



## Почему нужно управлять ресурсами серверов БД?

Приоритезация пользователей и их операций для  
гарантированного времени выполнения критичных  
операций



## ПОЧЕМУ НУЖНО УПРАВЛЯТЬ РЕСУРСАМИ СЕРВЕРА БД

Средняя нагрузка на сервер СУБД редко превышает **40-50%**

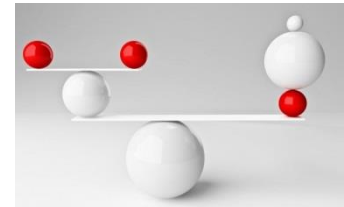


Аппаратные ресурсы мощные, а пользователи все равно не довольны

Пиковые нагрузки и время отклика пользовательских операций не укладываются в необходимые интервалы



Для СУБД все операции по умолчанию равнозначны с точки зрения выделения аппаратных ресурсов





## ЦЕЛИ И ЗАДАЧИ

### ЦЕЛЬ:

Обеспечить гибкое распределение ресурсов серверов БД информационной системы в моменты пиковых нагрузок для приоритетных информационных потоков



### ЗАДАЧИ:

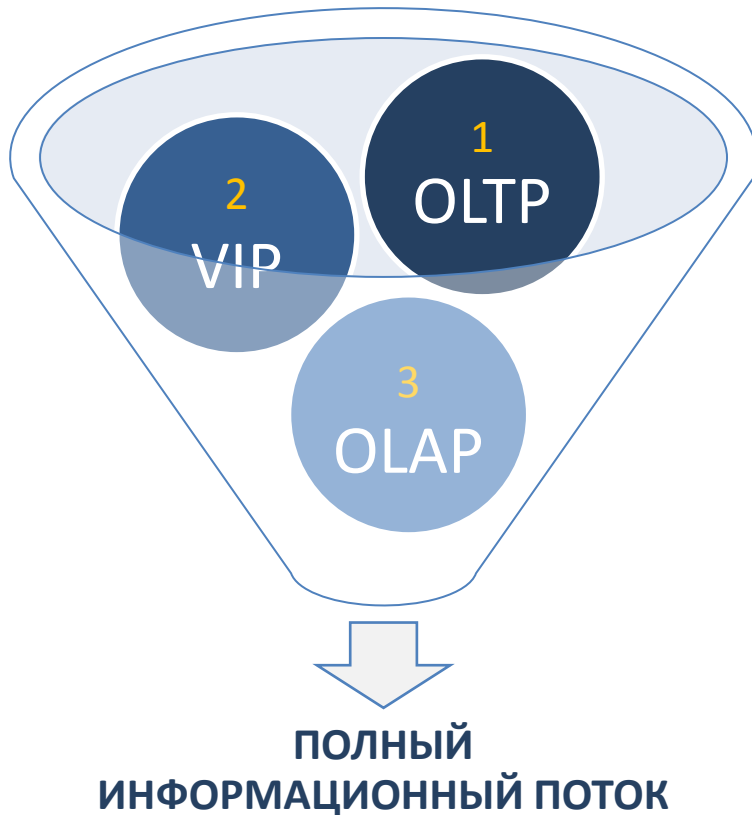
- Разделить общий информационный поток ИС на отдельные составляющие и приоритезировать их
- Распределить ресурсы серверов БД между потоками в соответствии с приоритетами





## НОВЫЙ ПОДХОД К УПРАВЛЕНИЮ РЕСУРСАМИ СЕРВЕРА

Разделение информационного потока к ИС на отдельные составляющие и установка им приоритетов



