

Консолидированная система представления состояния среды суперкомпьютерных вычислений ИММ УрО РАН

Д. А. Усталов, А. Ю. Берсенёв,
А. В. Созыкин, Д. Г. Ермаков

План доклада

- Введение
- Проблематика
- Подход
- Решение
- Оценка
- Заключение

НАУКА УРАЛА

МАЙ 2013 № 13 (1079)

Газета Уральского отделения Российской академии наук
выходит с сентября 1986 33-й год издания

Передний край

СУПЕРКОМПЬЮТЕР «УРАН»: НОВЫЙ ЭТАП РАЗВИТИЯ

17 мая в Институте математики и механики УрО РАН состоялось торжественное открытие суперкомпьютерного центра (СКЦ) после масштабной реконструкции. На церемонии были приглашены представители руководства УрО РАН, Института Уральского отделения, Уральского федерального университета, другие вузы.

Губернатор СКЦ приветствует гостей. М.Д. Гольдштейн рассказывает об особенностях развития суперкомпьютерного центра. В церемонии открытия участвуют представители руководства УрО РАН, Института Уральского отделения, Уральского федерального университета, другие вузы.

Губернатор СКЦ приветствует гостей. М.Д. Гольдштейн рассказывает об особенностях развития суперкомпьютерного центра. В церемонии открытия участвуют представители руководства УрО РАН, Института Уральского отделения, Уральского федерального университета, другие вузы.



В СКЦ расположены вычислительные и дисковые ресурсы. Такой масштабный ремонт не делался более 35 лет. Был проведена замена фальшпотолка, фальшстен, фальшпанелей, системы освещения. Все это выполнено из современных материалов, обеспечивающих необходимый уровень шумопоглощения и жароизоляции. Заметно улучшились и эстетические показатели интерьеров машинных залов. Работа по внедрению суперкомпьютера, что позволит добиться новых результатов в научных исследованиях.

Основным приоритетным заданием, которое стоит перед собой коллектив СКЦ — войти в список TOP-500 (самые мощные суперкомпьютеры мира), а также в европейские грейс-системы. Так называются формы распределенных вычислений, в которой виртуальный суперкомпьютер представляет собой кластер, соединенный с помощью сети слабых вычислительных ресурсов.

Академик
Я.И. АЛФЕРОВ:
«От вычисления —
к развитию»

— Стр. 4

Академик
А.Д. НЕКИПЕЛОВ:
«Открыть
академическую
форму»

— Стр. 5

СВЕТ
И РАДОСТЬ
ВСТРЕЧИ

— Стр. 8

Выборы президента РАН

АКАДЕМИКИ ПОЛУЧИЛИ ЛИДЕРА

После того как вчера на президентском заседании Российской академии наук вынесенный ее главой Юрием Осиповым список кандидатов на пост президента РАН на следующий срок, осталось три претендента: вице-президенты Игорь Алферов, Александр Некипелов и академик Владимир Фортов.

А затем уже на них президент РАН выбрал одного, чье имя поставит на главе списка и предложит Общему собранию РАН 29 мая избрать президентом РАН на следующий срок. Этого кандидата выдвинул Владимир Фортов. Кстати, ранее именно он выдвинул на кандидатуру нынешнего члена Отделения РАН Константина Косарева. Общее собрание не согласилось и поддержало другую кандидатуру, но, по мнению, что рекомендация президента дорога.

В своей программе Владимир Фортов, в частности, заявляет, что Академия должна стать генератором идей в экономике, занять системообразующую позицию в разработке и реализации стратегии развития России. В центре внимания всей системы РАН должны быть талантливые ученые. Ему надо создать условия, чтобы он с максимальной отдачей мог работать своей наукой в неделю и 24 часа в сутки.

Окончание на с.5

Введение

- Суперкомпьютер «**УРАН**» ИММ УрО РАН
 - №6 в Top50 СНГ,
№428 в Top500,
№385 в Green500,
№91 в Graph500.
 - Пик: 216,56 Тфлоп/с,
Linpack: 105,36 Тфлоп/с.



Введение (прод.)

- Состав суперкомпьютера «**УРАН**»:
 - 20 узлов на Xeon® X5675 и Tesla® M2050;
 - 14 узлов на Xeon® E5-2660 и Tesla® M2090;
 - 10 узлов на Xeon® X5675 и Tesla® M2090;
 - 160 узлов на Xeon® E5450.

Проблематика

- Гетерогенная среда: много разных вычислительных узлов для мониторинга.
- Мониторить по-отдельности — запросто!
— Zabbix, SLURM, LDAP, iLO, etc.
- **При эксплуатации СКЦ требуется вся картина о состоянии, а не отдельные отрывки информации.**
— Информационный киоск.

