры OpenSearch
размером 100+ хостов



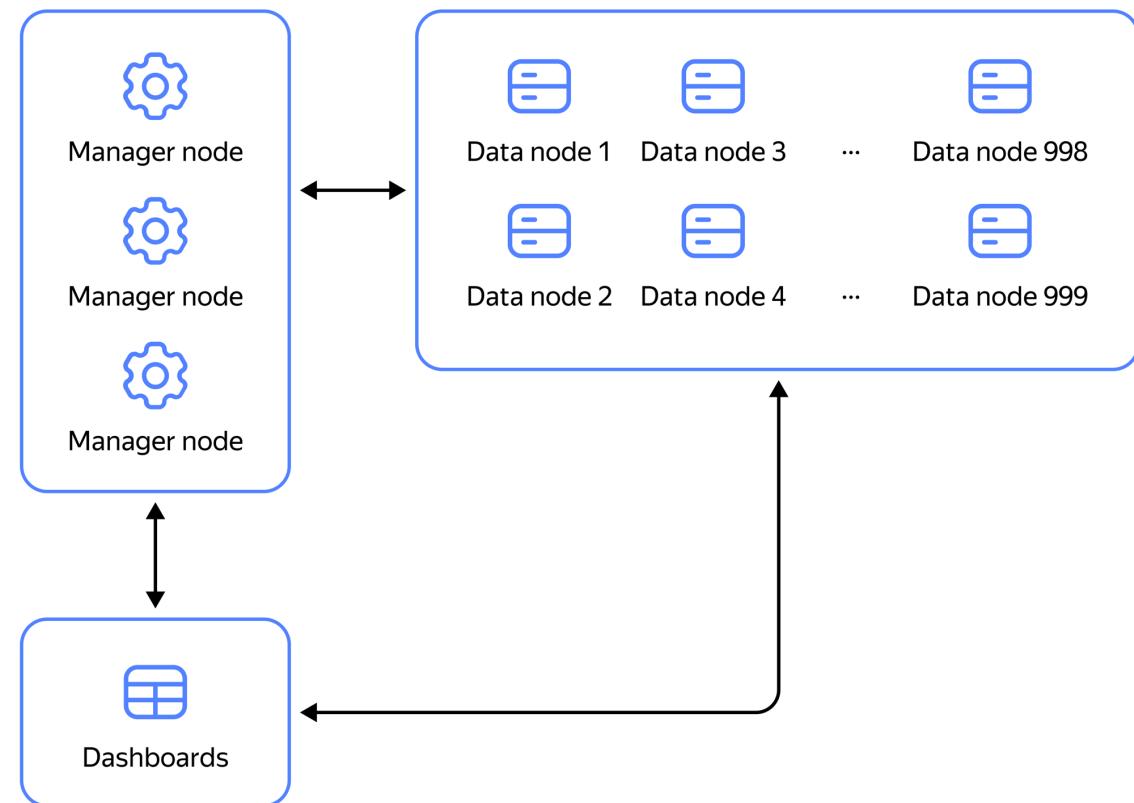
Команда разработки

Исправление багов
OpenSearch
FR: доработки
по запросам клиентов
Решение проблем
пользователей