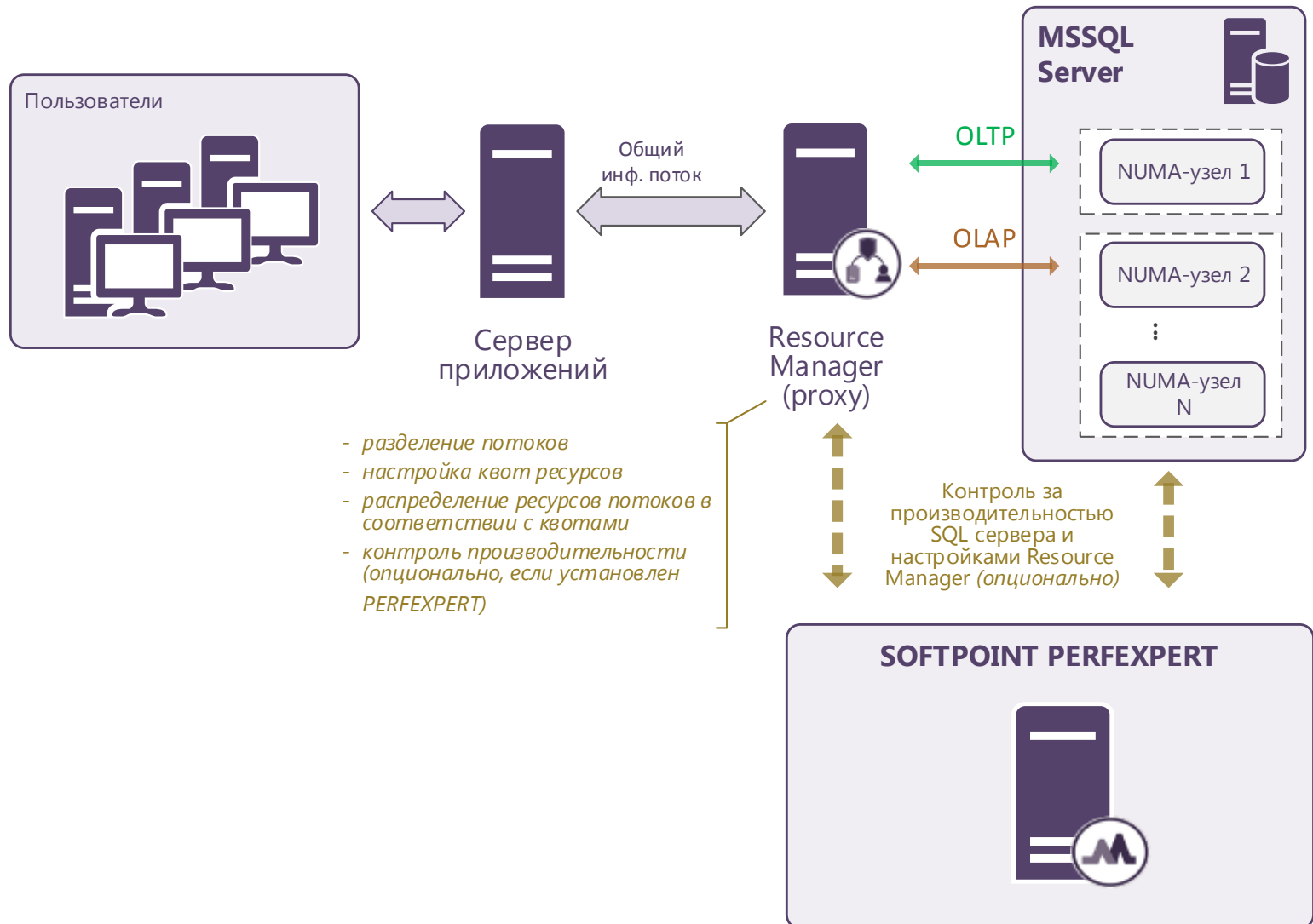
OLTP - транзакционные потоки  
*обработка транзакций в реальном времени*

