



Объединенная  
приборостроительная  
корпорация



# Опыт разработки программно-аппаратного комплекса для ЦОДов на основе отечественных технологий

**Александр Фролов**

Начальник отдела, АО «НИЦЭВТ»

А.Кручинин, С.Осипов, Е.Карев, П.Боронин, Н.Леонов,  
Ю.Гончарук М.Гильмендинов, А.Данилов, Е.Куштанов





## Цель проекта

- ▮ Создание решения на базе сертифицированных отечественных технологий для задач НРС и облачных сервисов

## Задачи проекта

- ▮ Выбор инженерной инфраструктуры
- ▮ Выбор серверной (аппаратной) инфраструктуры
- ▮ Выбор программного обеспечения
- ▮ Создание опытного образца (пилота)



# Опытный образец ПАК «Базис»

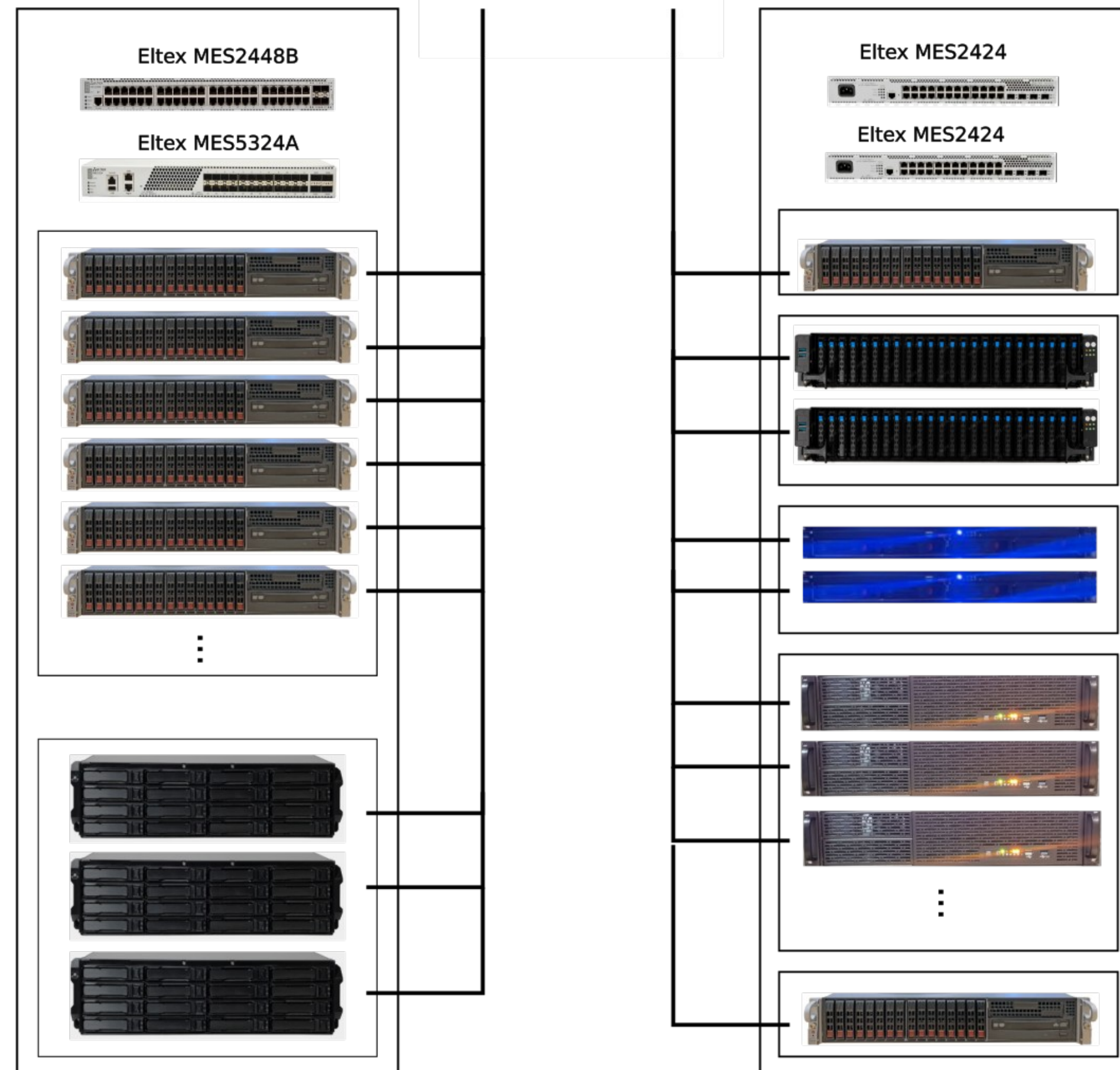




# Опытный образец ПАК «Базис»



## Коммутатор «Ангара»



### Серверы ОН / ВП НИЦЭВТ (9х)

2x AMD EPYC 7513 / 7742 / 7763  
512 Gb (1Tb)

