

*Суперкомпьютерные дни в России 2019
Семинар «Инструменты и технологии обеспечения эффективной работы суперкомпьютерных центров»*

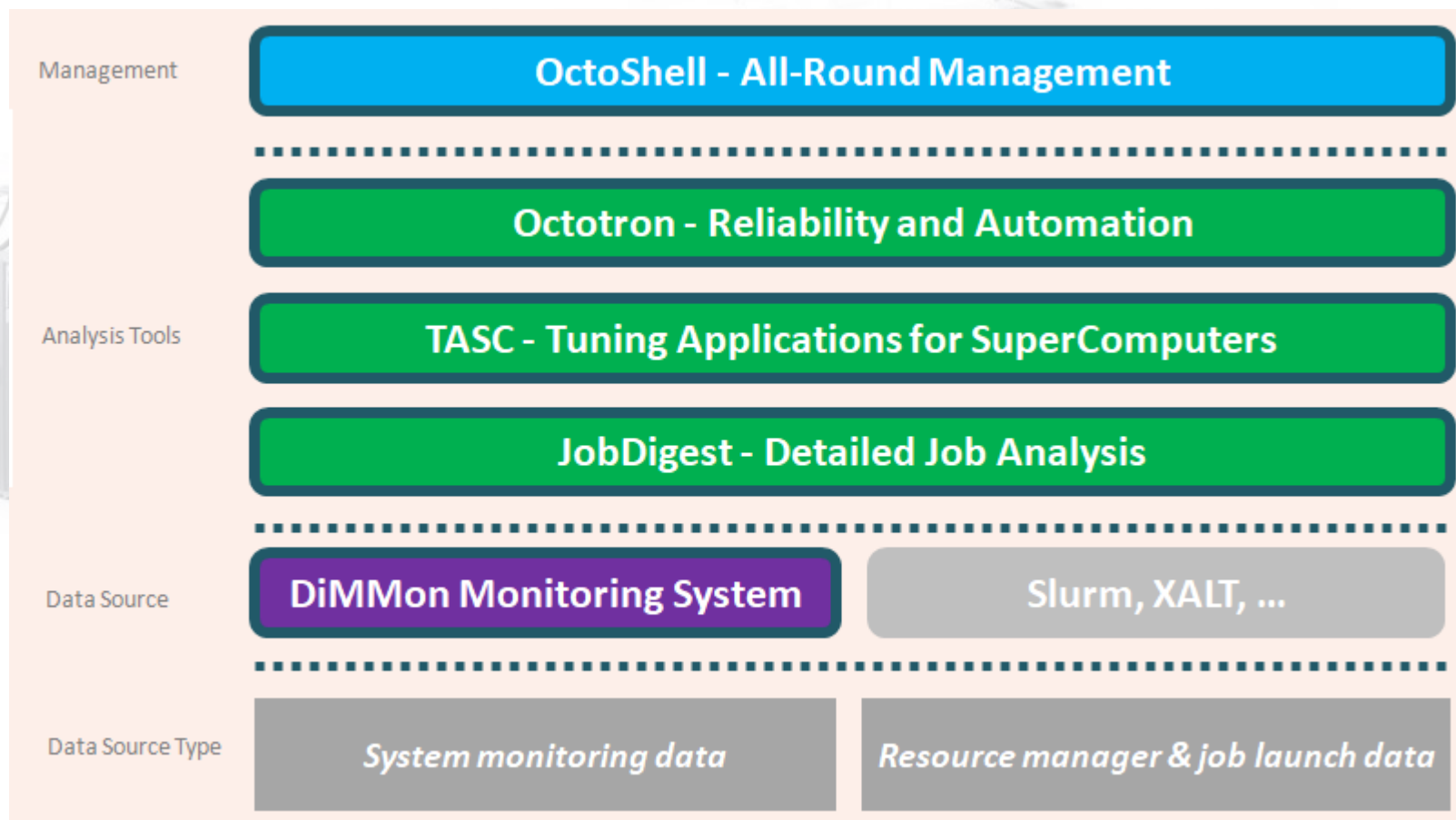
Всесторонний анализ эффективности функционирования суперкомпьютерных центров

Вадим Воеводин, Дмитрий Никитенко

*Научно-исследовательский вычислительный центр
Московского государственного университета имени М. В. Ломоносова*

23 сентября 2019, Москва

Инструменты для организации эффективной работы СКЦ, разрабатываемые в НИВЦ МГУ



Фрагмент постера «All-round open source tools for ensuring productive operation of HPC centers»