Проблематика (прод.)

- **Зачем?**

- Анализ операционной деятельности СКЦ с целью повышения качества обслуживания.
- Подготовка отчётов и документов для вышестоящих организаций.
- Ещё кое-что.

Подход

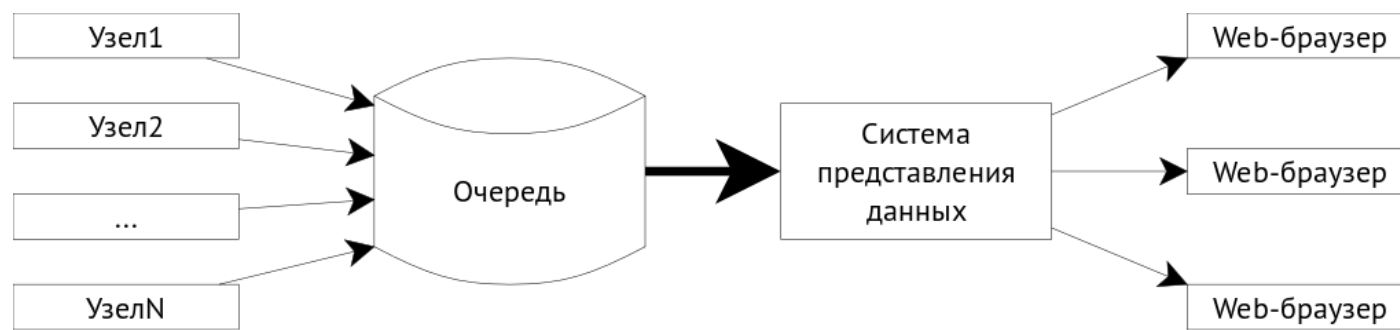
- **Мониторинг уже есть и как-то работает.**
 - Не стоит плодить сущности.
- Консолидируем данные из имеющихся систем.
 - Обеспечиваем их согласованность.
 - Предоставляем единый формат работы с ними.

Подход (прод.)

- **Источники данных:**
 - LDAP – пользователи;
 - Zabbix – температура и энергопотребление;
 - при помощи IPMI в iLO.
 - SLURM – очередь задач;
 - утилиты RHEL – нагрузка, память, и т. д.
 - и многое другое.

Решение

- Скрипты собирают данные, отправляют в очередь AMQP.
- На очередь подписываются слушатели и получают эти данные.
- Слушатели занимаются их обработкой.



Решение (прод.)

- Построена общая шина обмена сообщениями на основе AMQP.
- Данные с узлов и систем мониторинга поступают в формате JSON.
 - Написаны специальные скрипты на Python.
- Система визуализации доступна как Web-приложение.
 - Приложение разработано на Ruby.

Внешний вид

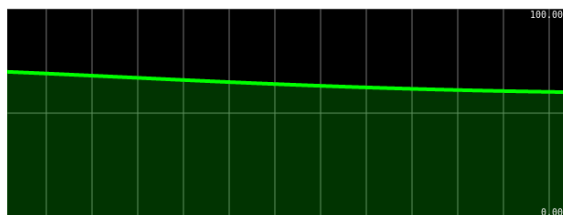
Кластер «УРАН»