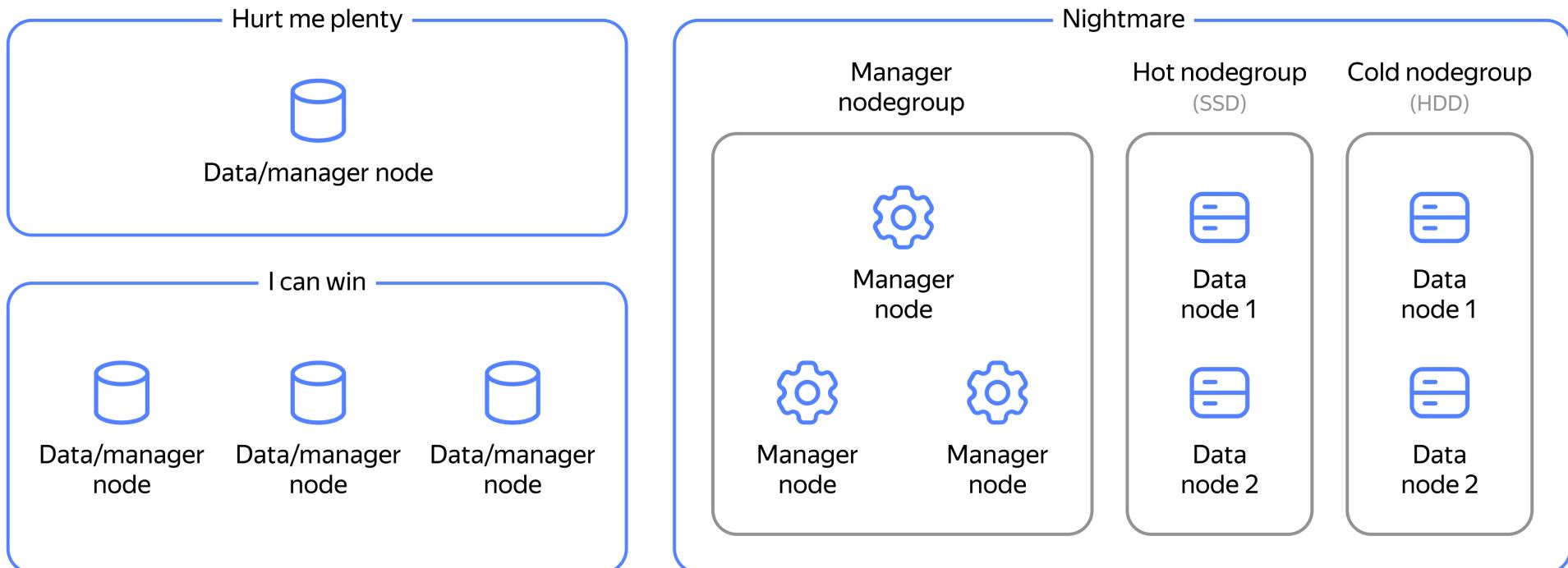


Компоненты кластера OpenSearch

- Manager nodes — координируют работу кластера, обеспечивают отказоустойчивость
- Data nodes — обрабатывают и хранят данные
- Dashboards — Web-интерфейс
- Для удешевления кластера, Manager nodes и Data nodes могут быть совмещены