2x 256 Gb SSD / 4x 1.8 Tb HDD

### СХД BITBLAZE Sirius E8000

3 контроллера «Эльбрус-8СВ»,  
64 Gb 16x 12 Tb

### Серверы СХД НИЦЭВТ

2x AMD EPYC 7252, 1 Tb 8x 1.8 Tb

### Серверы ASUS (2х)

RS720A-E11-RS24U  
AMD EPYC 7513, 2 Tb  
Nvidia A100 40 Gb

### Серверы ОН НИЦЭВТ (2х)

2x «Эльбрус-8СВ», 128 Gb

### Серверы ВП НИЦЭВТ (7х)

2x «Эльбрус-8СВ», 128 Gb

### Серверы СХД НИЦЭВТ

«Эльбрус-8СВ», 64 Gb 2x 8 Tb

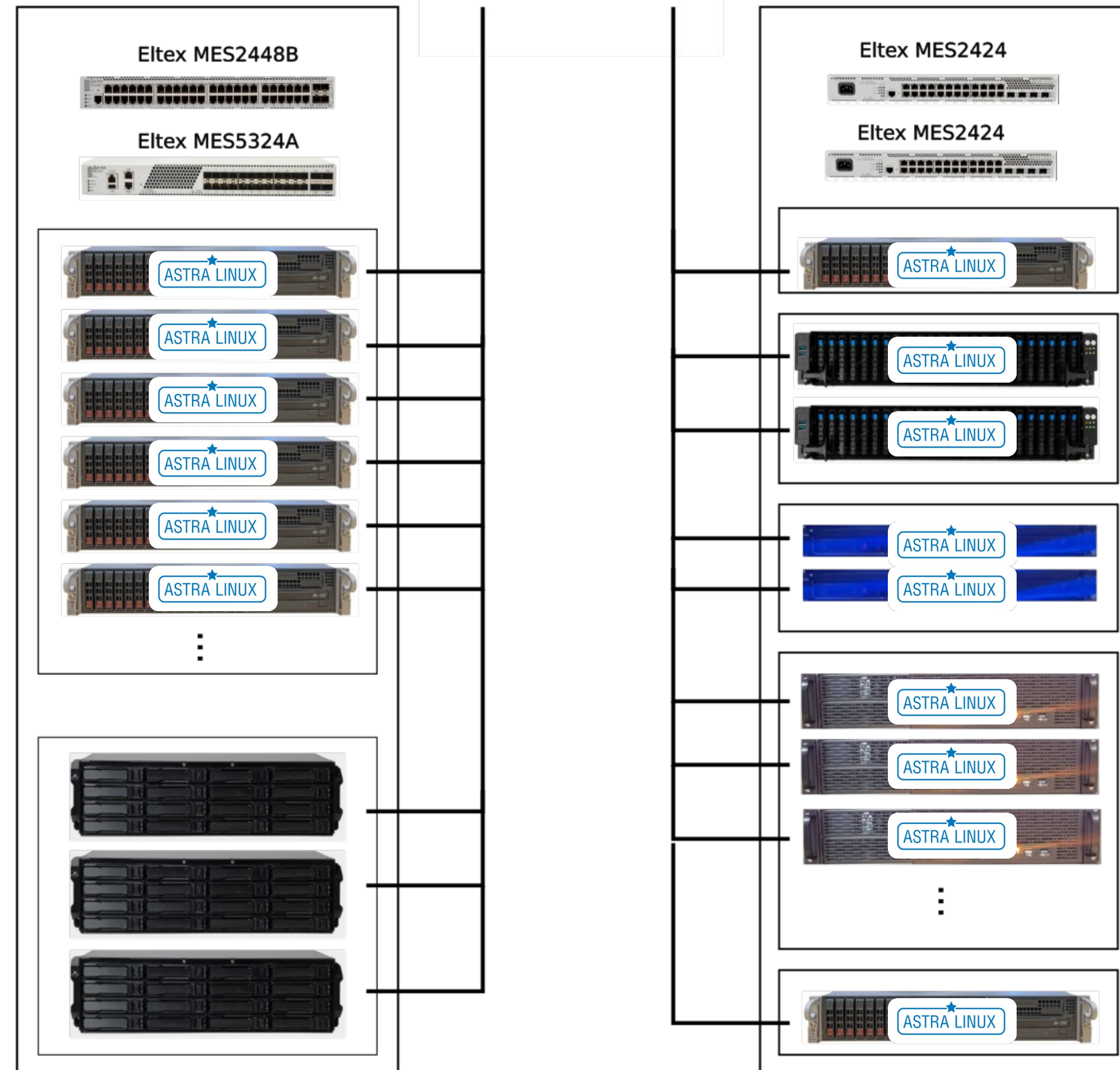




# Опытный образец ПАК «Базис»



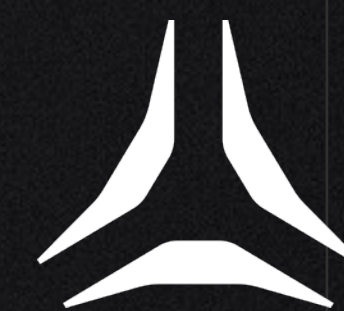
Коммутатор «Ангара»







Объединенная  
приборостроительная  
корпорация



АСТРА

# Аппаратная платформа



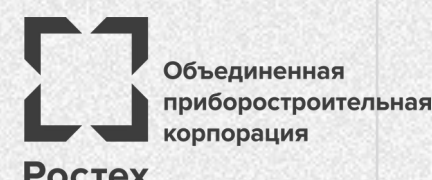


# Серверная линейка НИЦЭВТ с процессорами AMD EPYC (ПИРШ.467449.044)



## Характеристики сервера

- Высота 2U (19“)
- 2x AMD EPYC Naples / Rome / Milan
- До 1Tb DDR4 ОЗУ
- До 14x 2.5in HDD / SSD
- 4x PCIe 3.0/4.0 16x
- 1x USB 2.0, 3x USB 3.0
- 2x PSU
- BIOS (собственной сборки)
- BMC (Open BMC)
  - IPMI 2.0, Redfish API
  - KVM-консоль, Serial-over-LAN
  - Управление питанием, мониторинг датчиков
  - Поддержка SNMP
  - Поддержка LDAP / Active Directory





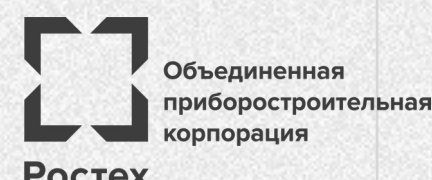
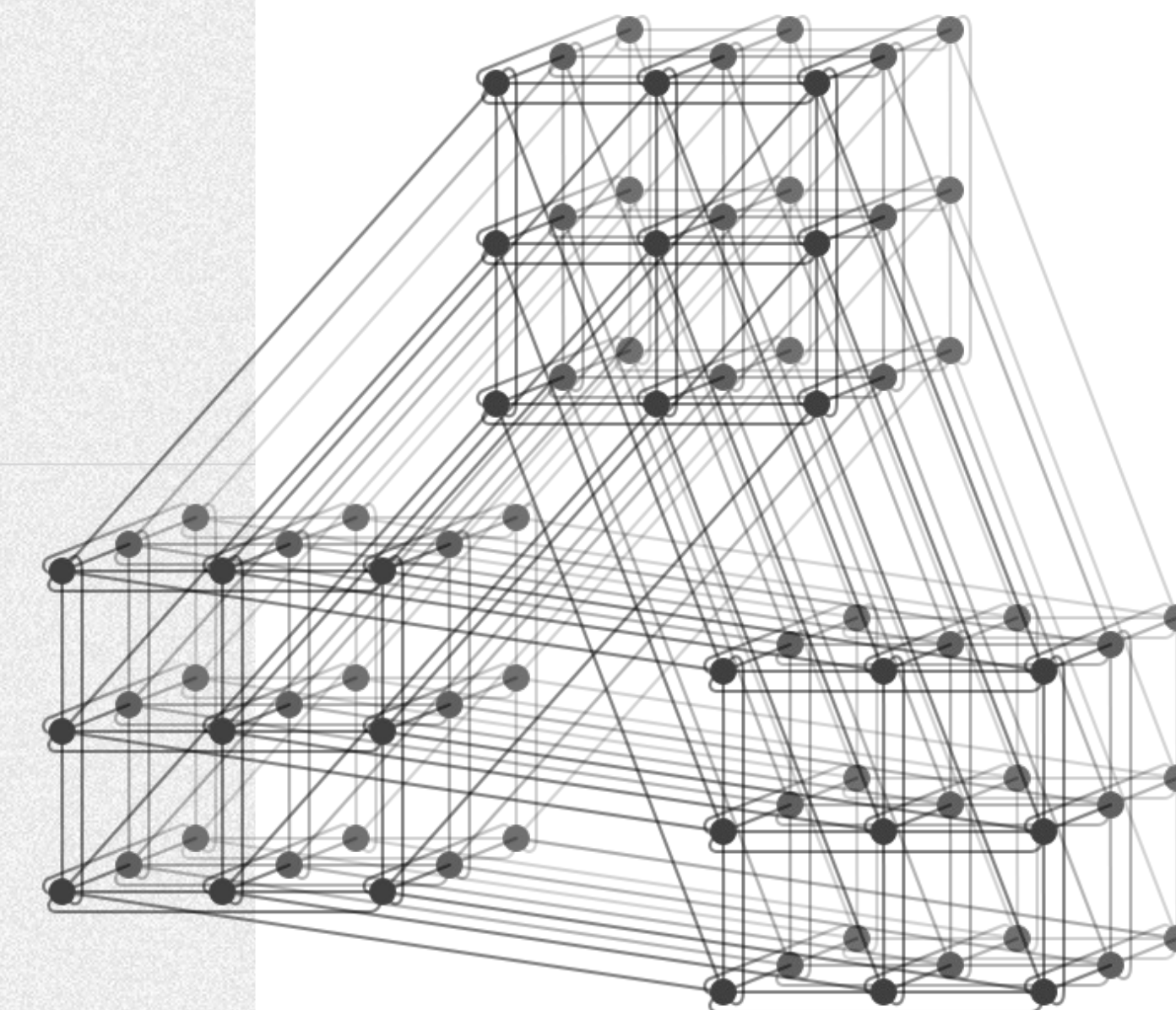
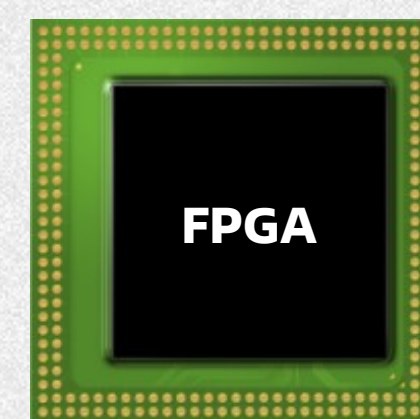
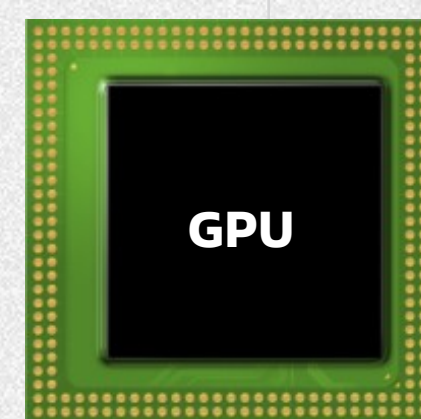
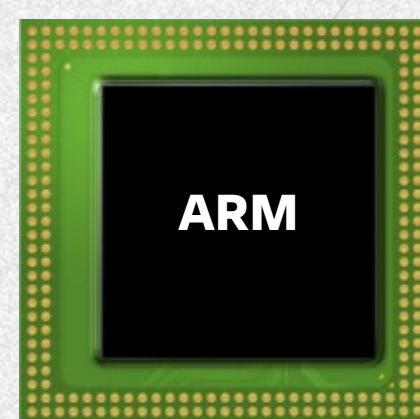
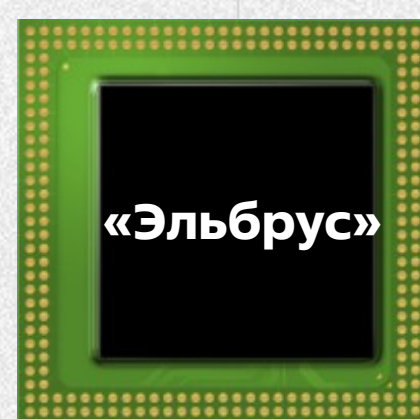
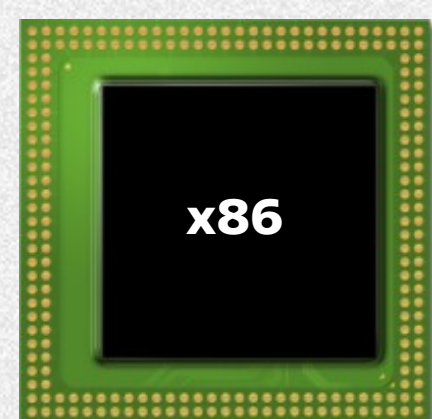
# Коммуникационная сеть «Ангара»

