

Универсальность решений РСК



Александр Московский,
Генеральный директор «РСК Технологии»,
Группа компаний РСК

Суперкомпьютерные дни в России,
Москва
26 сентября 2016 г.



Специализация

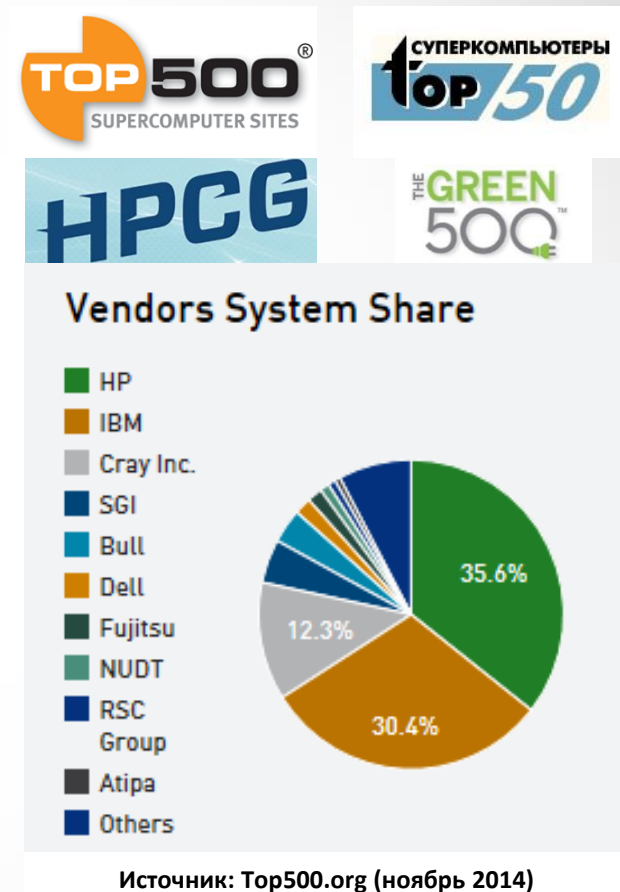
**Разработка инновационных энергоэффективных
высокопроизводительных и высокоплотных
вычислительных систем, обладающих уникальным
набором функций**

**Передовые суперкомпьютеры и вычислительные центры
для самых требовательных клиентов**



Лидерство на рынке

- Ведущий разработчик и интегратор HPC-решений в России и СНГ
- Ведущий российский производитель:
 - 9 систем в рейтинге **Top50** - Россия/СНГ,
 - 4 системы в **Top10** (#3 СПбПУ) - Россия/СНГ,
 - более 40%** всех российских систем в **Top500** и **Green500**,
 - более 70%** всех российских систем в рейтинге **HPCG**
- Самая энергоэффективная система в России (СПбПУ)
- Один из ведущих HPC-вендоров в регионе EMEA
- В **Top10** мировых производителей согласно **Top500** (#9)
- Элитный статус **Intel HPC Data Center Specialist**
- **Более 7 лет успешной работы** с ведущими российскими университетами, государственными учреждениями, промышленным сектором и др.





Сделано в России

- Российский разработчик и системный интегратор с 2009 г.
- Все центры разработки находятся в РФ
- Производится на российских предприятиях, интегрируется в России под ключ
- Уникальные технологии, не имеющие мировых аналогов или на уровне лучших в мире
- Защищено патентами РФ и международными (Германия, Китай, США, Япония): система прямого жидкостного охлаждения серверов, сверхплотная монтажная стойка
- Российская разработка: система охлаждения, система мониторинга и управления, интегрированный стек ПО «РСК БазИС», эффективное электропитание и управление серверными фермами
- Готовность использовать отечественные процессоры и коммуникационные сети
- Торговые марки PCK, RSC и PetaStream защищены в России, США, Китае, Японии, Корее и в странах ЕС





Технологии и решения

Кластерная архитектура
«**PSK Торнадо**»



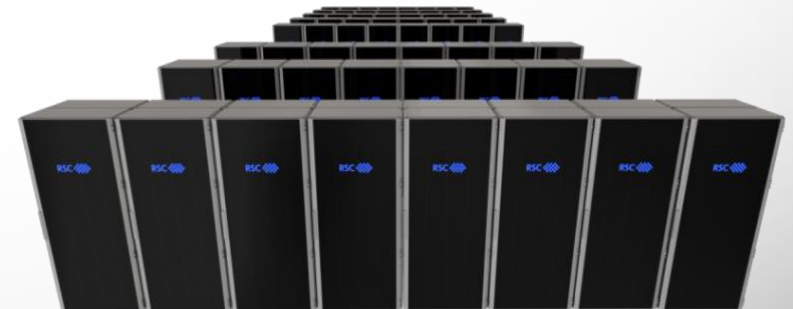
до **100 кВт** и **153 узла**,
528 ТФЛОПС
на шкаф 42U (1,28 м³)



Массивно-параллельная
архитектура **RSC PetaStream**



до **400 кВт** и **1024 узла**,
1,2 ПФЛОПС
на шкаф (2,0 м³)



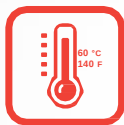
Мировые
рекорды

Вычислительная плотность
Энергетическая плотность
Компактность



«PSK ТОРНАДО» - ПРОГРАММНО-ОПРЕДЕЛЯЕМЫЕ РЕШЕНИЯ

МИРОВЫЕ РЕКОРДЫ



Вх. жидкость



100% жидкостное охлаждение



96% Eff.



