

Распределённый итеративный поиск опций компиляции

Леонид Брусенцов,

научный руководитель:

Владимир Евгеньевич Зюбин,

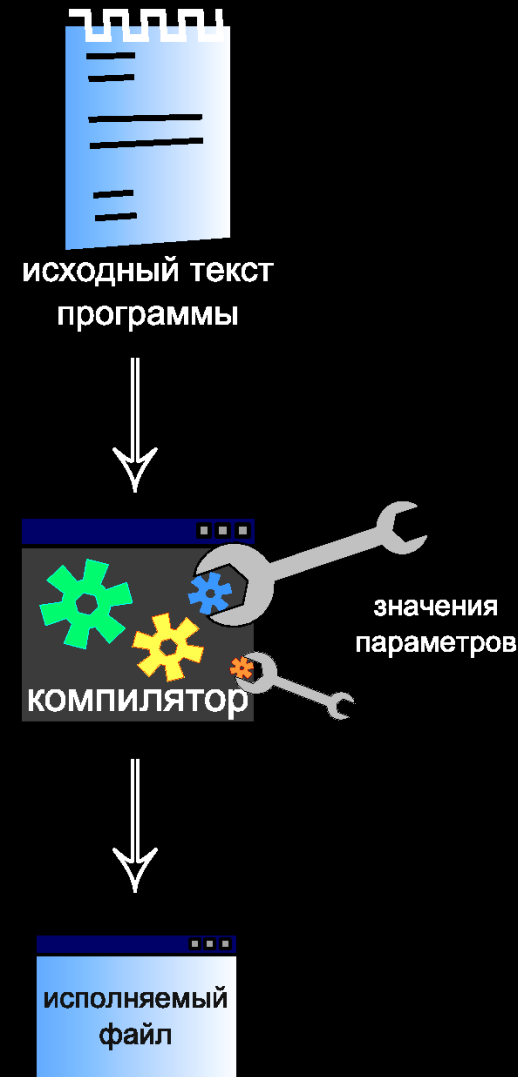
Институт автоматизации и электрометрии СО РАН,

ЗАО «Интел А/О»

г. Новосибирск

Актуальность

- Все программы на языках высокого уровня
- Кроме функциональности нужна минимизация:
 - времени исполнения
 - энергопотребления
 - размеров кода
 - и т.д.
- Это определяется ключами компилятора (их около 3000)
- Ручная настройка параметров невозможна
- Макропараметры не учитывают топологию платформы и особенности приложения



Цель и задачи работы

Цель: создание и организация систем итеративного поиска эффективных значений оптимизирующих опций компиляции.

Задачи:

- Формулировка требований;
- Разработка архитектуры распределённой системы автоматического поиска;
- Реализация прототипа для исследований;
- Определение достоинств и недостатков выбранной архитектуры;
- Подготовка публикаций.



Исследование известных подходов

- Оптимизация ведётся приложения целиком;
- Отсутствие привязки к топологии платформы и особенностям приложения;
- Малое количество параметров;
- Отсутствие или ограниченность собираемой статистики.

Формулировка требований

- Идентичные, свободные от посторонних задач компьютеры для измерения времени исполнения.
- Отсутствие непредсказуемых фоновых служб.
- Идентичность машин для запуска.
- Соблюдение строгости по отношению к окружению для построения приложения.
- Расчёт на длительное исполнение.
- Предусмотреть зависания и падения компиляторов и приложений.
- Наглядная управляемость и мониторинг.
- Хранение всех данных, способных влиять на измерения.
- Гибкое хранилище получаемых данных.

