



Ростех

Объединенная
приборостроительная
корпорация

Evaluation of the Angara Interconnect Prototype TCP/IP Software Stack: Implementation, Basic Tests and BeeGFS Benchmarks

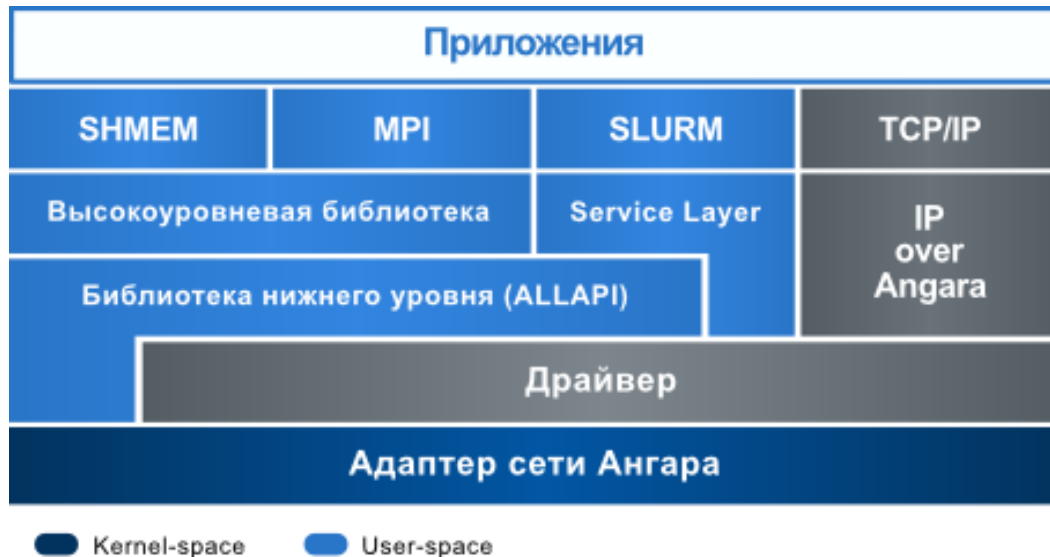
Yuri Goncharuk, Yuri Grishichkin, Alexander Semenov,
Vladimir Stegailov, Vasily Umrikhin

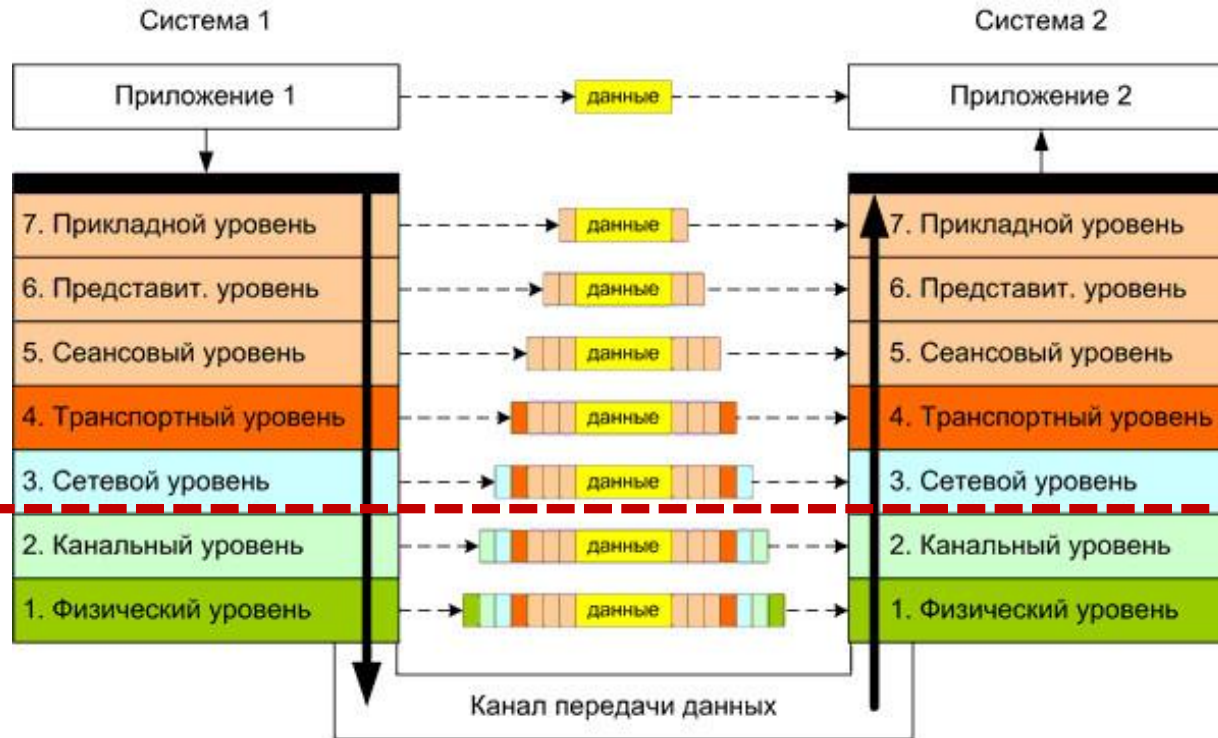
JSC NICEVT, JIHT RAS

27.09.2022

Поддерживаемое программное обеспечение

- Версии ядра Linux от 2.6.21 до 5.11
- ОС:
Astra Linux SE 1.3-1.7, ОС «Эльбрус»,
AltLinux 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
- Инженерные пакеты и приложения:
ANSYS 17.0, 18.2, 21r1, StarCCM+ 12.02,
16.06.008r8, Flowvision, OpenFOAM, VASP,
GROMACS, LAMMPS и многие другие
- Система виртуализации ECP Veil
- Стек протоколов TCP/IP v2.1
- Параллельные файловые системы:
Ceph, Lustre, GlusterFS, BeeGFS