плотность



"PSK Торнадо"

Узел на базе Intel Xeon

- Два Intel Xeon E5-2600 v4 (вкл. top bin)
- До 256 ГБ DDR4-2400 O3U
- EDR IB, Omni-Path, 10/40/100 GigE
- 2x SATA и 1x NVMe SSDs



"PSK Торнадо Phi"

Узел на базе Intel Xeon Phi

- Intel Xeon Phi 7200(F) (вкл. top bin)
- До 192 ГБ DDR4-2400 O3U + MCDRAM
- Omni-Path, EDR IB, 10/40/100 GigE
- 2x SATA и 1x NVMe SSDs



Унифицированный шкаф "PSK Торнадо"

- Гибкие опции наполнения:
- До 153 узлов "PSK Торнадо" [237 ТФЛОПС]
- До 153 узлов "PSK Торнадо Phi" [528 ТФЛОПС]
- Смешанное: "PSK Торнадо"/"PSK Торнадо Phi"
- До 9 полностью независимых доменов
- 0,64 м² площадь, 2 м высота шкафа



Пакеты расширения "PSK Торнадо"

- HPC, Big Data, VDI, Security, Машинное обучение
- и другие по запросу



Блоки питания "PSK Торнадо"

230-400 В AC/DC 12 кВт
230-12 В AC/DC 2,1 кВт

- 100% жидкостное охлаждение
- Гибкое управление, формат узла
- Резервирование от N+1 до N+N
- Эффективность до 96%

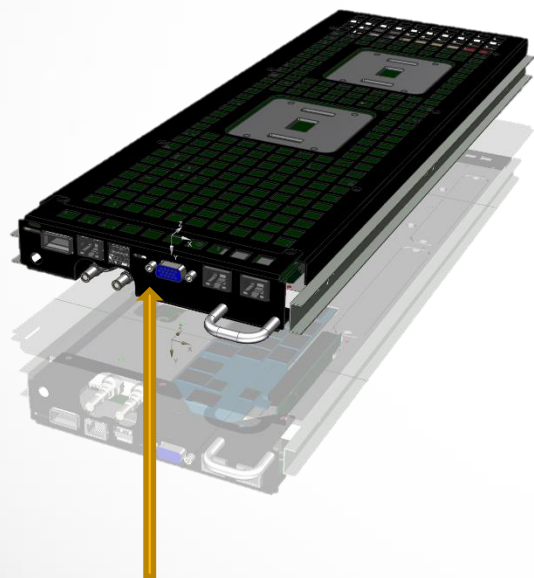


Компактный и легкий узел

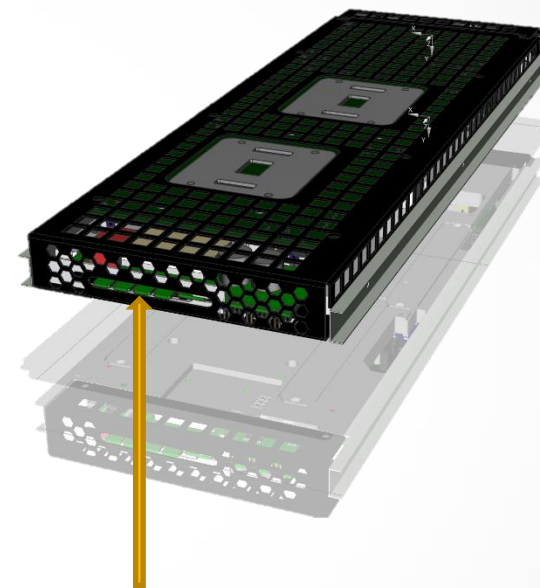
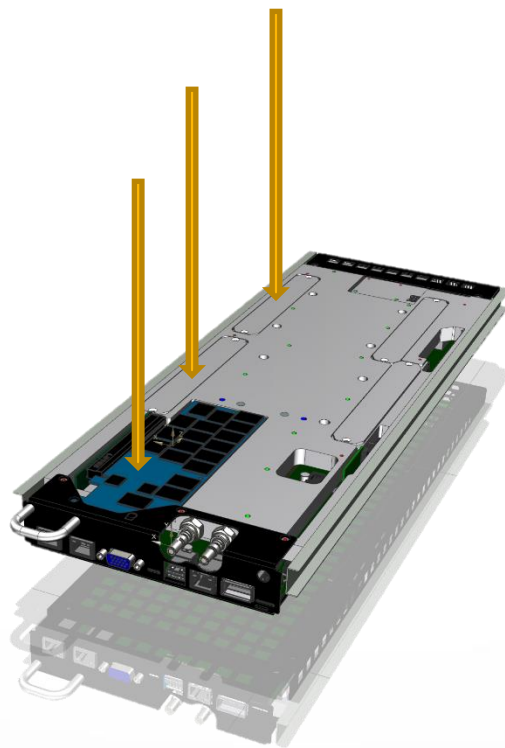
на основе процессоров Intel® Xeon 2699 v4 $\Rightarrow$ 1,54 ТФЛОПС



