



ЦОД ПИК

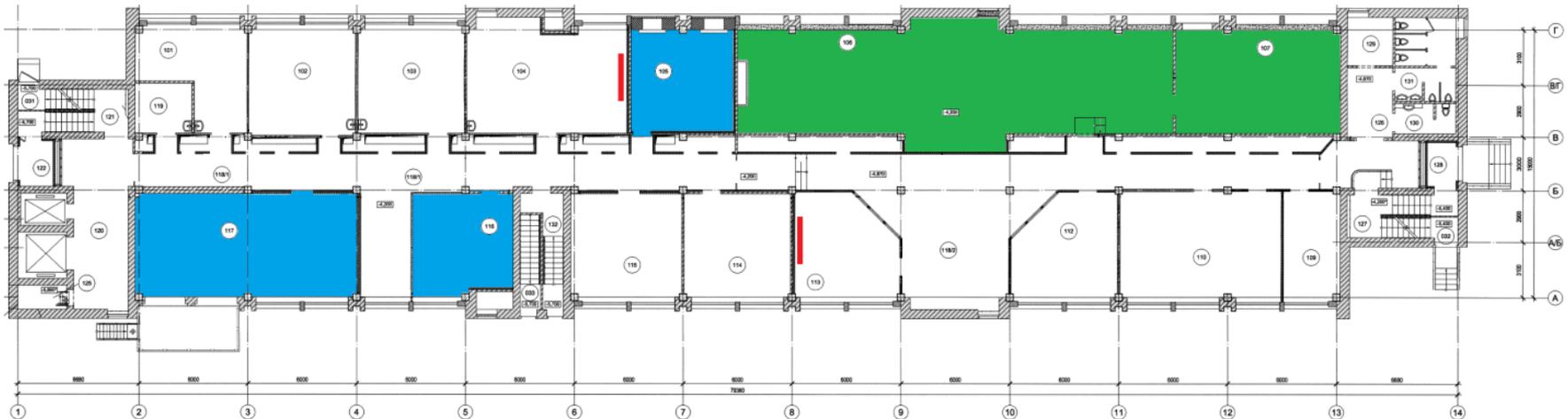
– Оборудование, возможности и текущее состояние





Общие сведения

- Центр обработки данных находится на первом этаже здания 105 РК ПИК
- Непосредственно оборудованием заняты помещения 106 (вычислитель) и 107 (ИБП)
- На площадке РК ПИК также располагаются ДГУ и элементы системы хладоснабжения (чиллеры, баки)





