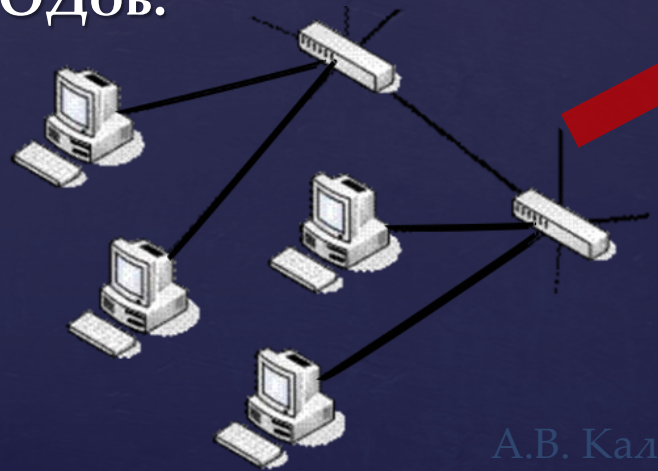
Программное управления и виртуализация - новые горизонты компьютерных сетей

член-корр. РАН Смелянский Р.Л.

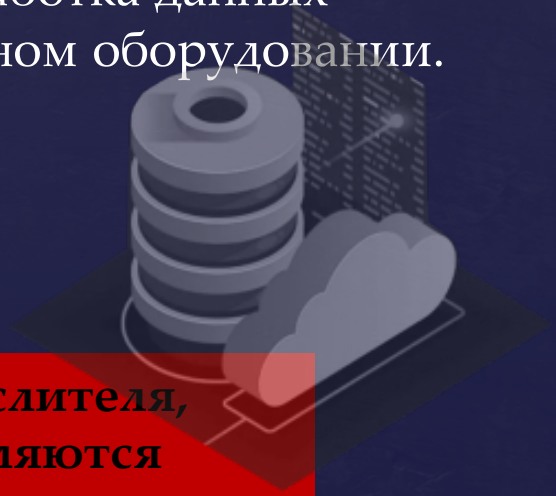
Эволюция организации вычислительной инфраструктуры

- конец 60-х - Вычислитель с пакетной организацией вычислений;
- 70-е - Вычислительный центр с mainframe и терминальной сетью;
- 80-е - Клиент-серверная инфраструктура с сетевым доступом;
- 90-е – Серверные фермы с Frontend сервером и локальной сетью доступа;
- 2000-е - ЦОД с высокоскоростной сетью;
- н/в - Сеть мини-ЦОДов.

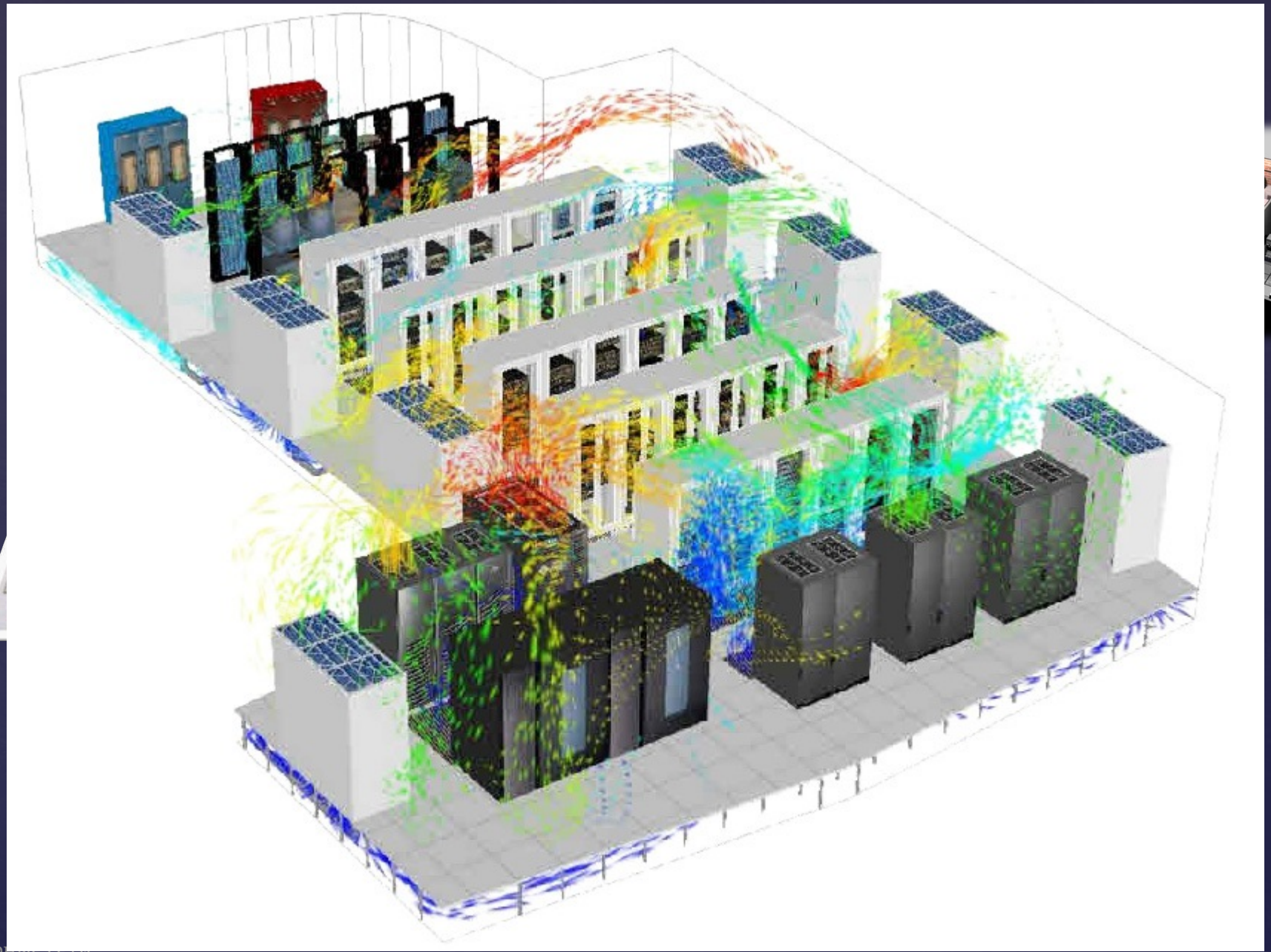


Потребности современного приложения:

- **Распределенность** – приложение – совокупность взаимодействующих функций/сервисов, которые работают параллельно на нескольких вычислителях, интегрируя и агрегируя распределенные данные;
- **Serverless** – программист/пользователь не обременен проблемами размещения, начальной загрузки и конфигурирования ресурсов и компонентов приложения;
- **Эластичность** – производительность приложения можно изменять без прерывания его работы;
- **Реальное время** – чувствительность к задержкам и времени отклика;
- **Кросс - платформенность** – приложение не зависит от программного и аппаратного окружения;
- **Взаимодействие и Синхронизация** — объединение результатов разных этапов вычислений вне зависимости от места их проведения, агрегация цепочек сервисов;
- **Обновления** – обновление приложения/сервисов не должно затрагивать пользователей;
- **Доступность ресурсов** – все виды ресурсов доступны всегда, включая данные и код, при любых отказах физической инфраструктуры
- **Ресурсы терминала пользователя не используются** - вся обработка данных производится на удаленном оборудовании.



Основной движущей силой развития архитектуры вычислителя, операционной системы, средств программирования являются потребности приложений!



[data centers design & consulting], Москва, 2013 г.

ЦЕЛЕВАЯ АУДИТОРИЯ ЦОД ПО TIER

ТИПЫ ПОТРЕБИТЕЛЕЙ УСЛУГ ЦОД И ИХ ТРЕБОВАНИЯ К КАЧЕСТВУ БИЗНЕС-ПРОЦЕССОВ

- организации, для которых не допустим любой перерыв в предоставлении ИТ-сервисов
- организации, успешность работы которых во многом зависит от непрерывности критически важных бизнес-процессов
- компании, допускающие временные перерывы в работе и предоставлении ИТ-сервисов
- небольшие «ремесленные» компании, работающие в off-line режиме

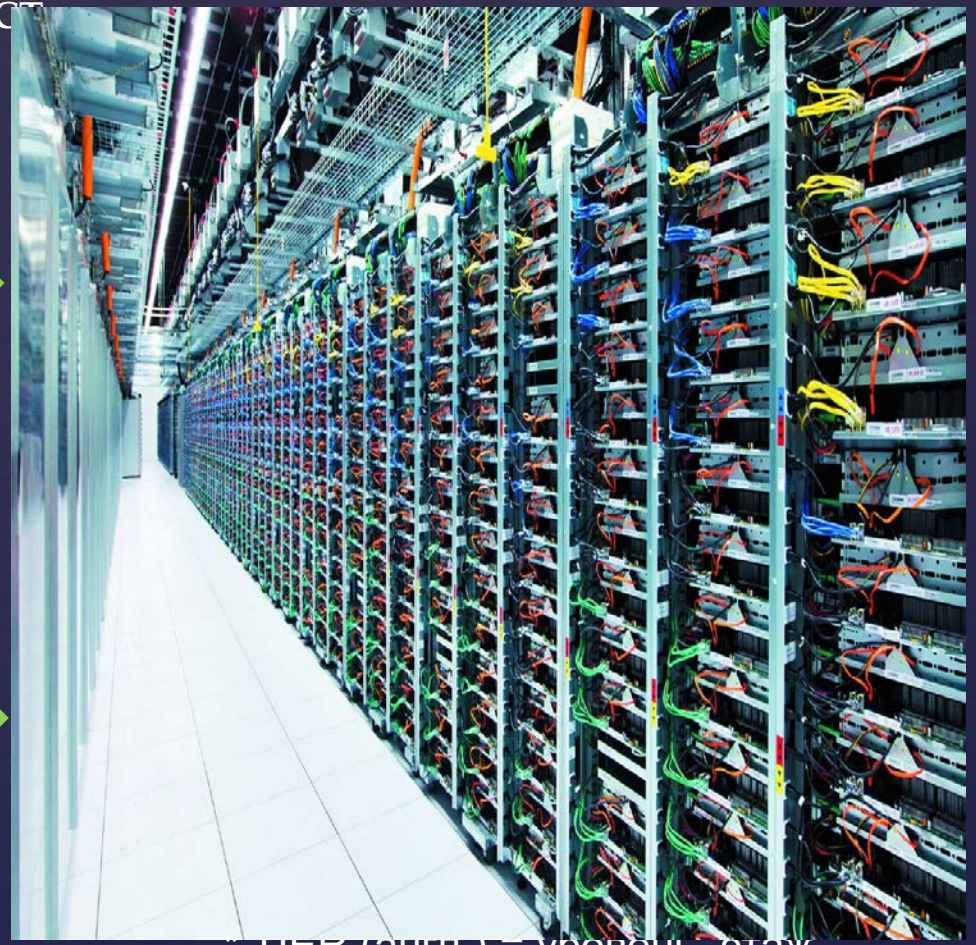
УРОВЕНЬ* НАДЁЖНОСТ И ЦОД

TIER IV

TIER III

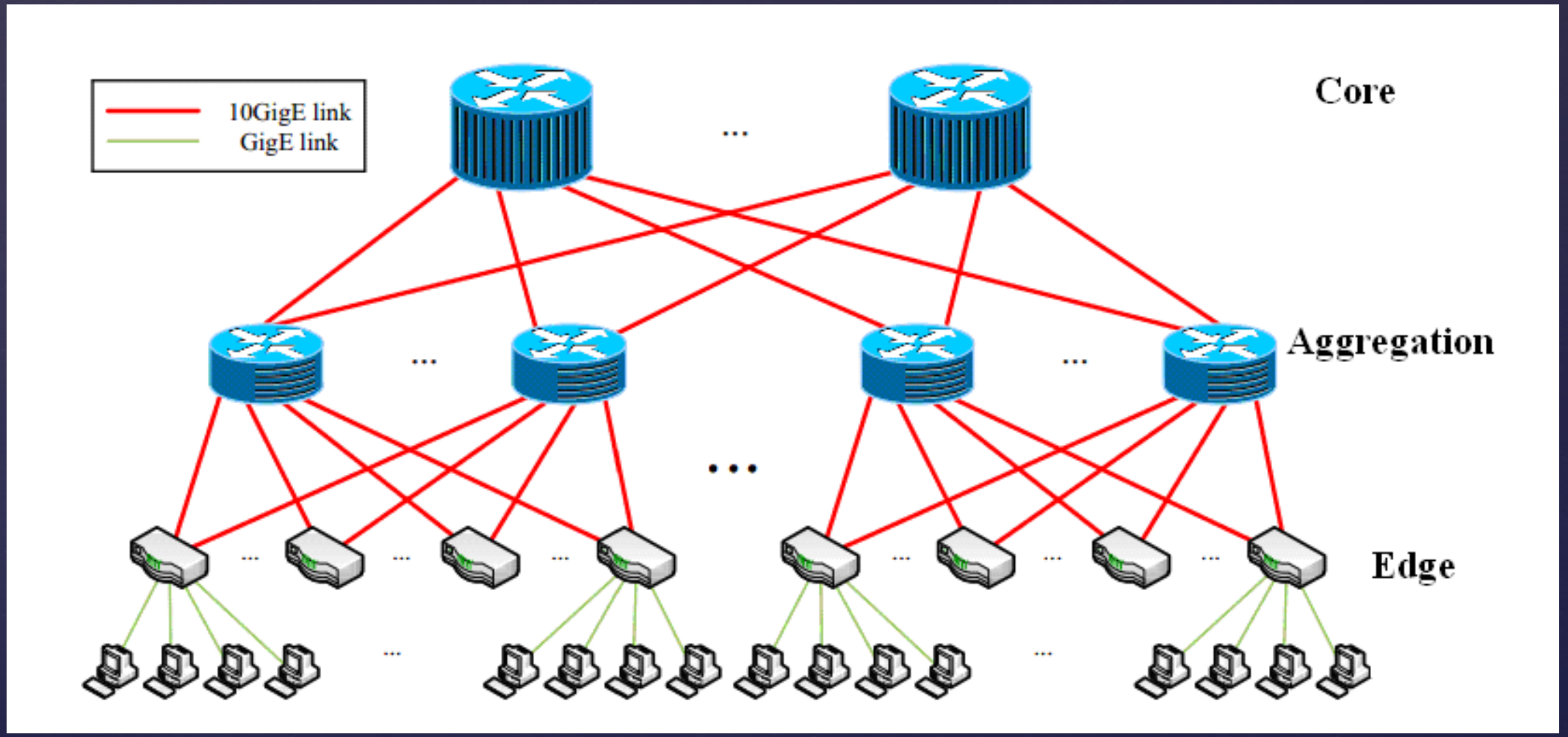
TIER II

TIER I



* TIER (англ.) = уровень, этаж

Типовая сетевая структура сети в ЦОД



Режимы работы облачной платформы

Облачные вычисления

Что такое облачные вычисления?