**Легкий доступ к платам
расширения и модулям памяти
Широкий спектр доступных
компонент**



Защитный кожух

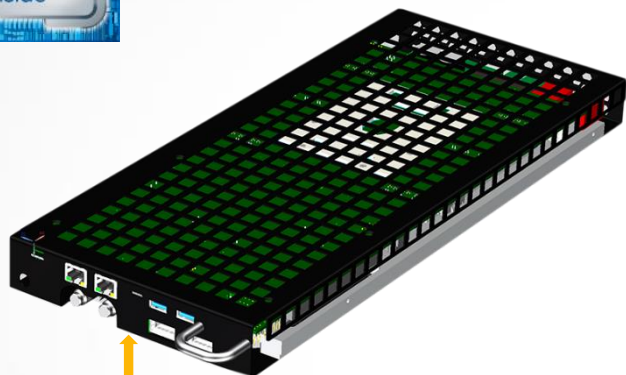


**Унифицированные
разъемы питания для
упрощения
модернизации или
реконфигурирования**



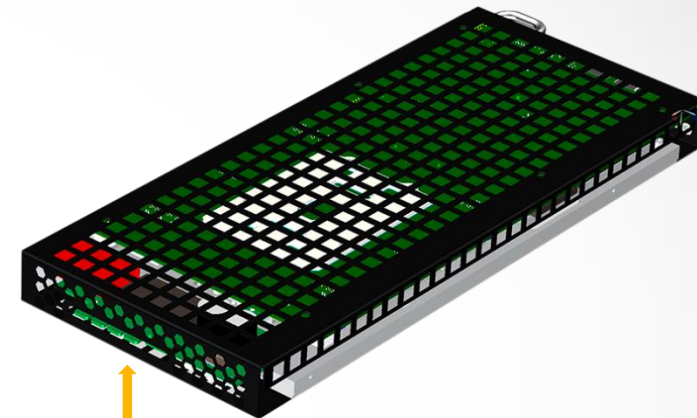
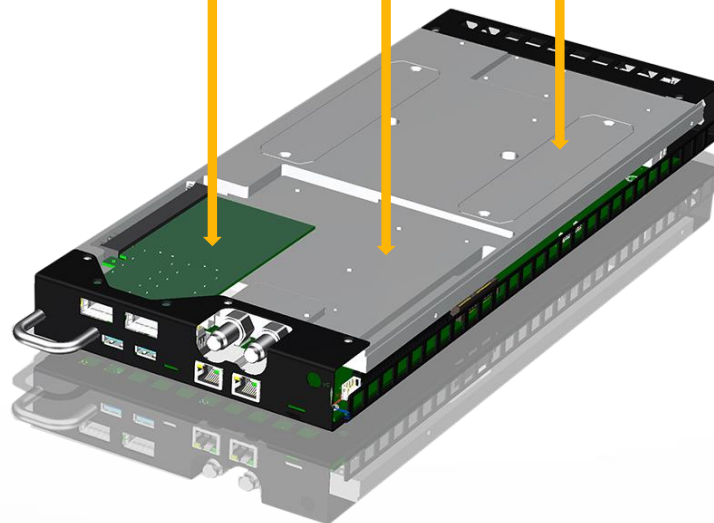
Компактный и легкий узел

на основе процессоров Intel® Xeon Phi™ 7290(F) $\Rightarrow$ 3,46 ТФЛОПС



Защитный кожух

Легкий доступ к платам
расширения и модулям памяти
Широкий спектр доступных
компонент



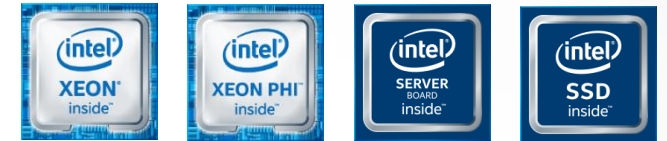
Унифицированные
разъемы питания для
упрощения модернизации или
реконфигурирования



Ресурсы МСЦ РАН

- Первый в России и СНГ проект на базе сопроцессоров Intel® Xeon Phi™ 7120D
- Суммарная пиковая производительность – более 656 ТФЛОПС
- Один из самых крупных в России вычислительных центров
- Многоцелевой комплекс с прямым жидкостным охлаждением

Состав суперкомпьютера



- **Суперкомпьютер МВС-10П:**
 - 208 узлов по 2,5 ТФЛОПС каждый
 - По 2 процессора Intel Xeon E5-2690
 - По 2 сопроцессора Intel Xeon Phi SE10X
 - На узле 64 ГБ памяти DDR3-1600 LV
 - FDR InfiniBand, fat-tree. NetApp & GPFS
- **Архитектура RSC PetaStream:**
 - Процессоры - Intel® Xeon Phi™ 7120D
 - Сервис-процессоры Intel® Xeon® E5-2600
 - Серверные платы Intel
 - Тверд. накопители Intel® SSD DC S3500
 - FDR InfiniBand
 - Система электропитания 400 В (DC)



РСК Расширение системы в МСЦ РАН (2016)

- Первый в России и СНГ проект на базе Intel® Omni-Path (100 Гбит/с)
- Универсальное решение на 150 узлов в вычислительном шкафу
- Постоянно работающий free-cooling 24x7x365 при температуре до 50°С
- Температура жидкости на входе в узлы 48-53°С
- Система хранения с поддержкой Intel® Omni-Path/InfiniBand FDR