TCP/IP driver

Ethernet over Angara
Network device driver

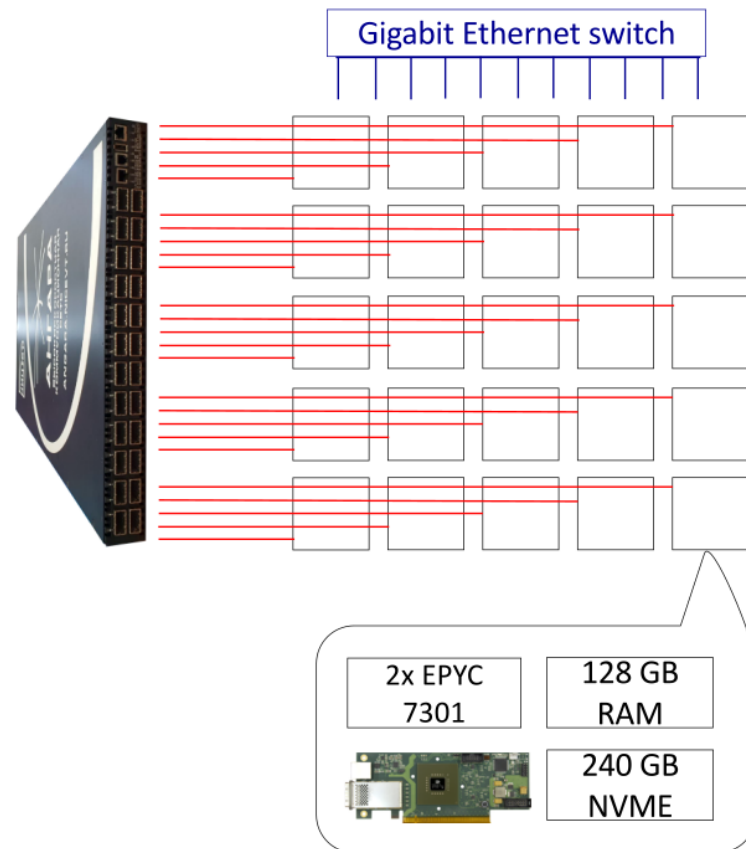
- Логическое соединение между уровнями
- Реализация передачи данных