Система хладоснабжения



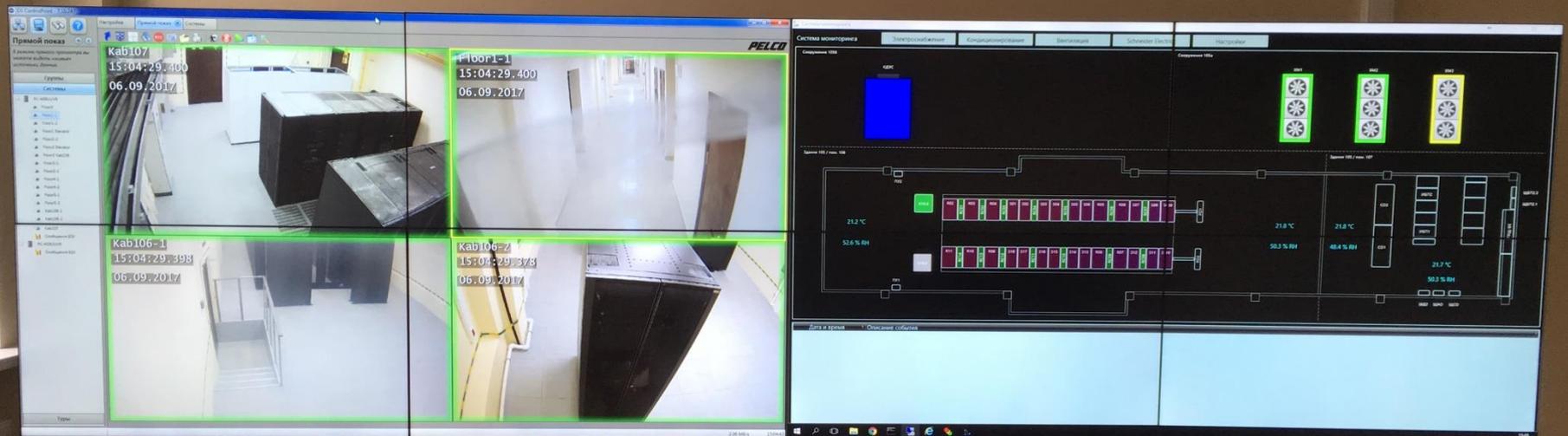


Инженерная инфраструктура

- Кондиционирование и электропитание рассчитано на максимальную потребляемую мощность 300 кВА
 - При такой мощности время работы вычислителя от ИБП ~10 мин
 - Система хладоснабжения имеет собственный ИБП и ДГУ для чиллеров. Время работы кондиционеров от ИБП около получаса.
- В настоящее время максимальное потребление ЦОД составляет порядка 120 кВА
 - Загрузка стоек оборудованием менее 50% (загружено 12 из 28 стоек)
 - Сетевая инфраструктура общего назначения построена из расчёта полной загрузки стоек оборудованием
 - Возможно расширение ЦОД без инженерных изменений с подключением оборудования к питающей и телекоммуникационной сетям



Система мониторинга и видеонаблюдения





Вычислительные мощности

- Теоретическая общая пиковая производительность составляет ~364 Тфлопс
 - На тестах достигнуты показатели ~204 Тфлопс на процессорах общего назначения и ~68 Тфлопс на векторных процессорах
- Вычислительный модуль состоит из следующих структурных элементов:
 - Система вычислительных узлов общего назначения: 160 узлов с 28 ядрами и 128 ГБ памяти на каждом (4,5 ГБ на ядро)
 - Система вычислительных узлов с векторными процессорами: 40 узлов с 68 ядрами (272 виртуальных)
 - Система вычислительных узлов с большой памятью:
 - 2 узла с 28 ядрами и 1,5 ТБ памяти на каждом
 - 16 узлов с 28 ядрами и 1 ТБ памяти + 1,6 ТБ энергонезависимой памяти (NVMe) на каждом



Помещение ЦОД





<http://top50.supercomputers.ru/>

6	Москва Центр обработки данных НИЦ "Курчатовский институт" 2015 г.	774/11082	узлов: 364 (2xXeon E5-2680v3 2.5 GHz 128 GB RAM) узлов: 23 (2xXeon E5-2680v3 [Acc: 3x Nvidia K80] 2.5 GHz 128 GB RAM) сеть: Infiniband FDR/Gigabit Ethernet/Gigabit Ethernet	374.13	500.55	SuperMicro, Борлас
7	Нижний Новгород Нижегородский государственный университет им. Н.И. Лобачевского 2014 г.	360/30760	узлов: 100 (2xXeon E5-2660 [Acc: 3x Kepler K20X] 2.2 GHz 65.536 GB RAM) узлов: 50 (2xXeon E5-2660 [Acc: 2x Fermi 2090] 2.2 GHz 65.536 GB RAM) узлов: 10 (2xXeon E5-2660 [Acc: 2x Phi 5110P] 2.2 GHz 65.536 GB RAM) узлов: 10 (2xXeon E5-2660 [Acc: 3x Fermi 2090] 2.2 GHz 65.536 GB RAM) узлов: 10 (2xXeon E5-2660 2.2 GHz 65.536 GB RAM) сеть: Infiniband QDR/Infiniband QDR/Gigabit Ethernet	289.50	573.00	Ниагара Компьютерс
8	Челябинск Южно-Уральский государственный университет 2013 г.	768/28032	узлов: 384 (2xXeon X5680 [Acc: Xeon Phi SE10X] 3.33 GHz 24.576 GB RAM) сеть: Infiniband QDR/Infiniband QDR/Gigabit Ethernet	288.20	473.64	Группа компаний РСК
9	Гатчина ОИВРТ НИЦ "Курчатовский институт" - ПИЯФ 2017 г.	496/9024	узлов: 160 (2xXeon E5-2680v4 2.4 GHz 128 GB RAM) узлов: 40 (Xeon Phi 7250 1.4 GHz 112 GB RAM) узлов: 30 (2xXeon E5-2680v4 2.4 GHz 256 GB RAM) узлов: 20 (2xXeon E5-2650v4 2.2 GHz 256 GB RAM) узлов: 16 (2xXeon E5-2680v4 2.4 GHz 1024 GB RAM) узлов: 2 (2xXeon E5-2680v4 2.4 GHz 1540 GB RAM) сеть: Infiniband EDR/Gigabit Ethernet/Gigabit Ethernet	200.44	362.38	NP-IT, Ниагара компьютерс



Помещение ИБП ЦОД





Передача и хранение данных

- Системы хранения данных
 - Система хранения данных для параллельных вычислений полным объёмом 2,9 ПБ (реальный объём 2,3 ПБ + 29 ТБ метаданных)
 - Система блочного хранения данных полным объёмом 2,5 ПБ
- Сетевая инфраструктура
 - Система передачи данных для параллельных вычислений со скоростью 100 Гбит/с (InfiniBand EDR)
 - Сеть общего назначения со скоростью от 1 до 10 Гбит/с
 - Сеть управления со скоростью 1 Гбит/с
 - Телекоммуникационный узел с возможностью подключения к внешним сетям на скорости 10 Гбит/с