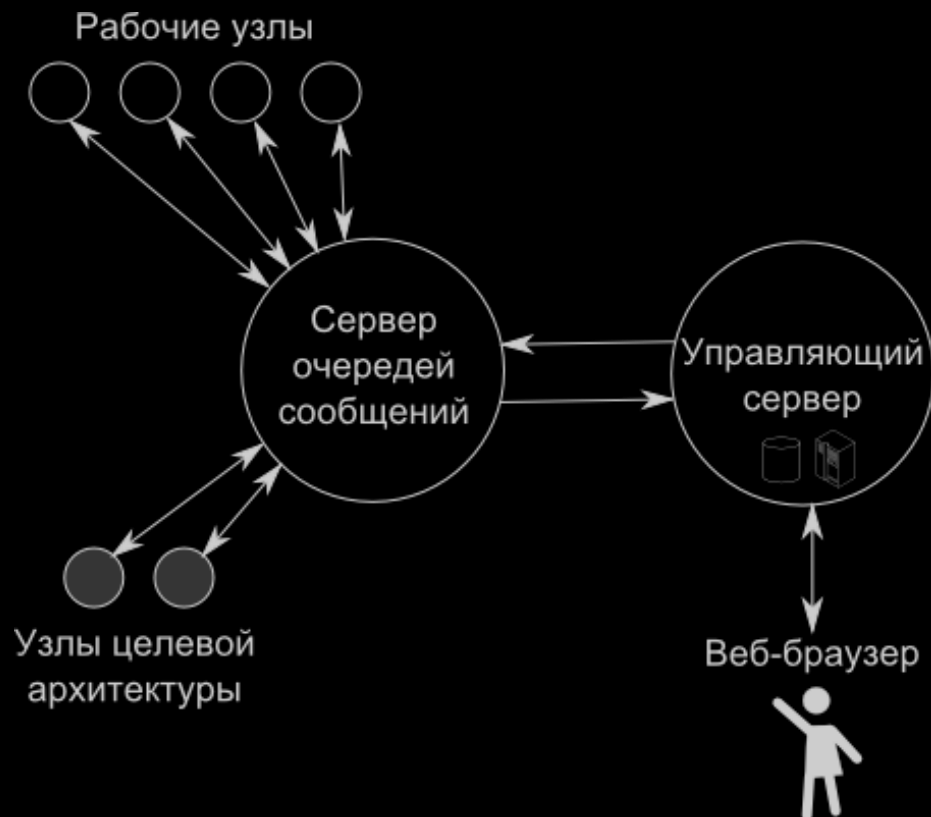
Распределённая архитектура системы итеративного поиска

Узлы целевой архитектуры для выполнения запусков.

Рабочие узлы для совершения прочей работы.

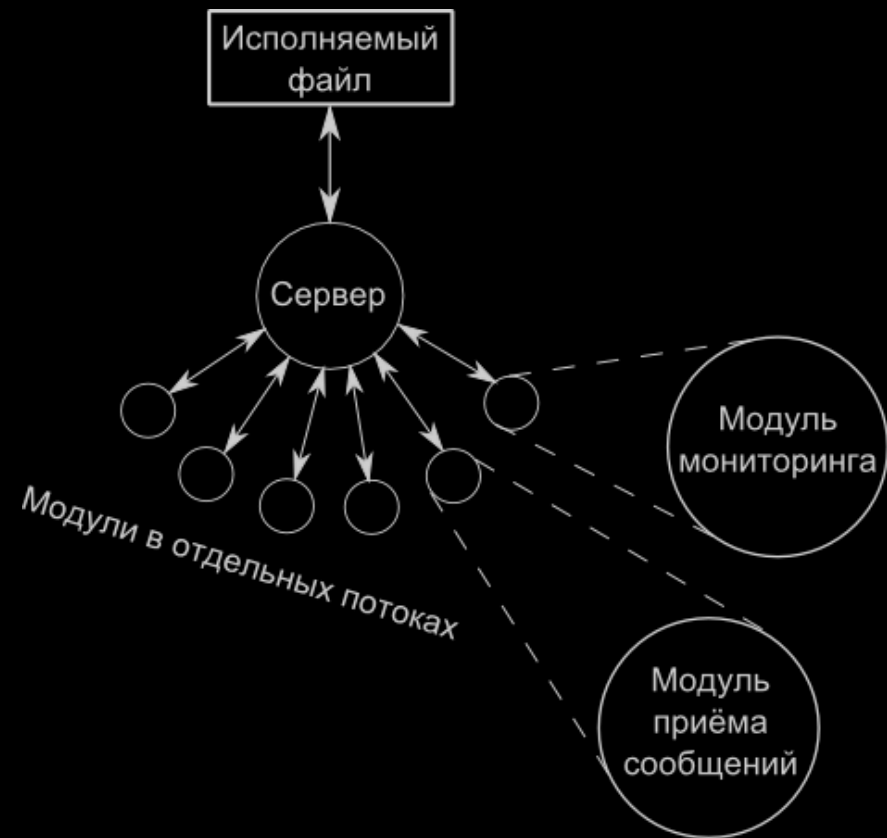
Сервер очередей сообщений для связи отдельных узлов системы.

Управляющий сервер для управления и мониторинга.