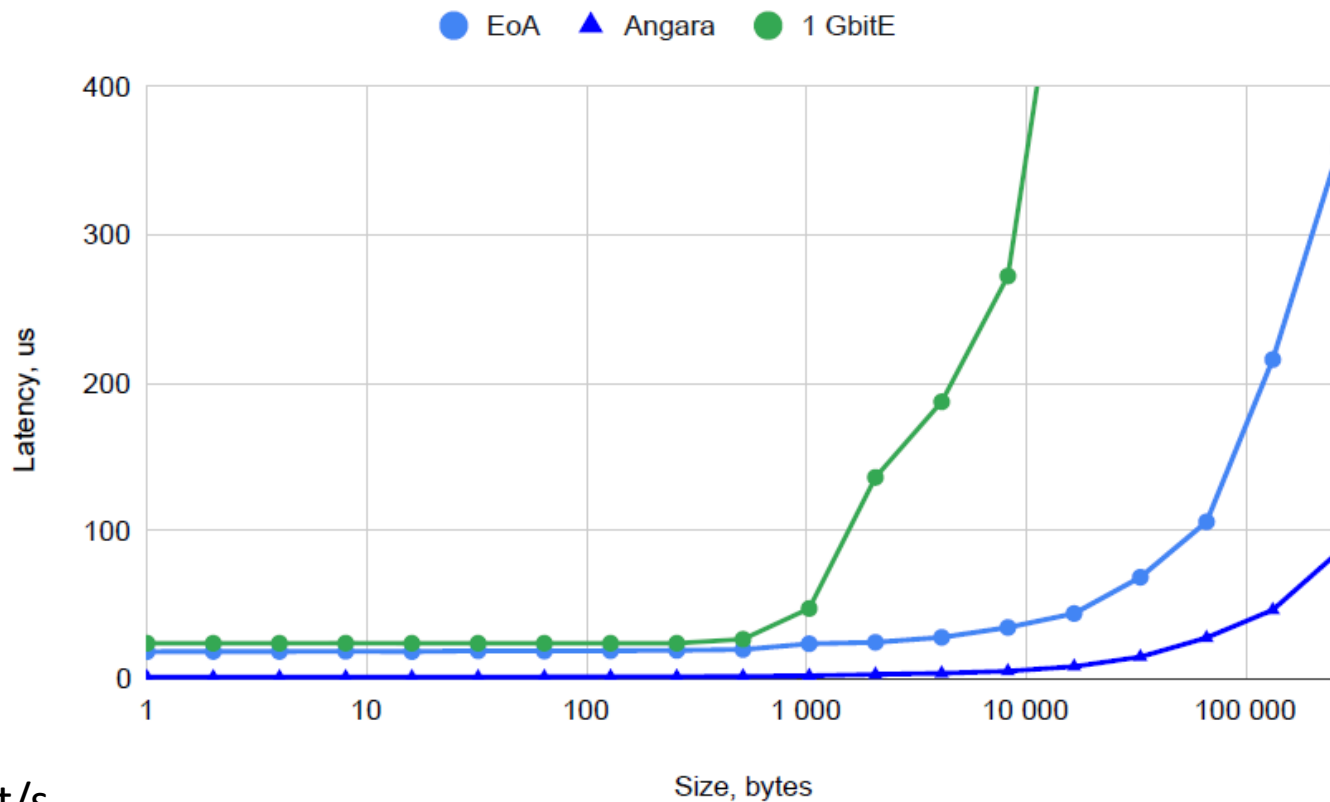
Hardware

Number of nodes	20
SSD M2 NVMe	Apacer AS2280P2

Software

OS	OpenSUSE Leap 15.2
Kernel	5.3.18-lp152.87- preempt
Ethernet over Angara	2.1
BeeGFS	7.2.3