VIP – информационные потоки,  
имеющие специализированные  
маркеры

OLAP – аналитические потоки в  
реальном времени и все остальные  
ПОТОКИ



## АРХИТЕКТУРА





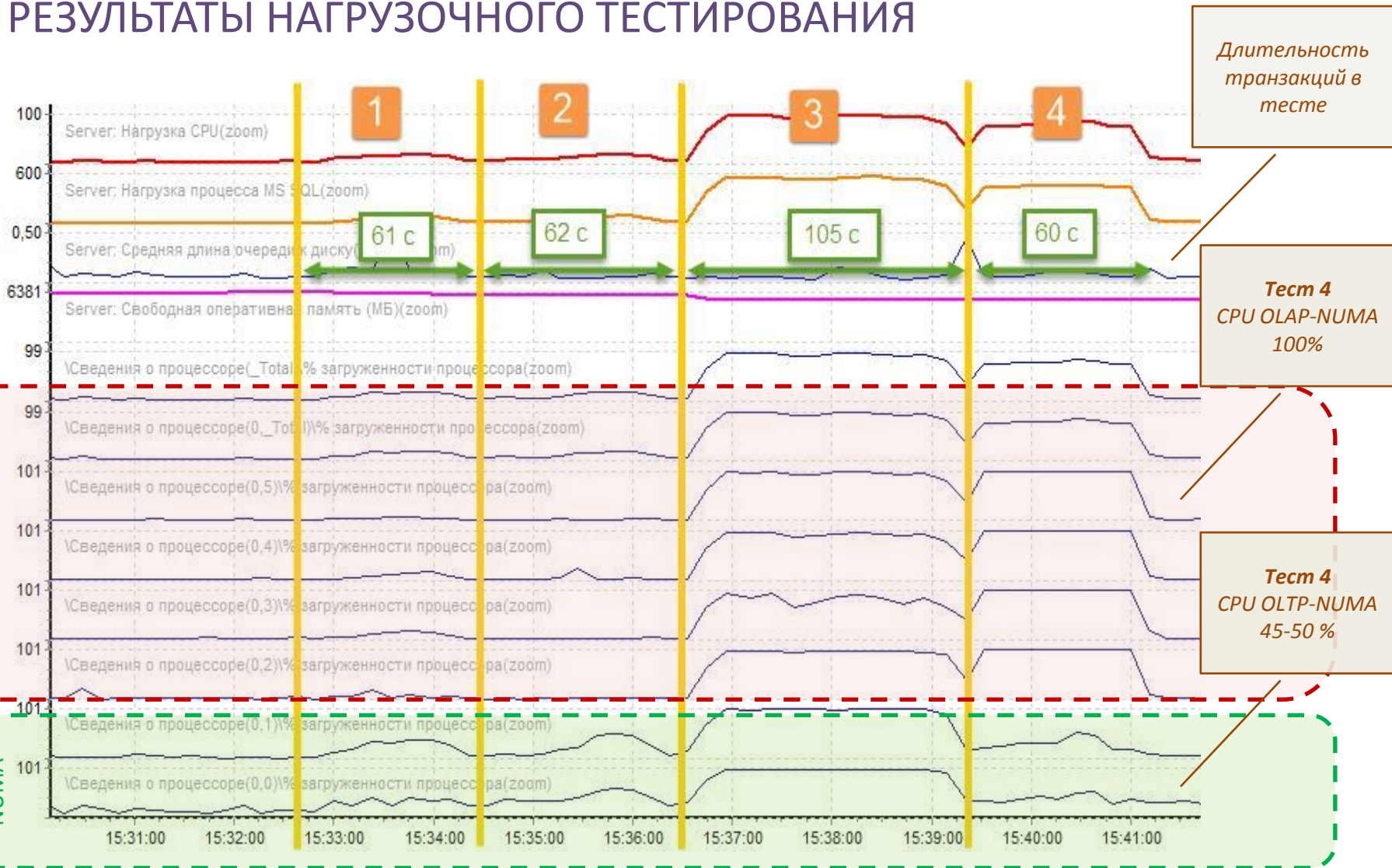
## ПРИМЕР РАСПРЕДЕЛЕНИЯ НАГРУЗКИ

Два потока в типичной ИС: OLTP и OLAP

- **Тест 1:** нагружаем ИС только OLTP-нагрузкой;
- **Тест 2:** нагружаем ИС только OLTP-нагрузкой, но выделяем ей один NUMA-узел (2 ядра CPU);
- **Тест 3:** нагружаем ИС OLTP+OLAP нагрузкой как есть без распределения потоков;
- **Тест 4:** нагружаем ИС OLTP+OLAP нагрузкой, но распределяем потоки по NUMA-узлам



# РЕЗУЛЬТАТЫ НАГРУЗОЧНОГО ТЕСТИРОВАНИЯ





RESOURCE MANAGER



## ЭФФЕКТ ПРИМЕНЕНИЯ

Гарантированная доступность системы для приоритетных пользователей и критичных бизнес-операций, даже в период пиковых нагрузок

Удобная и быстрая установка приоритетов между пользователями и их операциями для оптимального распределения серверных ресурсов;

Защита рисков недоступности системы для выделенной группы пользователей даже в условиях непродуманной и несбалансированной загрузки.





Спасибо за внимание!