Возможности Managed OpenSearch

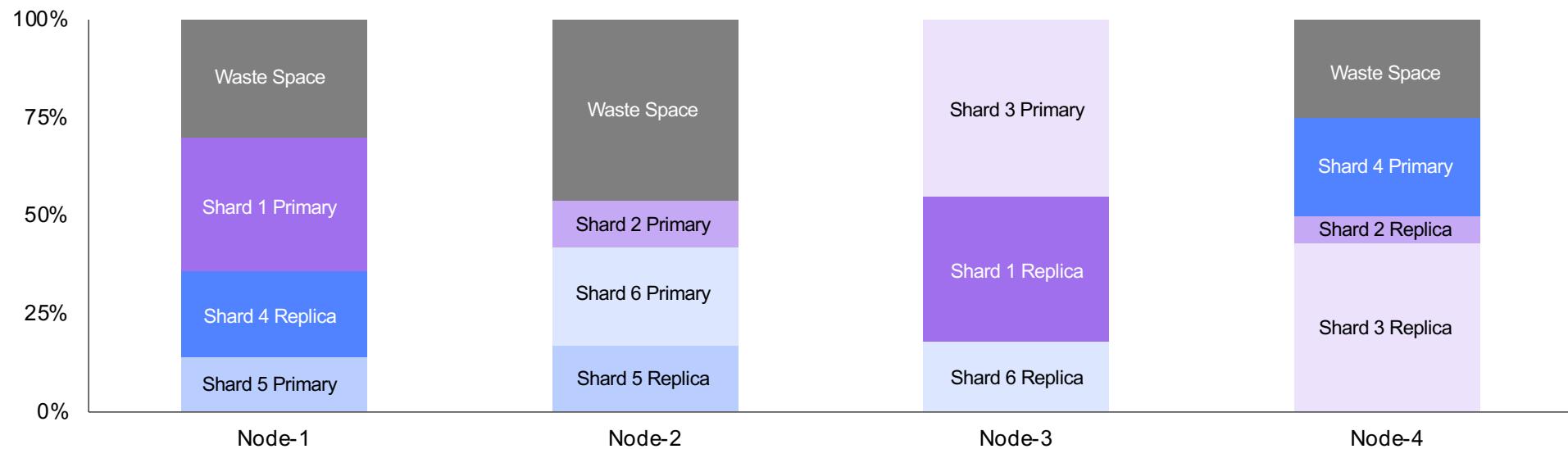


Балансировка: узкое место масштабирования

Стандартный алгоритм балансировки учитывает только количество шардов на ноде, игнорируя размер данных

Неэффективно работает с primary-шардами

Занятое пространство

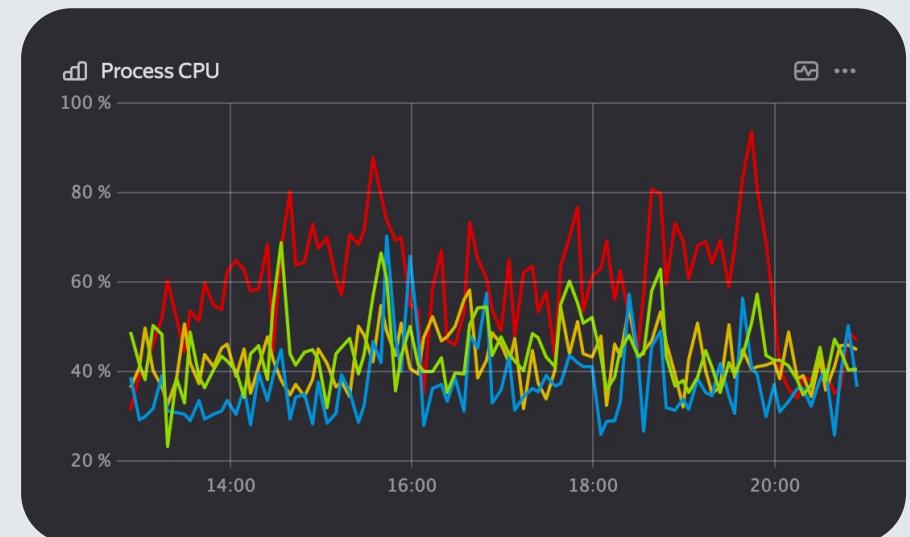
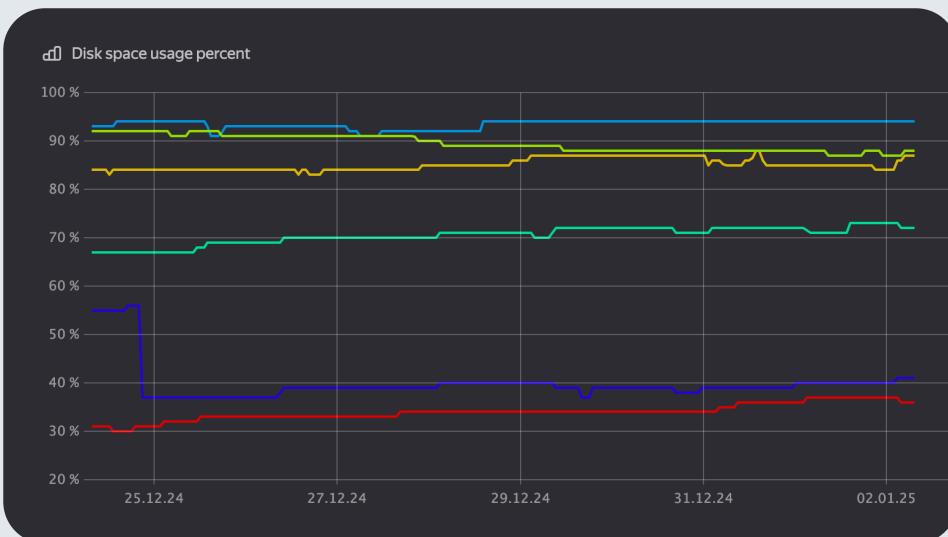