## Основные характеристики



### Ключевые особенности

- ▢ Топология сети 1D..4D-тор
- ▢ Адаптер на базе СБИС (65 нм, АО «НИЦЭВТ»)
- ▢ До 8 каналов связи с соседними узлами
- ▢ Прямой доступ в память удаленного узла (RDMA)
- ▢ Поддержка многоядерности
- ▢ Адаптивная передача пакетов
- ▢ Задержка на MPI ping-pong 0,85/1,54 мкс (x86 / «Эльбрус-8С»)
- ▢ Задержка на хоп 130 нс
- ▢ Масштабирование до 32К узлов
- ▢ Энергопотребление до 20 Вт
- ▢ Различные физические среды передачи данных



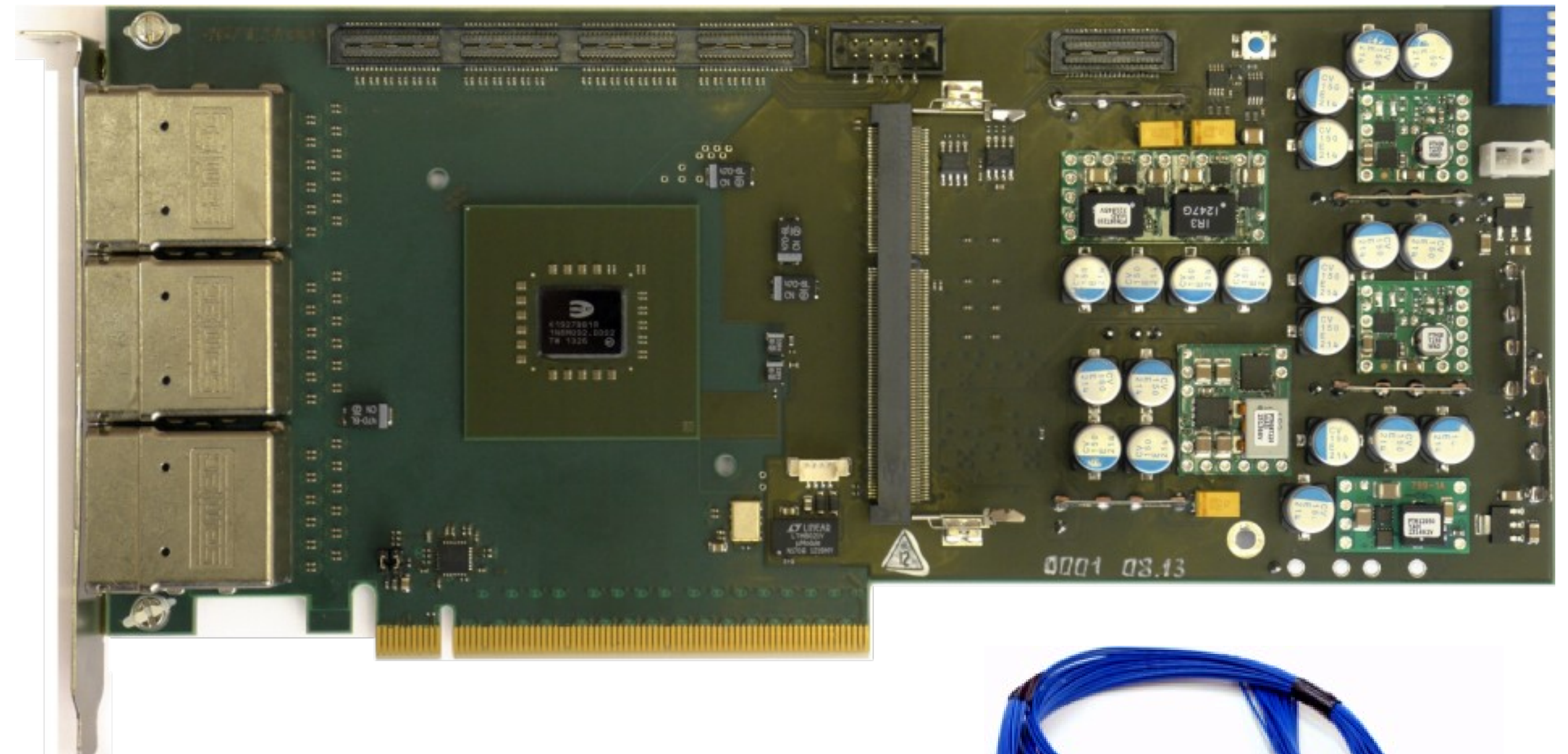


# Сетевое оборудование «Ангара» (1)



## Ключевые особенности

- ▮ Не требует коммутатора (switchless)
- ▮ Форм-фактор Full Height Full Length (FHFL)
- ▮ 75Gb/s на порт (bidirectional)
- ▮ 6 портов + 2 порта (через mezzanine-плату)
- ▮ Кабели Samtec HDLSP (до 2 метров)