TASC (Tuning Applications for SuperComputers)

Цель №1: помочь пользователям СКЦ обнаружить проблемы с эффективностью в их приложениях (если они есть) и рассказать, как можно устранить эти проблемы

Цель №2: помочь администраторам и руководству системы оценить эффективность работы СКЦ в целом

Подсистема 1: массовый анализ производительности

Цель №1: помочь пользователям СКЦ обнаружить проблемы с эффективностью в их приложениях (если они есть) и рассказать, как можно устранить эти проблемы

- Акцент на массовости анализа – исследуем все выполняемые задачи
- Анализ выполняется автоматически и «на лету»
- Входные данные для проведения анализа:
 - Данные о производительности от системы мониторинга (**DiMMon, JobDigest**)
 - Данные от менеджера ресурсов
 - Более «интеллектуальные» данные от других средств анализа, например, от ML методов поиска приложений с аномально низкой эффективностью

Реализация автоматического анализа. Общий список задач

Информация о завершенных задачах

На данной странице вы можете посмотреть информацию о ваших задачах, а также оценить эффективность их выполнения. Для отображения только интересующих задач укажите ниже требуемые фильтры.

Показаны задачи: 0..19 из 19

Нет скрытых проблем с производительностью.

Найденные проблемы	ID задачи	Начало счета	Конец счета	Статус	Число узлов	Время счета (часы)	Размер задачи (ЦПУ-часы)	Загрузка ЦПУ	Загрузка ГПУ	Load average	IPC	Получено байт по MPI (МБ/с)	Передано байт по MPI (МБ/с)
✖	697858	2018-07-25 13:12:09	2018-07-25 13:36:56	cancelled	1	0.4	5.8	86.7	0.0	25.4	0.36	0.0	0.0
✖	697859	2018-07-25 13:12:09	2018-07-25 13:36:56	cancelled	1	0.4	5.8	83.9	0.0	24.7	0.35	0.0	0.0
✖	697857	2018-07-25 13:12:06	2018-07-25 13:36:56	cancelled	1	0.4	5.8	55.2	0.0	18.7	0.36	0.0	0.0
✖	697867	2018-07-25 13:40:02	2018-07-25 14:23:28	completed	1	0.7	10.1	34.7	0.0	10.7	0.46	0.0	0.0

Заметный дисбаланс внутри узлов по использованию памяти.

Вкладка Эффективность в Octoshell, которая показывает результаты анализа эффективности по всем задачам (только для СК Ломоносов-2).

Базовая информация по отдельной задаче

Информация по задаче № [REDACTED]

Общая информация

Суперкомпьютер	lomonosov-2
ID задачи	[REDACTED]
Логин	[REDACTED]
Статус задачи	COMPLETED
Раздел суперкомпьютера	compute
Число ядер	14
Список узлов	1
Поставлена в очередь	2019-03-14 10:41:38 (MSK)
Начало счета	2019-03-14 10:41:45 (MSK)
Конец счета	2019-03-14 11:18:38 (MSK)
Время	0.6
Список узлов	n51427

Производительность

Метрика	Значение	Общая оценка
Средняя загрузка ЦПУ (%)	20.72	average
Среднее LoadAVG	8.73	high
Средняя загрузка ГПУ (%)	0.00	low
Использование MPI (МБ/с)	0.00	low
Интенсивность чтения из файловой системы (МБ/с)	146.90	high
Интенсивность записи в файловую систему (МБ/с)	313.12	high

Базовые свойства

Одноузловая задача
Задача имеет необычную локальность обращений в память
Слишком низкая загрузка процессора при данном уровне loadavg

Строка запуска: [REDACTED]

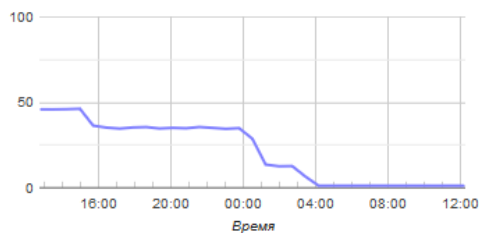
Данные мониторинга об эффективности выполнения программы

Динамика поведения задачи во время выполнения

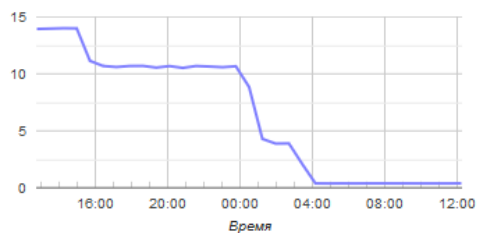
[Более подробное описание](#)