Балансировка: узкое место масштабирования

Невозможность полностью утилизировать дисковое пространство

Нестабильность кластера на больших объемах данных

Перекос нагрузки между нодами



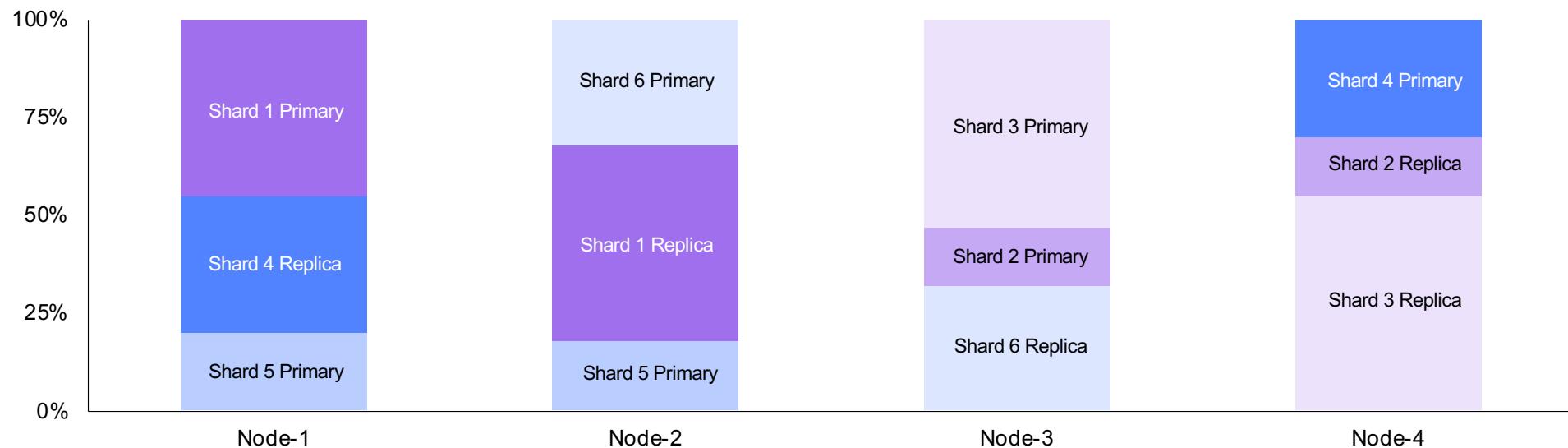
Решение: новый алгоритм балансировки

Революция в балансировке данных

Разработан новый алгоритм на основе
покоординатного спуска с хитрой
метрикой расстояния

Подробности на Data Internals 2025

Занятое пространство



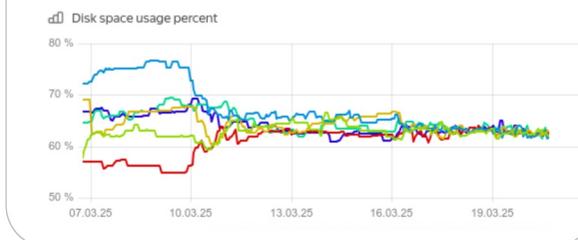
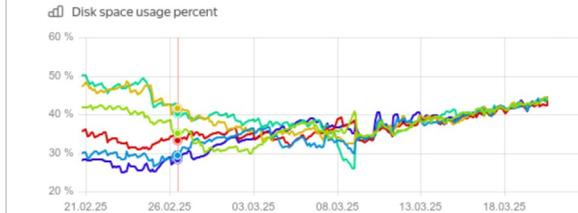
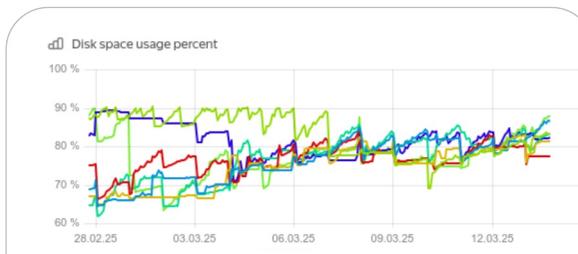
Новый алгоритм балансировки