Состав комплекса расширения

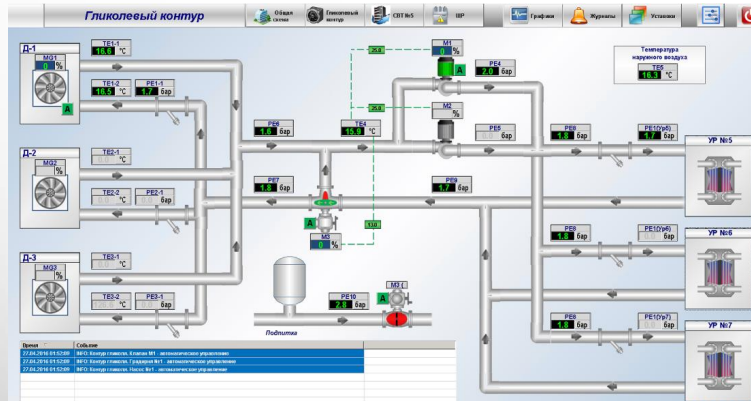
- 42 узла по 1,16 ТФЛОПС каждый
 - По 2 процессора Intel Xeon E5-2697 v3
 - На узле 128 ГБ памяти DDR4-2133
 - Intel® Omni-Path 100 Gbit/s fat-tree
 - Диски Intel® SSD S3510 120 ГБ
- Общая система хранения с существующей системой с поддержкой Intel® Omni-Path/FDR InfiniBand





Hot Water cooling infrastructure

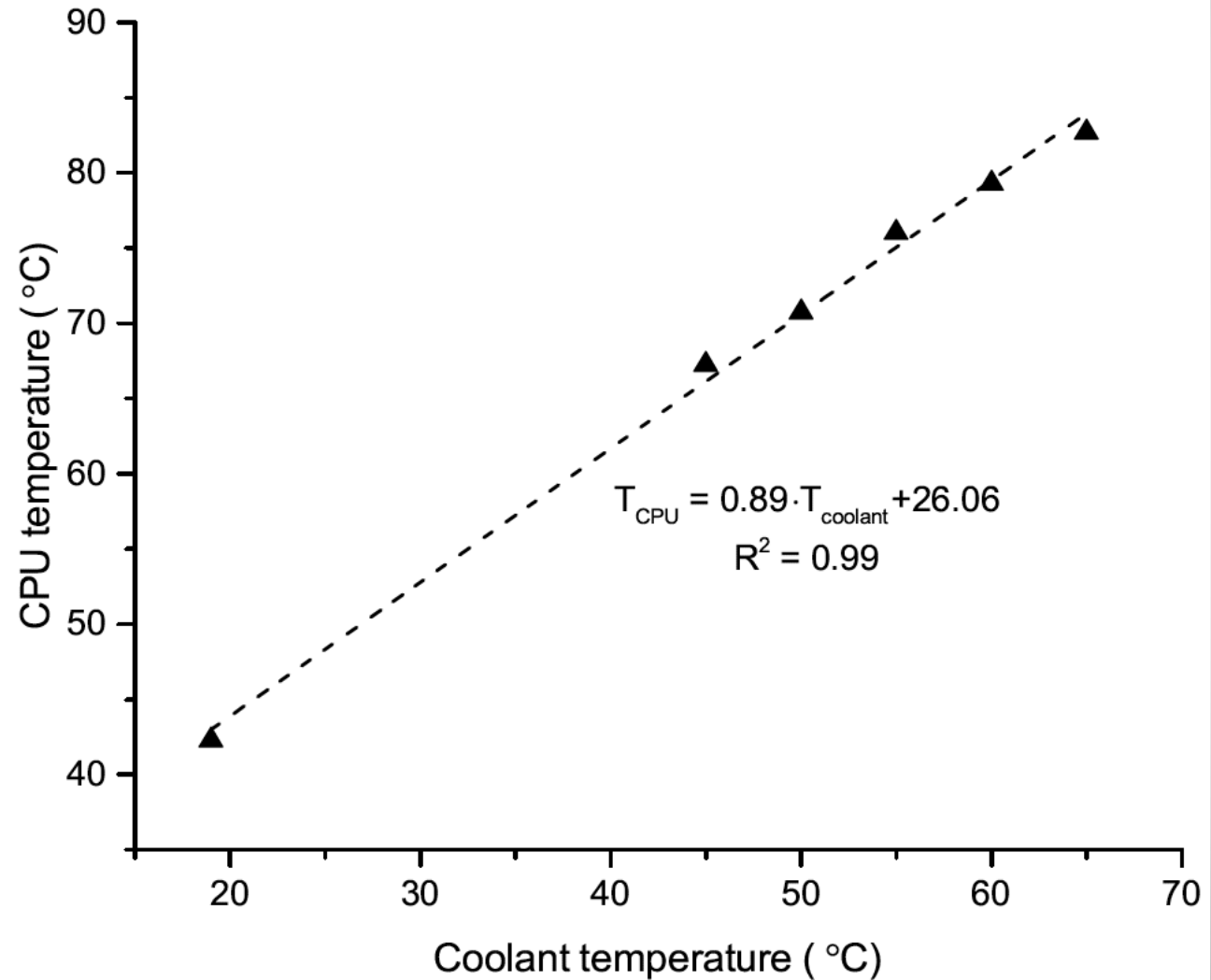
- RSC Tornado design, 220V AC power
- 1 rack, partially populated
- LINPACK for benchmark
- Omni-Path Interconnect
- Infrastructure for up to 450 nodes
- Drycooler only infrastructure (Moscow has 38^o C the highest registered air temperature ever)