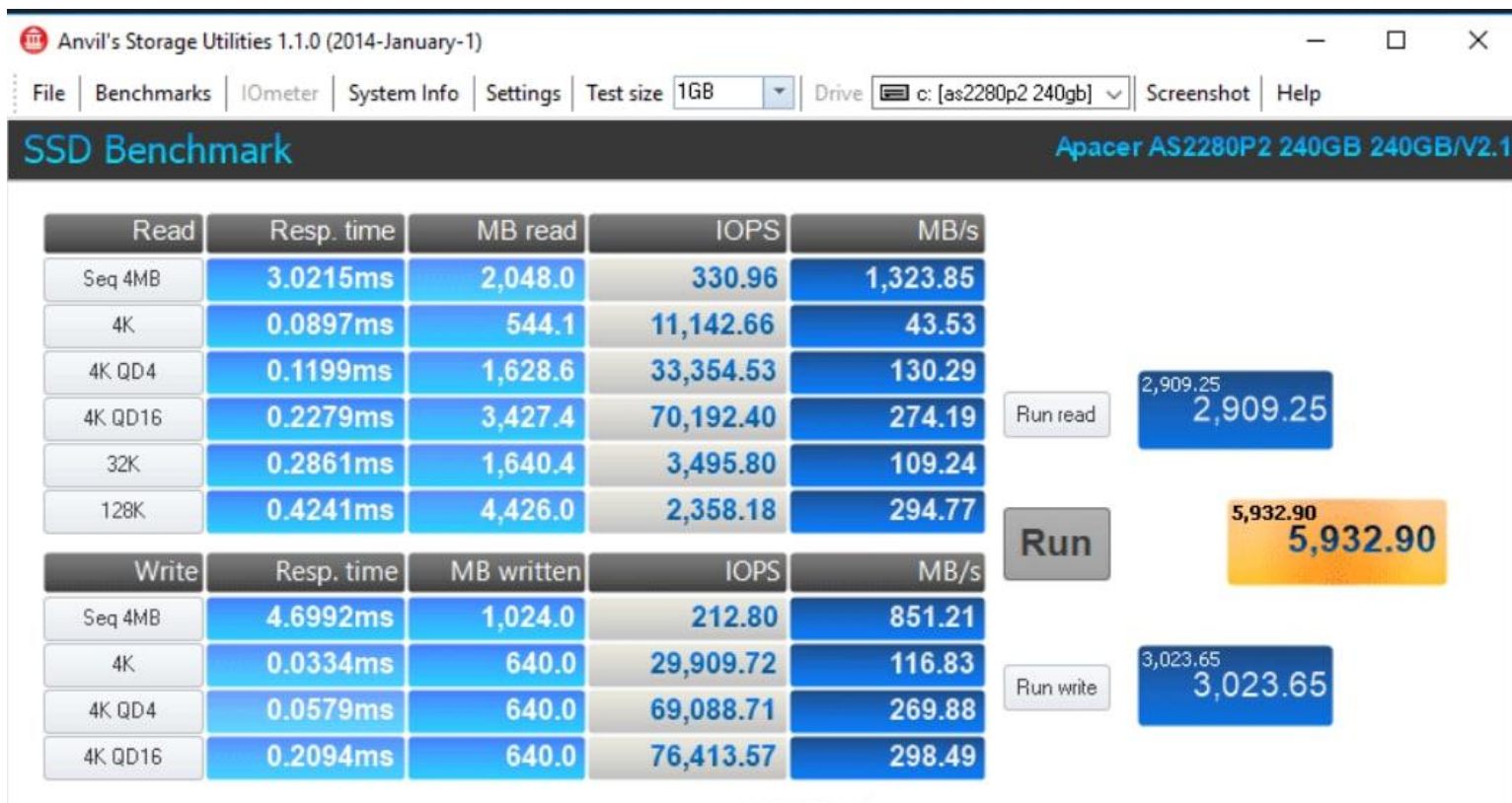


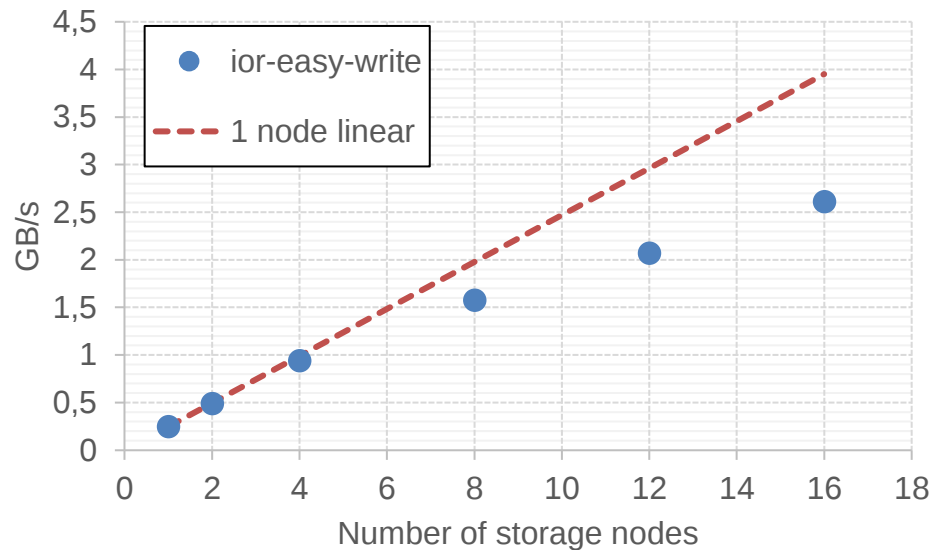
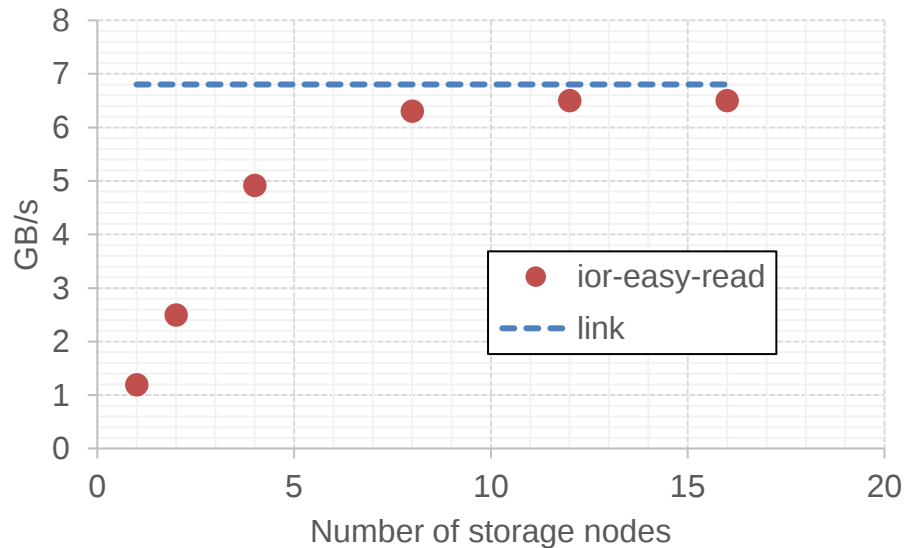