Учитывает размеры шардов

Стандартная балансировка



Новая балансировка



Новый алгоритм балансировки

Добавлена балансировка переключением Primary/Replica

Время балансировки Primary уменьшено с часов до минут



Новый алгоритм балансировки

Преимущества

Стабильность
и надёжность



Более
эффективное
использование
ресурсов



Сокращение
времени
на maintenance-
операции



Больше данных — меньше затрат

Сжатие данных



Доработали
Index State Manager
и Lucene Codecs

- Возможность перепаковки индекса с более высокой степенью сжатия
- Новый кодек сжатия на основе LZMA (до 30% экономии места)
- Идеально для hot-cold-архитектуры

Ускорение индексации данных

Изменения в алгоритме записи Translog для улучшения производительности на небольших батчах



Embedded
Checkpoints

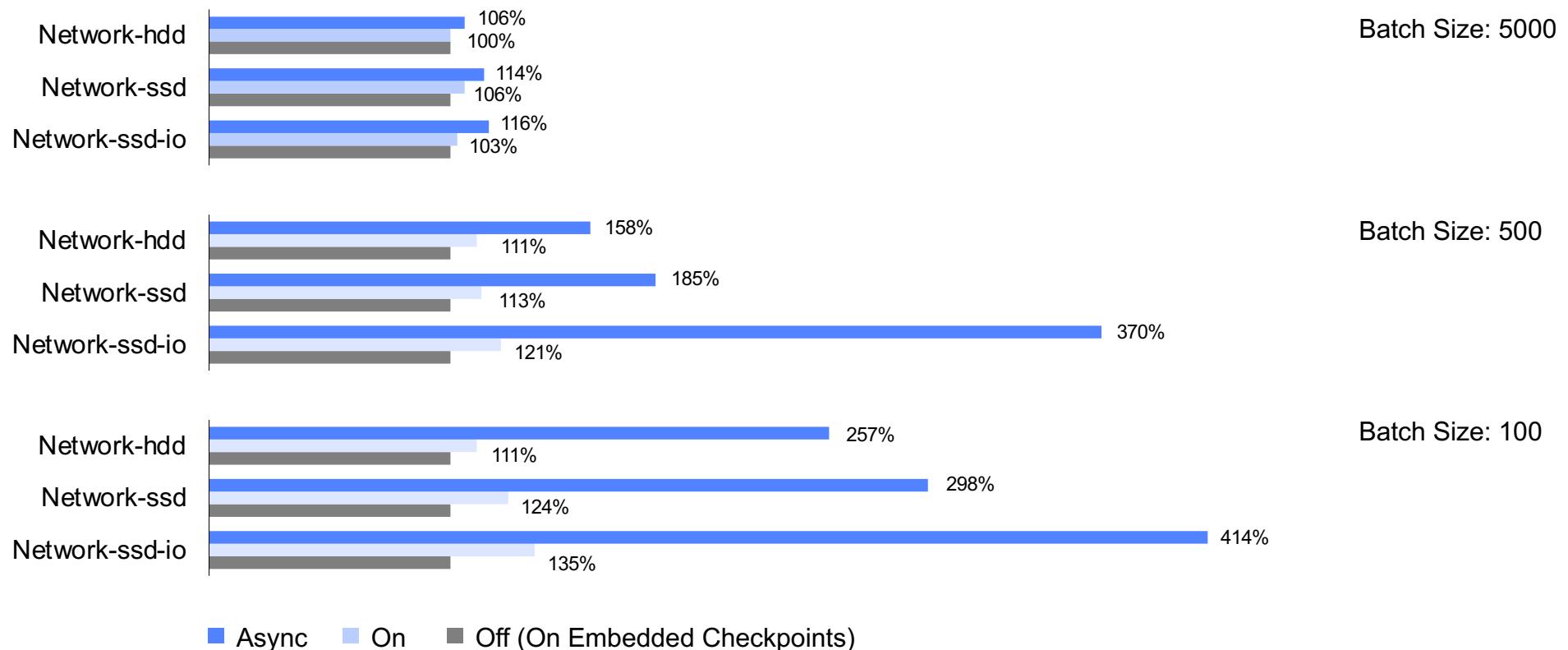


Async
Checkpoints Write

Особенно заметно на медленных дисках (HDD)

Ускорение индексации данных

Asynchronous Embedded Checkpoints. Ускорение индексации

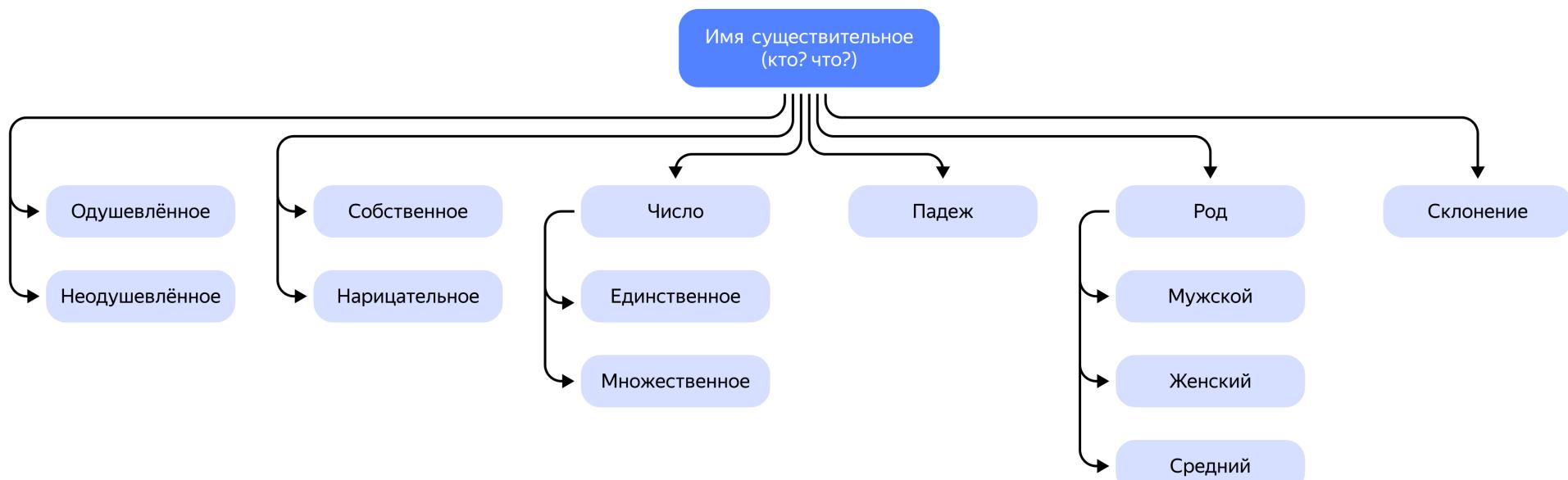


Поддержка морфологии

Качество поиска

[FR] Поддержка загружаемых словарей синонимов

[FR] Нативная поддержка славянских языков (+ казахский, татарский, турецкий)



Спасибо за внимание!



Владислав Таболин
Руководитель службы
Специальные СУБД Yandex Cloud