Сейчас

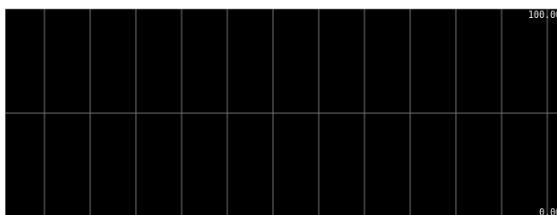
Очередь

Справка

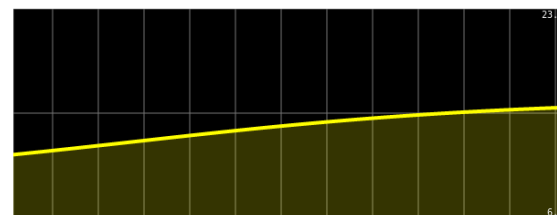
ЦПУ: 58.99%



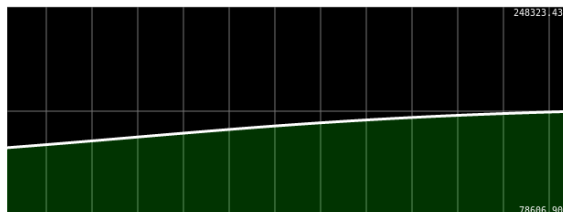
ГПУ: 0.00%



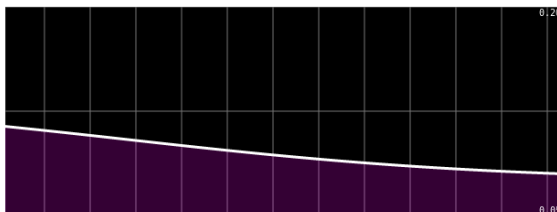
Мощность: 15.47 кВт



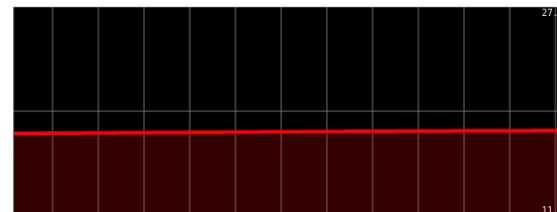
Память: 165548.96 ГБ



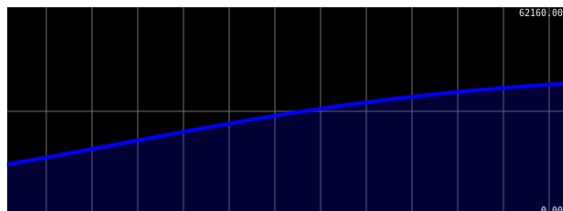
Память ГПУ: 0.08 ГБ



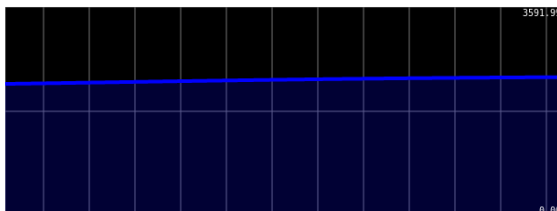
Температура: 18.02°C



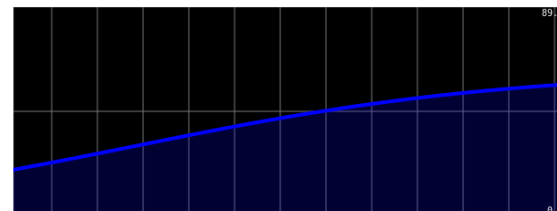
Сеть Ethernet: 41440.00 МБ



Сеть InfiniBand: 2394.66 ГБ



Операций NFS: 59.75



Консолидированная система мониторинга суперкомпьютера «УРАН» разработана при поддержке гранта УрО РАН №РЦП-13-И18.

© ИММ УрО РАН



Внешний вид (прод.)

Кластер «УРАН»

Сейчас	Очередь	Справка
--------	---------	---------

Количество выполняемых задач в настоящее время: 12. Задач в очереди: 109.

Выполняемые задачи

Пользователь	Узлы	Задача	Время запуска	Время работы	Лимит времени
	33	pw.x	20:26:00 30.05.2013	0:26	2:00:00
	22	mapi	07:59:00 30.05.2013	12:27:29	20:00:00
	22	mapi	07:43:00 30.05.2013	12:43:39	20:00:00
	20	sbatch	11:37:00 30.05.2013	8:49:19	20:00:00
	20	ch133	13:13:00 30.05.2013	7:13:29	20:00:00
	14	sbatch	11:37:00 30.05.2013	8:49:19	20:00:00
	14	sbatch	11:37:00 30.05.2013	8:49:19	20:00:00
	8	run-ompi	10:17:00 30.05.2013	10:09:39	20:00:00
	8	r2	17:37:00 30.05.2013	2:49:44	5:00:00
	4	lammps	18:39:00 30.05.2013	1:47:27	8:00:00
	1	mapi	07:40:00 30.05.2013	12:46:19	20:00:00
	1	sbatch	18:52:00 30.05.2013	1:34:06	4:00:00