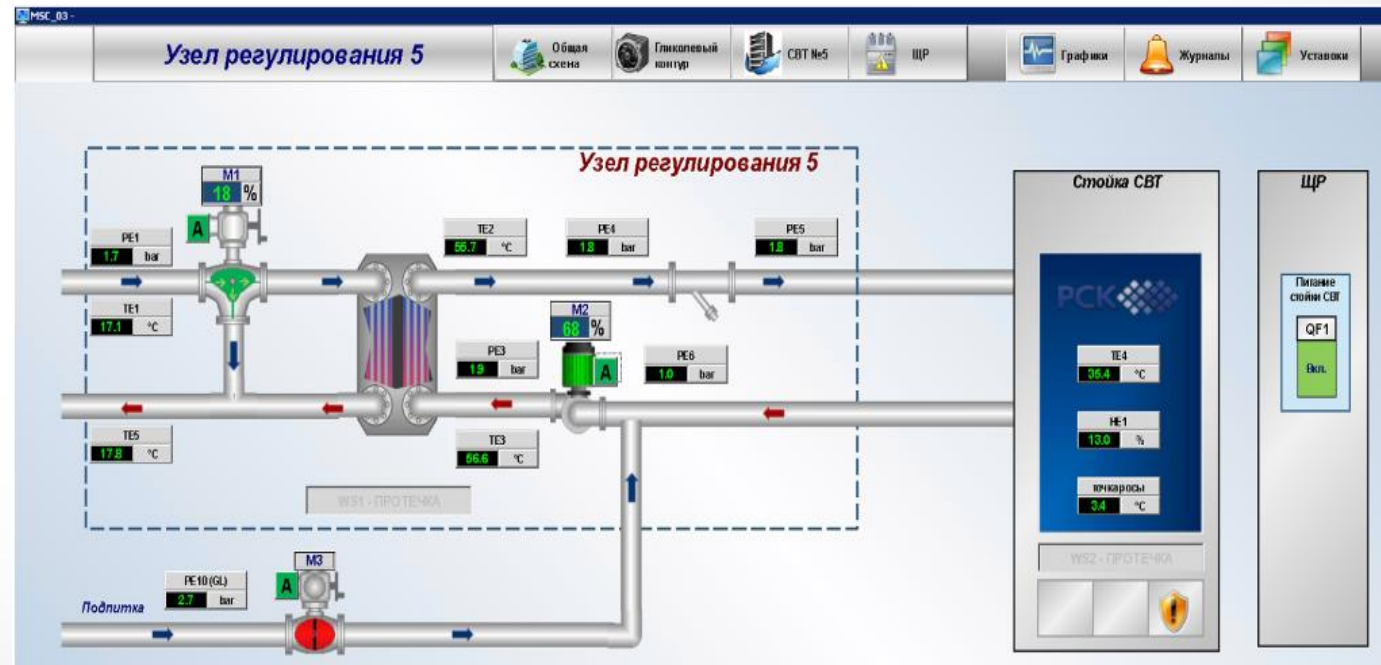
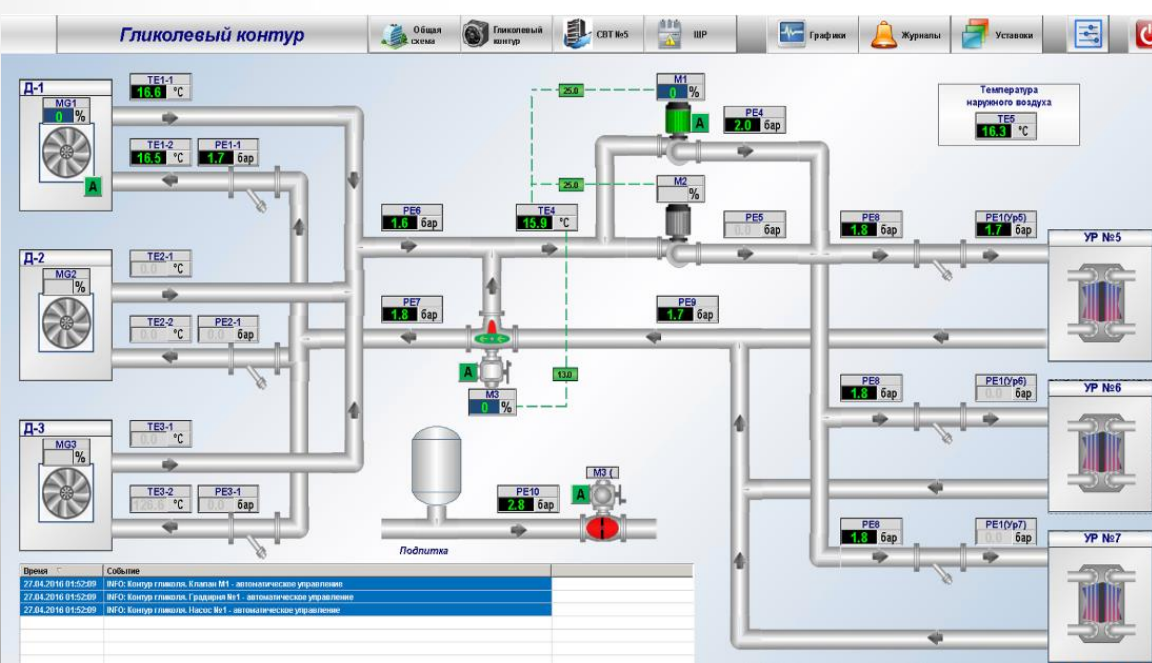


