

# RDMA over InfiniBand

Павел Бегунков  
*Научный руководитель:* Кирилл Коротаев

Санкт-Петербургский  
национальный исследовательский  
Академический университет  
Российской академии наук

Санкт-Петербург, 2017

- 1 Постановка задачи
- 2 InfiniBand
- 3 Итоги

Acronis Storage – универсальное безопасное эффективное надежное программно-определенное распределенное хранилище данных, которое позволяет управлять блоками данных, файлами и объектами.

- кросс-платформенность
- производительность
- масштабируемость
- репликация и коррекция ошибок



## Цель

Реализация поддержки RDMA over InfiniBand в хранилище Acronis

## Типичная операция чтения (read)



## Оптимизированная операция чтения (zero-copy read)

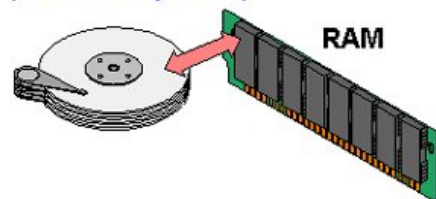


Прямой доступ к памяти (англ. direct memory access, DMA) – режим обмена данными между устройствами компьютера или между устройством и основной памятью (RAM) без участия центрального процессора.

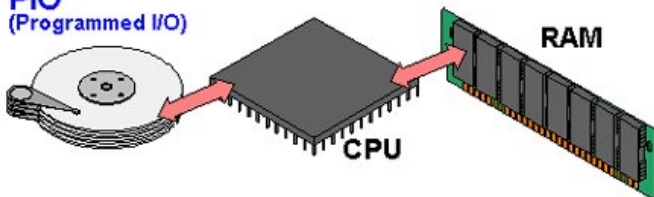
[https://en.wikipedia.org/wiki/Direct\\_memory\\_access](https://en.wikipedia.org/wiki/Direct_memory_access)

RDMA, (remote DMA) – DMA, позволяющий использовать напрямую память удаленного компьютера.

## DMA (Direct Memory Access)



## PIO (Programmed I/O)



InfiniBand – механизм/стандарт для высокопроизводительного сообщения между компьютерами. (узлами кластера). Поддерживает RDMA. Включает в себя сетевые карты, роутеры, драйвера, библиотеки и др.

## Программные интерфейсы

- ib verbs
- rdma cm (connection manager)
  
- RDMA, процессор не участвует в передаче
- zero-copy
- небольшая задержка операций (low-latency)
- уменьшения числа обращений к операционной системе



## Цель

Реализация поддержки RDMA over InfiniBand в хранилище Acronis

Использование RDMA для передачи данных между узлами кластера.

## Результат

- Изучена технология (infiniband verbs)
- Переработаны существовавшие обертки
- Реализован встроено динамическое распознавание поддержки InfiniBand и установка подключения между серверами.
- Начата работа над поддержкой передачей данных.

## Дальнейшая работа

- Реализация полной поддержки InfiniBand
- Тестирование на кластере с InfiniBand (используются эмуляторы)
- Улучшение производительности