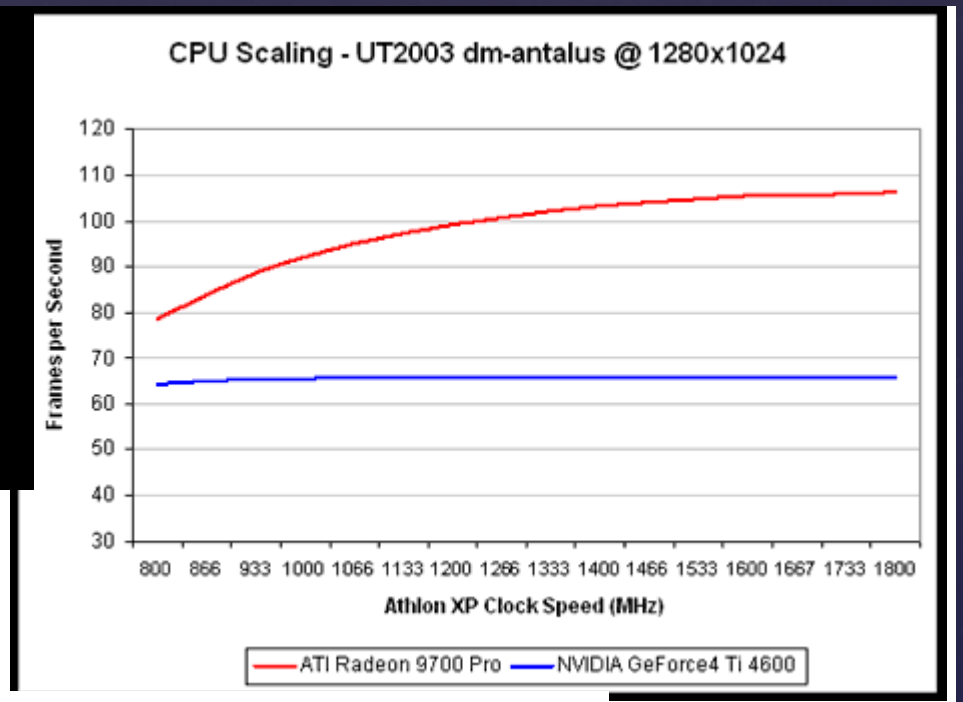
Облачная платформа – вычислительная инфраструктура, обеспечивающая виртуализацию и масштабирование сервисов.

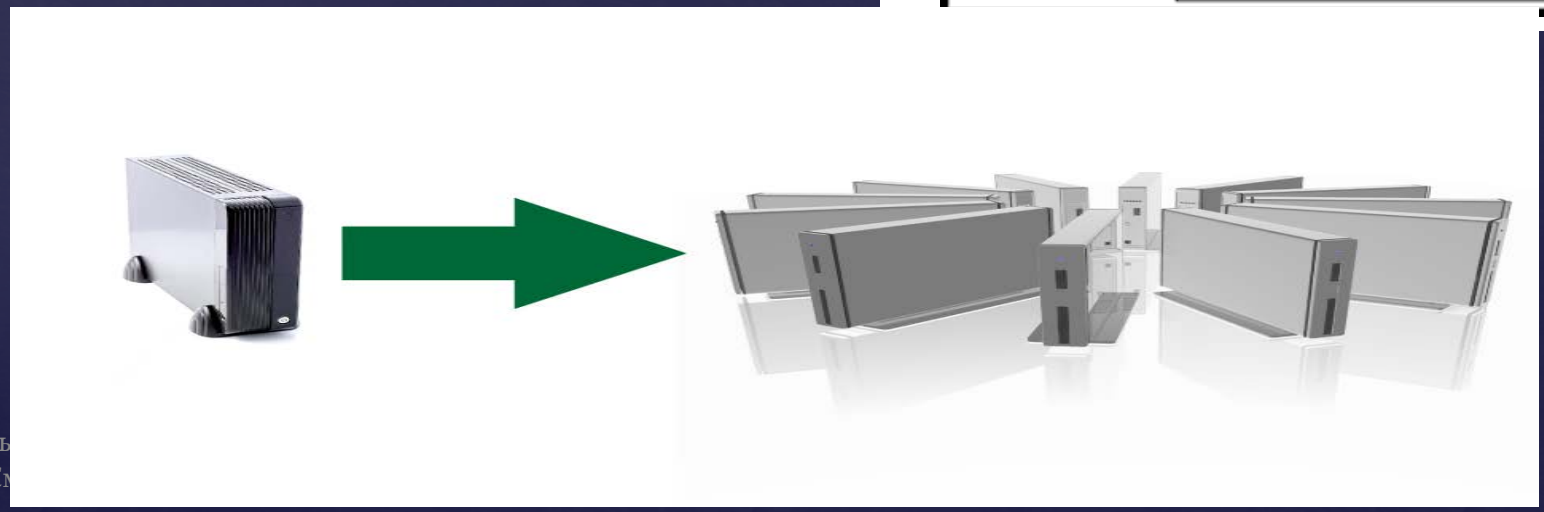
Что такое:

- **масштабирование**
- **виртуализация**
- **сервис**
- **serverless**

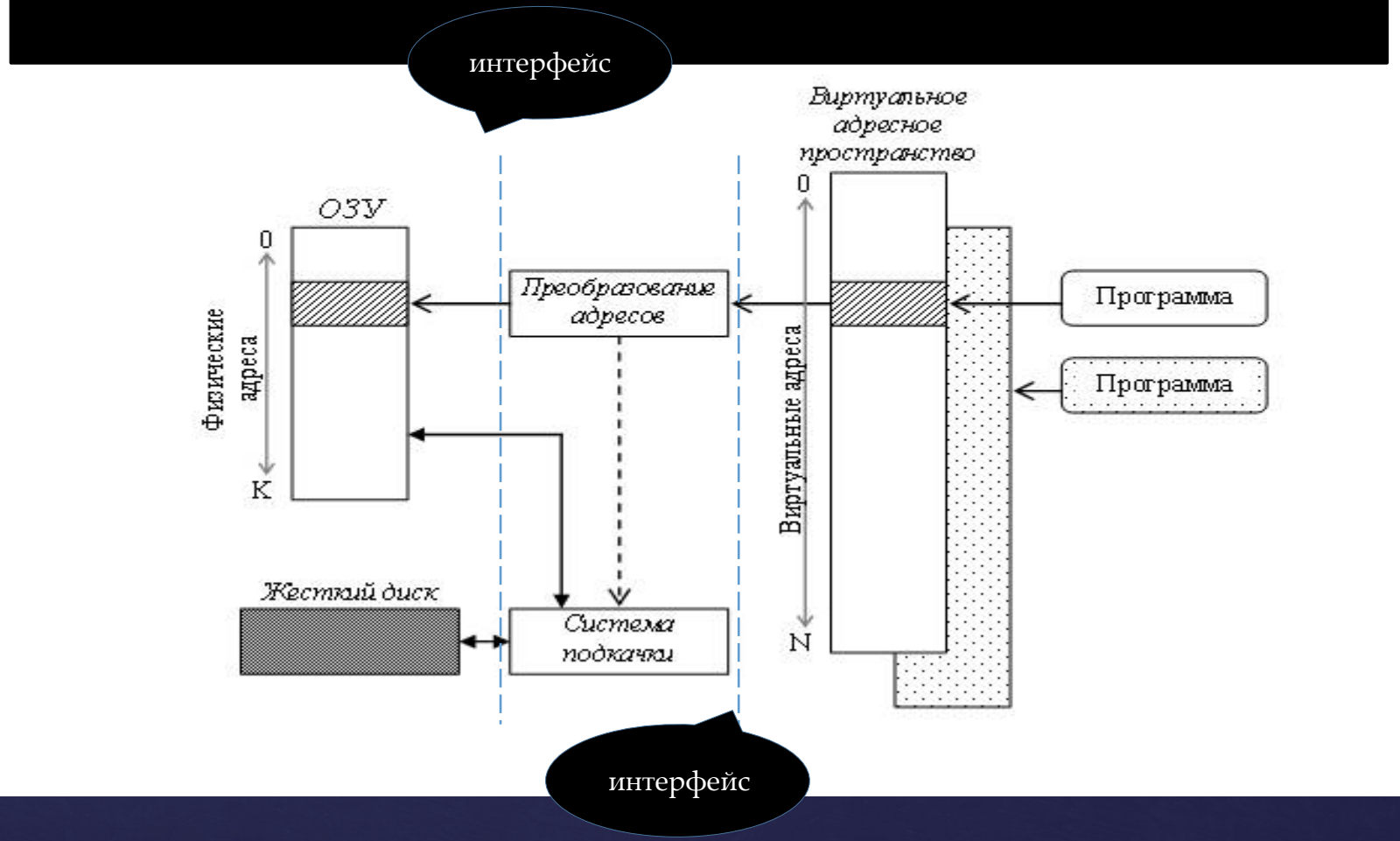
Масштабирование

Масштабирование
- вертикальное
- горизонтальное





Виртуальная память



Вычислительная система

Иерархия интерфейсов

App

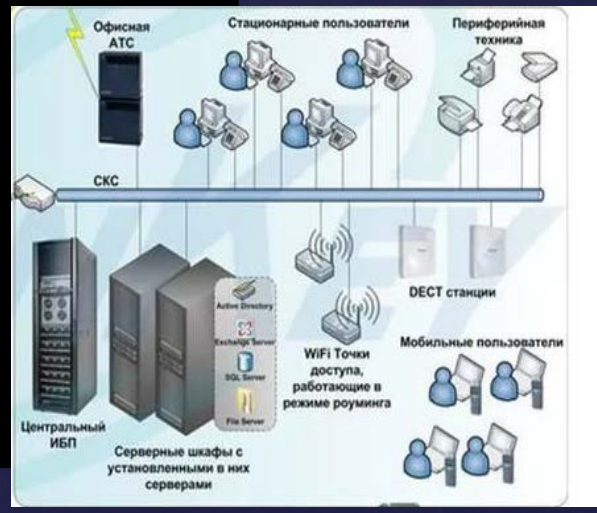
Интерфейс

OS

Взаимодействие приложения с внешними устройствами, памятью, процессором

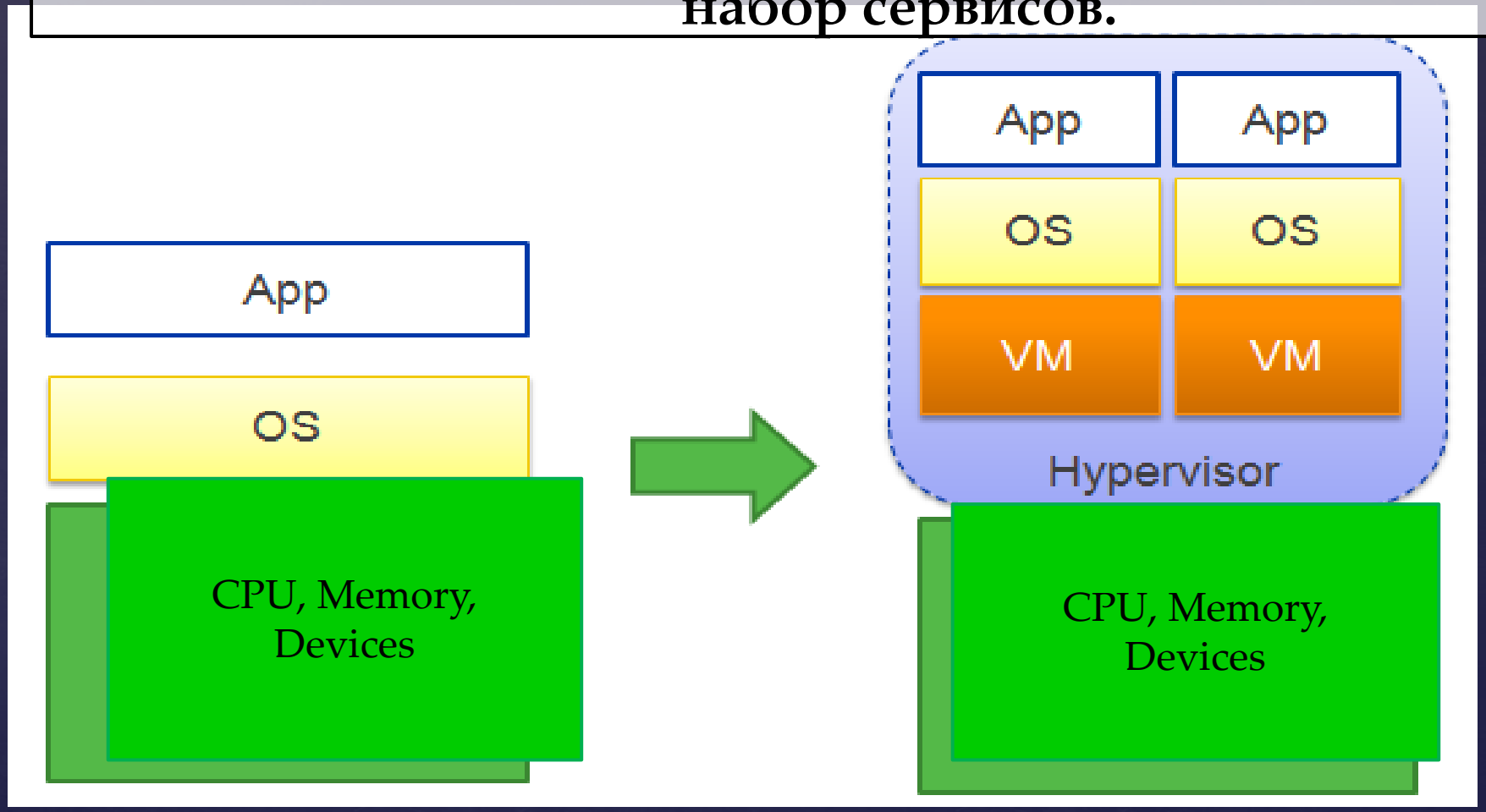
CPU, Memory, Devices

Интерфейс

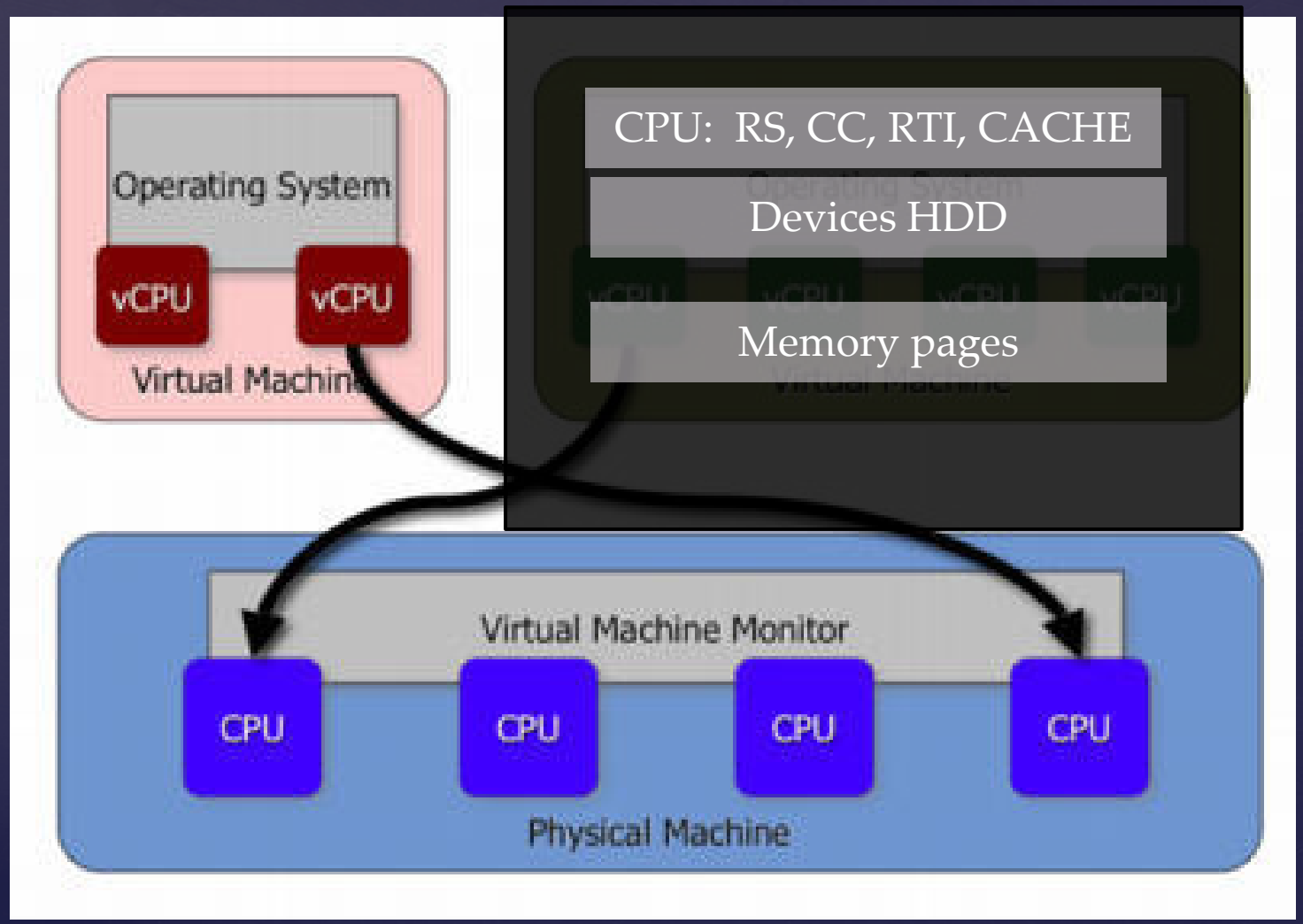


Виртуализация

Виртуализация – это динамическая подмена физической подсистемы с определенными сервисами, на объект, имеющий такой же интерфейс и реализующий тот же самый набор сервисов.