Node-level temperature dependency

- Linear dependency between external coolant temperature and CPU die temperature at maximum (TDP) load
- 3% performance degradation at T_{LC} in range 50-65 °C
- Power/performance decrease $\sim 2.2\%$ for every 10 °C of T_{LC}
- Overall drop in power/performance is 11.5% upon growth of coolant temperature from 19 to 65 °C



- Test run 2 hours:
 - +19° C inlet water $\Rightarrow$ LINPACK = 1.001TFLOPS per node
 - +45° C inlet water $\Rightarrow$ LINPACK = 1.003TFLOPS per node
 - +57° C inlet water $\Rightarrow$ LINPACK = 983TFLOPS per node
- Year around estimated PUE < 1.04

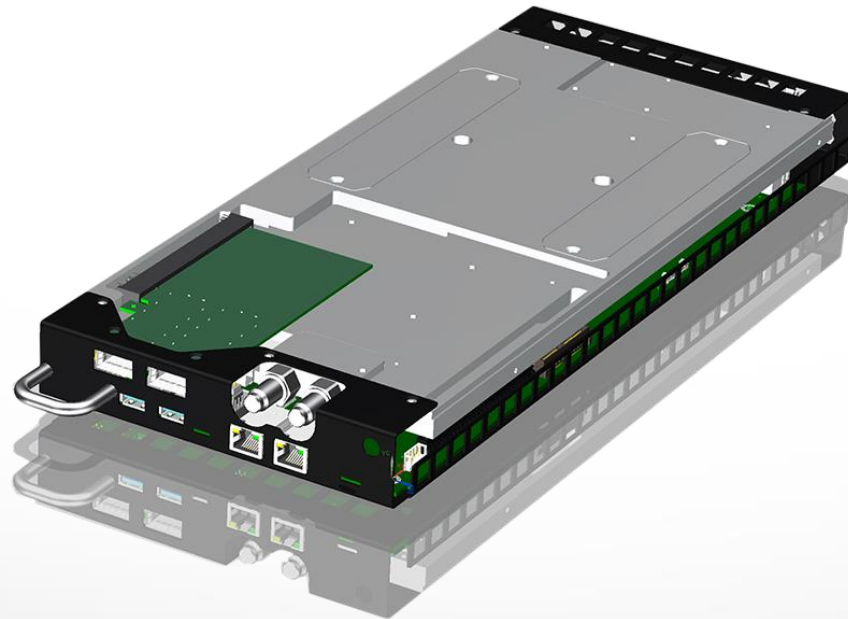




Real hot water cooling

Stress test for RSC Tornado Phi node

~63[°] C inlet temperature stress test

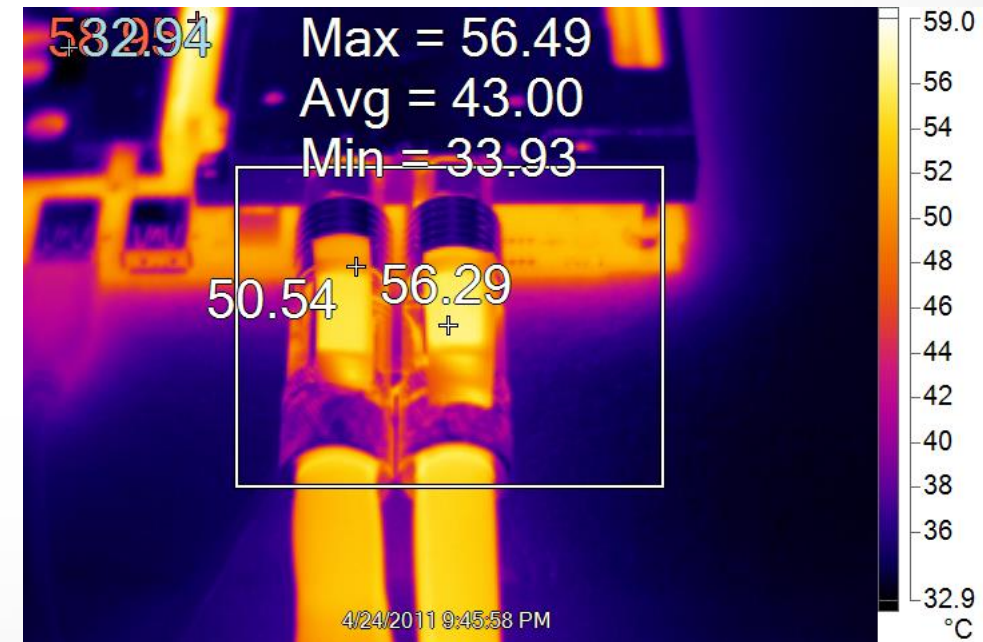




Experimental results

- Intel Xeon Phi 7250, Intel S7200AP server board, 192 GB RAM
- Measurements:
 - DGEMM performance
 - IPMI status
 - IR photos
 - Input current on 400V
- Results at 50 °C inlet:
 - No DGEMM GFLOPS variations with temperature
 - ~214.8W CPU power reported by platform with die less 68 °C
 - ~340W system inlet power measured
 - 20 °C headroom reported by platform temperature thresholds

```
22:20:45 done 1 DGEMM in 6.67s: 1437.77 GFLOPS Power: CPU 214.8W DRAM 14.2W Temp 67°C
22:20:52 done 1 DGEMM in 6.64s: 1443.85 GFLOPS Power: CPU 214.3W DRAM 13.2W Temp 68°C
22:20:59 done 1 DGEMM in 6.68s: 1435.88 GFLOPS Power: CPU 213.9W DRAM 14.2W Temp 68°C
22:21:06 done 1 DGEMM in 6.67s: 1437.88 GFLOPS Power: CPU 214.1W DRAM 14.8W Temp 68°C
22:21:12 done 1 DGEMM in 6.67s: 1438.48 GFLOPS Power: CPU 214.1W DRAM 14.1W Temp 67°C
22:21:19 done 1 DGEMM in 6.67s: 1437.71 GFLOPS Power: CPU 214.2W DRAM 14.1W Temp 68°C
22:21:26 done 1 DGEMM in 6.68s: 1435.62 GFLOPS Power: CPU 214.8W DRAM 14.3W Temp 66°C
22:21:32 done 1 DGEMM in 6.62s: 1447.68 GFLOPS Power: CPU 214.5W DRAM 13.3W Temp 66°C
