

Выпускная квалификационная работа

Тестирование и оптимизация программной поддержки NUMA в ОС «Эльбрус» для вычислительных комплексов Эльбрус 2С+

Выполнил: Картавцев А.Ю. 818 гр.

Научный руководитель: Кравцунов Е.М.

SMP и NUMA платформы

- SMP (Symmetric Multiprocessing) - это многопроцессорные системы, в которой процессоры соединены общей шиной.
 - Время доступа к памяти не зависит от расположения процессоров.
- NUMA (Non Uniform Memory Access) – это многопроцессорные системы, в которой процессоры группируются в узлы, каждый из которых связан с соответствующей областью памяти через отдельную шину.
 - Время доступа к памяти зависит от узлов, в которых располагаются процессоры и память.

Постановка задачи

- Разработать тесты для выявления аппаратных ошибок когерентности кэшей.
- Оценить работоспособность политик выделения памяти.
- Реализовать метод синхронизации - иерархические спинлоки.
- Сравнить результаты тестов производительности на машине NUMA с двумя узлами, используя три метода синхронизации: mutex, spinlocks, HBO spinlocks.

Разработанные тесты для выявления аппаратных ошибок

Тест RBTREE

- Предпосылки:
 - Ошибка возникала в подсистеме выделения страниц buddy-allocator.
 - При переводе ядра ОС «Эльбрус» на новую версию ядра Linux(2.6.33) ошибка стала проявляться в подсистеме планирования.
- Проблема:
 - Вывод всей системы из строя.
- Алгоритм
 - Создание M деревьев (M узлов NUMA).
 - Создание N структур типа task_struct.
 - Работа с задачами.
- Результат:
 - Несличение результата. Ошибка найдена.

Разработанные тесты для выявления аппаратных ошибок

Тест Lists

- Предпосылки:
 - Аппаратный комплекс зависал при работе с примитивами синхронизации.
- Проблема:
 - Частая очистка кэша через load.
- Алгоритм теста
 - Создание потоков чтения и блокировки на узлах NUMA.
 - Проверка целостности своих данных после каждой операции.
- Результат:
 - Ошибка найдена.

Разработанные тесты для выявления аппаратных ошибок

Тест RW

- Предпосылки:
 - Ошибки чтения\записи в память.
- Проблема:
 - Вывод всей системы из строя.
- Алгоритм:
 - Создание потоков чтения и записи.
 - Потоки привязываются к определенному процессору.
 - Запись\чтение номера итераций.
 - Массив с ячейками подобран таким образом, чтобы данные разных потоков попадали внутрь одной строки кэша L2.
- Результат:
 - Ошибка найдена.

Разработанные тесты для выявления аппаратных ошибок

Тест UCODE

- Предпосылки:
 - Зависания при запуске приложений, слинкованных динамически с библиотеками.
- Проблема:
 - Зависание работы процессора.
- Алгоритм:
 - Код подготавливается с уникальным смещением.
 - Динамическое размещение кода в памяти и передача управления.
 - В конце кода находится инструкция возврата.
- Результат:
 - Ошибка найдена.

Политики выделения памяти на NUMA машинах

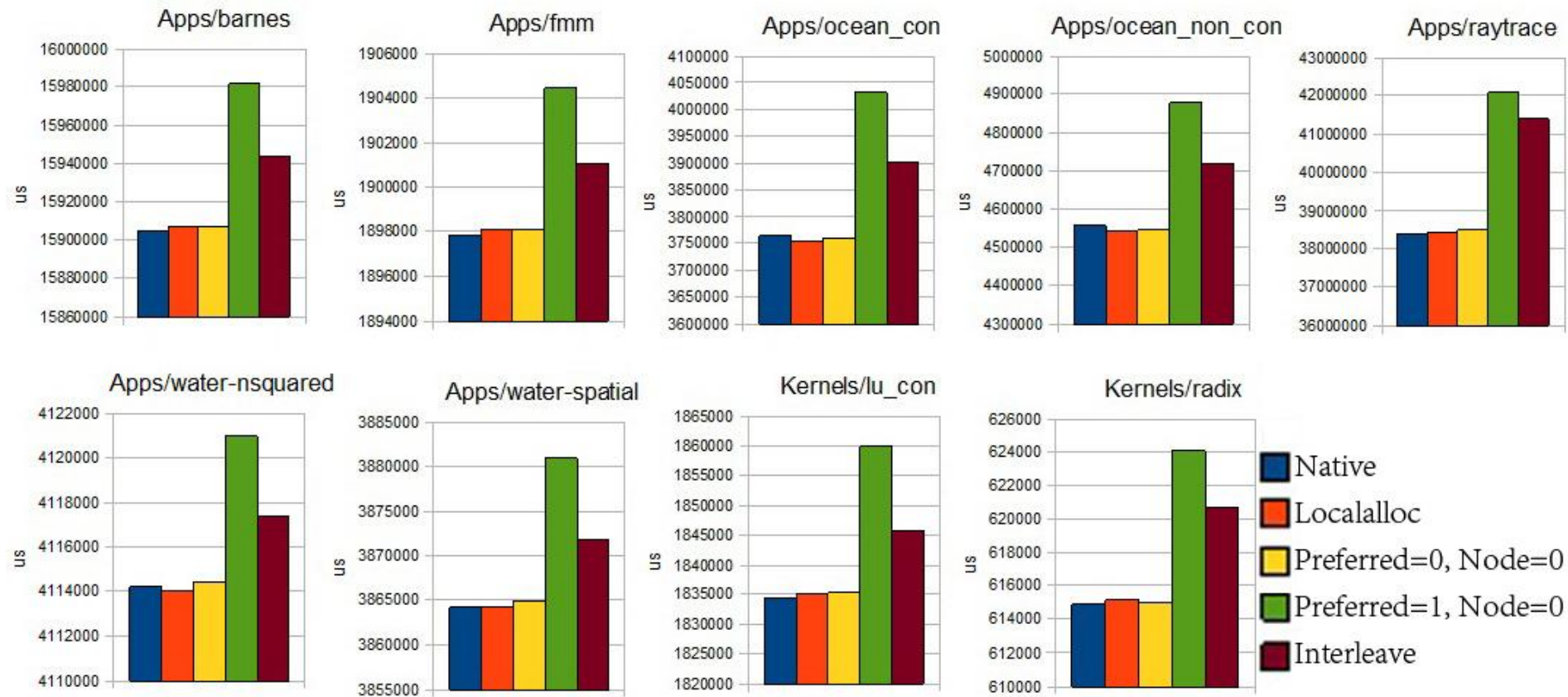
- `default` (политика по умолчанию) — выделение памяти на локальном узле.
- `bind` — выделение памяти на заданном наборе узлов.
- `interleave` — распределение выделяемых страниц памяти по всем имеющимся узлам равномерно.
- `preferred` — выделение памяти преимущественно на локальном узле.