Виртуализация вычислителя

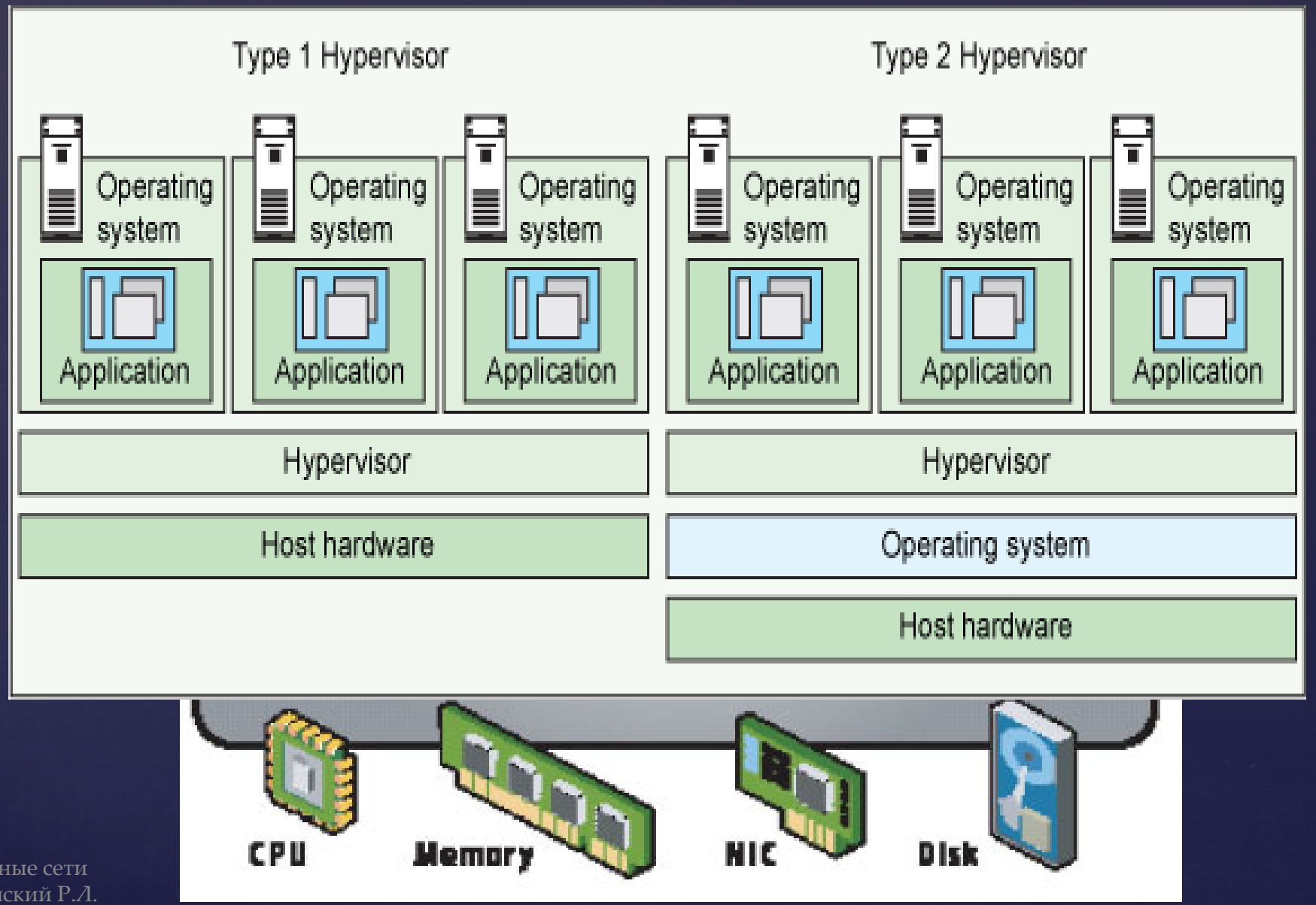


Виртуализация прямая

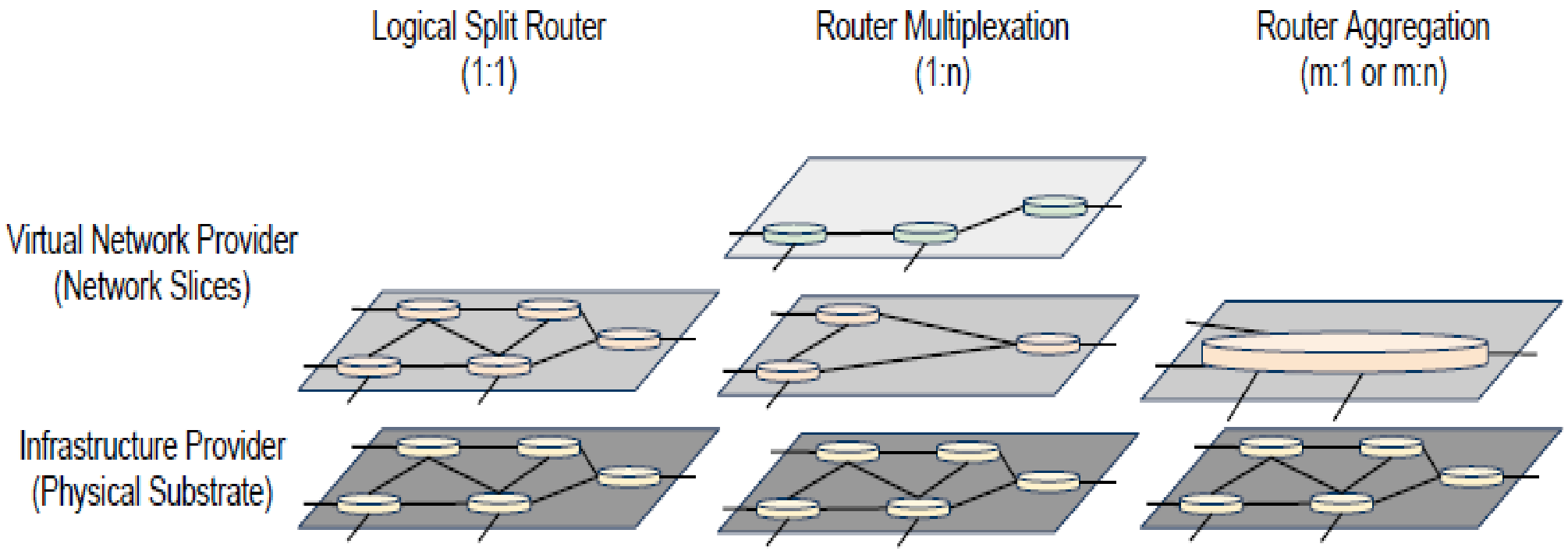
Паравиртуализация

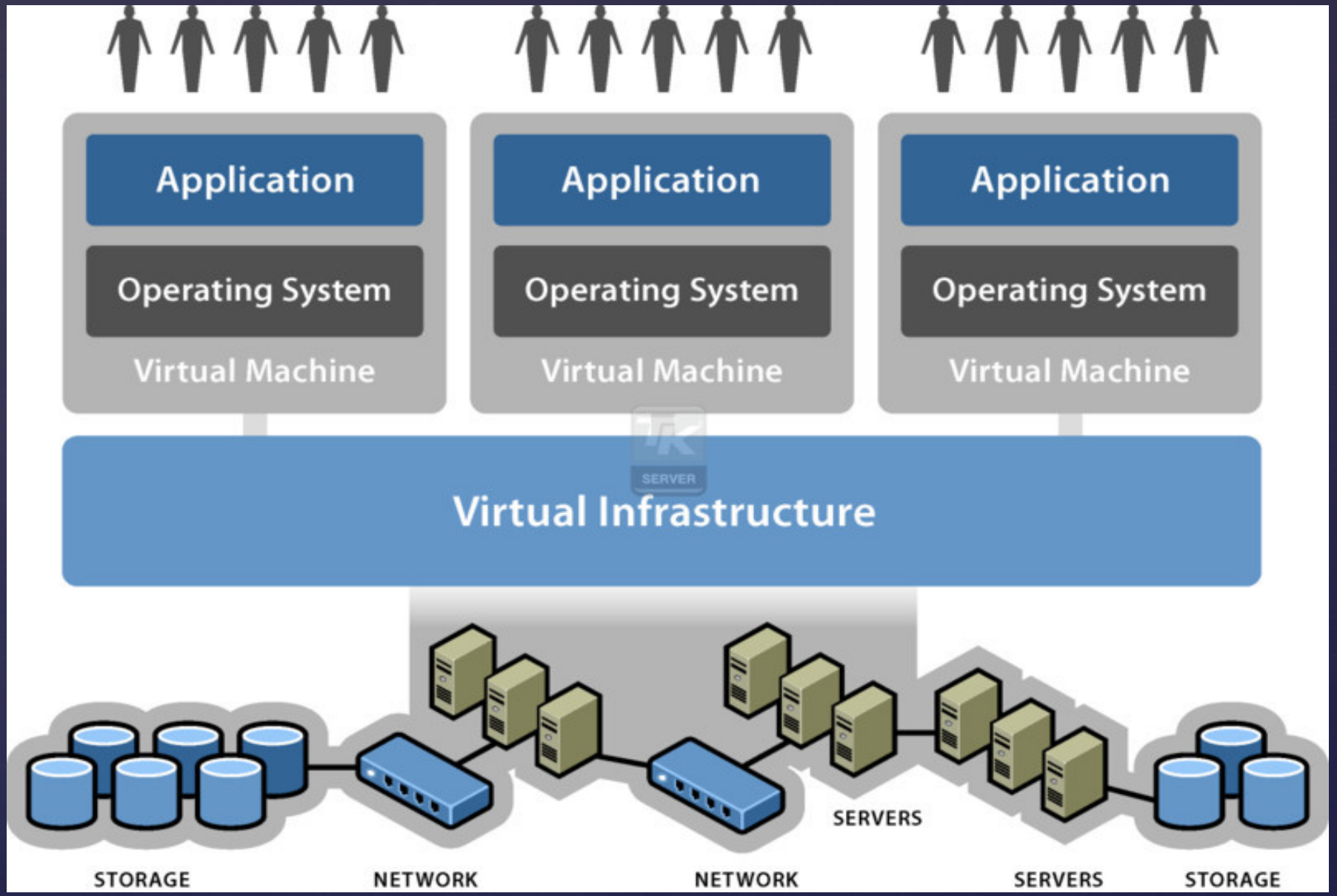
Виртуализация
с аппаратной поддержкой

Виртуализация

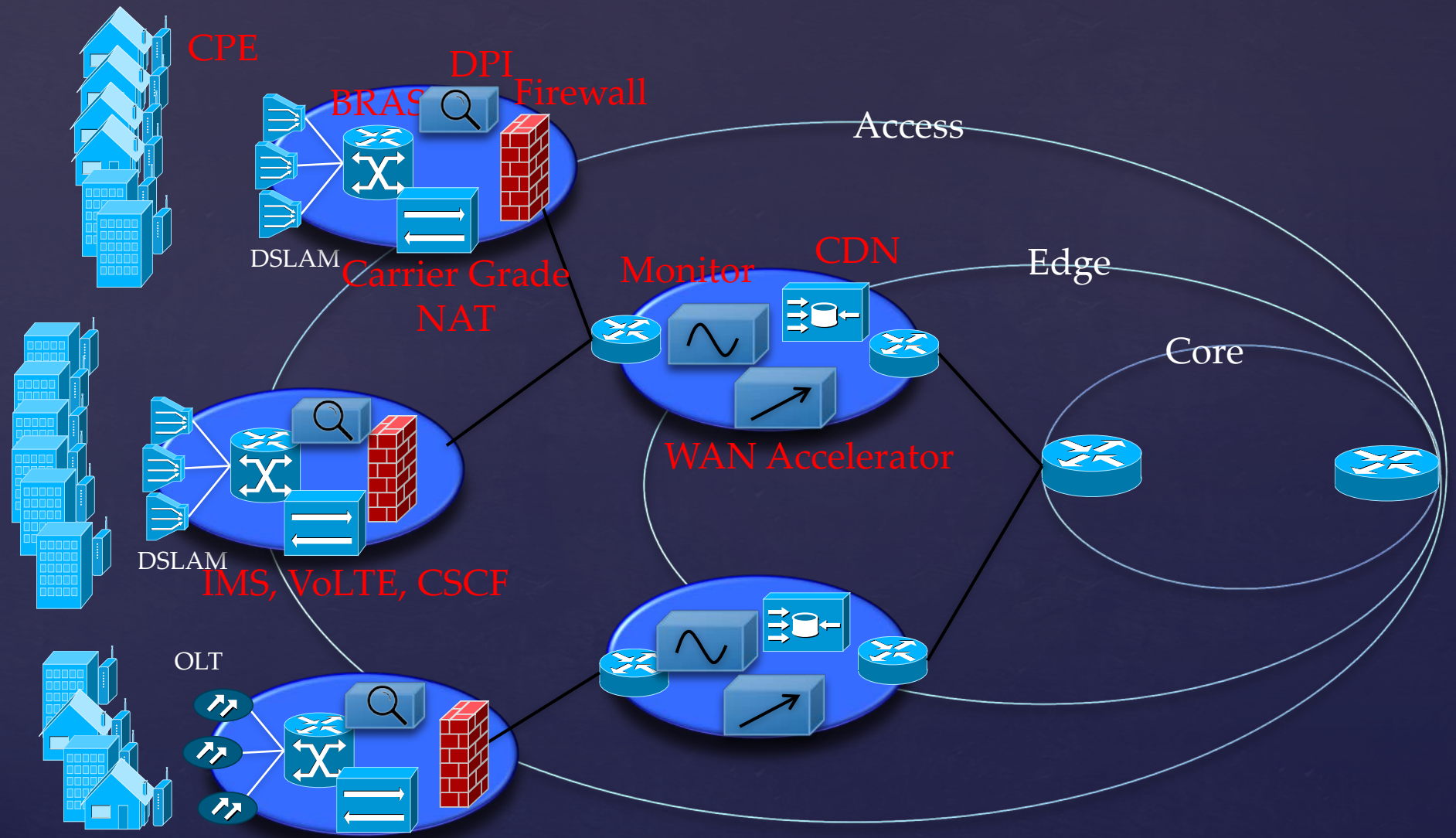


Виртуализация сети




Application
Operating System
Virtual Machine
Application
Operating System
Virtual Machine
Application
Operating System
Virtual Machine
Virtual Infrastructure
SERVERS
STORAGE
NETWORK
NETWORK
SERVERS
STORAGE

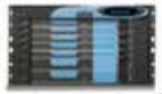
Сеть оператора с BCC



Виртуальные Сетевые Сервисы (NFV)


Message Router


CDN


Session Border Controller


WAN Acceleration


ewall


Carrier Grade NAT


Tester/QoE monitor


SGSN/GGSN

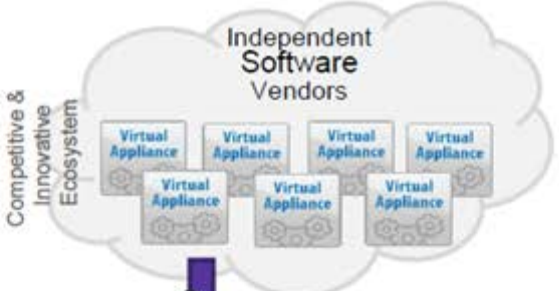

PE Router


BRAS


Radio Network Controller

Специализированное аппаратное обеспечение.
Требует физического размещения в каждом месте.
Невозможность быстрого развития и инноваций.

Традиционный подход к размещению сетевых функций.




Standard High Volume x86 Servers


Standard High Volume Storage


Standard High Volume Ethernet Switches

Это похоже на ...