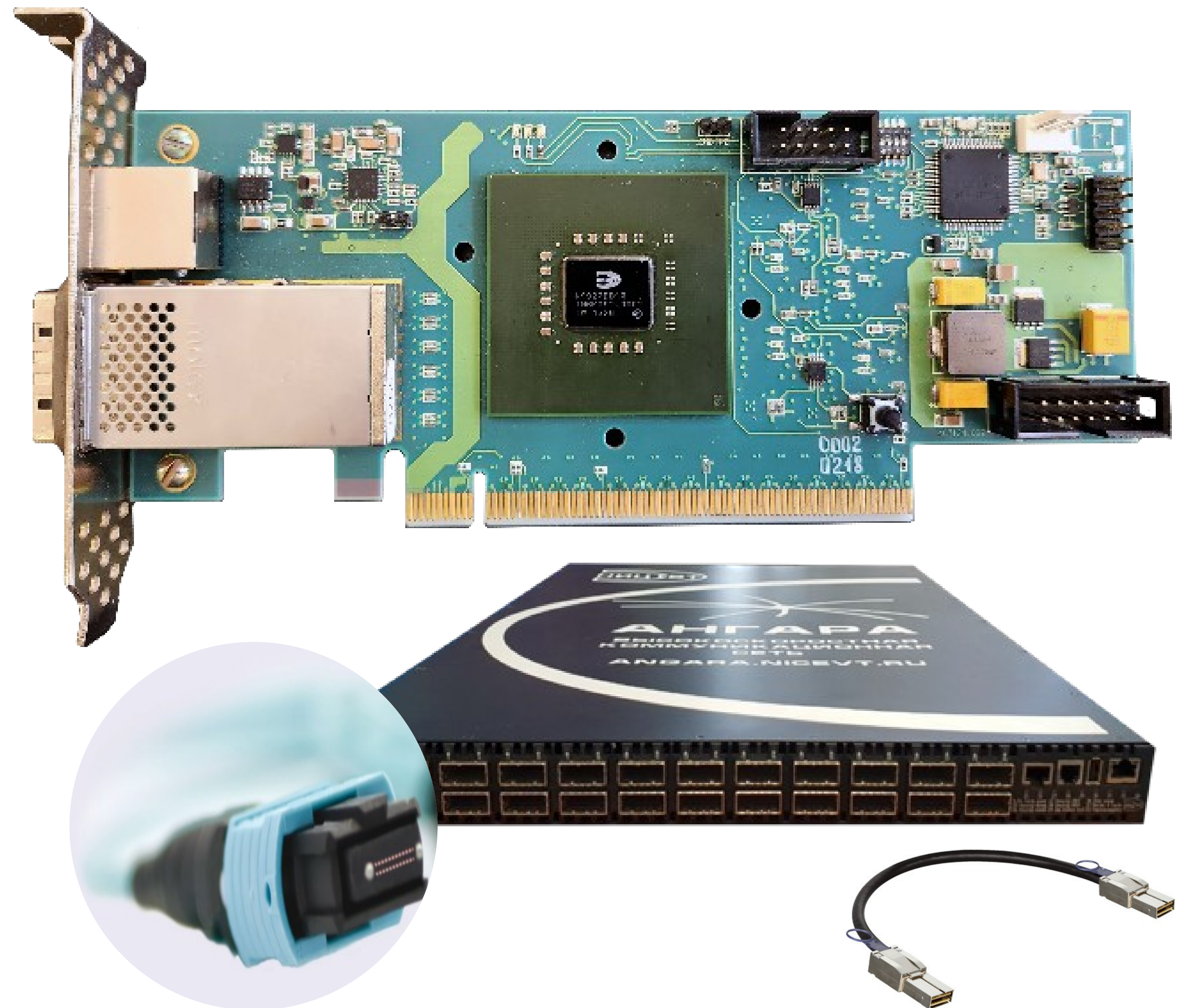


# Сетевое оборудование «Ангара» (2)



## Ключевые особенности

- ▢ 24-портовый коммутатор (4x СБИС «Ангара»)
- ▢ Однопортовый адаптер
- ▢ Форм-фактор Low-Profile
- ▢ 75Gb/s на порт (bidirectional)
- ▢ Кабели медные СХР (до 2 метров)
- ▢ Поддержка оптических кабелей (МТР/МРО)
- ▢ Дублирование питания (1+1)





# Сетевой адаптер EC8451.01



## Характеристики



«Ангара-2»



Не требуется коммутатор



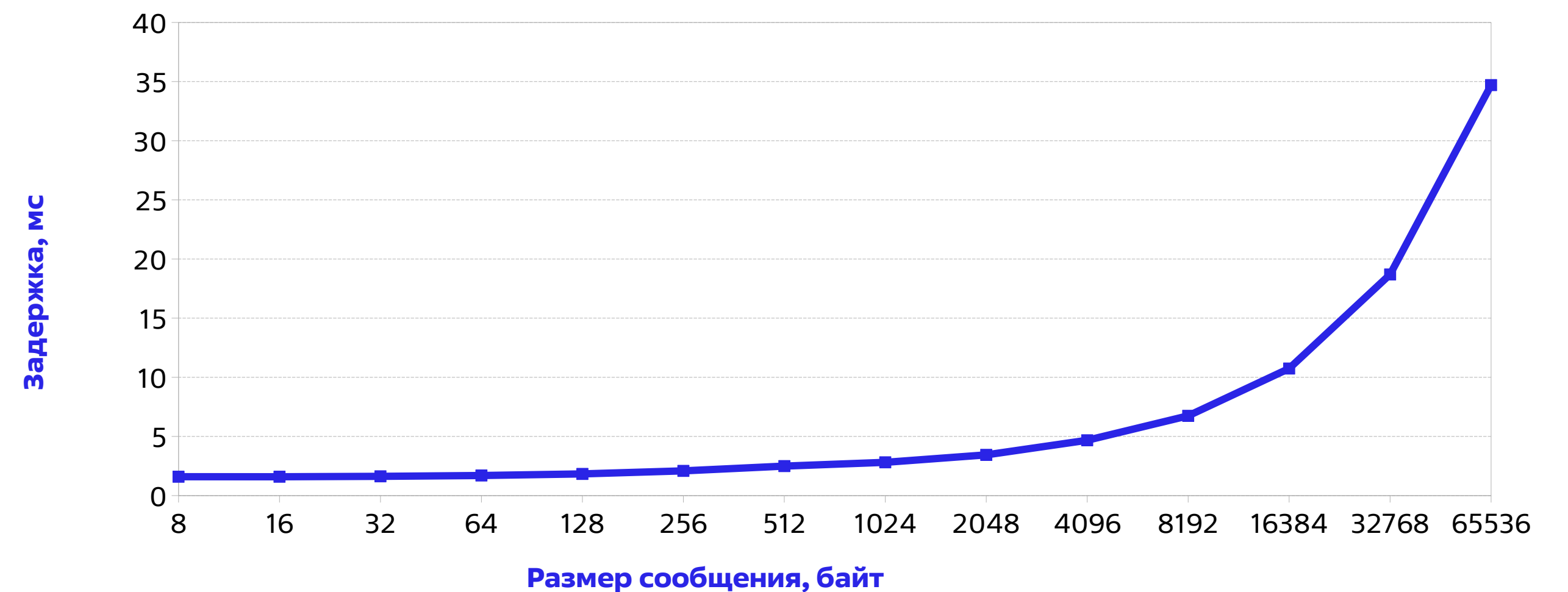
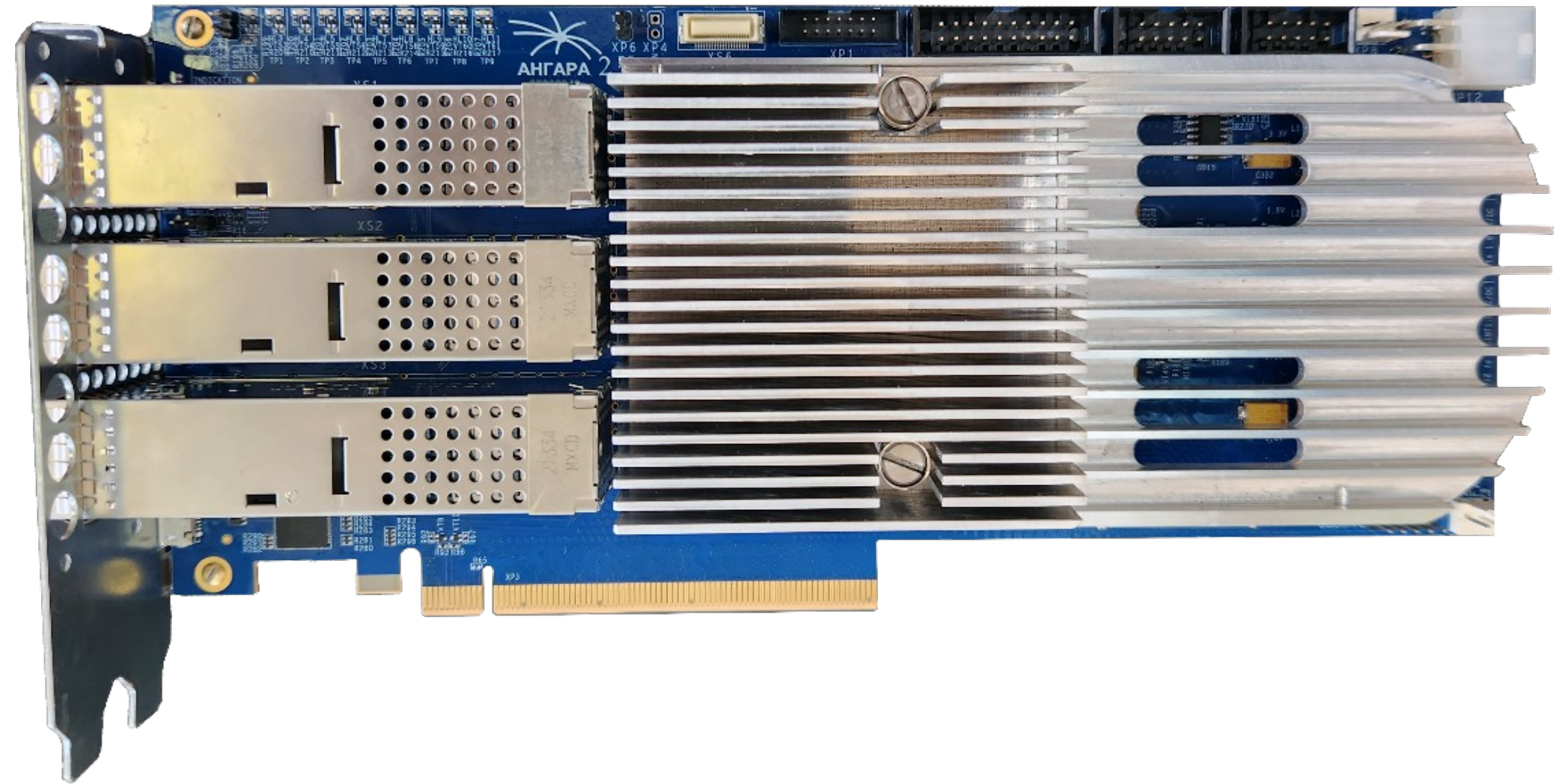
QSFP-DD