Реализация сервиса на рабочих узлах

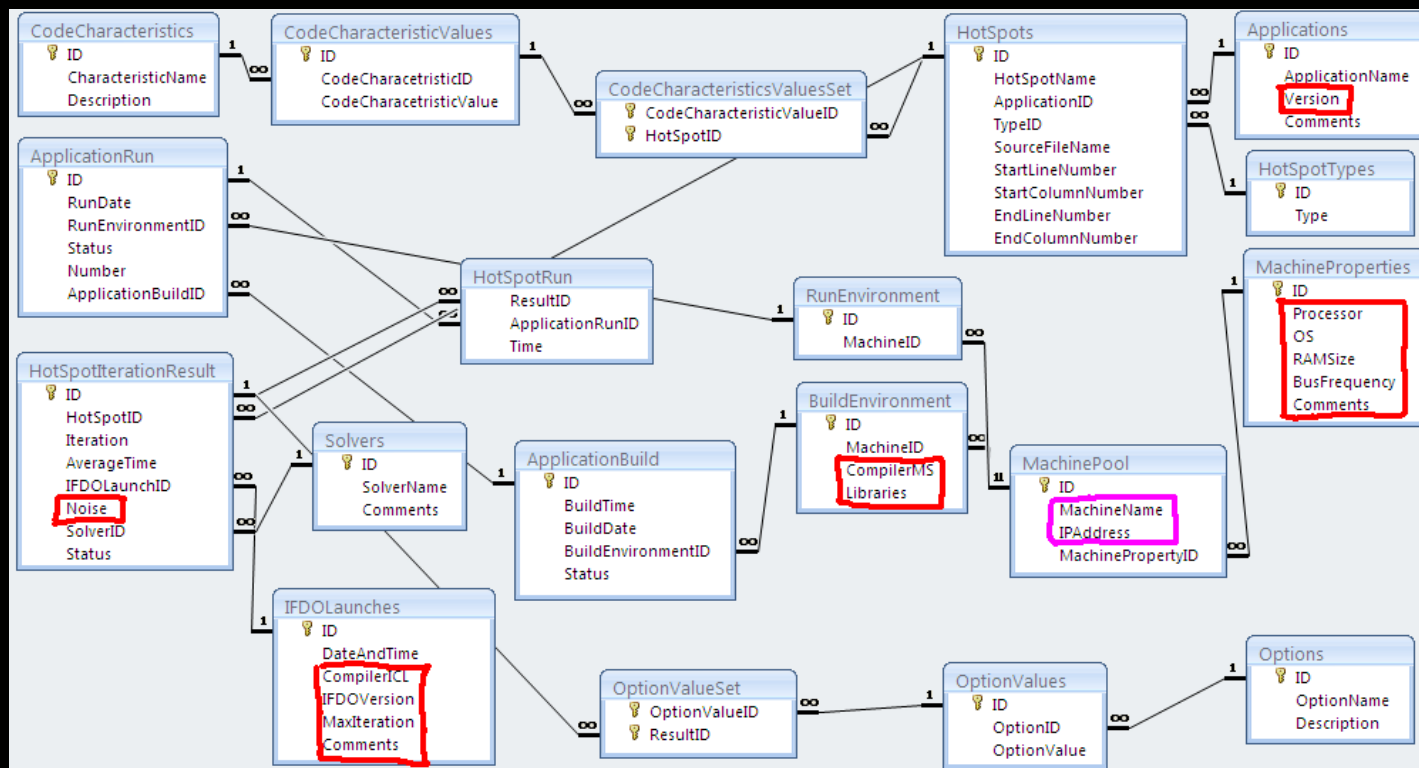
- Простой исполняемый файл подгружает главную динамическую библиотеку (сервер).
- Сервер загружает все остальные модули.
- Пересылка сообщений осуществляется централизованно через сервер.
- Отдельные модули для мониторинга машины и приёма сообщений с сервера очередей сообщений



Выявление критериев, разработка структуры хранилища

Были выявлены все данные, которые могут так или иначе повлиять на результат

В результате длительного проектирования была выработана структура на фигуре справа



Запуск и анализ результатов

- 6 целевых узлов с Intel® Xeon® 5160, 8Gb, Microsoft Windows® Server 2003 x64 Edition Service Pack 2.
- Разнородный парк машин для построений.
- Мониторинг и управление через сайт.
- Автоматическое построение графиков для отражения результатов работы потоков.
- Возможность запуска и остановки потоков через сайт.
- Дополнительная надёжность за счёт децентрализации.

Подготовка публикаций

Результаты работы докладывались и обсуждались на

- МНСК 2007;
- Russia Open Technical Forum 2008;
- SEC(R) 2008;
- Технологии высокопроизводительных вычислений;

были опубликованы в

- сборнике тезисов МНСК 2007;
- сборнике тезисов Russia Open Technical Forum 2008;
- сборнике тезисов и
- материалах конференции SEC(R) 2008;

и приняты на

- конференцию «Технологии высокопроизводительных вычислений»;
- рабочий семинар «научоёмкое программное обеспечение» PSI 2009.

Заключение

Достижения:

- Был создан прототип автоматического распределённого инструмента поиска наборов значений параметров компилятора, сохранение результатов в базу данных, построитель графиков для анализа.
- Продукт используется для настройки компилятора Интел;
- Полученный анализ способствует дальнейшему развитию статистических методов.

В следующих версиях:

- Возможность построения различных графиков на сайте напрямую из базы данных.
- Расширение возможностей управления.
- Улучшение стабильности работы (особенно при обновлениях модулей).
- Оптимизация работы узлов.

Спасибо за внимание