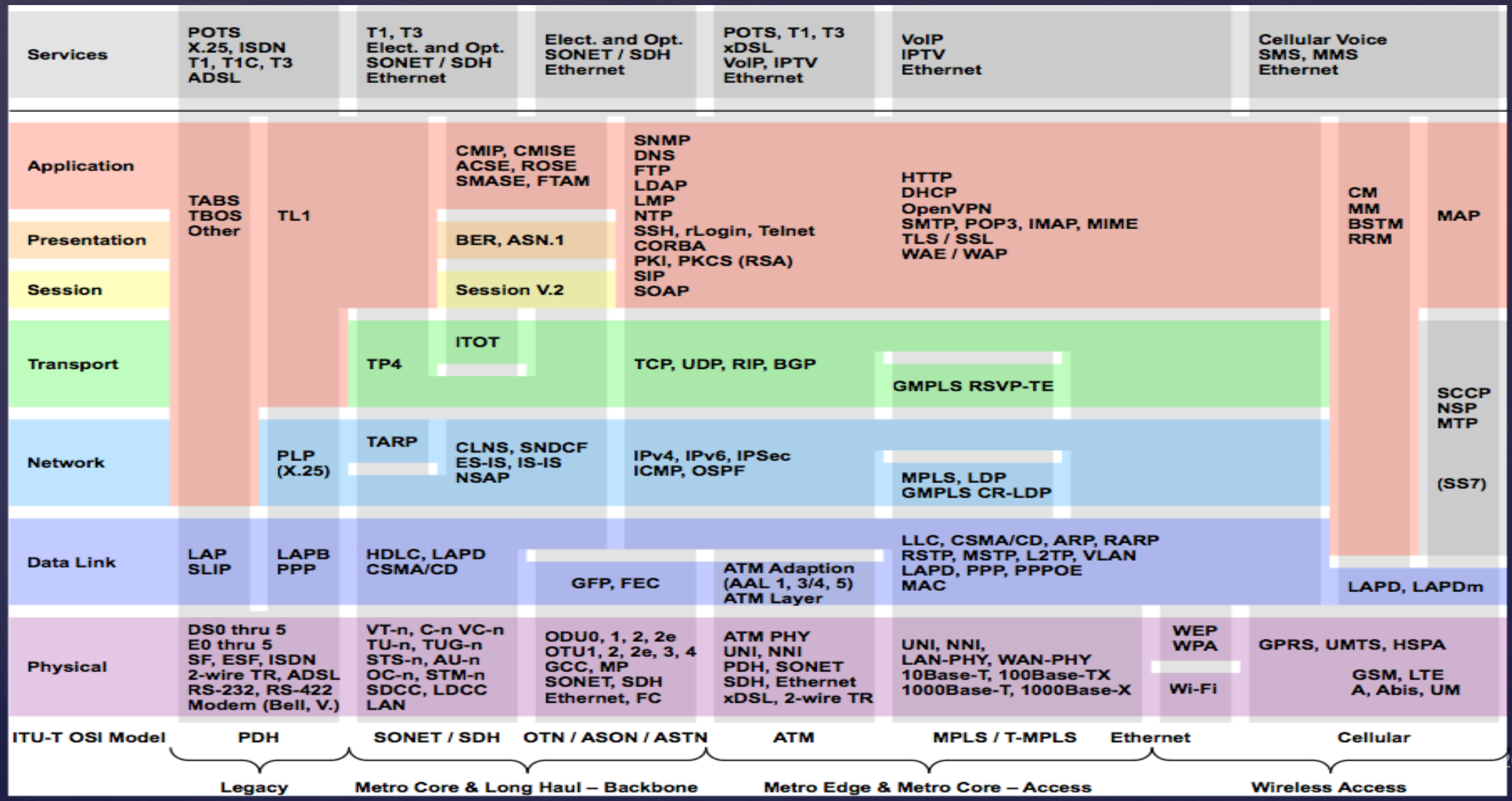

Калькулятор

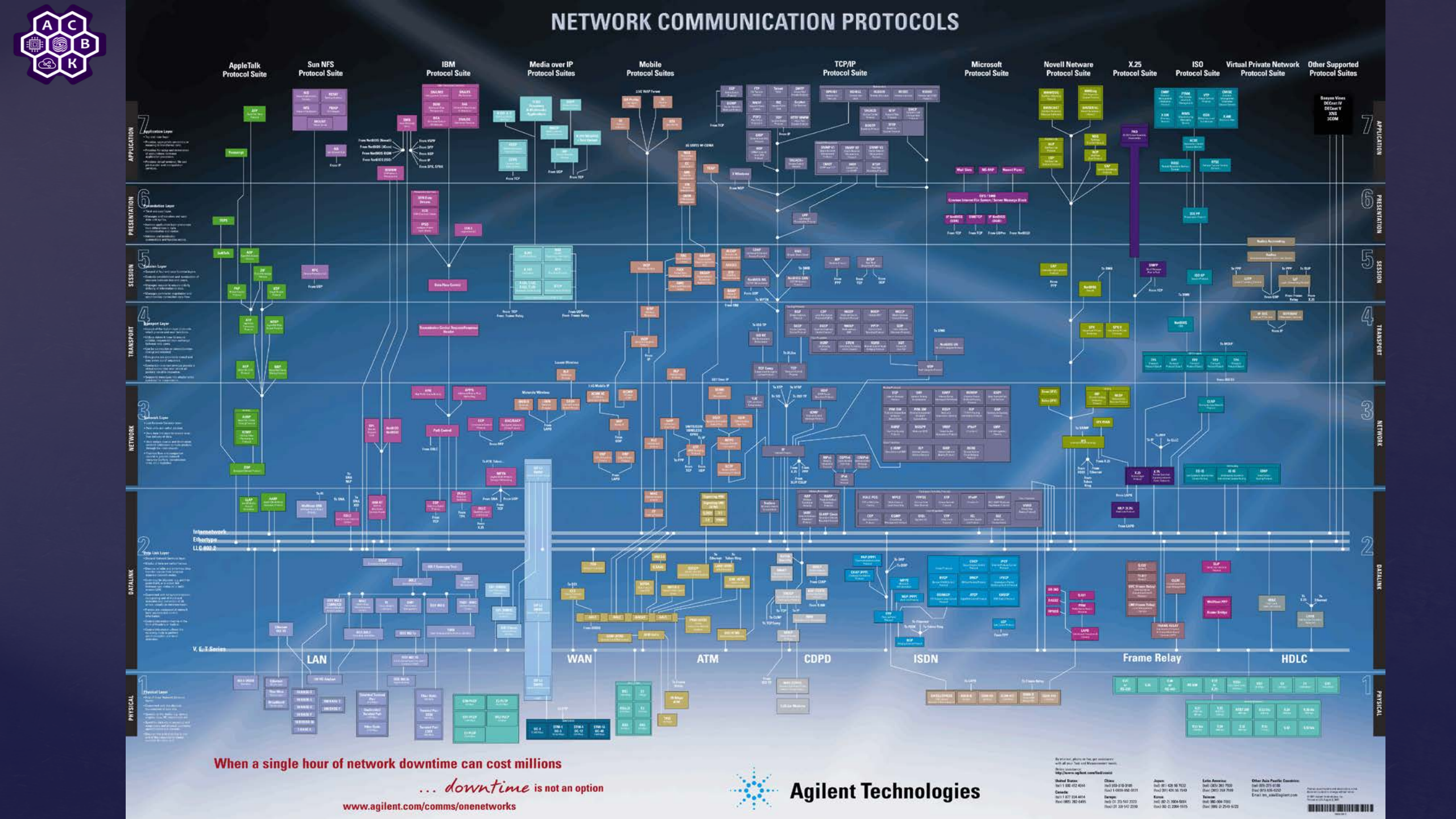

Пишущая машина


Печатный станок


Факс.аппарат

Модель стека протоколов не совершенна



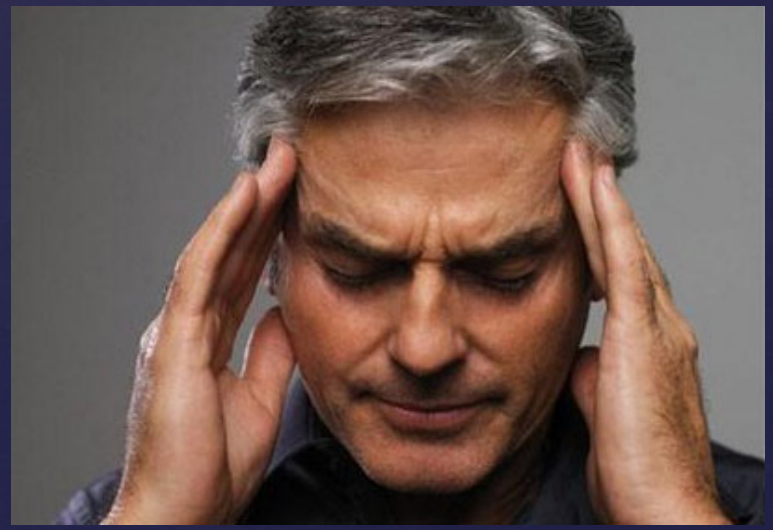

NETWORK COMMUNICATION PROTOCOLS
AppleTalk Protocol Suite
Sun NFS Protocol Suite
IBM Protocol Suite
Media over IP Protocol Suites
Mobile Protocol Suites
TCP/IP Protocol Suite
Microsoft Protocol Suite
Novell Netware Protocol Suite
X.25 Protocol Suite
ISO Protocol Suite
Virtual Private Network Protocol Suite
Other Supported Protocol Suites
APPLICATION
PRESENTATION
SESSION
TRANSPORT
NETWORK
DATALINK
PHYSICAL
Internetwork
Ethertype
LLC 802.2
V, E, T Series
LAN
WAN
ATM
CDPD
ISDN
Frame Relay
HDLC
When a single hour of network downtime can cost millions
... downtime is not an option
www.agilent.com/comms/onenetworks
Agilent Technologies

Проблемы современных компьютерных сетей

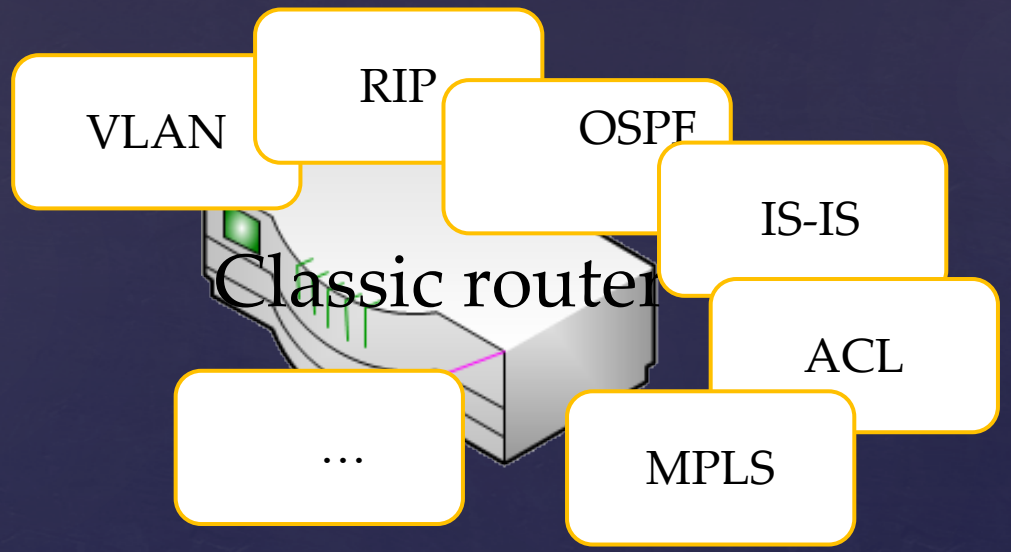
- Современные сети проприетарны: диктат производителя
- Закрытость для инноваций: внесение любых изменений трудоемко, дорогостояще, длительно по срокам (увеличение срока ROI)
- Сложность: свыше 600 используемых протоколов, более 10 000 RFC
- Число middle boxes растет постоянно
- Нет надежных решений по безопасности
- Невозможно контролировать и надежно предвидеть поведение таких сложных объектов, как глобальные компьютерные сети (ping, traceroute)

Выводы

- Необходимо устранить зависимость вывода сервиса от вендора сетевого оборудования
- Уметь быстро создавать те сервисы, которые востребованы пользователями
- Отделять функциональную часть сервиса от железа (программируемость)
- Оптимизация затрат на предоставление сервисов (запускать там, столько и тогда сколько востребовано пользователями - виртуализация)
- Повысить уровень автоматизации управления сетью – администратор говорит что надо, сеть это делает!

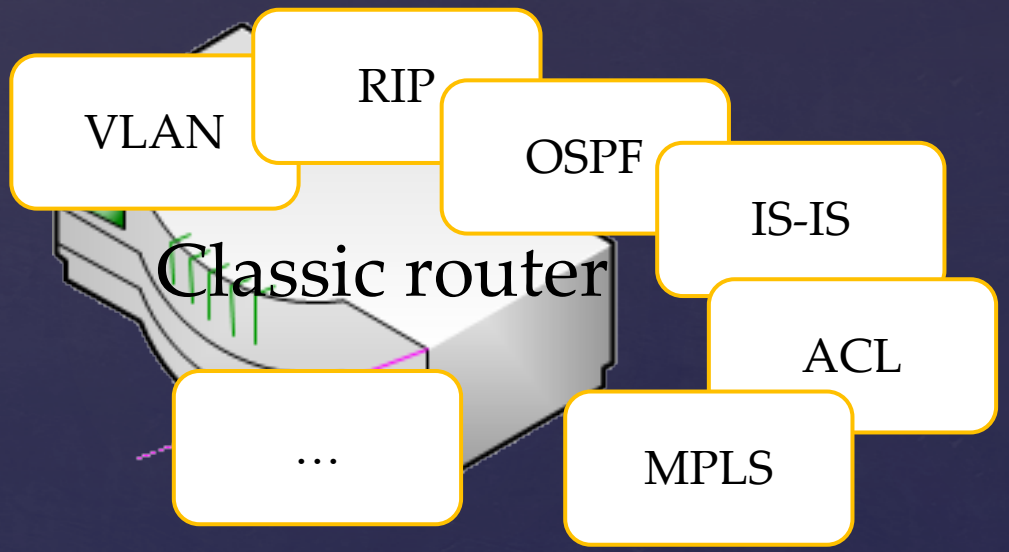


ПКС на базе OF

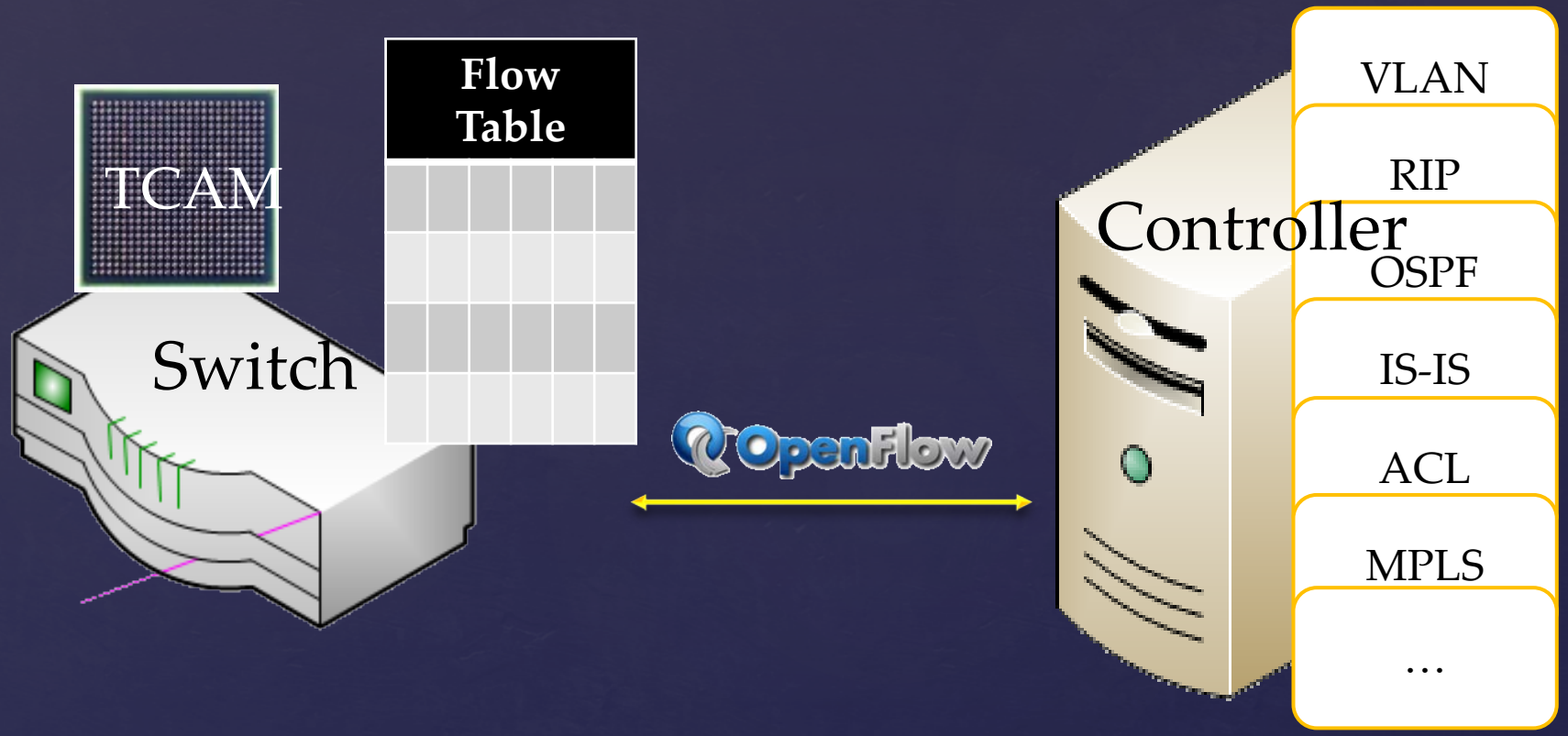


ПКС на базе OF

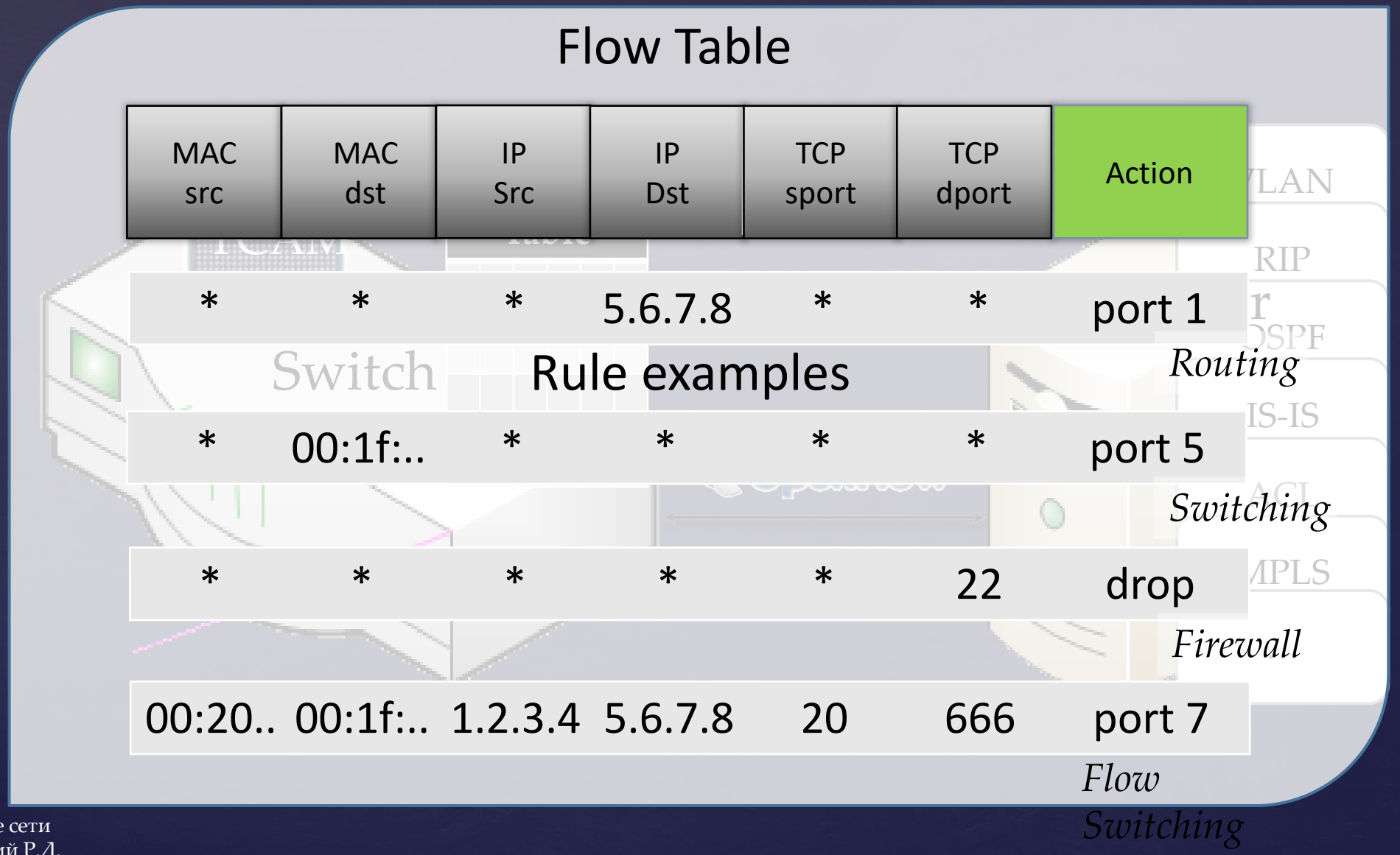
ПКС на базе OF



ПКС на базе OF



ПКС на базе OF



ПКС на базе OF

Flow Table

MAC src	MAC dst	IP Src	IP Dst	TCP sport	TCP dport	Action
*	*	*	5.6.7.8	*	*	port 1
*	00:1f:..	*	*	*	*	port 5
*	*	*	*	*	22	drop
00:20..	00:1f:..	1.2.3.4	5.6.7.8	20	666	port 7