6 портов по 40 Гбит/с



Следующая редакция: 68 Гбит/с



Объединенная приборостроительная корпорация

Ростех



# Программное обеспечение сети «Ангара»



## Поддерживаемое программное обеспечение

### Операционные системы

Astra Linux SE 1.3-1.7, ОС «Эльбрус»,  
Alt Linux 9, OpenSUSE/SLES 11 SP3/4,  
CentOS 6.0-7.7, ОС «Нейтрино» 6.5

### Компиляторы и языки программирования

C/C++ (GNU, Intel), Fortran 77/90/95 (GNU, Intel)

### Библиотеки параллельного программирования

MPICH 3.0.4, 3.2, 3.3, 3.4, OpenMPI 1.10.2, 4.0.5  
MVAPICH2 1.9, IntelMPI 2021.5.1, NCCL

### Инженерные пакеты и приложения

ANSYS 17.0, 18.2, 21r1, StarCCM+ 12.02, 16.06.008r8, Flowvision,  
OpenFOAM, VASP, GROMACS, LAMMPS и многие другие

### Стек протоколов TCP/IP

### Параллельные файловые системы и SDS

Ceph, Lustre, GlusterFS, BeeGFS



Kernel-space      User-space



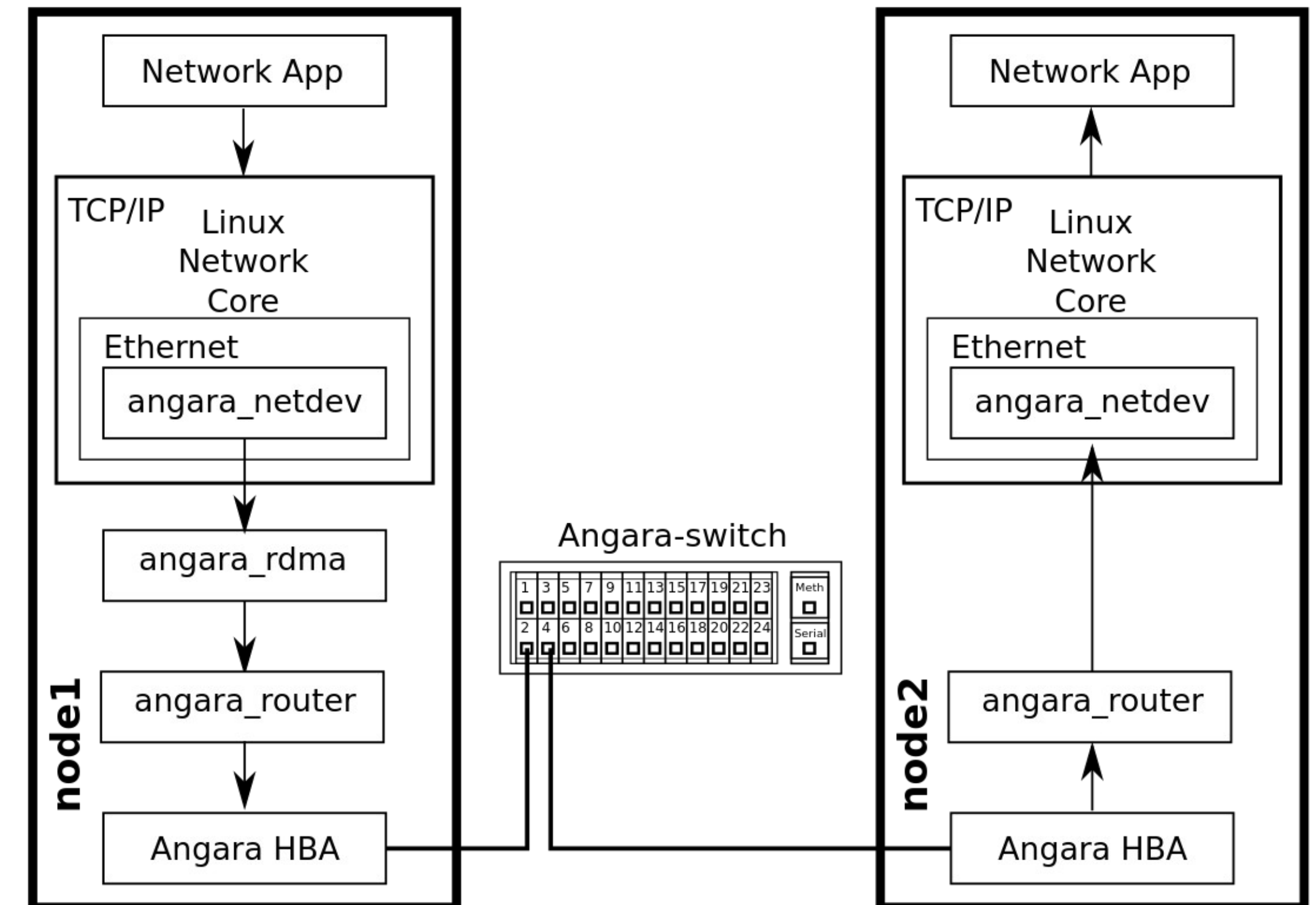


# Поддержка режима эмуляции Ethernet в сети «Ангара»



## Особенности реализации

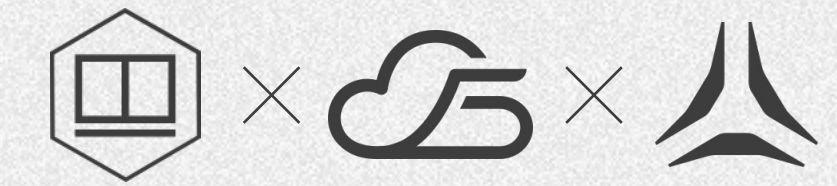
- Linux Network Device (TCP/IP)
  - iproute2, ifconfig, ethtool, route...
  - Программная поддержка ARP
- Активный поллинг кольцевых буферов (для первого поколения)
- Поддержка прерываний (для второго поколения)
- Поддержка NAPI
- Поддержка виртуализации, контейнеров
- Совместимость с linux bridge, macvtap, open vswitch
- производительность:
  - iperf (x86): 10 Gbit/s и >15 Gbit/s (тюнинг)
  - iperf (e2k): 5 Gbit/s и 7 Gbit/s (тюнинг)



```
9: angara0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 4076 qdisc pfifo_fast
state UNKNOWN group default qlen 1000
link/ether 00:15:d5:00:00:03 brd ff:ff:ff:ff:ff:ff
inet 172.16.1.14/24 scope global angara0
valid_lft forever preferred_lft forever
```