Ссылка на отчет JobDigest – Еще больше подробностей о динамике выполнения

Загрузка ЦПУ (0% .. 100%)



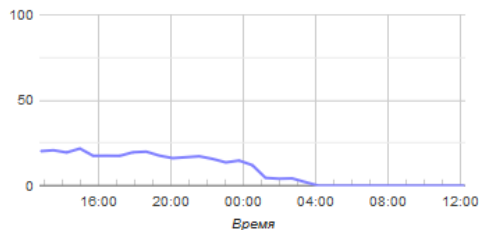
LoadAVG



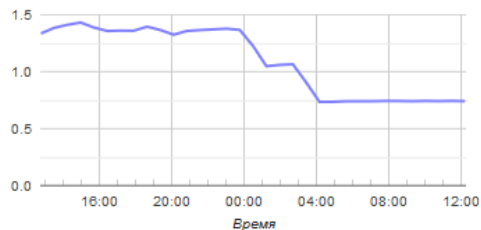
Использование MPI (МБ/с)



Загрузка ГПУ (0% .. 100%)



IPC















Использование файловой системы (МБ/с)




Информация по найденным потенциальным проблемам с эффективностью

Всего мы обнаруживаем порядка 30 проблем – неэффективная работа с MPI, явный дисбаланс нагрузки, подозрительно низкая утилизация ресурсов и т.д.

Тип	Описание	Предположение	Рекомендация
   	Задача запущена в разделе для GPU задач, однако практически не использует графические процессоры.	Неправильно выбран раздел для задачи.	Рекомендуется сменить раздел.
   	Низкая активность работы задачи при низкой локальности обращений в память.	Работы с памятью, вероятно, организована неэффективно, требуется оптимизация.	Анализ работы с памятью -> Эффективность
   	Задача выполняется аномально неэффективно.	Задача работает некорректно или зависла.	Рекомендуется проверить корректность запуска и при необходимости отменить его.



Профиль Поддержка Проекты Пере-regистрации  Пакеты Статистика **Эффективность** Комментарии

Проблемные задачи: ✕

Если выполняющаяся задача явно зависла или некорректно запущена, оперативно высылаем оповещение!

Подсистема 2: анализ качества использования СКЦ

Цель №2: помочь администраторам и руководству системы оценить эффективность работы СКЦ в целом

- Предоставление «срезов» – удобных отчетов о всевозможных аспектах работы СКЦ
- В будущем – оповещение о возможных критических ситуациях
- Акцент на интеграции множества различных входных данных:
 - Данные о производительности от системы мониторинга (DiMMon, JobDigest)
 - Данные от менеджера ресурсов
 - Более «интеллектуальные» данные от других средств анализа, например, от ML методов поиска аномалий
 - Данные о проектах, организациях, предметных областях (Octoshell)
 - Данные о использованных компиляторах, прилинкованных библиотеках
 - Данные о работоспособности компонент суперкомпьютера (Octotron)
 - Данные о производительности файловой системы

Анализ качества использования СКЦ. На какие вопросы мы хотим получить ответы?