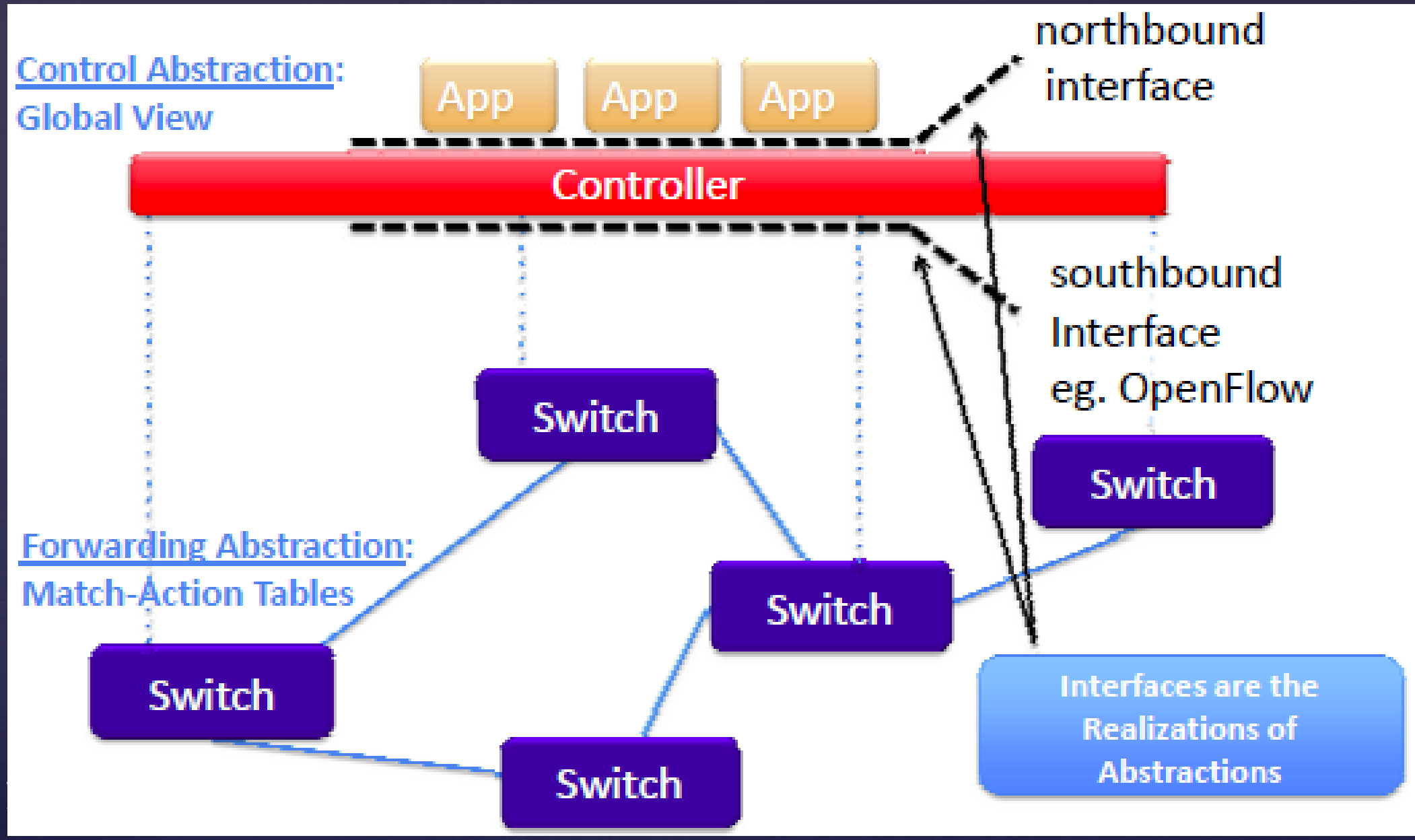
Rule examples

Routing

Switching

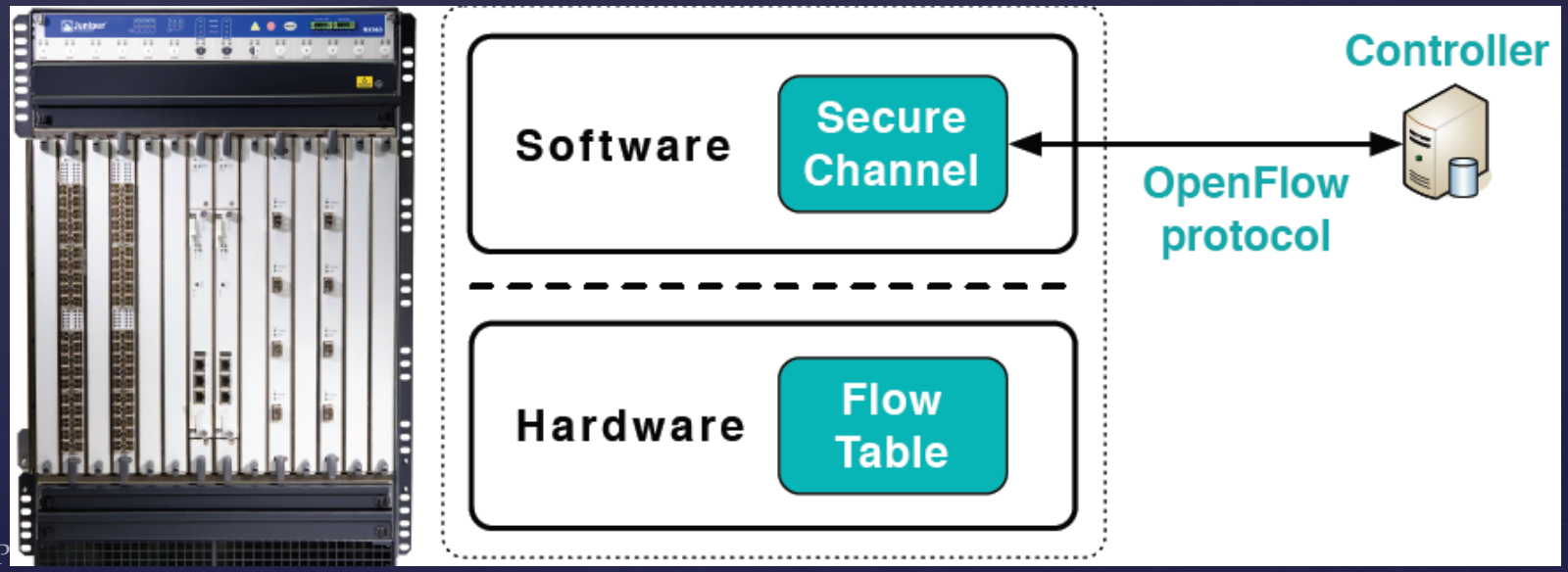
Firewall

Flow Switching


Control Abstraction:
Global View
App
App
App
northbound
interface
Controller
southbound
Interface
eg. OpenFlow
Switch
Switch
Forwarding Abstraction:
Match-Action Tables
Switch
Switch
Switch
Interfaces are the
Realizations of
Abstractions

OpenFlow коммутатор (v1.0)

- Таблица потоков – определяет, как коммутатор будет обрабатывать каждый поток
- Защищенный канал – соединяет коммутатор с удаленным контроллером
- OpenFlow protocol – стандарт для взаимодействия коммутатора с контроллером



Запись в OpenFlow таблице

Запись в таблице переходов:

Просматриваемые поля	Счётчики	Инструкции

Ingress Port	Ether src	Ether dst	Ether type	VLAN PCP (*6)	VLAN id	IP src	IP dst	IP proto	TCP/UDP src port	TCP/UDP dst port

* Просматриваемые поля: входной порт, заголовок пакета, метаданные
* Инструкции:
  * Изменение пакета
  * Продвижением пакета по конвейеру
  * Добавление новых действий в Набор действий (Action Set)
* Счётчики: количество байтов и пакетов, время соединение

ПКС: промежуточный итог

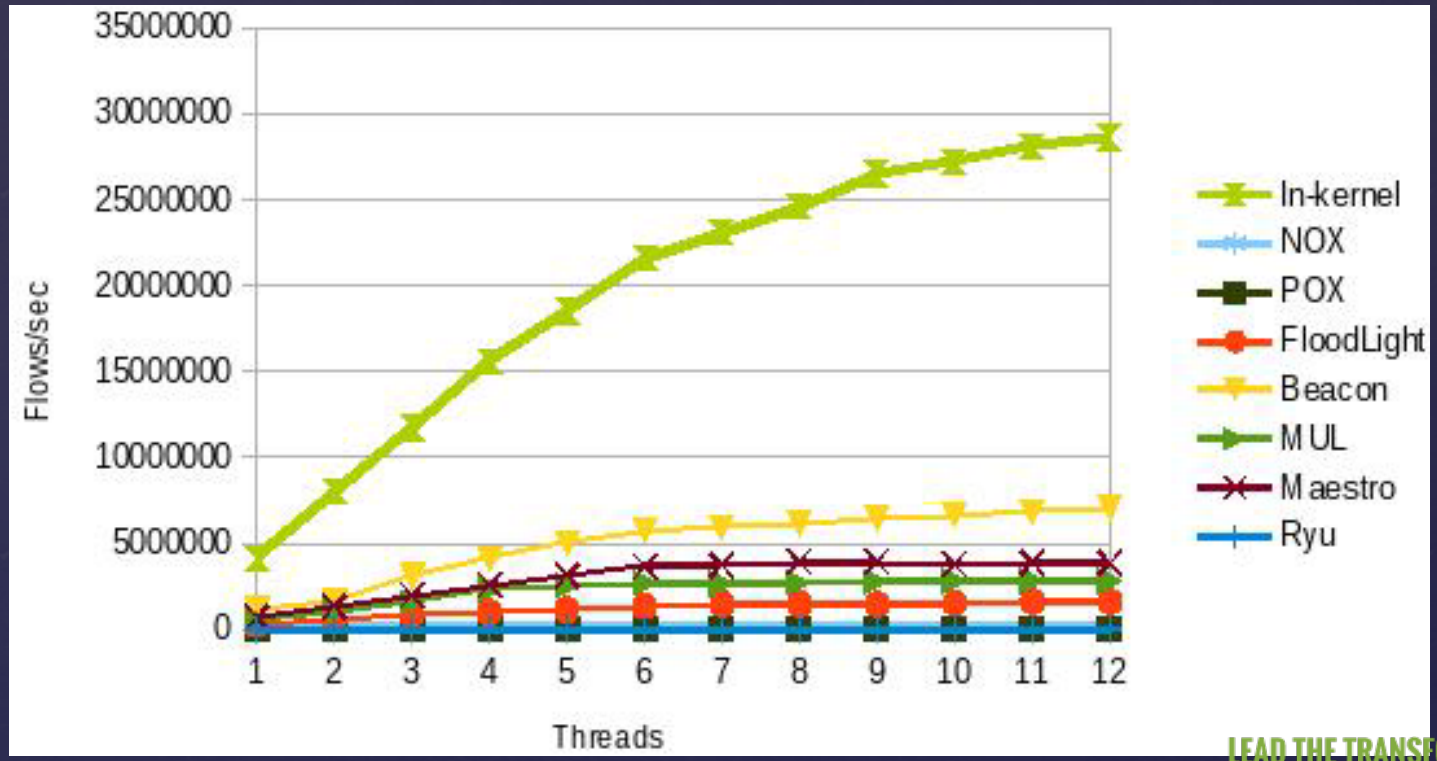
- Изоляция контура управления от контура передачи данных
- Унифицированный интерфейс для приложений управления
- Унифицированный интерфейс для контура передачи данных
- Иерархии заголовков пакета в контуре управления отсутвует
- Централизация управления:
  - понятие состояния сети
  - резкое сокращение времени сходимости
  - маршрутизатор = коммутатор

RUNOS

RUNOS = **RU**ssian **N**etwork **O**perating **S**ystem

- It is a series of SDN/OpenFlow controllers
  - **In-kernel** – super fast, hard to develop apps
  - **Fusion** – userspace memory control interface to the kernel controller
  - **Easy** – fully userspace controller with high functionality, easy to develop your apps, relatively high performance comparing to cotemporary userspace controllers
  - **Distributed** – HA version of the userspace controller
- The project is in the open source github.com/arccn/runos

Performance (in_kernel)

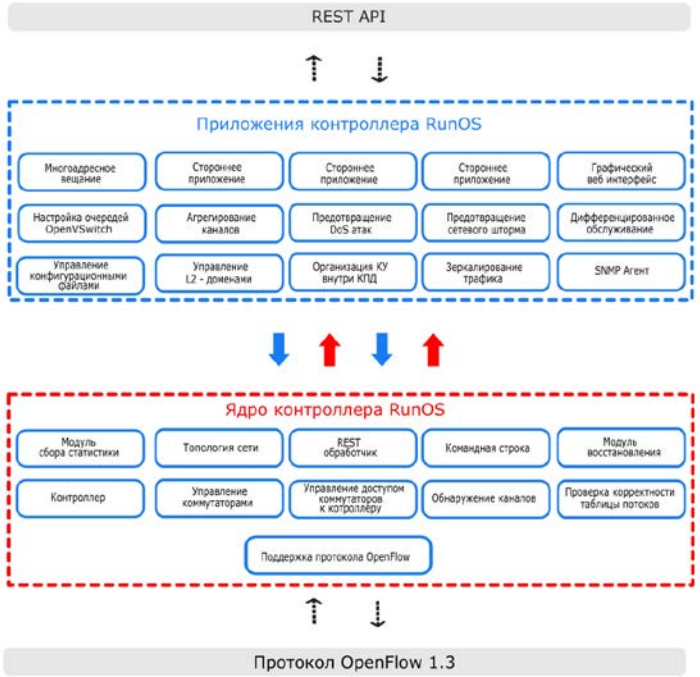


- Performance is **30M fps**
- Latency is **45us**

Контроллер RunOS

RunSDN

Российский программный продукт для программно-конфигурируемых сетей (ПКС), предназначен для использования в сетях операторов связи, сетевой инфраструктуре центров обработки данных, в корпоративных сетях с высокими требованиями к гибкости и масштабированию.