# Коммуникационная сеть Ethernet



## Коммутаторы Eltex MES2424 / 2448B

- ▮ 24 (48) порта 1GbE, 4 порта 10GbE
- ▮ Неблокируемая коммутационная матрица
- ▮ Коммутатор уровня L2+, L3



## Eltex MES5324A

- ▮ 24 порта 10GbE, 4 порта 40GbE
- ▮ Пропускная способность 800 Гбит/с
- ▮ Неблокируемая коммутационная матрица
- ▮ Коммутатор уровня L3





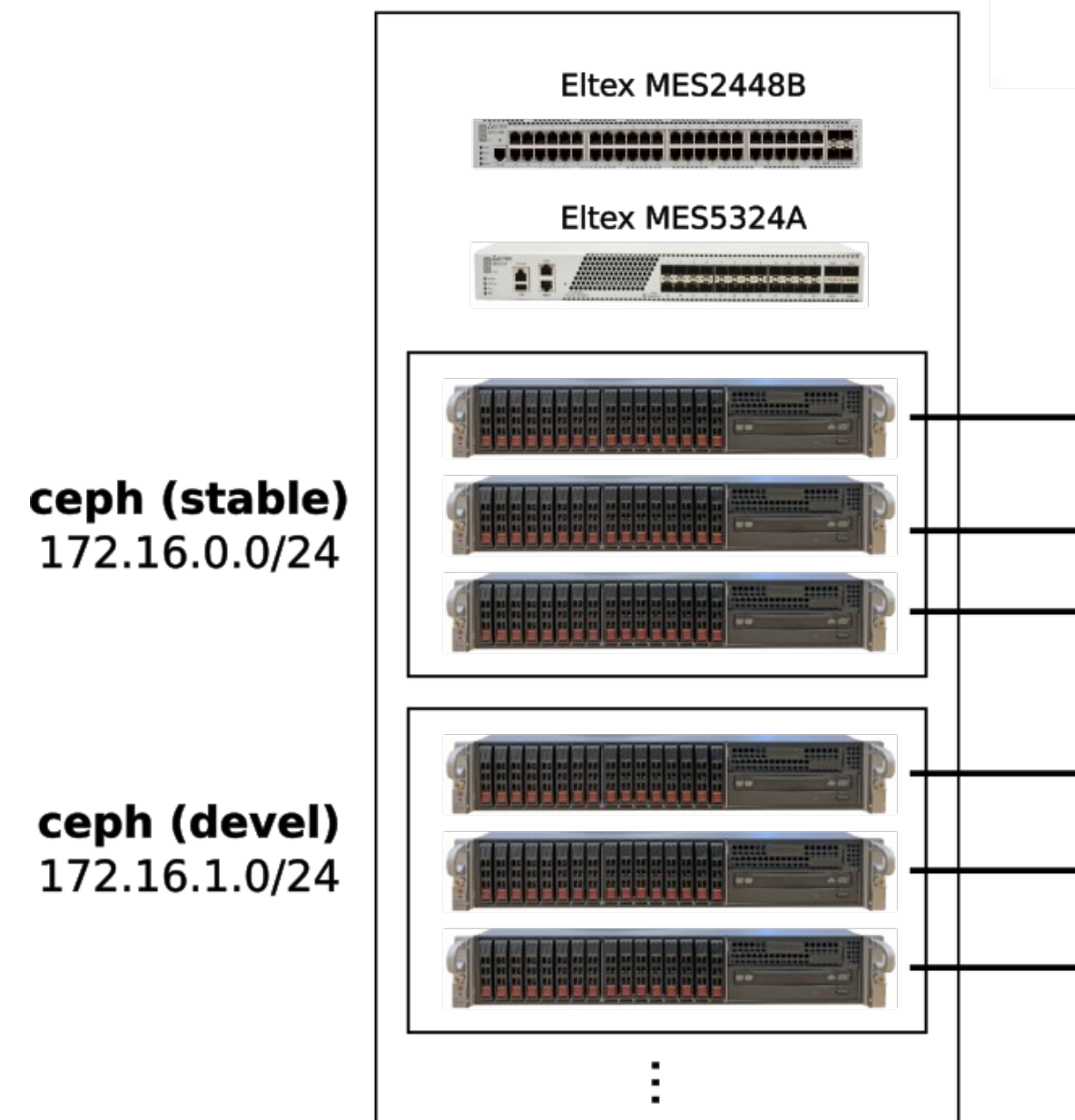
# Система хранения данных (1)



## Два Серв-кластера (по ~5Тб «грязными»):

- sp3-host[1,2,3] — stable
- sp3-host[4,5,6] — devel (angara)

Коммутатор Ангара



```
astra@sp3-host1:~$ sudo ceph -s
cluster:
  id:      4d67e740-ce62-44d3-9ab6-1dee5d9c9913
  health: HEALTH_OK

services:
  mon: 3 daemons, quorum sp3-host1,sp3-host2,sp3-host3 (age 41h)
  mgr: sp3-host2(active, since 41h), standbys: sp3-host1
  osd: 27 osds: 27 up (since 41h), 27 in (since 41h)

data:
  pools:  2 pools, 513 pgs
  objects: 441.04k objects, 1.4 TiB
  usage:   4.0 TiB used, 1.3 TiB / 5.3 TiB avail
  pgs:    513 active+clean

io:
  client:  21 KiB/s rd, 1.3 MiB/s wr, 105 op/s rd, 189 op/s wr
```

```
astra@sp3-host4:~$ sudo ceph -s

cluster:
  id:      6e8822e0-7810-40bd-bb38-5571184936fd
  health: HEALTH_OK

services:
  mon: 3 daemons, quorum sp3-host4,sp3-host5,sp3-host6 (age 5d)
  mgr: sp3-host5(active, since 5d), standbys: sp3-host4, sp3-host6
  osd: 27 osds: 27 up (since 2h), 27 in (since 2h)

data:
  pools:  2 pools, 33 pgs
  objects: 23.87k objects, 88 GiB
  usage:   242 GiB used, 5.0 TiB / 5.3 TiB avail
  pgs:    33 active+clean
```