22:21:39 done 1 DGEMM in 6.69s: 1434.19 GFLOPS Power: CPU 213.5W DRAM 13.7W Temp 65°C
22:21:45 done 1 DGEMM in 6.65s: 1441.38 GFLOPS Power: CPU 214.3W DRAM 13.2W Temp 67°C
22:21:52 done 1 DGEMM in 6.67s: 1437.88 GFLOPS Power: CPU 213.9W DRAM 14.8W Temp 67°C
22:21:59 done 1 DGEMM in 6.68s: 1435.79 GFLOPS Power: CPU 214.2W DRAM 13.9W Temp 67°C
22:22:05 done 1 DGEMM in 6.64s: 1444.28 GFLOPS Power: CPU 214.3W DRAM 14.1W Temp 68°C
22:22:12 done 1 DGEMM in 6.63s: 1445.94 GFLOPS Power: CPU 214.3W DRAM 14.1W Temp 68°C
22:22:19 done 1 DGEMM in 6.66s: 1439.39 GFLOPS Power: CPU 214.8W DRAM 14.8W Temp 68°C
22:22:25 done 1 DGEMM in 6.66s: 1448.49 GFLOPS Power: CPU 214.3W DRAM 14.2W Temp 68°C
22:22:32 done 1 DGEMM in 6.67s: 1438.46 GFLOPS Power: CPU 213.6W DRAM 13.6W Temp 67°C
22:22:39 done 1 DGEMM in 6.66s: 1439.66 GFLOPS Power: CPU 214.8W DRAM 14.4W Temp 66°C
22:22:45 done 1 DGEMM in 6.68s: 1452.71 GFLOPS Power: CPU 214.5W DRAM 14.1W Temp 67°C
22:22:52 done 1 DGEMM in 6.68s: 1435.69 GFLOPS Power: CPU 213.7W DRAM 13.6W Temp 68°C
22:22:59 done 1 DGEMM in 6.71s: 1438.14 GFLOPS Power: CPU 214.8W DRAM 13.4W Temp 68°C
22:23:05 done 1 DGEMM in 6.78s: 1431.85 GFLOPS Power: CPU 214.2W DRAM 14.1W Temp 68°C
22:23:12 done 1 DGEMM in 6.68s: 1436.66 GFLOPS Power: CPU 214.8W DRAM 14.3W Temp 68°C
22:23:19 done 1 DGEMM in 6.67s: 1437.52 GFLOPS Power: CPU 214.8W DRAM 14.3W Temp 68°C
22:23:25 done 1 DGEMM in 6.65s: 1441.11 GFLOPS Power: CPU 214.3W DRAM 14.2W Temp 66°C
22:23:32 done 1 DGEMM in 6.63s: 1447.53 GFLOPS Power: CPU 213.7W DRAM 13.6W Temp 66°C
22:23:39 done 1 DGEMM in 6.64s: 1444.37 GFLOPS Power: CPU 214.4W DRAM 14.3W Temp 67°C
22:23:45 done 1 DGEMM in 6.65s: 1442.33 GFLOPS Power: CPU 214.1W DRAM 14.2W Temp 68°C
22:23:52 done 1 DGEMM in 6.62s: 1448.68 GFLOPS Power: CPU 214.4W DRAM 14.1W Temp 67°C
22:24:05 done 1 DGEMM in 6.62s: 1448.52 GFLOPS Power: CPU 214.4W DRAM 14.1W Temp 67°C
```