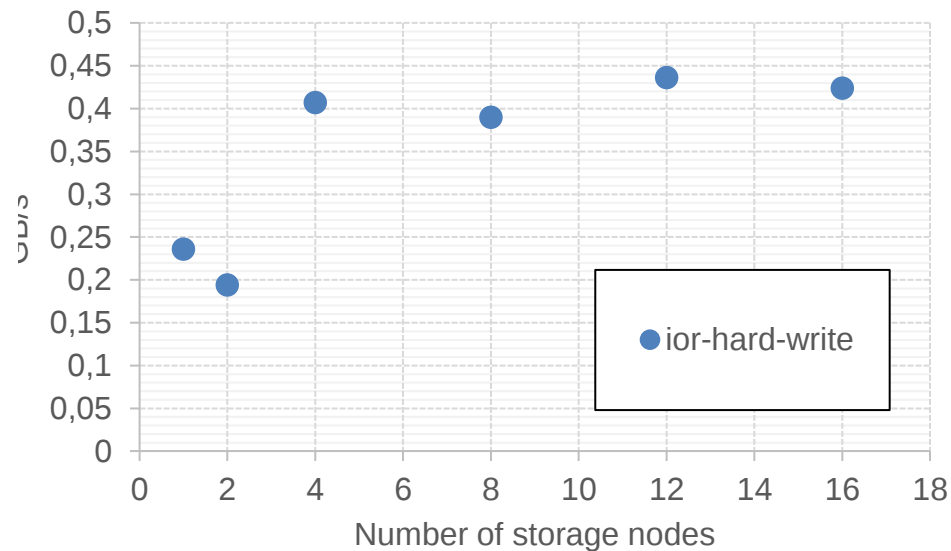
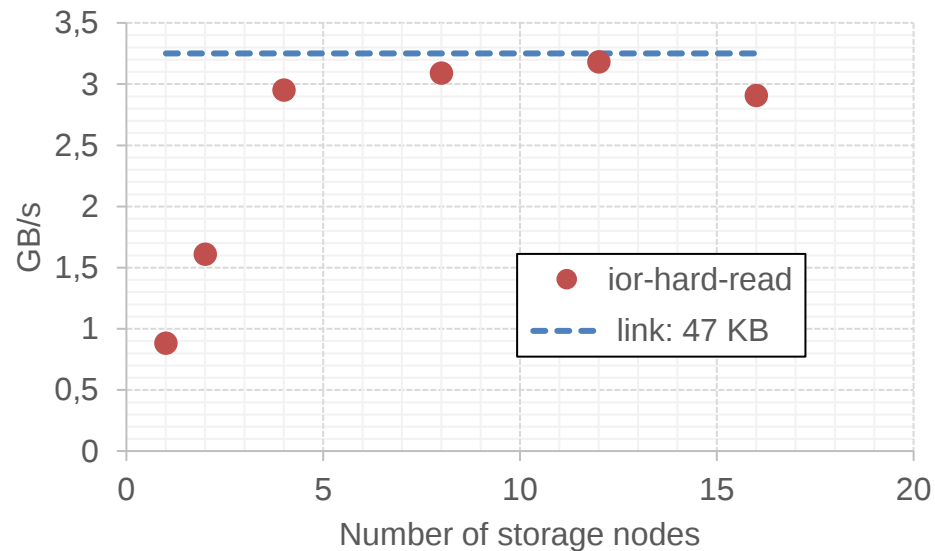
iperf3 = 15 Gbit/s

Компонент	Тесты	Описание
IOR 'easy'	ior_easy_write, ior_easy_read	Пропускная способность для больших последовательных IO операций
IOR 'hard'	ior_hard_write, ior_hard_read	Пропускная способность для невыровненных (47001 байт) операций от каждого клиентского процесса к одному файлу
mdtest 'easy'	mdtest_easy_delete, mdtest_easy_stat, mdtest_easy_write	Операции с метаданными файлов 0 размера при помощи отдельной директории для каждого MPI-процесса
mdtest 'hard'	mdtest_hard_delete, mdtest_hard_stat, mdtest_hard_write, mdtest_hard_read	Операции с метаданными файлов небольшого размера (3901 байт) в разделяемой директории
Find	Find	Поиск файлов при помощи прохода