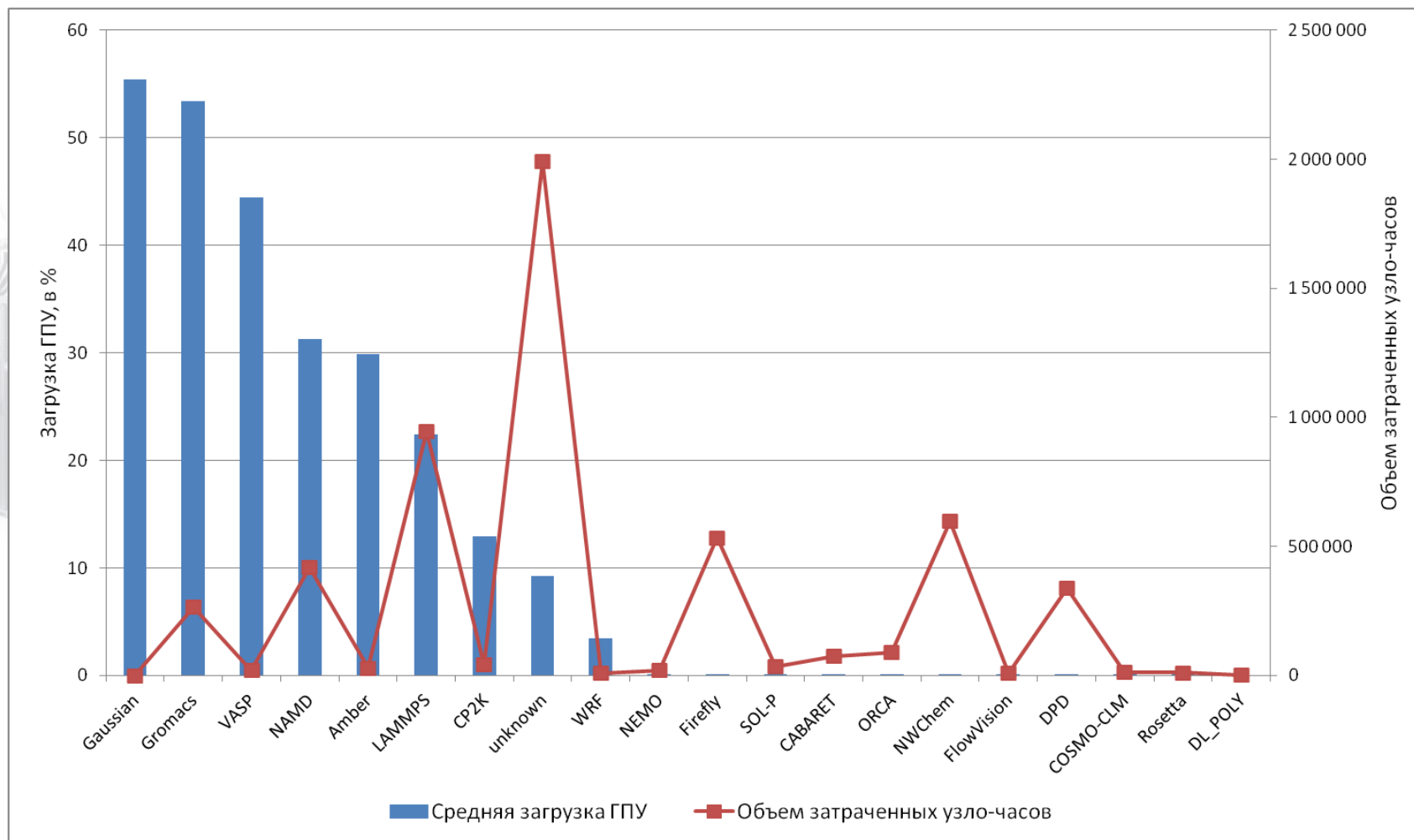
- С какой интенсивностью приложения, построенные на основе пакета GROMACS, используют коммуникационную сеть Infiniband?
- В каких предметных областях чаще всего запускают задачи большого размера (с большим числом процессоров)?
- Насколько часто в проектах задействуются технологию OpenMP совместно с MPI, и каково распределение используемого числа потоков на узлах?
- Какой из компиляторов (и какая версия данного компилятора) наиболее распространен среди проектов по механике сплошных сред?
- Часто ли пользователи запускают самостоятельно собранные прикладные пакеты или библиотеки? В какой предметной области таких пользователей больше всего? Насколько эффективны такие запуски?
- Пользователи из какой предметной области активнее всего используют VASP (по числу запусков, объему занимаемой памяти)?
- Какие прикладные пакеты активнее всего задействуют графические ускорители?
- Какая предметная область больше всего пострадала от сбоев на вычислительных узлах?
- Каково распределение активности использования ФС по проектам (суммарно за последний месяц)?
- В какой предметной области наибольший процент пользователей с числом некорректных запусков задач (со знаком ☒)?
- Какой значимый (с объемом работ более XXX ядро-часов) некоммерческий проект наименее эффективно утилизирует ЦПУ?
- ...

В каких предметных областях чаще всего запускают задачи большого размера?



Распределение доли затраченных узло-часов для различных предметных областей, данные с СК Ломоносов-2 за первое полугодие 2019 года

Какие прикладные пакеты активнее всего задействуют графические ускорители?



Средняя загрузка ГПУ для различных прикладных пакетов, данные с СК Ломоносов-2 за первое полугодие 2019 года

*Суперкомпьютерные дни в России 2019
Семинар «Инструменты и технологии обеспечения эффективной работы
суперкомпьютерных центров»*

Спасибо за внимание!

Воеводин Вадим

vadim@parallel.ru

23 сентября 2019, Москва