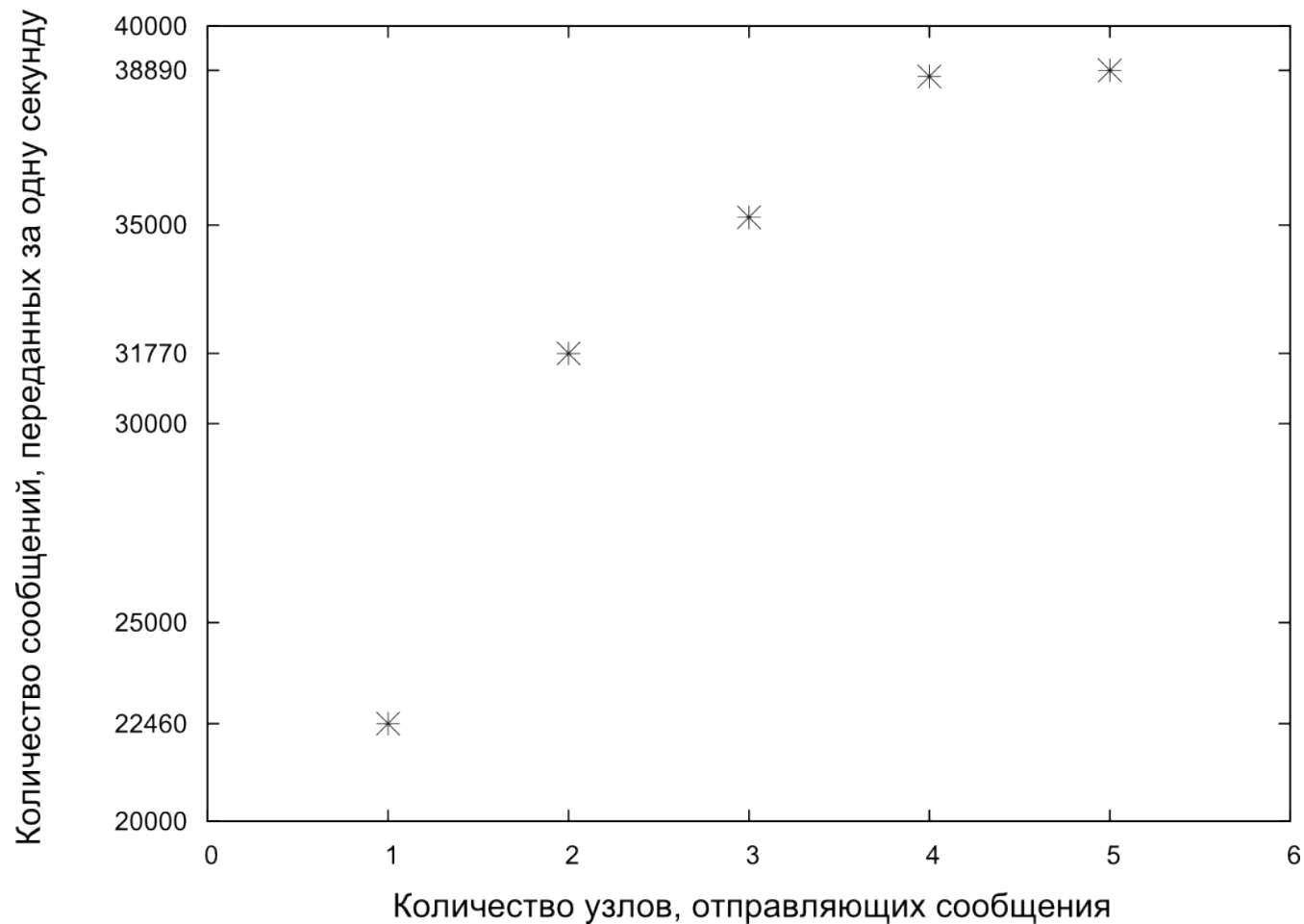
Задачи в очереди

Пользователь	Узлы	Задача	Лимит времени
	50	pw.x	2:00:00
	50	pw.x	2:00:00
	14	sbatch	20:00:00
	14	sbatch	20:00:00
	14	sbatch	20:00:00
	14	sbatch	20:00:00
	14	sbatch	20:00:00
	14	sbatch	20:00:00
	14	sbatch	20:00:00
	14	sbatch	20:00:00
	14	sbatch	20:00:00
	14	sbatch	20:00:00
	14	sbatch	20:00:00
	14	sbatch	20:00:00
	14	sbatch	20:00:00
	14	sbatch	20:00:00