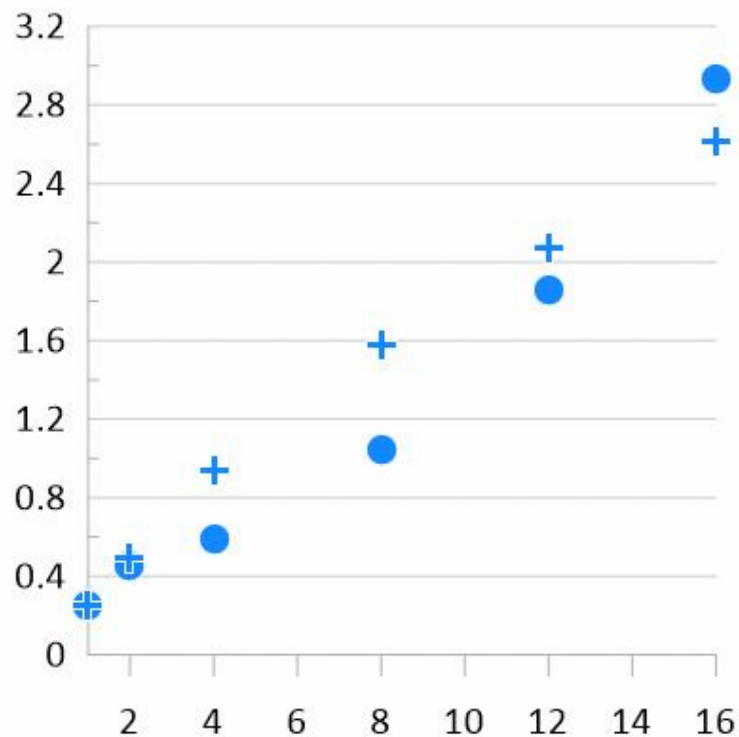




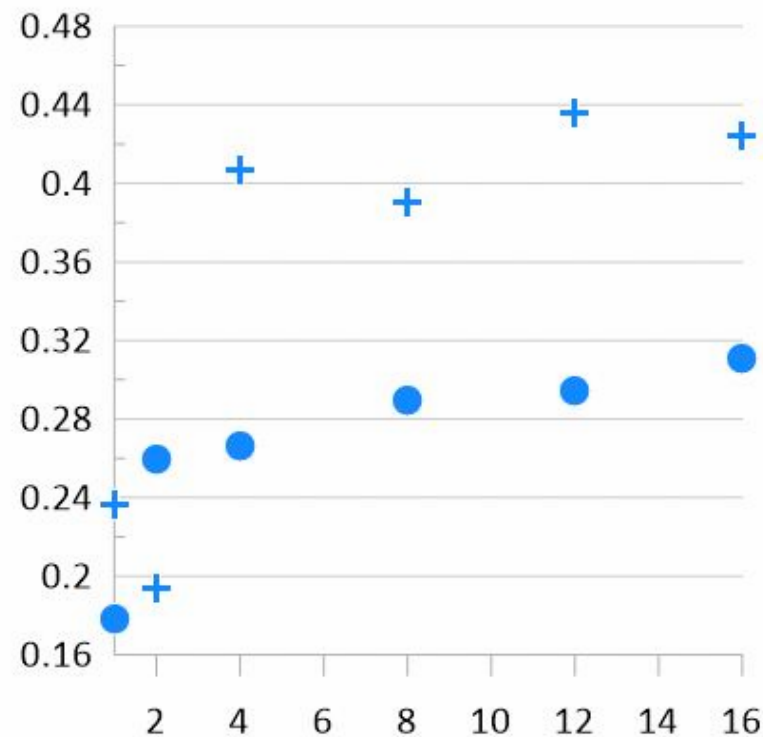


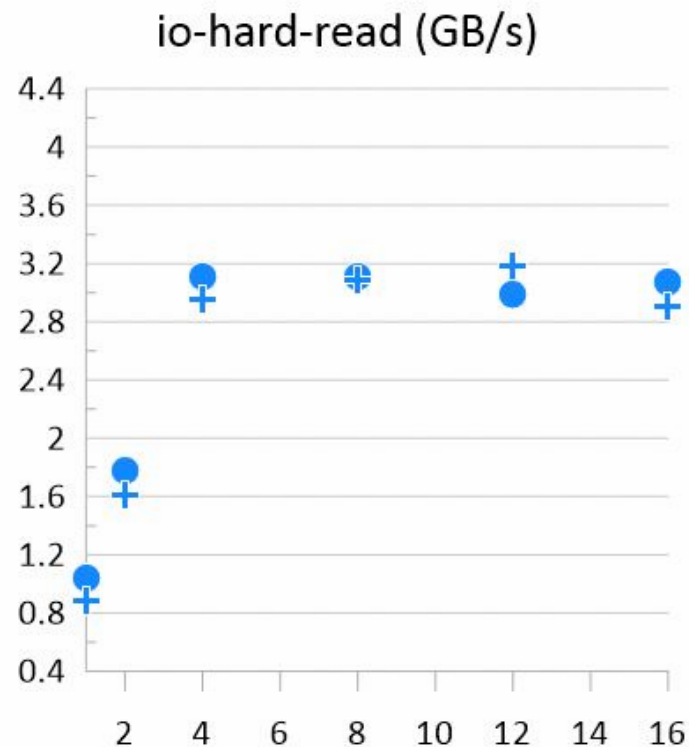
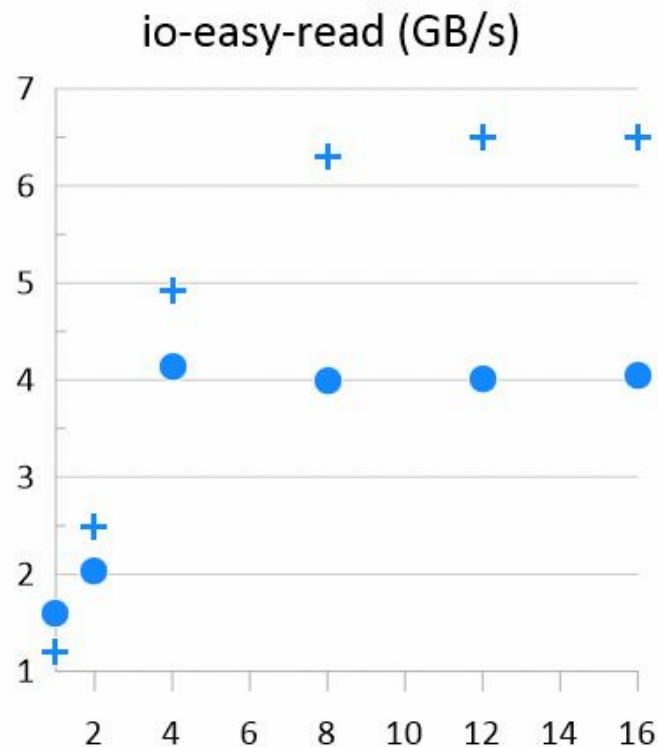


io-easy-write (GB/s)