Функциональные возможности

Компоненты контроллера - это многопоточное ядро и пользовательские приложения. Пользовательские приложения могут быть как собственными приложениями контроллера, так сторонними.

Пользовательские приложения реализуют политики управления сетью:

- ✓ многоадресное вещание,
- ✓ зеркалирование трафика,
- ✓ организацию контура управления (КУ) внутри контура передачи данных (КПД),
- ✓ агрегирование каналов,
- ✓ создание L2-доменов,
- ✓ управление очередями.

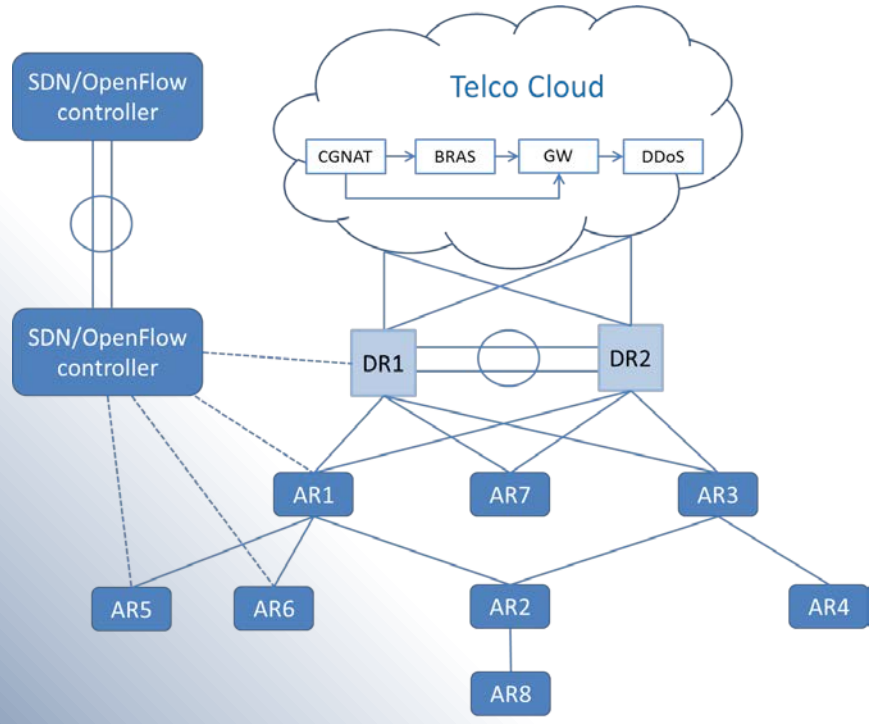
В ядре реализованы функциональности, которые используют приложения:

- ✓ разведка топологии,
- ✓ управление с командной строки,
- ✓ обеспечение высокой доступности,
- ✓ обеспечение работы по протоколу OpenFlow/NETCONF,
- ✓ предоставление информации о коммутационных устройствах,
- ✓ сбор статистики с коммутационных устройств.

Применение

Управление сетью Интернет-провайдера

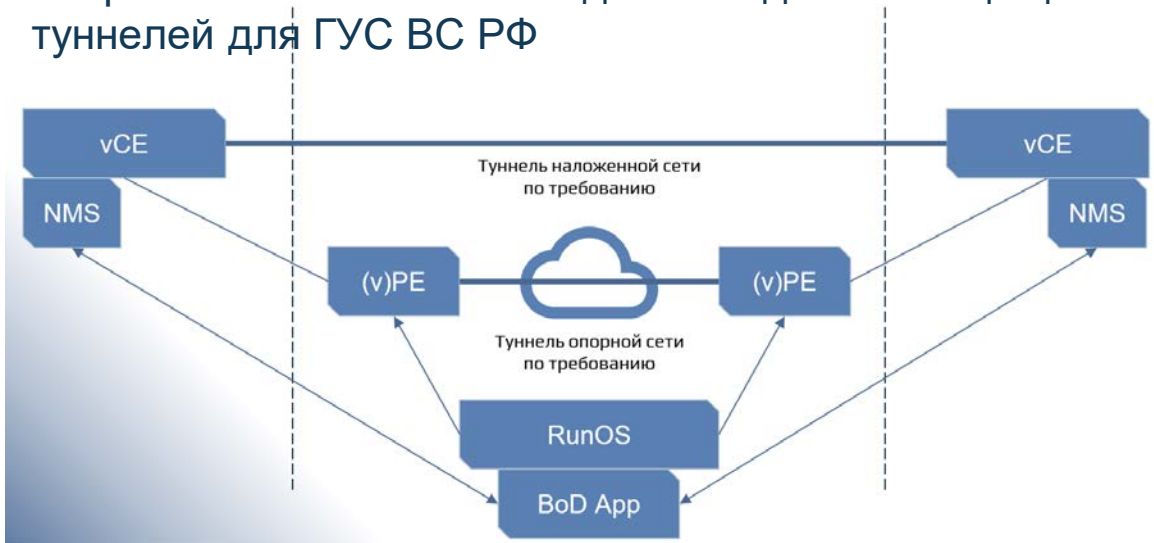
- ✓ Полная версия RunoS прошла апробацию в сетях ПАО "Ростелеком“
- ✓ Для разработчика телекоммуникационного оборудования АО "НИИ "Масштаб" на базе коммерческой версии был разработан и прошел успешные испытания тестовый стенд.



Пропускная способность по требованию (BoD)

Принципиально новый подход к организации передачи данных, который позволяет заменить аренду канала с гарантированной полосой пропускания, рассчитанной на максимальную интенсивность трафика, предоставлением L3-канала с пропускной способностью, отвечающей текущим потребностям клиента, без потери качества обслуживания.

- ✓ Успешно завершен пилотный проект предоставления сервиса BoD между ЦОДами в Новосибирске (Cortel) и в Москве (Облакотека)
- ✓ Разработан тестовый стенд по созданию защищенных туннелей для ГУС ВС РФ



ПКС контроллер RunOS и библиотека приложений

RunSDN

(https://github.com/ARCCN/runos)

RunOS - отечественный ПКС контроллер с открытым API для написания приложений сторонними разработчиками. Прошел апробацию в лаборатории Ростелекома

Технические характеристики RunOS

- ✓ обработка 30 миллионов потоков в секунду;
- ✓ время на установку нового соединения 45 мкс;
- ✓ поддержка 1000 коммутаторов;
- ✓ управление из графического интерфейса.

RunOS с приложениями совместим с коммутаторами многих вендоров:

BROADCOM, Mellanox TECHNOLOGIES, HUAWEI, NoviFlow SDN made smarter, Extreme networks, CISCO, ELTEX SOLUTIONS

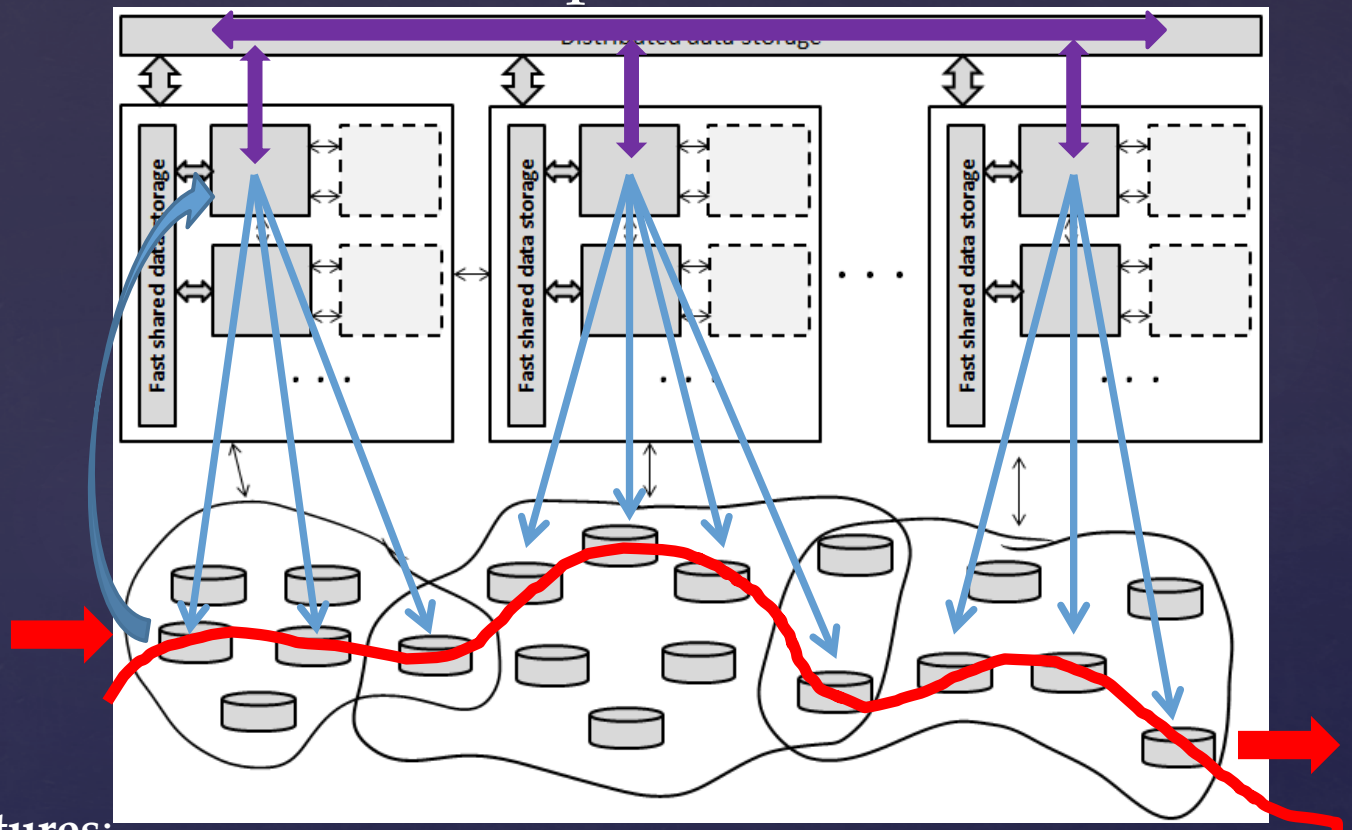
Отечественный программный коммутатор на x86 серверах

Сервисная модель RunOS для телеком оператора

- ✓ Поддержка произвольной топологии;
- ✓ Поддержка режимов управления «Out of band» и «In band»;
- ✓ Поддержка Active/Stand by модели резервирования SDN платформы управления RunOS;
- ✓ Обеспечение выхода в Интернет для B2C (поодержка PPPoE и IPoE);
- ✓ Организация VPN сервисов (P2P, P2MP, MP);
- ✓ Гибкое управление политикой маршрутизации для каждого сервиса (автоматически / вручную);
- ✓ Автоматический выбор маршрута по заданным критериям;
- ✓ Передача Multicast и VoIP трафика;
- ✓ Поддержка Broadcast «Storm-Control»;
- ✓ Динамическое управление QoS (PQ, WRR, Policiers;
- ✓ Поддержка LAG (включая LACP);
- ✓ Зеркалирование трафика по гибкому набору критериев с возможностью отправки на контроллер;
- ✓ Сбор статистики на контроллере, с возможностью вывода на внешние аналитически системы.

Сервисы	Режимы работы	Управление QoS
✓ B2C, VPN (P2P, MP), B2B; ✓ Multicast; ✓ Storm Control; ✓ LAG/LACP; ✓ InBand /OutOfBand; ✓ Маршрутизация по параметрам; ✓ Зеркалирование.	✓ Active/StandBy; ✓ Поддержка произвольной топологии; ✓ Fast Failover; ✓ Резервирование ✓ Ручное и автоматическое резервирование маршрутов.	✓ Priority Queuing, WRR; ✓ Rate-Policy, Ingress QoS, metering; ✓ Гибкое управление очередями.