Оценка

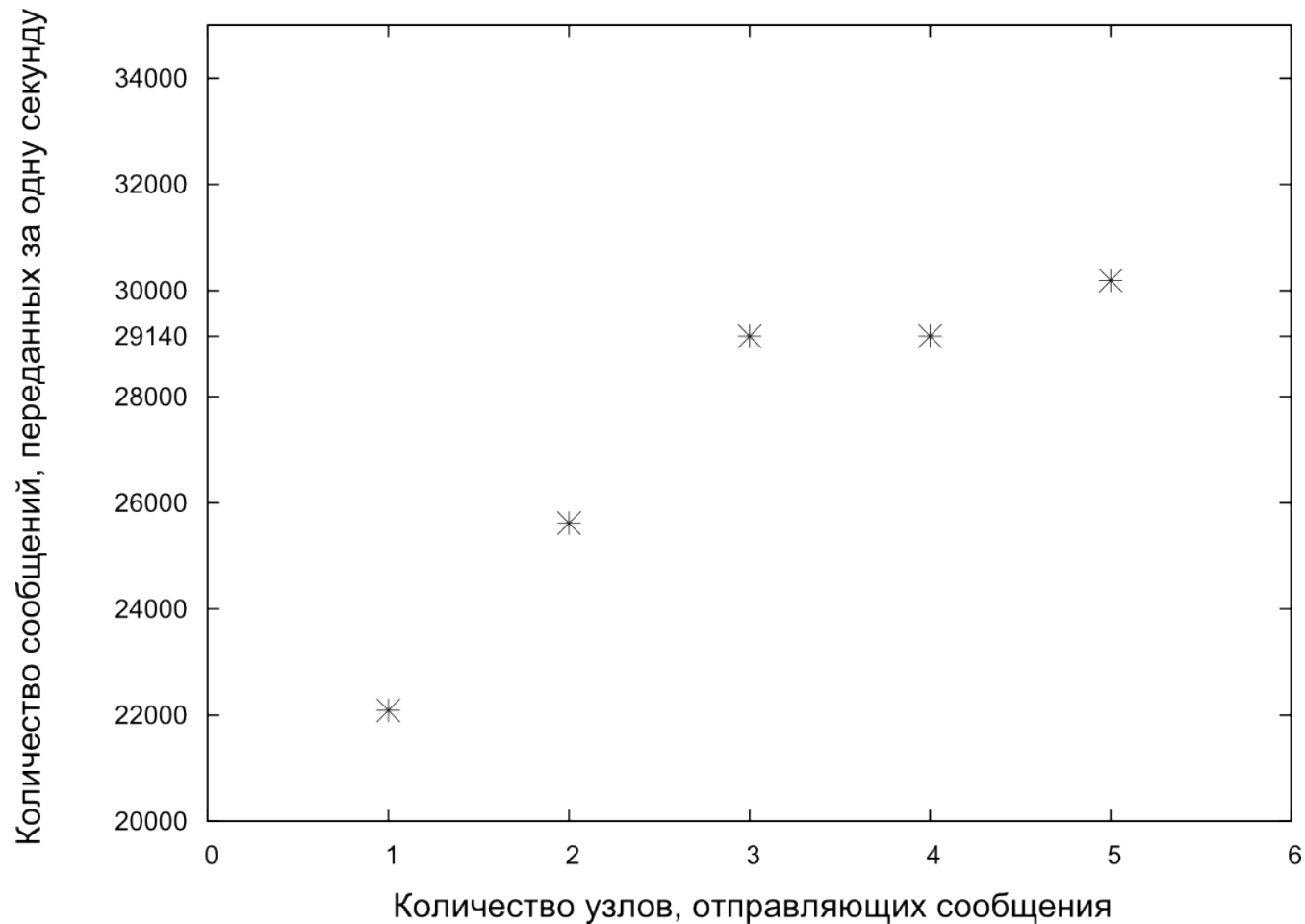
- Интересно было бы знать производительность AMQP-очереди.
- Симулируем высокую нагрузку:
 - узел очереди: Xeon®, 512 МБ ОЗУ;
 - узлы отправки: Xeon® E5520, 18 ГБ ОЗУ;
 - узлы приёма: Xeon® E5-2660, 100 ГБ ОЗУ;
 - Intel® 82576 Gigabit Ethernet.

Оценка (прод.)



Получателей: 1.

Оценка (прод.)



Получателей: 2.

Заключение

- Представленный подход позволяет быстро интегрировать новые источники данных.
- Система работает в СКЦ ИММ УрО РАН.
 - <https://github.com/dmchk/uranwatch>
- Протокол AMQR позволяет строить масштабируемые и производительные информационные системы легко и удобно.

Что дальше?

- Качество данных и новые функции:
 - тепловая карта СКЦ;
 - данные с сетевого оборудования;
 - обобщиться от специфики СКЦ ИММ УрО РАН.

Благодарности

- Работа выполнена в рамках проекта УрО РАН №РЦП-13-И18.

Вопросы?

- **Дмитрий Усталов**

- ведущий программист ИММ УрО РАН

- <http://www.imm.uran.ru/>

- Сайт: <http://eveel.ru/>

- Электронпочта: dau@imm.uran.ru