Набор тестов SPLASH-2

Тестовый пакет параллельных приложений, предназначенный для исследования производительности распределенных и централизованных многопроцессорных систем.

- Включает в себя приложения
 - рассчитывающие взаимодействия тел и частиц,
 - сортировку.

Проверка работоспособности политик выделения памяти



Вывод: Политики выделения памяти работают корректно.

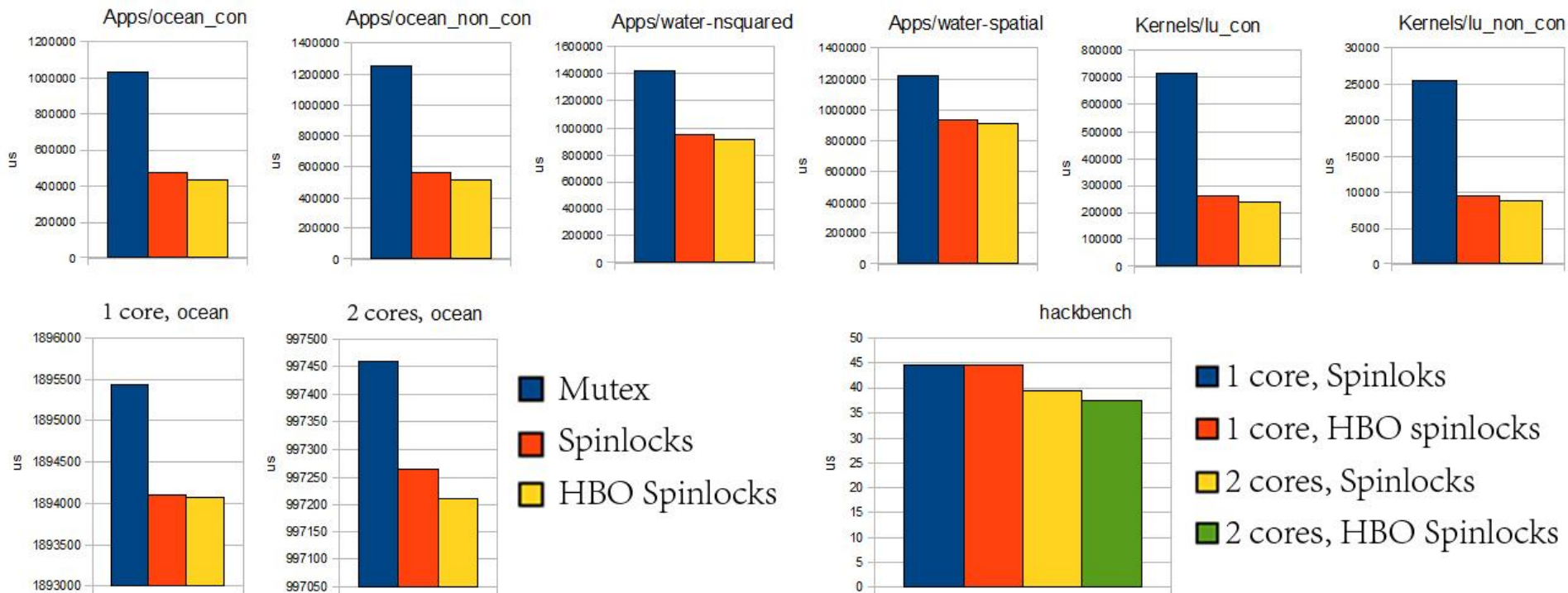
Сравнение методов синхронизации

Методы:

- Mutex
 - Если критическая секция занята, то процесс засыпает.
 - При выходе из критической секции процесс будит ожидающий процесс.
- Spinlocks
 - Если критическая секция занята, то процесс находится в цикле ожидания.
- HVO spinlocks
 - Если критическая секция занята, то процесс находится в цикле ожидания, но с задержками, которые определяются узлами NUMA.
 - Преимущество по сравнению с Spinlocks:
 - Меньшая нагрузка на шину данных.

Сравнение методов синхронизации

Результаты



Вывод: Показана эффективность иерархических спинлоков.

Результаты

- Разработан набор тестов, с помощью которых выявлен ряд аппаратных ошибок, связанных с когерентностью.
- Проведено тестирование различных политик выделения памяти.
- Реализованы НВО spinlocks и показана их эффективность работы по сравнению с mutex и spinlocks.

Спасибо за внимание!