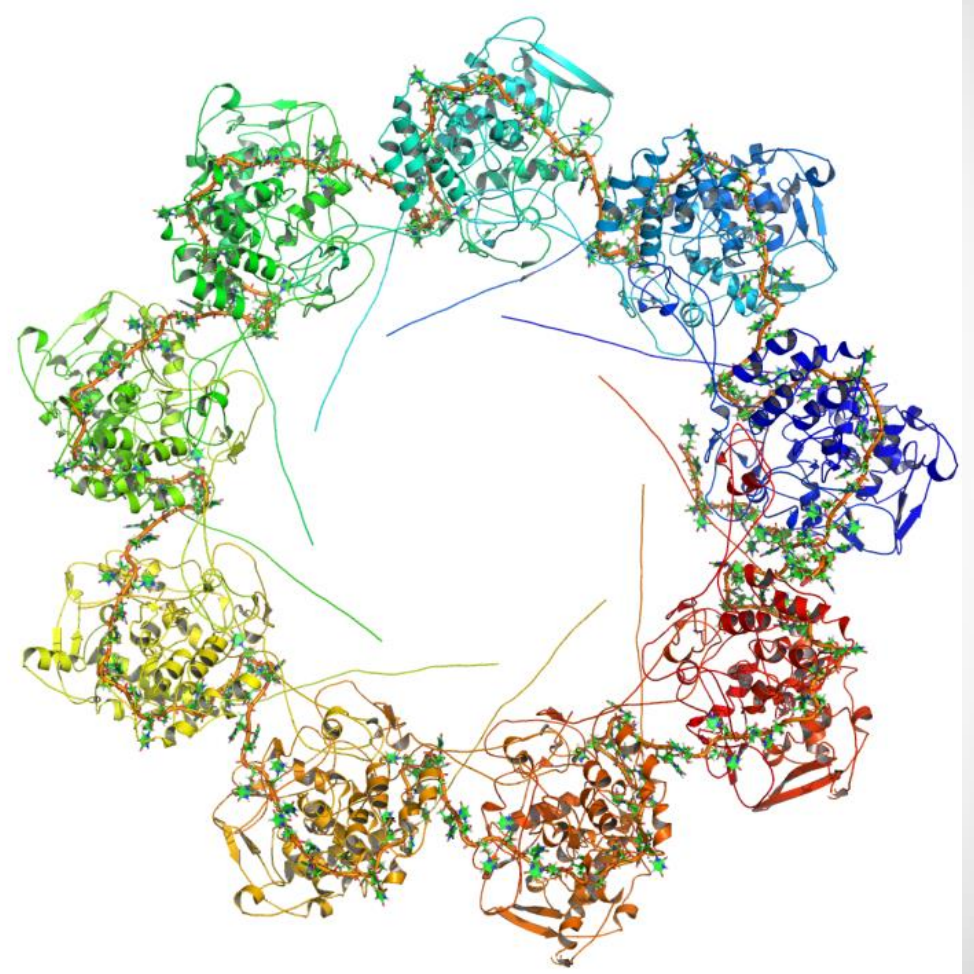


The screenshot displays the Intel Integrated BMC Web Console interface. The browser address bar shows the URL 10.40.3.140/index.html. The console has a navigation bar with tabs for System Information, **Server Health**, Configuration, Server Diagnostics, and Remote Control. The Server Health section is active, showing a list of sensor readings and power statistics. The status of each sensor is indicated by a color-coded box: OK (green), Warning (yellow), or Error (red).

Category	Sensor Name	Status	Value
Sensor Readings	BB P1 VR Temp	Normal	OK 64 degrees C
	BB Mem VR Temp	Normal	OK 57 degrees C
	P1 Status	reports the processor's presence has been detected	OK 0x0000
	P1 Therm Margin	Normal	OK -17 degrees C
Power Statistics	P1 Therm Ctrl %	Normal	OK 0
	CPU ERR2	All deasserted	OK 0x0000
	IERR	All deasserted	OK 0x0000
	CPU Missing	All deasserted	OK 0x0000
	P1 DTS Therm Mgn	Normal	OK -7 degrees C
	VRD Hot	All deasserted	OK 0x0000
	DIMM Therm Mgn 1	Normal	OK -14 degrees C
	DIMM Therm Mgn 2	Normal	OK -22 degrees C
	Auto Shutdown	All deasserted	OK 0x0000
	P1 Mem Therm Trip	All deasserted	OK 0x0000
	BB +12.0V	Normal	OK 11.882 Volts
	BB +3.0V VBAT	Normal	OK 2.929 Volts
	Voltage Fault	reports it has been asserted	Warning 0x0002
	NM Capabilities	All deasserted	OK 0x0001
	MTT CPU1	Normal	OK 0

At the bottom of the console, there are buttons for 'Refresh' and 'Show Thresholds', and a field to 'Set auto-refresh in seconds (0 to disable)'.

- Standard tests:
 - DGEMM – 2100 GFLOPS
 - LINPACK – 2000 GFLOPS
 - HPCG – 48.1 GFLOPS
- Partner codes on KNL
 - GROMACS ~2x bytes/flop over Haswell
 - SeisSol ~2x FLOPS over Haswell
 - GAMESS ~1.5 speed over





Семинар PCK на Russian Supercomputing Days

26 сентября, 16:20 – 18:20, зал «Чистые пруды»

«Работа реальных приложений на процессоре Intel® Xeon Phi™ 7200: подход, опыт и результаты»

Результаты получены на 4-х узловом кластере «PCK Торнадо Phi»

Конфигурация узла:

- процессор Intel® Xeon Phi™ 7250
- 192 ГБ DDR4-2400 + 16 ГБ MCDRAM
- Intel® SSD DC S3510 480 ГБ