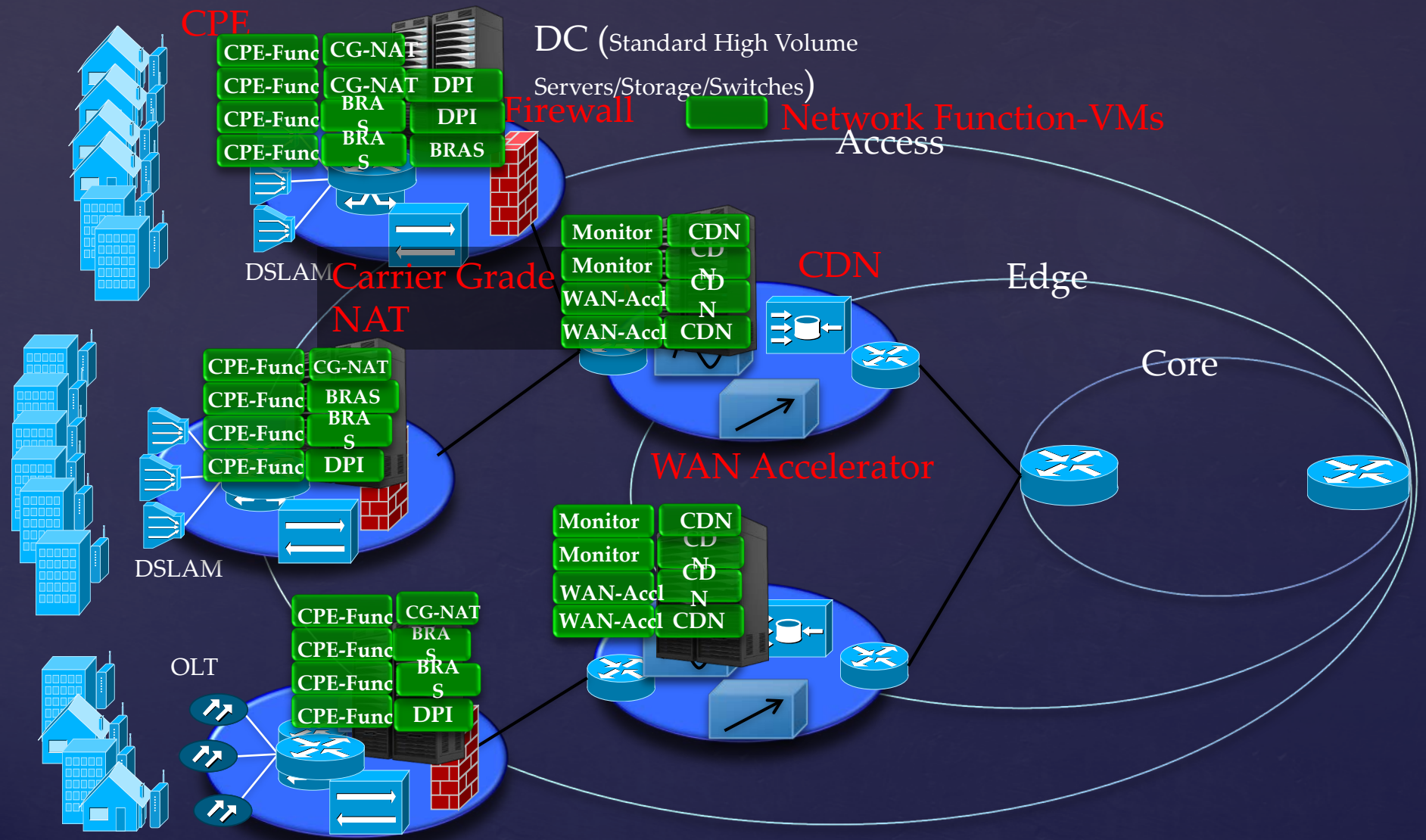
Distrib: Distributed OpenFlow controller



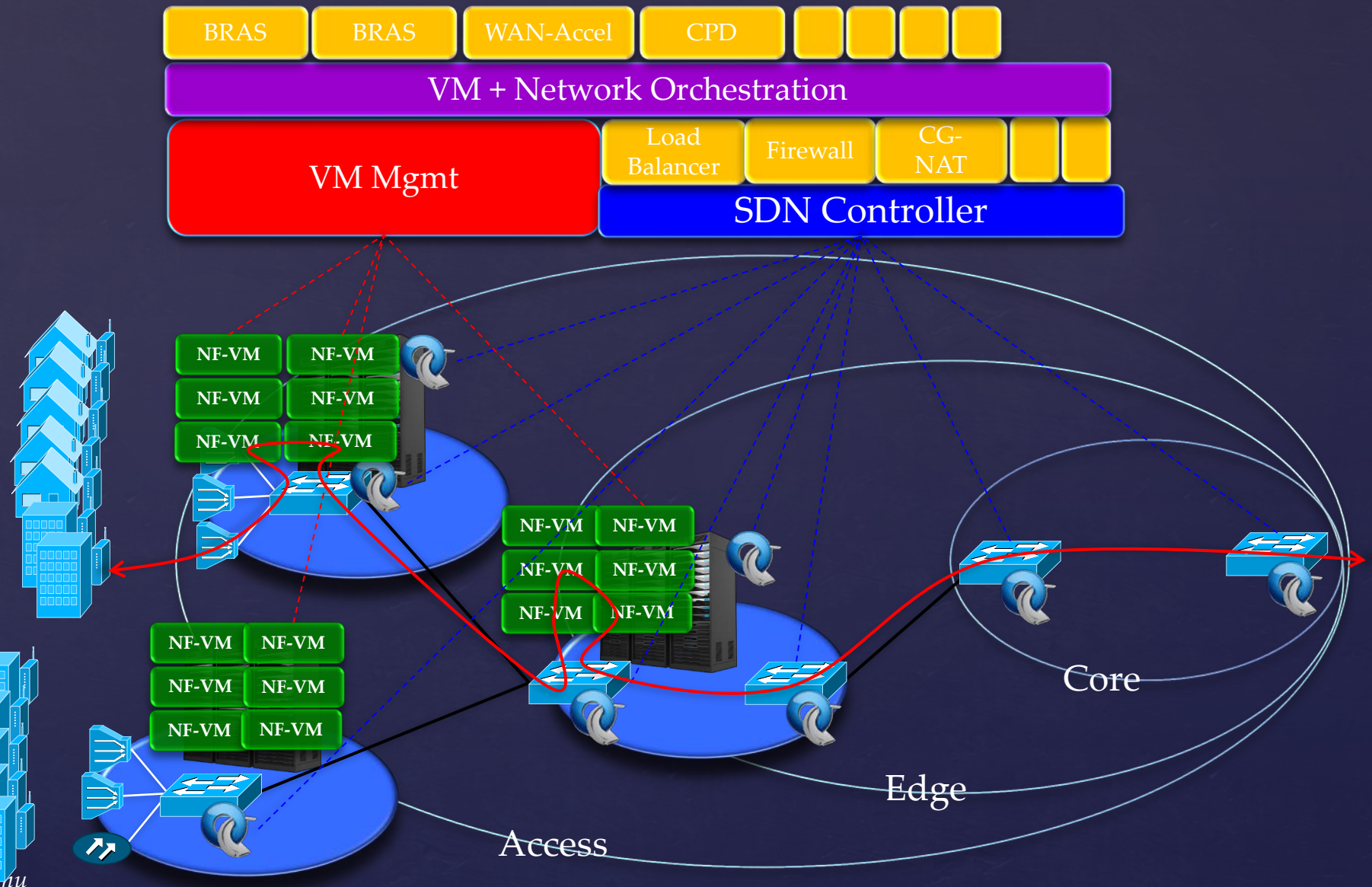
Main features:

- Reliability and fault tolerance (redundancy within a cluster and between clusters)
- Load balancing (adding new nodes to the controller, depending on the load)
- Coordinated management and vision of the entire network
- Working with distributed network applications
- Safety and counteraction to external loads

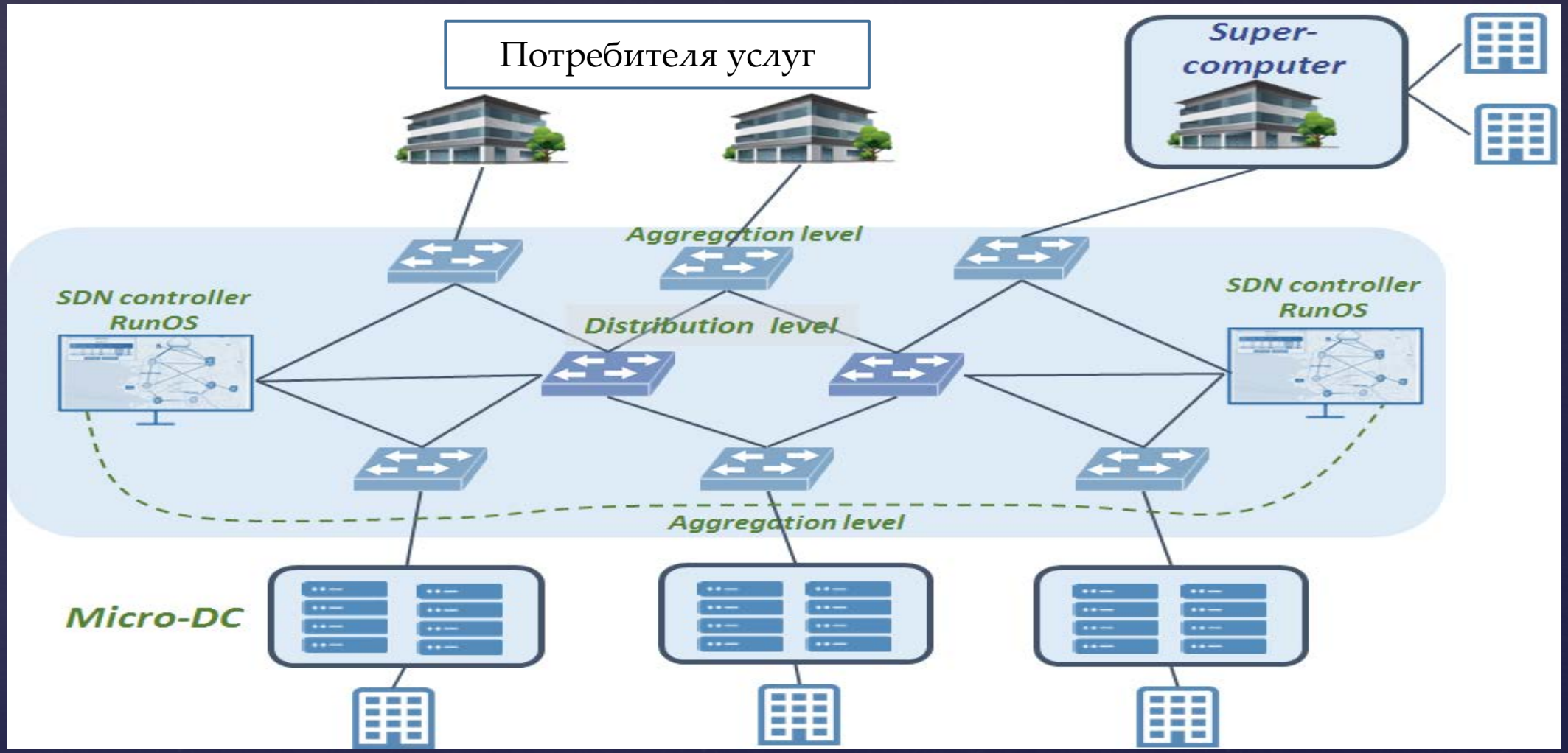
Сеть оператора с NFV



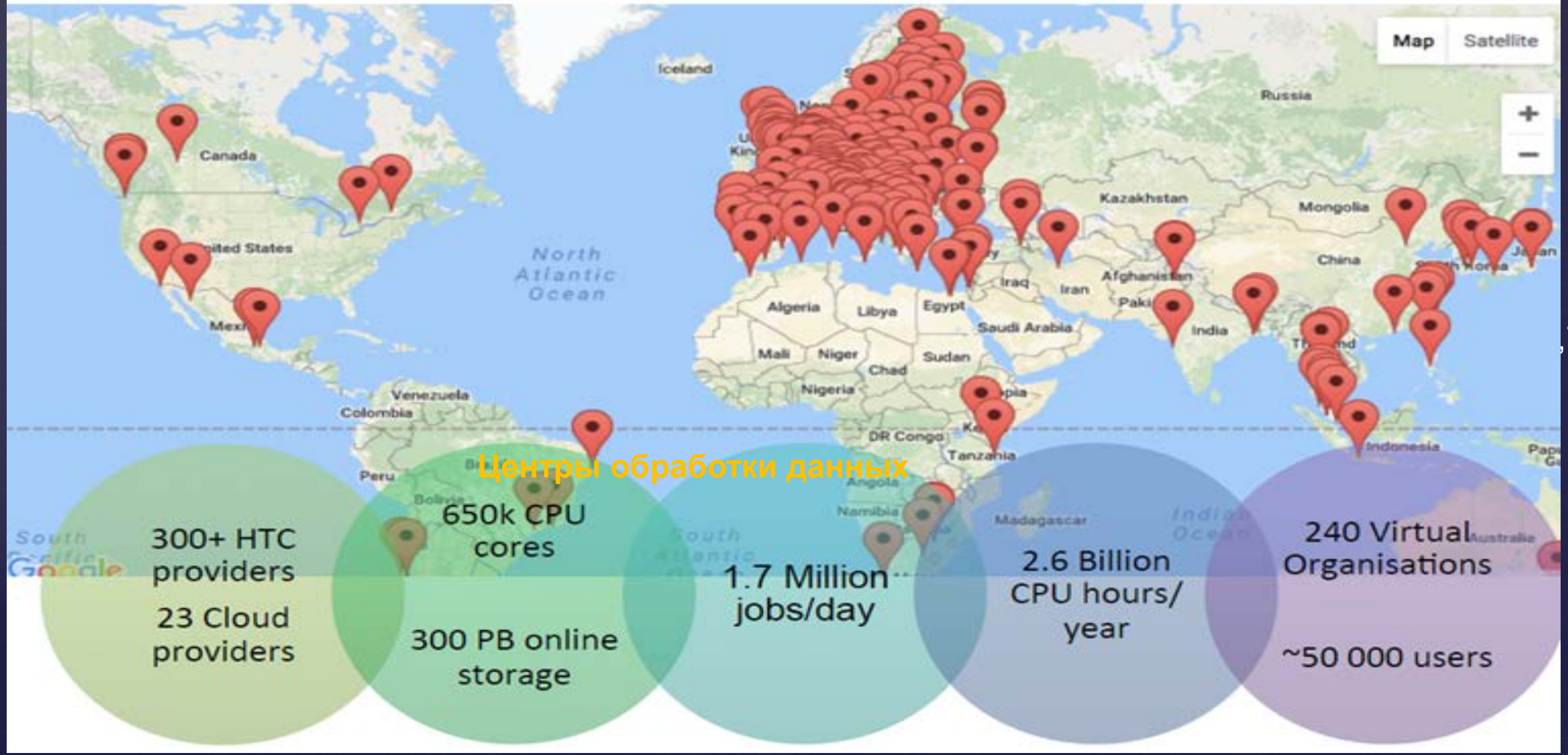
NFV с плоскостью управления SDN



SD – WAN технология



EGI Federated Infrastructure



Заключение

- Программируемое управление сетью
- Разделение контура управления и контура передачи данных
- Централизация управления – повышение уровня автоматизации управления сетью
- Упрощение конструкции коммутаторов = снижение стоимости = повышение производительности
- Совместимость (конвергентность) с традиционными сетями
- Программная настройка сетевых коммутаторов – сеть не зависит от стека протоколов
- Виртуализация
- NFV – в сети передачи данных есть только сервисы