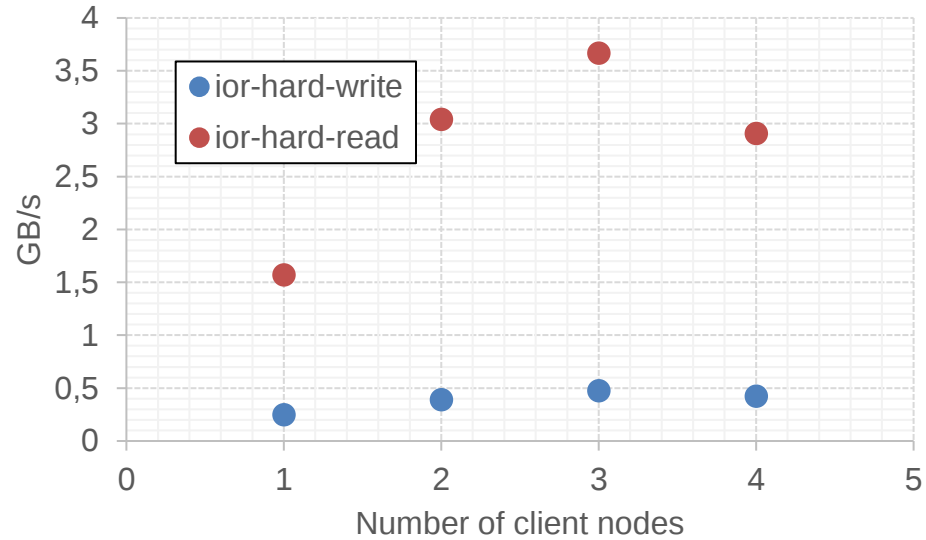
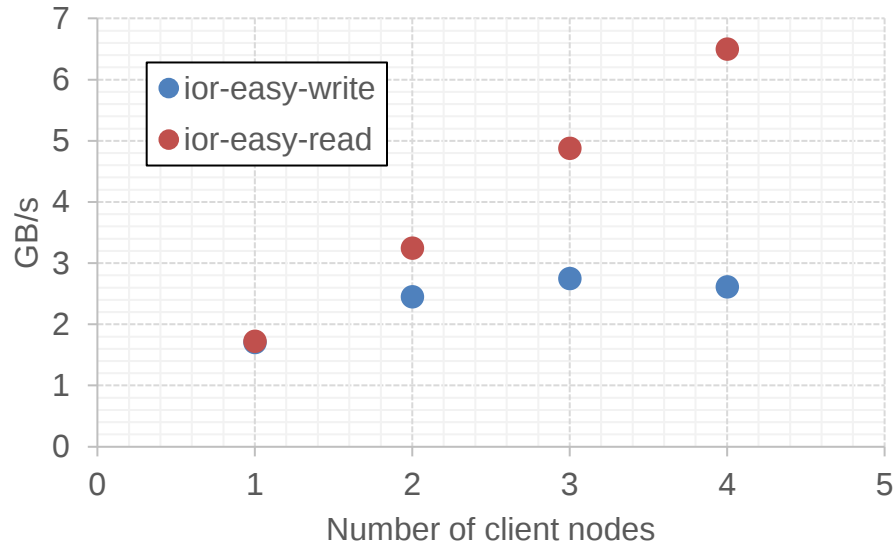


io-hard-write (GB/s)