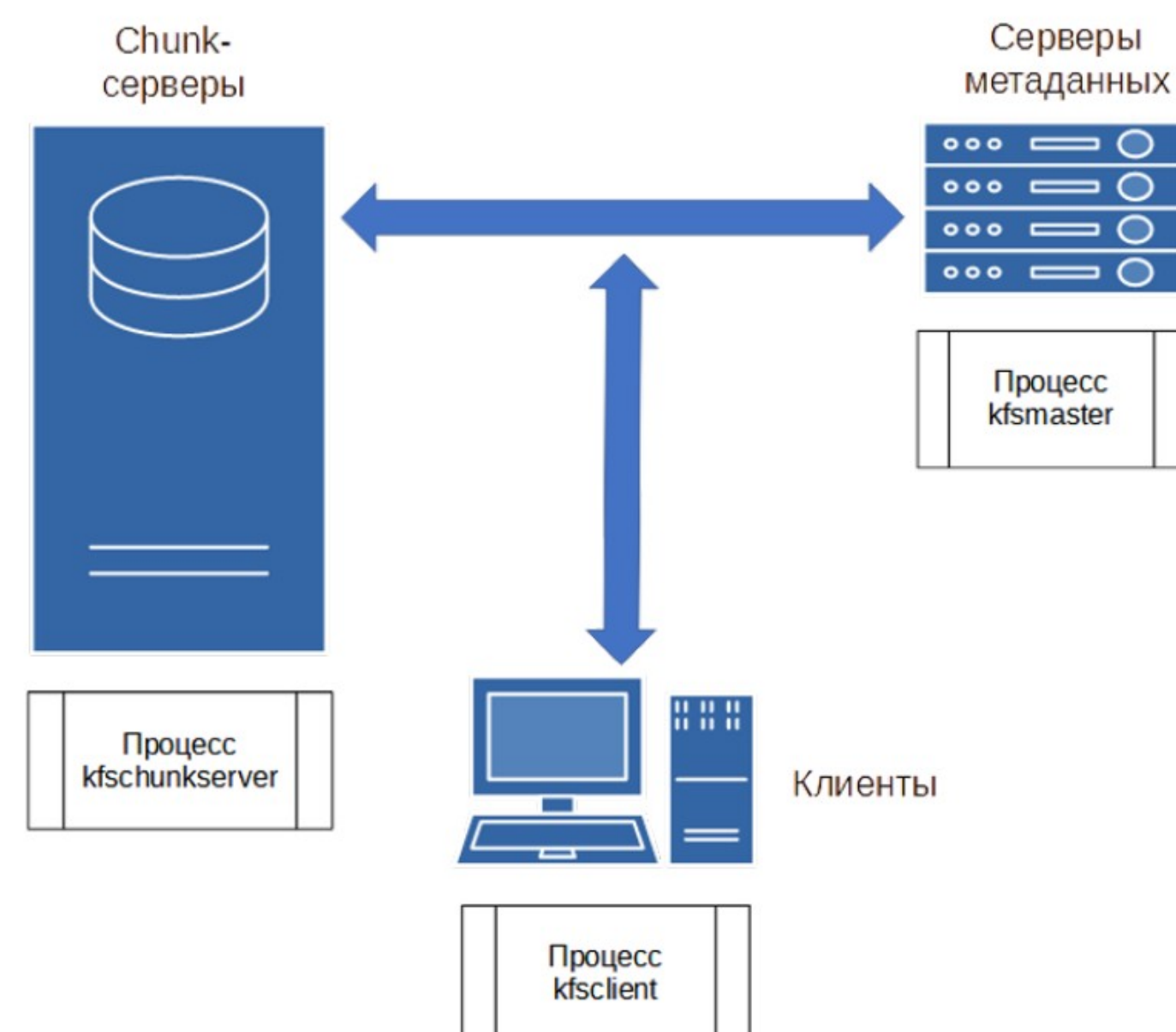


# Система хранения данных (2)



## BITBLAZE Sirius E8000 (ООО «Промобит»)

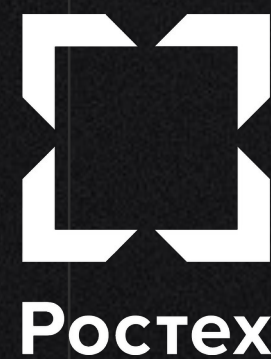
- 3 сервера хранения («Эльбрус-8СВ»)
- 16x 12T HDD диска / сервер
- Двухпортовый 10GbE
- Адаптер «Ангара»\*



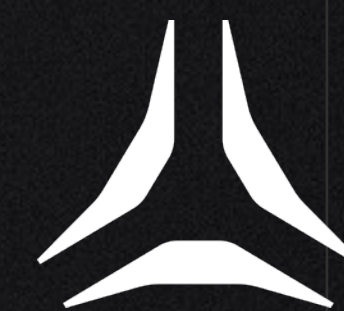
## Управляющее ПО

- BITBLAZE KFS
  - Распределенная горизонтально масштабируемая отказоустойчивая СХД
  - Доступ к данным файловый
  - Программные возможности: масштабируемость, репликация
  - Поддержка клиентов в т.ч. для Astra Linux 1.7 и Astra Linux 8.1 («Ленинград»)
- BLAZEX (доступно со 2 квартала 2024 г.)
  - Доступ к данным блочный, файловый
  - Программные возможности: снимки, тонкие тома, виртуализация, поддержка двухконтроллерного режима
  - Производительность до 5 000 000 IOPS





Объединенная  
приборостроительная  
корпорация



АСТРА

# Программная платформа





## Сегмент x86\_64 (AMD EPYC)

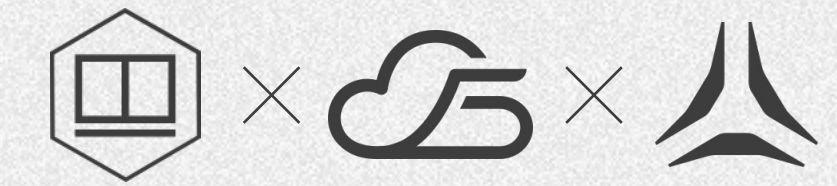
- ▮ ОС Astra Linux 1.7 (x86)
- ▮ ПК СВ «Брест» 3.2 (3.3)
- ▮ Termidesk 4.2
- ▮ RuBackup 2.0.49
- ▮ ПО сети «Ангара» (драйвер, библиотеки, утилиты)
- ▮ MPICH v3.4.3
- ▮ SLURM v18.08.8
- ▮ Клиент KFS (kfs-mount v1.3.4.10)

## Сегмент E2K («Эльбрус-8С/СВ»)

- ▮ ОС Astra Linux 8.1.4.22 (e2k)
- ▮ Компилятор lcc 1.23.27
- ▮ ПО сети «Ангара» (драйвер, библиотеки, утилиты)
- ▮ MPICH v3.4.3
- ▮ SLURM v18.08.8
- ▮ Клиент KFS (kfs-mount v1.3.2.13)



# Кластер виртуализации (ПК СВ «Брест» 3.2)



## 3 мастера (RAFT)

- sp3-host1, sp3-host2, sp3-host3

## 5 узлов виртуализации

- sp3-host1, sp3-host2, sp3-host3
- asus-host1, asus-host2

## Инфраструктурные VM

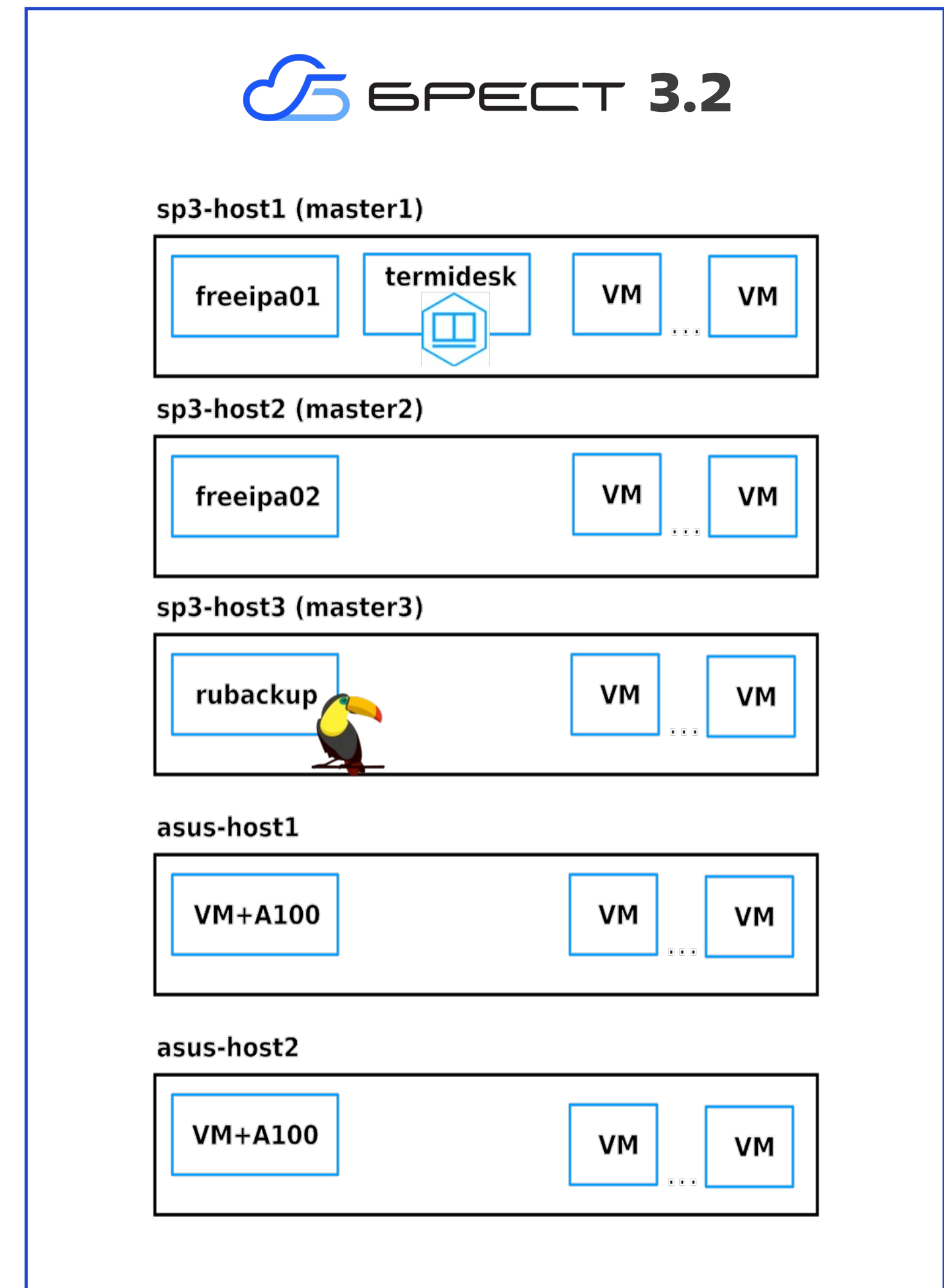
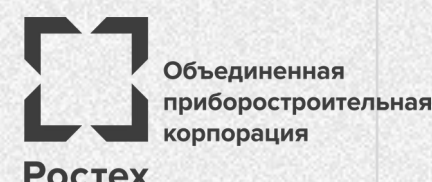
- freeipa01, freeipa02
- termidesk (v4.2)
- rubackup (v2.0.49)
- zabbix