Устройство системы хранения





Программное обеспечение

- Основная программная платформа: CentOS 7, 64-bit
 - Планировщик заданий: Slurm
 - Мониторинг: Ganglia
 - Библиотеки MPI: Open MPI, Intel MPI, Platform MPI
 - Компиляторы: GCC, Intel Compiler
- В настоящее время для всех задач используется бесплатное программное обеспечение
- Для более эффективного использования инфраструктуры можно закупить платное системное ПО. Этот вопрос в настоящий момент прорабатывается.



Коммутатор InfiniBand и СХД



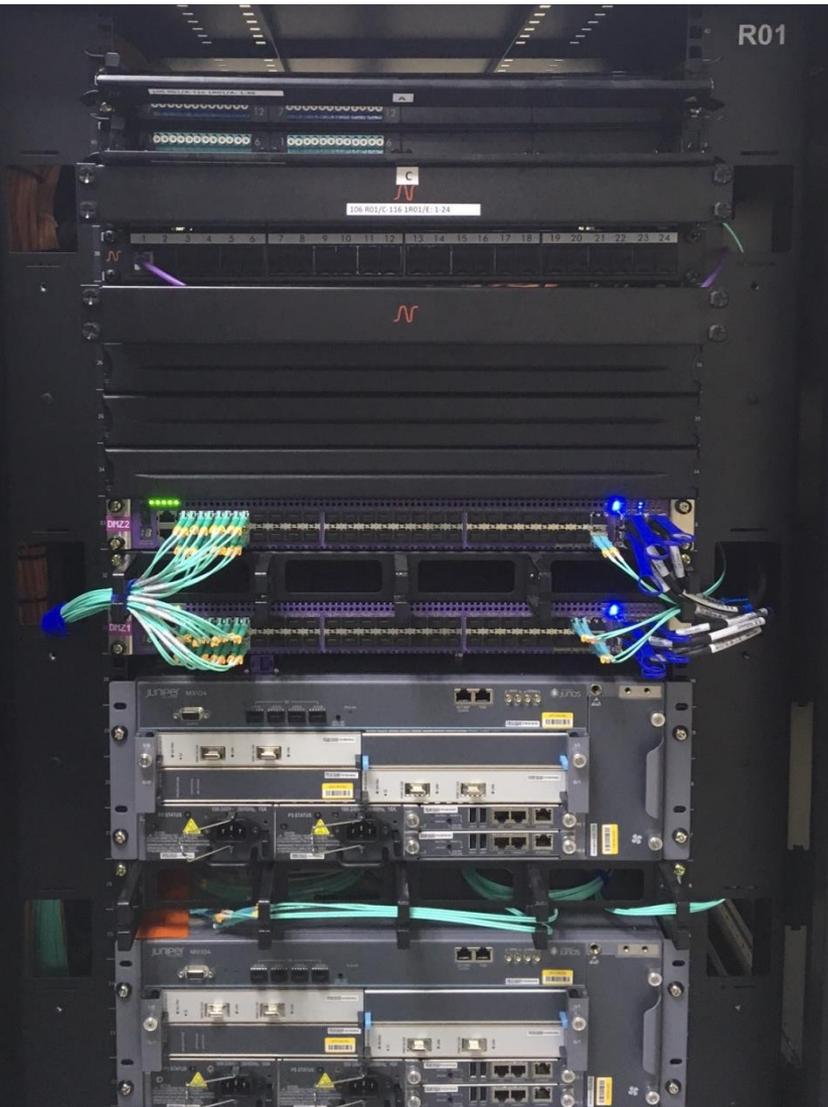


Текущее состояние

- В настоящий момент оборудование ЦОД полностью смонтировано и прошло испытания
- Произведена прокладка оптоволоконного кабеля для связи ЦОД с телекоммуникационными узлами на к. 9 (АТС) и к. 7
- Основное программное обеспечение установлено и производится его настройка
- Заканчиваются строительные работы в здании ЦОД
- К началу 2018-го года ЦОД будет полностью готов для выполнения всех функций



Телекоммуникационный узел





Проблемный совет по научной программе ЦОД ПИК

- Научная программа ЦОД
 - Потребности научных коллективов в интенсивных вычислениях
 - Как и где сейчас решаются вычислительные задачи?
- Использование нестандартных ресурсов
 - Большая память
 - Векторные вычисления
- Прикладное программное обеспечение
 - Что требуется развернуть?
 - Что нужно закупить, под какими лицензиями?
- Хранение научных данных
 - Объём?



Спасибо за внимание!