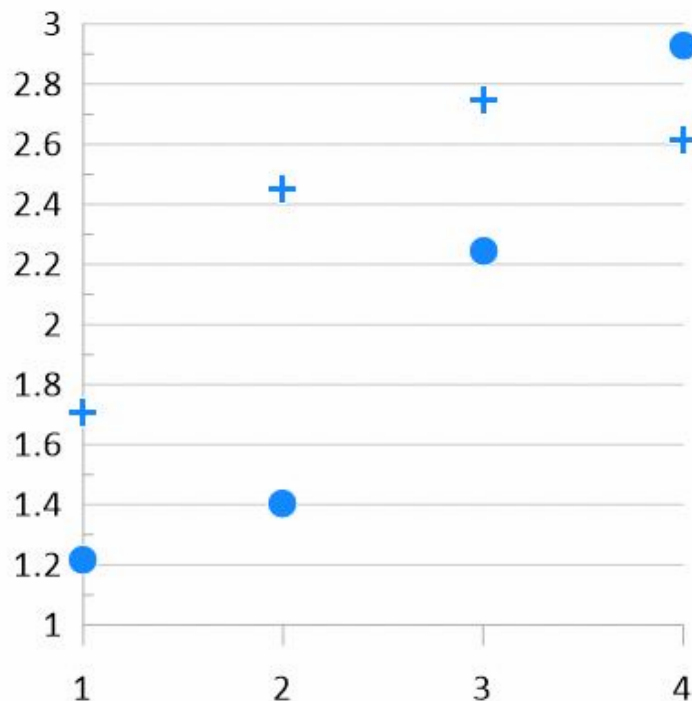




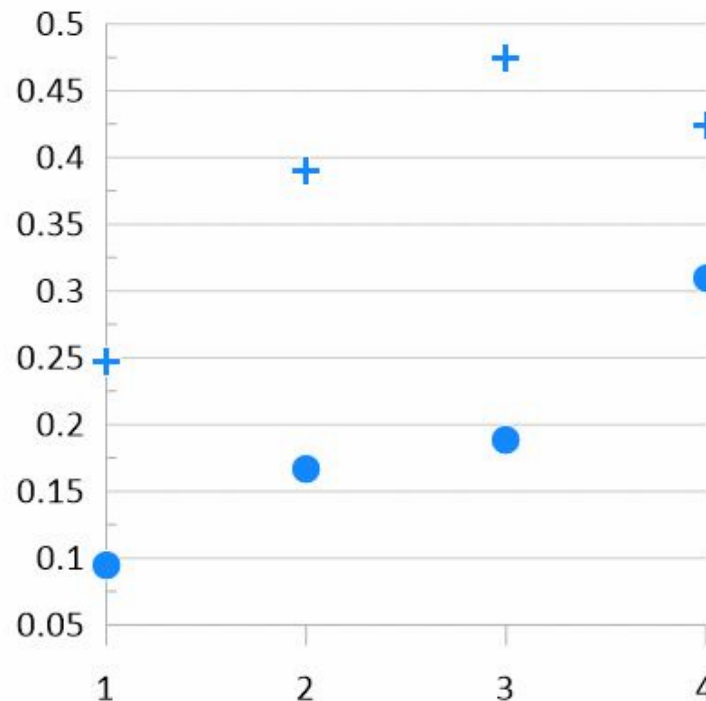
Number of storage nodes



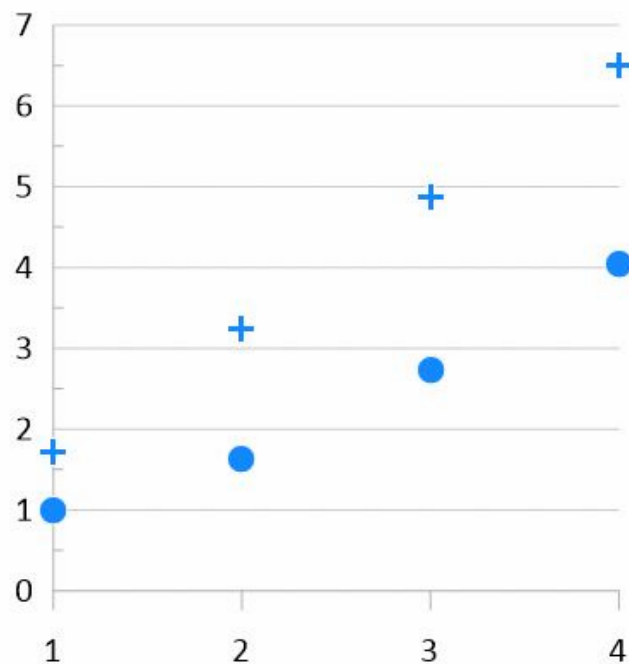
io-easy-write (GB/s)



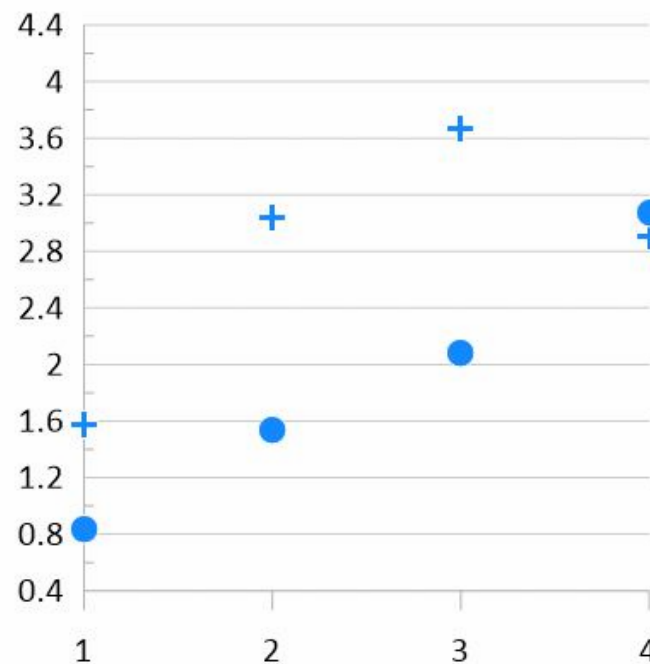
io-hard-write (GB/s)



io-easy-read (GB/s)



io-hard-read (GB/s)



Number of client nodes

Тест	Характеристика	1 Gbit Ethernet	ЕоА	ЕоА / 1 Gbit Ethernet
ior-easy-write	GiB/s	0,459	2,612	5,69
mdtest-easy-write	KIOPS	15,180	17,170	1,13
ior-hard-write	GiB/s	0,171	0,419	2,44
mdtest-hard-write	KIOPS	5,101	3,808	0,75
find	KIOPS	175,189	124,549	0,71
ior-easy-read	GiB/s	0,448	6,498	14,50
mdtest-easy-stat	KIOPS	74,599	73,687	0,99
ior-hard-read	GiB/s	0,456	2,908	6,38
mdtest-hard-stat	KIOPS	67,311	70,303	1,04
mdtest-easy-delete	KIOPS	14,203	10,181	0,72
mdtest-hard-read	KIOPS	14,592	15,033	1,03
mdtest-hard-delete	KIOPS	4,187	5,952	1,42
SCORE				
Bandwidth	GiB/s	0,356	2,132	5,99
IOPS	KIOPS	22,205	21,042	0,95
Total		2,812	6,698	2,38

- Продемонстрировано стабильное функционирование TCP/IP (ЕоА) на суперкомпьютере из 20 узлов
- Дальнейшие планы:
 - Оптимизация реализации драйвера ЕоА
 - Задержка передачи сообщений
 - Пропускная способность на пакетах небольшого размера
 - Уже получена скорость 25 Гбит/с, ориентир – 50 Гбит/с

Контакты:

117587, Москва, Варшавское ш, 125

angara@nicevt.ru