## Датасторы

- image/system (ceph\_ds)
- image/system (qcow2)

## Виртуальные сети

- vnet100 (/24), vnet101 (/24),...

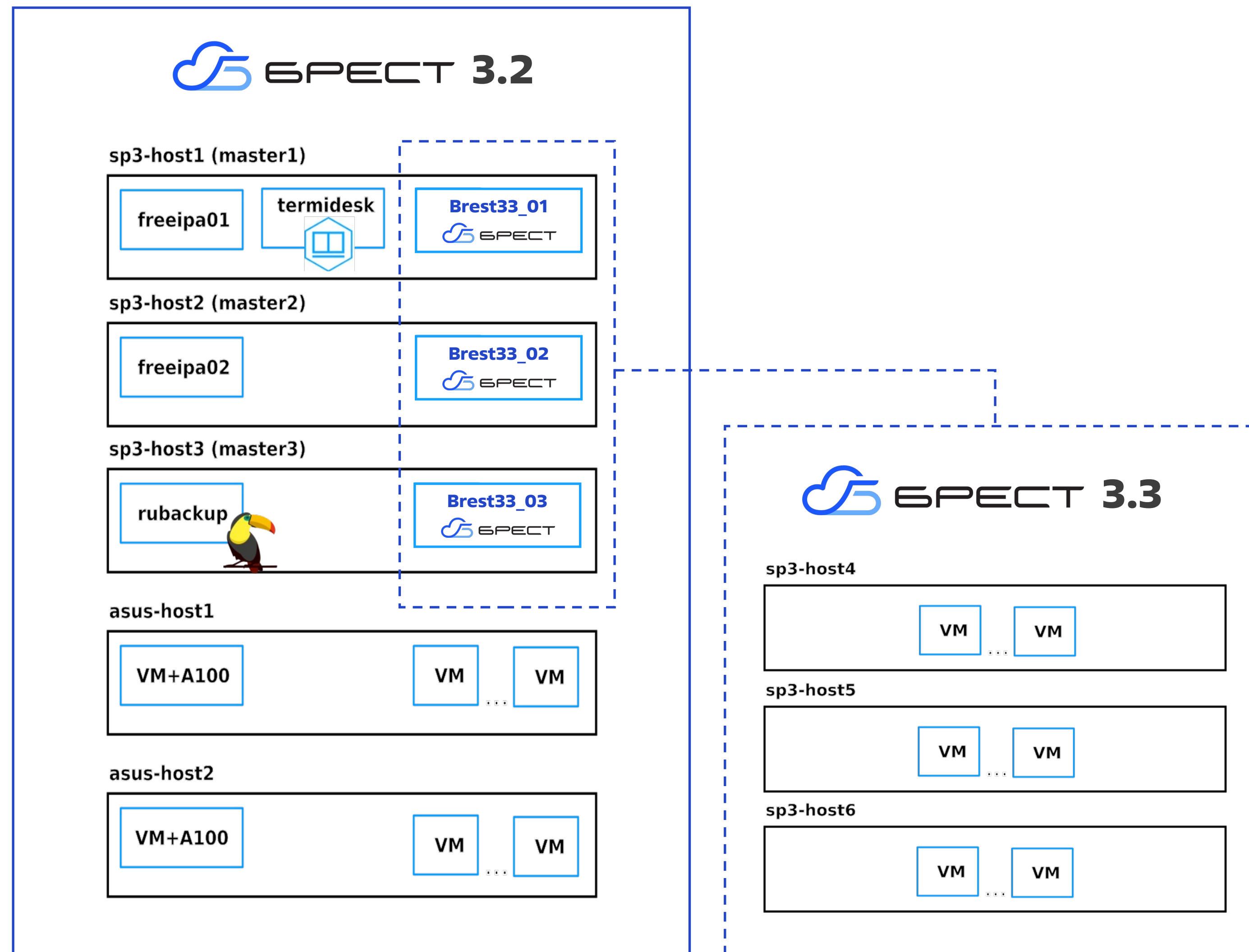




# Кластер виртуализации

## (тестирование ПК СВ «Брест» 3.3)

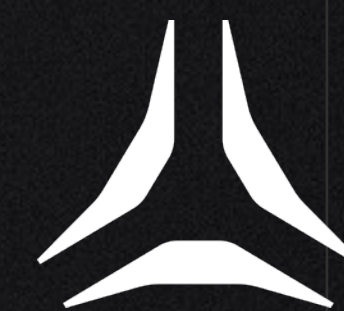
- ▮ **3 мастера (RAFT)**
  - brest33\_01, brest33\_02, brest33\_03
  
- ▮ **3 узла виртуализации**
  - sp3-host4, sp3-host5, sp3-host6
  
- ▮ **Датасторы**
  - image/system (ceph\_ds)
  
- ▮ **Виртуальные сети**
  - vnet110 (/24), vnet111 (/24),...







Объединенная  
приборостроительная  
корпорация



АСТРА

# Тестирование





# Тестирование VDI (vdi\_test\_helper)



## Утилита vdi\_test\_helper

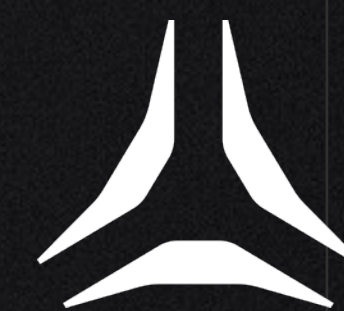
- Интерактивный CLI-интерфейс
- Размещение окон приложений в виде сетки на рабочем столе
- Управление множеством клиентов (termidesc-client)
  - Подключение к серверу
  - Запрос BPM
- Управление множеством SPICE / VNC-клиентов (remote-viewer)
  - Выполнение логина в гостевую ОС
  - Запуск fly-term
  - Отправка команд через fly-term
  - Подключение к агенту и отправка команд через shell







Объединенная  
приборостроительная  
корпорация



АСТРА

# Демонстрация







Объединенная  
приборостроительная  
корпорация



# Спасибо за внимание!

## Вопросы?

**Александр Фролов**

Начальник отдела, АО «НИЦЭВТ»

**Контакты:**

117587, Москва, Варшавское ш, 125

[angara@nicevt.ru](mailto:angara@nicevt.ru)

[frolov@nicevt.ru](mailto:frolov@nicevt.ru)

